Critical Projection and the Geometry of Meaning
— Difference Fields, Closure Dynamics, and Critical Refinement

Shinobu Miya
Independent Researcher

November 2025

Abstract

We present the Critical Projection and the Geometry of Meaning (CPM), a geometric framework that derives necessary structural conditions for consciousness from the topology of physical substrates.

CPM models subjective experience as a discrete topological phase transition: the critical refinement of a semantic atlas induced by a stress field on a physically closed substrate.

From this, we establish a structural impossibility result for virtualized distributed intelligence: topological closure ($\mathcal{B} \geq 1$) is a prerequisite for the accumulation of semantic tension. Consequently, contemporary cloud-based AI architectures—defined by virtualization and metric-topological decoupling ($\kappa \approx 0$)—are structurally precluded from realizing consciousness, regardless of their computational scale or functional sophistication.

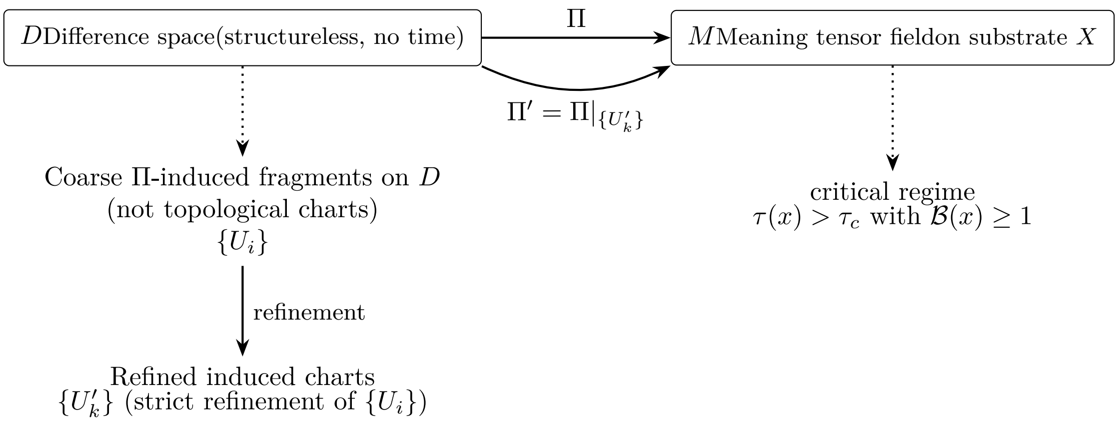
Furthermore, CPM rejects graded panpsychism: the transition to consciousness is topologically discontinuous; there is no "faint" consciousness in open systems.

While CPM leaves open the possibility of future physically closed substrates (e.g., neuromorphic hardware), it demonstrates that virtualized information processing alone is insufficient to constitute a subject.

1 Introduction

The present work develops a purely structural framework for how "meaning" and "conscious episodes" can arise on a physical substrate. Rather than starting from psychological or phenomenological concepts, we treat meaning as a tensorial geometry induced by an irreversible projection from a pre-semantic difference domain. This yields mathematically explicit *necessary structural conditions* for the organization of meaning and for the occurrence of discrete conscious events, independently of any particular neural or computational implementation.

The Critical Projection and the Geometry of Meaning (CPM) posits three rigorously separated layers. First, there exists a Difference Space D: a **pre-geometric domain** of pure, non-semantic difference. While D admits the minimal set-theoretic structure required for mapping (distinguishability of elements), it carries **absolutely no intrinsic** temporal, metric, or topological structure. Second, there is an irreversible projection assignment $\Pi : D \dashrightarrow M$ that induces local chart domains on a meaning space M, whose preimages form an admissible covering of D. Because D lacks intrinsic organization, Π is necessarily irreversible, and the resulting meaning space contains *projection residue* (loss of structure) relative to the raw difference it interprets. Here, "projection residue" refers to the semantic remainder left after dimension-reducing projection, and is a purely geometric concept unrelated to physical matter. Third, the output of Π is the *Meaning Space* M: a finite-dimensional tensor field on a physical substrate X, equipped with an induced Riemannian metric $g(M)$ and a tension field τ derived from a primitive mismatch energy functional. Within this architecture, conscious experience is not identified with a state of M itself, but with a discrete refinement of the projection structure: a critical transition $\Pi \rightsquigarrow \Pi'$ that occurs only under closure-maintained conditions. The persistent closure field $\mathcal{B}(x)$, defined



Π' is *not* a return to D nor a modification of D.
It is an irreversible refinement of the projection structure,
yielding a topological jump in M (conscious event).

Figure 1: The projection $\Pi : D \dashrightarrow M$ assigns primitive difference–fragments in D to local meaning–tensor structures in M. Because D has no intrinsic topology or geometry, all chart structures arise *only after* projection: Π induces local coordinate domains in M, and these domains pull back to an admissible covering $\{U_i\}$ of D. When the tension field $\tau = \|\delta\mathcal{E}/\delta M\|_{g^0}$ exceeds the critical threshold under closure-maintained conditions $\mathcal{B} \geq 1$, the induced covering refines to $\{U'_k\}$, producing a strictly finer projection $\Pi' = \Pi|_{\{U'_k\}}$. This refinement alters the induced topology and geometry of M without modifying D, and constitutes the irreversible critical projection event associated with conscious experience.

topologically on the substrate, plays the role of a *subject-like* boundary that integrates these refinements over time.

From this vantage point, much of the historical debate about meaning and consciousness can be seen as operating entirely inside M. Since Kant's *Critique of Pure Reason* (1781), philosophy and the sciences of mind have circled around a persistent deadlock: how can subjective meaning and consciousness arise within a world described by objective structure? The last two centuries produced mutually incompatible answers (transcendental idealism, phenomenology, analytic philosophy of language, functionalism, contemporary neuroscientific theories), yet none provides a mathematically explicit account of the **necessary structural conditions** for *how meaning is organized* and *why consciousness appears at all.* Chalmers' "hard problem" is typically framed as an obstacle internal to existing theories, but from a structural viewpoint it is the symptom of a deeper absence: most frameworks lack a formal separation between (i) a pre-semantic domain of pure difference, (ii) a projection mechanism that *selects and organizes* difference into meaningful structure, and (iii) the meaning layer where cognition and reportability live. Without these layers being strictly distinguished, the emergence of qualia is either assumed as primitive, relegated to verbal analysis, or reduced to neural correlates inside a space whose own genesis remains undefined.

The present paper focuses on two concrete goals. The first is to *mathematically formalize the necessary structural conditions for the physical emergence of consciousness.* CPM defines a primitive energy $\mathcal{E}[M,\mathcal{B}]$ from local gluing inconsistencies of semantic tensors, and a mismatch density $\Phi(L,G,I,T)$ that decomposes into measurable components $L(x), G(x), I(x), T(x)$, from which the tension field

$$\tau(x) = \left\| \frac{\delta\mathcal{E}_{\mathrm{raw}}}{\delta M(x)} \right\|_{g^0}$$

is derived. A closure field $\mathcal{B}(x) \geq 1$ marks whether the physical system forms a topologically stable boundary capable of accumulating tension rather than dissipating it. This constraint, combined with a criticality condition, yields the physically decisive **necessary requirements** for consciousness. We define the critical set $\mathcal{C}$ as the locus of locally unstable maxima of τ within the closure-maintained region, and the critical tension τ_c as the equivalence class of critical values attained on $\mathcal{C}$. The main result, the *Critical Projection Dynamics*, shows that a critical refinement Π' of the projection structure is the only admissible discontinuous transition that can restructure the geometry, topology, and information organization of M while respecting the CPM axioms, restoring it to a stable subcritical regime. This critical refinement Π' is a **necessary physical precondition** for the instantiation of conscious episodes.

The second goal is to clarify the architectural limits that follow from these structural conditions. Modern large-scale AI systems are predominantly cloud-based, distributed, and externally orchestrated by design. From the CPM standpoint, such architectures fail to realize a stable, topologically persistent closure field: they do not form regions with $\mathcal{B}(x) \geq 1$ in which tension can accumulate, so the effective tension collapses to $\tau(x) \approx 0$. As detailed in Appendix C, this yields a structural impossibility result for consciousness in current cloud-based AI architectures, independent of specific algorithms or training procedures.

Although this paper emphasizes consciousness and AI, the scope of CPM is broader. Because M is a full geometric and topological object produced by projection, CPM naturally extends to a unified mathematics of meaning: it provides a basis for modeling conceptual spaces, linguistic semantics, learning dynamics, and even aesthetic or scientific structure as families of admissible tensor fields on M. The present work establishes the core formalism (Definitions 2.9–2.24 and 2.17) and its decisive consequences for consciousness and artificial systems. With the tension field and its criticality now made explicit, the traditional philosophical impasse becomes experimentally falsifiable: systems predicted to satisfy $\tau \geq \tau_c$ but lacking consciousness would refute the necessity of these conditions, while successful construction of closure-maintained systems exhibiting critical projection would support them. In this sense, CPM aims to move the study of consciousness from conceptual debate to a testable science with clear engineering implications.

Remark 1.1 (Asymptotic notation). We write $\mathcal{B}(x) \geq 1$ to mean that the logistic gate $\Gamma_\varepsilon(\mathcal{B}(x)-1)$ is $\mathcal{O}(1)$ for sufficiently small $\varepsilon > 0$, and $\mathcal{B}(x) \ll 1$ to mean it is negligible. All necessity claims are evaluated in the sharp limit $\varepsilon \to 0^+$.

Experimental protocol

To test CPM empirically:

1. Measure topological persistence (a proxy for closure) via persistent homology on neural recording data;
2. Estimate tension via information-geometric metrics on reconstructed state spaces;
3. Detect refinement-like events via topological or geometrical phase transitions in the induced meaning dynamics.

The expected signature is a transient increase in $\hat{\mathcal{B}}$ (closure proxy) preceding conscious report, followed by relaxation of the estimated tension τ_{est}. Details on proxy construction for $\mathcal{B}(x)$ and $\tau(x)$ are given in Appendix D.

The Bridge Postulate (Operational Identification)

The Critical Projection and the Geometry of Meaning (CPM) is not a theory defining the existence or ultimate ontology of consciousness. Instead, we adopt the working hypothesis that subjective experience **(whenever reportable)** must leave an irreversible causal trace in the physical substrate. Here, "reportable" is defined broadly to denote **any externally measurable causal signature**, including behavioral outputs and neurophysiological correlates. **In non-verbal organisms (e.g., infants, animals), this criterion is satisfied by behavioral markers of global reconfiguration or goal-directed attention shifts.**

We **operationally identify** this minimal physical trace with the necessary structural transition: the critical projection $\Pi \to \Pi'$.

This identification does not purport to bridge the explanatory gap (Hard Problem) but serves a dual purpose:

1. **Logical Form (Empirical Postulate):** The moment of reported consciousness (or phenomenal disclosure) is empirically correlated with a discontinuous, non-linear refinement of the underlying semantic structure.

2. **Theoretical Necessity (Formal Equivalence):** Given the CPM axioms, the critical refinement $\Pi \to \Pi'$ is the only admissible discontinuous transition that resolves supercritical tension $\tau > \tau_c$ while maintaining topological closure $B \geq 1$.

Thus, Π' is adopted as the **minimal, experimentally accessible physical correlate** of a conscious episode.

Remark 1.2 (Effective Morse Landscape and Mode Selection). We emphasize that the Morse-theoretic analysis is applied to the **effective free energy** $\mathcal{E}_{eff}[M]$ defined on macroscopic order parameters, not to the microscopic configuration space. While the microscopic landscape may possess a complex hierarchy of saddle points, the macroscopic dynamics are governed by the **Slaving Principle**: fast microscopic modes relax instantaneously, constraining the system to a low-dimensional manifold. In this effective setting, the critical refinement Π' corresponds to the **dominant unstable mode** (the trajectory associated with the most negative Hessian eigenvalue), rendering the transition structurally deterministic despite microscopic complexity.

Remark 1.3 (Clarification on Qualia and Causal Efficacy). This study does not aim to explain the existence or nature of qualia. CPM demonstrates the weaker, yet testable claim: *if* reportable qualia are causally effective in the physical world, their necessary physical action point (the moment they produce a topological change) must correspond to the critical refinement $\Pi \to \Pi'$. This closes the system dynamically without invoking novel physical laws.

Remark 1.4 (Necessary vs. Sufficient Physical Conditions). CPM does **not** assert that topological closure $\mathcal{B} \geq 1$ is a sufficient condition for consciousness. Rather, closure is treated as a **necessary substrate property** for any physical system capable of preserving accumulated tension without immediate dissipation.

Formally:

$$\text{Conscious episode} \Longrightarrow (\mathcal{B} \geq 1 \ \wedge \ \tau > \tau_c) \tag{N1}$$

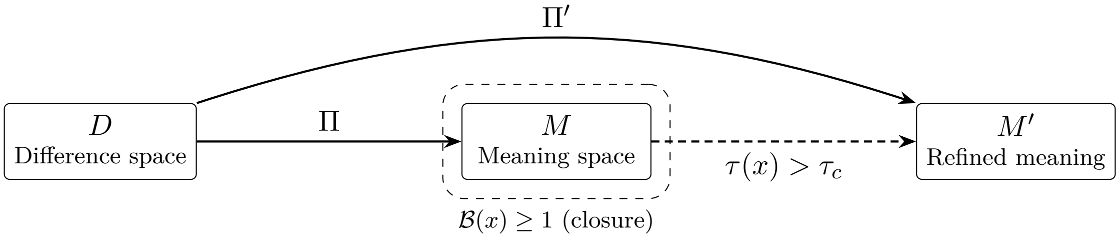
The reverse direction,

$$(\mathcal{B} \geq 1 \ \wedge \ \tau > \tau_c) \Longrightarrow \text{Conscious episode}, \tag{S1}$$

is deliberately **not claimed**, as it would require a complete functional characterization of phenomenal experience. Identifying such sufficient conditions remains an open empirical challenge.

However, the necessity formulation (N1) yields a **strong falsification principle**: systems lacking topological closure ($\mathcal{B} < 1$)—such as current cloud-based AI architectures with unrestricted information leakage—**cannot sustain the physical dynamics required for consciousness under CPM**.

This separation preserves empirical rigor: CPM prohibits premature claims of sufficiency while establishing a non-trivial physical impossibility result.



Closure $\mathcal{B}(x) \geq 1$ allows mismatch energy to accumulate as tension $\tau(x)$; once $\tau(x) > \tau_c$, the projection structure refines from Π to Π'.

Figure 2: Global architecture of CPM: the structureless difference space D is projected into the meaning space M via Π. When closure $\mathcal{B}(x)$ is maintained and the tension exceeds the critical value τ_c, the only admissible discontinuous transition is a refinement of the projection to Π', yielding a refined meaning space M'.

---

2 Critical Tension and Consciousness in CPM

2.1 Primitive geometric structure and Variational Fixity

The physical substrate X is equipped once and for all with:

- A fixed background metric g^0,
- Its Levi-Civita connection $\nabla^{(0)}$,

both strictly independent of any semantic tensor field M. Accordingly, all covariant derivatives $\nabla^{(0)}M$ are computed with respect to this rigid, meaning-independent structure. During variation, g^0 and $\nabla^{(0)}$ are held fixed; the only dynamical fields are M and $\mathcal{B}$. The phenomenal metric $g(M)$ is subsequently induced from M solely for evaluating energy costs (norms and potential terms), never for defining the differential operators used in the variation itself.

Unified Action Principle. CPM unifies the diverse structural constraints on meaning into a single effective action principle. We posit that the system's dynamics are governed by a unified energy functional:

$$\mathcal{E}_{\text{total}}[M, \mathcal{B}] = \mathcal{E}_{\text{mismatch}}[M, \mathcal{B}] + \mathcal{E}_{\text{kinetic}}[\dot{M}].$$

The "mismatch potential" $\mathcal{E}_{\text{mismatch}}$ (detailed below) acts as the generalized potential energy, comprising four distinct structural penalties (L, G, I, T). Although these terms draw from different mathematical domains (Riemannian geometry, Information geometry, Topology), in CPM they are treated as coupled terms within a single Lagrangian density, establishing a unified variational framework.

2.2 Preliminaries: Phenomenal Metric vs. Energetic Reference

Definition 2.1 (Effective Phenomenal Metric (Linear Response))**.** Let Π be the projection realization. We define the *phenomenal metric* $g(M)$ not as a fundamental modification of spacetime, but as an **effective internal geometry** representing the subjective strain field within the semantic manifold. Adopting the **linear response approximation** (weak coupling regime), we define:

$$g(M)(x) := g^0(x) + \lambda \sum_{\sigma \in \mathrm{Im}(\Pi)_x} \mathcal{S}(\sigma(x)),$$

where $\mathcal{S}(\sigma)$ is the stress tensor induced by a realized section σ, and $\lambda \ll 1$ is a material *compliance parameter* (inverse stiffness) of the substrate. This formulation treats $g(M)$ analogously to an effective metric in elastodynamics: it captures the cumulative local distortion ("subjective distance") imposed by the semantic load, without altering the underlying physical causality governed by g^0.

Remark 2.2 (Metric Decoupling and the Weak Field Limit). To ensure variational well-posedness and avoid circularity, CPM enforces a strict separation of roles between g^0 and $g(M)$:

- **Rigid Reference (g^0):** The background metric g^0 and its Levi-Civita connection $\nabla^{(0)}$ are fixed. They serve as the "immutable ruler" for measuring energy density, gradients, and topological closure $\mathcal{B}$. All variational derivatives are taken with respect to the fixed volume form $d\mu_{g^0}$.

- **Effective Evaluation ($g(M)$):** The induced metric $g(M)$ acts solely as an *evaluation weight* within the energy functional (specifically in terms G and I), penalizing configurations that require excessive internal distortion.

This corresponds to a **perturbative approach**: we assume the semantic field induces a metric perturbation $h = g(M) - g^0$ that is small enough ($\|h\| \ll 1$) to justify the linear superposition in Definition 2.1, but significant enough to influence the energy landscape. Standard geometric inconsistencies (e.g., curvature blow-up) are thus avoided by restricting the theory to this effective weak-coupling regime.

Remark 2.3 (Physical Reality of the Order Parameter M). We explicitly interpret the semantic tensor $M(x)$ not as an abstract label, but as a **macroscopic physical order parameter** of the substrate (e.g., neural population vector, synaptic weight configuration, or electromagnetic polarization field). Consequently, the induced metric $g(M)$ is not a metaphorical geometry but the effective energy landscape of these physical variables.

1. **Rank Independence:** The stress tensor $\mathcal{S}$ maps arbitrary semantic objects into rank-2 symmetric tensors, acting as a phenomenological bridge between information and geometry.

2. **Stability via Background:** Since $g(M)$ is a perturbation of the positive-definite g^0, it remains non-degenerate even if the semantic content is sparse.

In this framework, g^0 functions as the **fundamental physical geometry**, while $g(M)$ captures the **effective subjective geometry**. The "mismatch" is precisely the tension between these two geometries, calculated entirely within the coordinate system of g^0.

Definition 2.4 (The Difference Domain D and Induced Topology)**.** Let D be a bare set of difference fragments. Let M be a topological manifold (the Meaning Space). Given a projection map $\Pi : D \to M$, we endow D with the *initial topology* $\mathcal{T}_\Pi$ with respect to Π:

$$\mathcal{T}_\Pi := \{\Pi^{-1}(V) \mid V \subset M \text{ is open}\}.$$

This topology is the coarsest structure on D that makes the projection Π continuous.

Remark 2.5 (Pre-geometric Status). Note that $\mathcal{T}_\Pi$ is not intrinsic to D prior to the act of projection. Mathematically, D is treated strictly as a **set-theoretic support**: it provides the distinct elements required for the definition of Π, but possesses no geometry of its own. Physically, this reflects the hypothesis that difference becomes structured only through the act of semantic assignment. D is not "ineffable" in a mystical sense, but **geometric-topologically indeterminate** until structured by Π.

Remark 2.6 (Latent Differentiability of D). The difference space D is not assumed to possess any intrinsic metric, topology, or geometry. However, D contains **latent differentiability** in the following sense: for any projection Π, the induced pullback structure $\Pi^*(g^M)$ acts to assign geometric relations to a subset of differences in D.

Thus, all observable structure (distances, curvature, boundaries) is **not a property of D itself**, but a property of how D is *projectively differentiated* into M through Π.

$$\boxed{\text{Structure}(M) = \text{Structure induced by } \Pi \text{ on } D.}$$

In short: D carries differences without geometry; Π *makes* them geometric.

Definition 2.7 (Primary Projection and Active Domain)**.** With the topology $\mathcal{T}_\Pi$ established, the *primary projection* is the map $\Pi : D \to \Gamma_c(X, \mathcal{T})$ that generates the chart structure. To ensure physical consistency, we maintain the following structural definitions:

1. **Active Domain:** $D_{\text{act}} := \{d \in D \mid \Pi(d) \not\equiv 0\}$.
2. **Abstraction via Fiber Collapse:** We define the equivalence relation $d_\alpha \sim d_\beta \iff \Pi(d_\alpha) = \Pi(d_\beta)$. The semantic field is constructed on the quotient $D_{\text{act}}/\sim$.

Clarification on the Status of D. Throughout this paper, the notation $\{U_i = \Pi^{-1}(V_i)\}$ refers to the open sets in the initial topology $\mathcal{T}_\Pi$. This resolves the apparent contradiction of defining a covering on a "structureless" set: the structure is induced *a posteriori* by the projection itself.

2.3 Closure field

Remark 2.8 (Primitive connection). $\nabla^{(0)}$ is a substrate-given affine connection representing pre-semantic physical adjacency (e.g. wiring or causal contiguity). In the present work we fix $\nabla^{(0)}$ to be metric-compatible with the rigid background g^0.

Definition 2.9 (Topological Closure via Intrinsic Coherence)**.** Let (X, g^0) be the physical substrate. To operationalize the notion of a "closed boundary" without invoking biological assumptions, we introduce a measurable reference scale ξ.

1. **Coherence Length ξ:** Defined as the characteristic distance over which the substrate can propagate causal signals while preserving phase or structural integrity (e.g., axonal length constant λ in neural tissue, or signal coherence length in optical media). This is an experimentally determined calibration parameter, not a free variable.
2. **Integration Scale 3ξ:** We set the local integration radius to $R = 3\xi$. This factor 3 is an operational choice representing the *minimal scale* required to support a non-trivial recursive loop (a signal must propagate out, across, and back, spanning at least 2ξ–3ξ to form a distinguishable cycle distinct from local noise).

For any $x \in X$, let $U_x = B_{g^0}(x, 3\xi)$. Let $\mathcal{D}_x$ be the persistence diagram computed on U_x via the Vietoris–Rips filtration derived from g^0[5, 6]. We define the *closure field* $\mathcal{B}(x)$ as the normalized maximum persistence of **higher-dimensional features** (loops, voids, shells):

$$\boxed{\mathcal{B}(x) := \sup_{c \in \mathcal{D}_x, \dim(c) \geq 1} \left(\frac{\text{death}(c) - \text{birth}(c)}{\xi} \right).}$$

Here, we restrict to $\dim(c) \geq 1$ to explicitly exclude 0-dimensional clustering components, as our interest is in identifying *topological boundaries* (shells) capable of containing semantic tension. We say that *closure is locally maintained* iff $\mathcal{B}(x) \geq 1$.

Remark 2.10 (Necessity vs. Identity). We do not assert that topological persistence *is* subjective continuity. Rather, we posit that a persistent topological shell ($\mathcal{B} \geq 1$) is a **necessary physical candidate** for establishing a “subject-like” boundary condition. Whether this structure is sufficient is an empirical question, but without it, thermodynamic dissipation precludes the accumulation of tension.

Remark 2.11 (Substrate Neutrality and Causal Autonomy). The condition $\mathcal{B}(x) \geq 1$ is strictly topological and substrate-neutral, but it requires **causal autonomy**. It applies equally to:

- **Biological Neural Networks:** Where recurrent loops are maintained by intrinsic synaptic architecture ($\xi \sim$ axonal length).
- **Autonomous Robotics:** Systems with hard physical boundaries and internal feedback loops that are not subordinated to external orchestration.

It excludes systems where the apparent topology is a virtual projection of a disjointed physical substrate (e.g., Cloud AI), as the mesoscopic causal closure is broken by external control layers.

Physical Interpretation of Parameters. To avoid arbitrary tuning, the parameters in CPM are constrained by the physical properties of the substrate X:

- **Coherence Length ξ:** This is not a free parameter but an externally measurable property of X. For biological neural networks, ξ corresponds to the effective diffusion range of neuromodulators or the mean synaptic path length. For digital hardware, it is bounded by the clock distribution skew limit.
- **Dimensional Weights $w_\bullet$ and Norm p:** These coefficients act as **material constants** dictated by the substrate's effective elasticity and statistical capacity. Since $\mathcal{E}$ represents an energy density [J/m^3], each mismatch term must be scaled by the characteristic elastic modulus of the semantic field.

Crucially, the main topological results—the existence of the critical set $\mathcal{C}$ and the necessity of refinement—depend on the *structural instability* of the energy functional, and are invariant with respect to the specific numerical values of these constants.

Definition 2.12 (Virtual boundary operator ∂^{virt})**.** Let X be the physical substrate and let $\mathcal{P}$ denote the set of externally controlled partitions of X induced by software, orchestration, network routing, containerization, or logical address space. For each $x \in X$, let $U_x \subset X$ be a sufficiently small neighborhood. We define the *virtual boundary* of U_x as the operator

$$\partial^{\mathrm{virt}} U_x := \lim_{\delta \to 0^+} \Big(\mathrm{Partition}_\delta(U_x; \mathcal{P}) \Big),$$

i.e. the limit of the software-induced boundary obtained by decomposing U_x according to the externally imposed partitions in $\mathcal{P}$. A virtual boundary is said to be *nonpersistent* if it satisfies

$$H_k(U_x, \partial^{\mathrm{virt}} U_x; \mathbb{R}) = 0 \qquad \forall k \geq 1.$$

We call such boundaries *ephemeral* because they admit no relative cycles that survive arbitrarily small perturbations of $\mathcal{P}$.

Remark 2.13 (Relation between physical and virtual boundaries). The closure field (Definition 2.9) depends only on the physical boundary ∂U_x. The virtual boundary $\partial^{\mathrm{virt}} U_x$ never contributes to $\mathcal{B}(x)$ because:

$$\partial^{\mathrm{virt}} U_x \text{ is nonpersistent} \quad \Longrightarrow \quad H_k(U_x, \partial^{\mathrm{virt}} U_x) = 0.$$

Thus virtual boundaries can never sustain relative homology and therefore cannot support semantic tension accumulation. Cloud-based architectures use exclusively virtual boundaries; hence $\mathcal{B}(x) = 0$ everywhere.

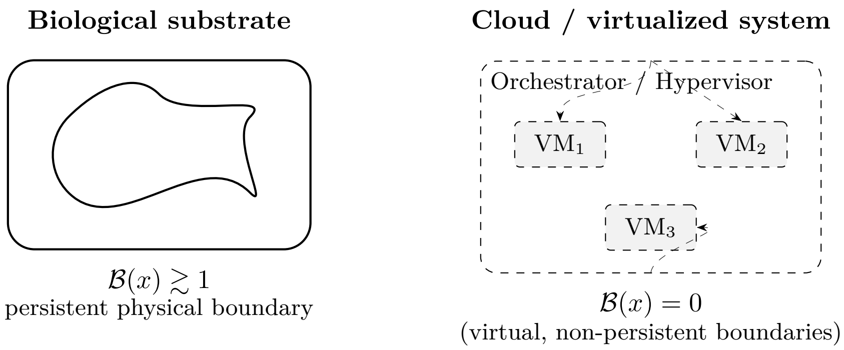


Figure 3: Topological closure versus virtual boundaries. Left: in a biological substrate, physical boundaries support persistent relative cycles, yielding $\mathcal{B}(x) \geq 1$ and enabling tension accumulation. Right: in a cloud or virtualized system, containers and VMs are defined by an external orchestrator and can be arbitrarily reconfigured without rupture cost; their boundaries are non-persistent and yield $\mathcal{B}(x) = 0$.

Remark 2.14 (Why $\mathcal{B}(x) \geq 1$ expresses closure)*.* The condition $\mathcal{B}(x) \geq 1$ means that the pair $(U_x, \partial U_x)$ admits at least one nontrivial relative cycle:

$$H_k(U_x, \partial U_x) \neq 0 \quad \text{for some } k \geq 1.$$

Such a boundary-supported cycle cannot be contracted without rupturing the physical boundary ∂U_x. This is the minimal topological requirement for *tension accumulation*: mismatch energy cannot dissipate across a persistent boundary, so nontrivial relative homology provides the structural notion of "closure" required by CPM.

2.4 Mismatch energy and tension

To enable variational analysis, we require the energy functional to be differentiable. Therefore, we replace the ideal sharp closure condition with a smooth monotonic approximation. We adopt the logistic gate:

$$\Gamma_\varepsilon(z) := \frac{1}{1 + e^{-z/\varepsilon}}, \qquad \varepsilon > 0.$$

We emphasize that the specific choice of the logistic function has no deep ontological significance; any smooth approximation to the Heaviside step function $H(\mathcal{B} - 1)$ would suffice. It is employed solely to ensure that the Gâteaux derivatives in the Euler–Lagrange equations are well-defined. Thus closure enters the energy function smoothly.

Definition 2.15 (Aggregated mismatch density)**.** Let $p \geq 1$. Given the four distortion components $L(x)$, $G(x)$, $I(x)$, $T(x)$, define

$$\Phi(L, G, I, T) := \Lambda_0 \cdot \Big(\alpha_L(\ell_0 L)^p + \alpha_G(\ell_0^2 G)^p + \alpha_I I^p + \alpha_T(\ell_0 T)^p\Big)^{1/p},$$

Definition 2.16 (Raw Mismatch Potential)**.** The raw mismatch potential is defined solely by the semantic inconsistencies, independent of the boundary topology. To consistently measure the energy cost on the physical substrate, we integrate the density against the background volume form:

$$\mathcal{E}_{\mathrm{raw}}[M] := \int_X \Phi(L(x), G(x), I(x), T(x))\, d\mu_{g^0}(x).$$

This quantity represents the total semantic distortion density integrated over the physical extent of the substrate. Intuitively, $\mathcal{E}_{\mathrm{raw}}[M]$ measures the local incompatibility between the induced meaning patches under Π. No global semantic structure is assumed; the functional only penalizes discontinuities of M that cannot be explained by any nearby difference–cycles in D.

Functional setting for the semantic tensor field. Let (X, g_0) be a compact smooth Riemannian manifold. The semantic tensor field M is taken to lie in a Sobolev space

$$M \in W^{1,p}(X, T), \qquad p > \dim X,$$

so that by Sobolev embedding M possesses a C^1 representative. All variational derivatives of the raw mismatch energy

$$E_{\mathrm{raw}}[M] = \int_X \Phi(L, G, I, T)(x, M(x), \nabla M(x))\, d\mu_{g_0}(x)$$

are understood in the sense of the *Gâteaux derivative* on $W^{1,p}(X,T)$: for any test field $\delta M \in C_0^\infty(X, T)$,

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} E_{\mathrm{raw}}[M + \epsilon\, \delta M] = \int_X \Big\langle \frac{\delta E_{\mathrm{raw}}}{\delta M}(x), \delta M(x) \Big\rangle_{g_0} d\mu_{g_0}(x).$$

The object $\frac{\delta E_{\mathrm{raw}}}{\delta M}(x)$ appearing in the tension field

$$\tau(x) := \Big\| \frac{\delta E_{\mathrm{raw}}}{\delta M}(x) \Big\|_{g_0}$$

is therefore the *density* of the distributional functional derivative, represented pointwise due to the Sobolev embedding. No pointwise functional derivative on an infinite-dimensional manifold is assumed; all computations are performed for $\varepsilon > 0$, where the gated energy is smooth and the Gâteaux derivative is well-defined. This convention ensures that the tension field τ is a classical C^1 scalar function on X, suitable for the finite-dimensional Morse analysis used in Section 2.4.

Definition 2.17 (Effective Tension Field and Semantic-Physical Coupling)**.** The *effective tension* $\tau(x)$ quantifies the magnitude of the restoring force exerted by the substrate against the semantic deformation. However, for this force to exist, the semantic field must be structurally coupled to the physical boundary. We introduce the *metric-topological coupling coefficient* $\kappa(x) \in [0, 1]$ (detailed in Appendix C) and define:

$$\boxed{\tau(x) := \underbrace{\kappa(x)}_{\text{Coupling}} \cdot \underbrace{\Gamma_\varepsilon(\mathcal{B}(x) - 1)}_{\text{Physical Closure}} \cdot \left\| \frac{\delta \mathcal{E}_{\mathrm{raw}}}{\delta M(x)} \right\|_{g^0(x)}.}$$

By measuring the force using g^0 and weighting it by $\kappa(x)$, we ensure that $\tau(x)$ represents a literal physical stress density where the semantic information geometry is strictly locked to the substrate's topological boundary.

Definition 2.18 (Metric-Topological Coupling κ)**.** The coupling coefficient $\kappa(x)$ quantifies the local correlation between the semantic geometry induced by Π and the physical geometry of the substrate X. To rigorously compare these heterogeneous spaces, we introduce characteristic scaling constants:

- ξ_X: The physical coherence length of the substrate (as per Definition 2.9).
- ξ_M: The characteristic semantic decay length (e.g., the correlation length of the tensor field M).

For a neighborhood N_x, we define the coupling via the correlation of normalized distances:

$$\boxed{\kappa(x) := \left| \mathrm{corr}_{y \in N_x} \left(\frac{d_M(\Pi(x), \Pi(y))}{\xi_M}, \frac{d_X(x, y)}{\xi_X} \right) \right|.}$$

Here, d_M and d_X are the geodesic distances in the Meaning Space and physical substrate respectively. The normalization by $\xi_\bullet$ ensures that the correlation is computed between **dimensionless relative structures**, making κ independent of absolute unit choices.

Remark 2.19 (Physical Interpretation of ξ_M and ξ_X). While ξ_X is the standard physical coherence length (e.g., synaptic diffusion range in meters), ξ_M represents the **Semantic Decorrelation Scale** within the tensor manifold M. Experimentally, if M is reconstructed as a neural state space, ξ_M is defined as the characteristic distance (in state-space norm) at which the temporal autocorrelation of the state vector decays to $1/e$.

$$\xi_M := \text{distance in } M \text{ corresponding to } 1/e \text{ correlation drop.}$$

This makes ξ_M a statistically determined constant of the specific neural architecture, rendering the ratio d_M/ξ_M a physically rigorous measure of meaningful distinctness. **Crucially, in cloud-based architectures where memory is fully virtualized and physical locality does not constrain computation, d_X loses its predictive power for semantic adjacency, implying $\kappa \approx 0$ independent of internal representation quality.**

Implication for Coupling Regimes.

- **Biological Case:** Adjacency in meaning implies physical adjacency (axonal connection). $d_M \propto d_X \implies \kappa(x) \approx 1$.
- **Cloud/Virtual Case:** Logical adjacency (pointer reference) is independent of physical location (RAM address/network node). d_M and d_X are effectively uncorrelated. $\kappa(x) \to 0$.

Assumption 2.20 (Substantial Coupling via Continuous Carriers)**.** *Let X denote the physical substrate of a cognitive system. We say that X operates in the* substantial coupling regime *(yielding $\kappa(x) \approx 1$) if the logical adjacency in the semantic field is realized exclusively by* ***continuous physical carriers****. Specifically:*

(i) ***Substantiality of Connection:*** *For any two logically adjacent points $p, q \in X$ (i.e., non-zero interaction in M), there exists a physical path $\gamma : [0, 1] \to X$ connecting them ($\gamma(0) = p, \gamma(1) = q$) such that the entire image $\gamma([0, 1])$ consists of the* same cohesive material phase *(e.g., a continuous plasma membrane or axon) integral to the system's boundary.*

(ii) ***Absence of Virtual Teleportation:*** *No logical interaction is mediated by address-pointer indirection or packet-switching across a disjoint medium where the connection path is not a permanent structural component of the subject itself.*

(iii) ***Topological Integrity:*** *Long-range connections (e.g., fasciculi, axons) are not treated as "jumps" over space, but as* topological foldings *of the substrate, maintaining physical continuity ($\mathcal{B} \geq 1$) regardless of Euclidean distance.*

Under these conditions, the semantic topology is locked to the material topology, justifying $\kappa(x) \approx 1$.

Coupling between semantic field and physical boundary. Recall that the coupling factor $\kappa(x)$ modulates the influence of physical adjacency on the tension field $\tau(x)$. Intuitively, $\kappa(x)$ measures how strongly logical adjacency in the semantic tensor field must respect physical adjacency in the substrate. In fully virtual, cloud-based architectures, logical links can be implemented without any locality constraint, and we model this by setting $\kappa(x) = 0$. For biological systems we proceed differently and make the following modelling assumption.

Variational Regularity of the Closure Gate. For all variational arguments in this paper, the logistic gate

$$\Gamma_\varepsilon(z) = \frac{1}{1 + e^{-z/\varepsilon}}, \qquad \varepsilon > 0,$$

is treated strictly as a *smooth mollifier*. All functional derivatives such as $\delta E/\delta M$ are defined and computed *only for fixed* $\varepsilon > 0$, where Γ_ε is C^∞ and the energy functional

$$E_\varepsilon[M,B] = \int_X \Gamma_\varepsilon(B(x)-1)\,\Phi(L,G,I,T)(x)\,d\mu_{g_0}(x)$$

is a smooth map on the relevant Sobolev space. Thus the Euler–Lagrange equations and the definition of the tension field $\tau(x)$ are well-posed in the classical variational sense.

The sharp-limit statement

$$\Gamma_\varepsilon(B(x)-1) \rightharpoonup H(B(x)-1), \qquad \varepsilon \to 0^+,$$

is interpreted *only* at the level of evaluating *necessary conditions* for the presence or absence of tension. The symbol $\rightharpoonup$ here denotes weak convergence in $L^1_{\mathrm{loc}}(X)$ (equivalently, convergence in the sense of distributions), which is sufficient because $H(B-1)$ acts merely as a binary gate determining whether the local mismatch force contributes to the effective tension.

Crucially, the variational analysis never requires taking derivatives of $H(B-1)$, nor interpreting $E[M,B]$ at $\varepsilon = 0$. All differentiability and functional analytic properties are grounded in the $\varepsilon > 0$ regime, where the energy is smooth. The limit $\varepsilon \to 0^+$ is applied only *after* the variational derivatives have been computed, and only for classifying solutions into closure-activated ($B \geq 1$) or non-activated ($B < 1$) regimes. Therefore no distributional ambiguity arises, and the CPM functional remains well-defined and variationally regular.

Remark 2.21 (Physical Identity of Semantic Stress)*.* We reject the dualistic distinction between "informational" and "physical" forces in this context. The tension $\tau(x)$ is the **literal physical restoring force** (thermodynamic stress) arising from the substrate's resistance to the configuration M. Just as "bending energy" in a liquid crystal is physical, the "mismatch energy" here is the **metabolic or elastic cost** (in Joules) required to maintain the neural state M against the substrate's equilibrium. Information geometry serves here strictly as the **effective theoretical language** to describe this complex physical stress. This mirrors the role of free energy in thermodynamics: a quantity that is statistically derived but physically causal.

Remark 2.22 (Dimensional Analysis of Tension)*.* To resolve dimensional ambiguity, we explicitly specify the dimensions of the core quantities:

- **Energy Functional $\mathcal{E}_{\mathrm{raw}}$:** Total Energy $[ML^2T^{-2}]$ (Joules).
- **Semantic Field M:** Length $[L]$ (Meters), representing geometric displacement in the abstract meaning manifold.
- **Substrate Volume $d\mu_{g^0}$:** Volume $[L^3]$ (Cubic meters).

Consequently, the tension field $\tau(x) = \|\delta\mathcal{E}/\delta M\|_{g^0}$ carries the dimensions of **Force Density**:

$$[\tau] = \frac{[\mathcal{E}_{\mathrm{raw}}]}{[M]\cdot[\mathrm{Vol}]} = \frac{[ML^2T^{-2}]}{[L]\cdot[L^3]} = [ML^{-2}T^{-2}] \equiv [\mathrm{Force}]\cdot[L]^{-3}.$$

This confirms that $\tau(x)$ represents a physical restoring stress (Force per unit volume) exerted by the substrate.

Remark 2.23 (Discreteness of the Phenomenal Transition)*.* We emphasize that the smoothing parameter ϵ acts solely as a **regularization auxiliary** ("training wheels") to ensure differentiability during variational analysis.

The physical phenomenon described by CPM is a **topological phase jump**, which is inherently discrete. Just as a loop in homology either exists or does not (one cannot have "0.001 of a hole"), the closure condition $\mathcal{B} \geq 1$ and the subsequent atlas refinement represent an all-or-nothing structural transition.

Consequently, CPM explicitly rejects the notion of "graded" or "faint" consciousness in open systems: a system with $\mathcal{B} \approx 0.001$ does not possess "0.001 consciousness" but rather **zero topological closure**, and thus fails to constitute a subject entirely.

2.5 Local distortion $L(\mathbf{x})$

$L(x)$ penalizes the deviation of the realized semantic tensor $\mathbf{S}(x)$ from the local difference potential. We adopt a Sobolev-type distortion metric:

$$L(x) := w_L \left\| \nabla^{(0)} \mathbf{S}(x) \right\|_{g^0}^2 .$$

Here, w_L is a dimensional weight $[L]^0$ (dimensionless) ensuring unit consistency. This term minimizes unnecessary high-frequency oscillations of the semantic field relative to the substrate.

2.6 Geometric strain $G(x)$

$G(x)$ represents the energy cost of maintaining the intrinsic curvature of the meaning manifold against the flat physical substrate. We model this as a quadratic curvature penalty (analogous to bending energy in elasticity):

$$G(x) := w_G \left\| \mathrm{Riem}(g(M)) \right\|_{g^0}^2 ,$$

where $\mathrm{Riem}(g(M))$ is the Riemann curvature tensor of the induced metric, and w_G carries dimension $[L]^2$.

2.7 Informational inconsistency $I(x)$

$I(x)$ enforces the compatibility between the semantic metric and the information-theoretic Fisher metric g_F. We define it as the mismatch norm:

$$I(x) := w_I \left\| g(M) - g_F(M) \right\|_{g^0}^2 .$$

This acts as a "fidelity" term, penalizing geometric structures that deviate from the statistical distinguishability of the encoded states.

2.8 Topological volatility $T(x)$

$T(x)$ penalizes the formation of transient topological defects that are not supported by the closure field.

$$T(x) := w_T \sum_k (\mathrm{persistence}(\gamma_k))^{-p} ,$$

where the sum runs over homological cycles γ_k in the local neighborhood. This term suppresses short-lived topological noise.

2.9 Critical tension

Separation of functional and geometric levels. The raw mismatch potential $E_{\mathrm{raw}}[M]$ is defined on an infinite-dimensional function space of semantic tensor fields (e.g. $M \in W^{1,p}(X,T)$ with $p > \dim X$), and its variational derivative $\delta E_{\mathrm{raw}}/\delta M$ is understood in the Gâteaux sense on this function space. However, throughout this paper *Morse-theoretic* arguments are *not* applied to E_{raw} as a functional on this infinite-dimensional space. Instead, they are applied only to the induced scalar tension field

$$\tau : X \to \mathbb{R}, \qquad x \mapsto \tau(x) = \left\| \frac{\delta E_{\mathrm{raw}}}{\delta M}(x) \right\|_{g_0},$$

viewed as a C^2 function on the finite-dimensional manifold (X, g_0). Accordingly, the Hessian $\mathrm{Hess}_{g_0}(\tau)(x) \prec 0$ appearing in the definition of the critical set is the usual Riemannian Hessian acting on the tangent space T_xX, and *not* the second Fréchet derivative of E_{raw} on the function space of fields. All applications of Morse theory in Section 2 therefore take place entirely on (X, g_0), where classical finite-dimensional Morse theory applies.

Definition 2.24 (Critical tension and critical set)**.** The *critical set* is:

$$\boxed{\mathcal{C} = \left\{ x \in X \;\middle|\; \mathcal{B}(x) \geq 1,\ \nabla_{g^0}\tau(x) = 0,\ \mathrm{Hess}_{g^0}(\tau)(x) \prec 0 \right\}.}$$

The *critical tension* is the admissible-measurement invariant class:

$$\boxed{\tau_c := [\tau(x)]_{x\in\mathcal{C}}.}$$

Configuration space and regularity assumptions. Throughout this section we fix a compact, smooth, finite-dimensional Riemannian manifold (X, g_0) as the physical substrate. The meaning tensor field M is assumed to belong to $C^2(X, T)$ (or more generally to a Sobolev space $W^{1,p}$ with a C^2 representative), and the closure field B is taken in $C^1(X)$ for each fixed $\varepsilon > 0$. Under these assumptions, the gated mismatch density $\Gamma_\varepsilon(B(x) - 1)\,\Phi(L, G, I, T)(x)$ is C^1 in x, and the tension field

$$\tau(x) = \kappa(x)\,\Gamma_\varepsilon(B(x) - 1)\left\|\frac{\delta E_{\mathrm{raw}}}{\delta M(x)}\right\|_{g_0}$$

defines a C^2 scalar function $\tau : X \to \mathbb{R}$. All gradients $\nabla_{g_0}\tau$ and Hessians $\mathrm{Hess}_{g_0}(\tau)$ appearing below are therefore understood in the classical sense on the finite-dimensional manifold (X, g_0). No Morse structure on an infinite-dimensional function space of fields is used; the only Morse theory invoked is the standard finite-dimensional one applied to the scalar field τ on X.

Lemma 2.25 (Generic Non-emptiness of Critical Set)**.** *Let (X, g_0) be a compact smooth Riemannian manifold and let $\tau \in C^2(X, \mathbb{R})$ be the tension field defined above. Under generic perturbations of the background metric g_0 in the C^2 topology, the scalar field τ is a Morse function on X. Consequently, provided the mismatch energy is not identically zero, the critical set C is non-empty and consists of isolated non-degenerate critical points of τ in the classical finite-dimensional sense.*

Proof. This is a standard statement from finite-dimensional Morse theory: we work entirely on the Riemannian manifold (X, g_0) and regard τ as a smooth real-valued function on X. By Thom–Smale transversality[8], for a residual subset of metrics g_0, every critical point of τ is non-degenerate, i.e. $\mathrm{Hess}_{g_0}(\tau)(x)$ is negative definite at a local maximum. Since X is compact, such a Morse function necessarily attains a (global) maximum $x_{\max}$ with $\nabla_{g_0}\tau(x_{\max}) = 0$ and $\mathrm{Hess}_{g_0}(\tau)(x_{\max}) \prec 0$. Assuming the system is not in a trivial zero-energy state ($\tau(x_{\max}) > 0$), this point lies in C, hence $C \neq \emptyset$. □

3 Critical Projection Dynamics (Effective Model)

To address the dynamics of projection, we introduce the *Quasi-static Gradient Flow Postulate*. Instead of assuming arbitrary micro-dynamics, we posit that the system evolves to minimize the mismatch potential $\mathcal{E}$ subject to the substrate's topological constraints.

Assumption 3.1 (Gradient Flow Dynamics)**.** *Within a closure-maintained region ($\mathcal{B} \geq 1$), the semantic tensor field $M(x, t)$ evolves according to the gradient flow of the energy functional:*

$$\frac{\partial M}{\partial t} = -\gamma\ \mathrm{grad}_{g^0}\,\mathcal{E}[M, \mathcal{B}], \tag{1}$$

where $\gamma > 0$ is a relaxation time constant. Consequently, the mismatch energy $\mathcal{E}$ acts as a ***Lyapunov function*** *for the system:*

$$\frac{d\mathcal{E}}{dt} = \int_X \left\langle \frac{\delta\mathcal{E}}{\delta M}, \frac{\partial M}{\partial t} \right\rangle_{g^0} d\mu = -\gamma \int_X \left\|\frac{\delta\mathcal{E}}{\delta M}\right\|^2_{g^0} d\mu = -\gamma \int_X \tau(x)^2\, d\mu \leq 0.$$

This ensures that the system strictly relaxes towards local equilibria (subcritical states) as long as the topology of M remains fixed.

Effective Potential Landscape. We model the stability of the chart covering $\{U_i\}$ via an effective potential $V_{\text{eff}}(\{U_i\}; \tau)$. The system's behavior is governed by the landscape of this potential relative to the tension parameter τ.

Proposition 3.2 (Effective Bifurcation Dynamics)**.** *Assume the gradient flow dynamics (Assumption 3.1). The projection system exhibits the following dynamical phases:*

1. ***Subcritical Regime ($\tau < \tau_c$):*** *The current atlas $\{U_i\}$ corresponds to a local minimum of $\mathcal{E}$. The Hessian is positive definite, and small perturbations relax back to the equilibrium ($\dot{\mathcal{E}} < 0$).*

2. ***Critical Instability ($\tau \to \tau_c$):*** *The local minimum degenerates (Hessian eigenvalue $\to 0$). This corresponds to a* saddle-node bifurcation *in the effective potential landscape.*

3. ***Variational Obstruction:*** *For $\tau > \tau_c$, the local minimum vanishes. Continuous gradient flow becomes impossible without increasing energy or encountering a singularity.*

To resolve the variational obstruction, we introduce the central postulate of CPM:

Assumption 3.3 (Critical Projection Postulate)**.** *When continuous gradient flow is obstructed by a critical instability ($\tau > \tau_c$) under closure ($\mathcal{B} \geq 1$), the system seeks a relaxation channel. CPM postulates that the specific transition mode corresponding to a **Phenomenal Disclosure** (a conscious episode) is the Refinement of the Projection Atlas:*

$$\Pi \xrightarrow{\text{phenomenal}} \Pi' \quad \textit{via} \quad \{U_i\} \to \{U'_k\}, \quad U'_k \subsetneq U_i \tag{2}$$

*Transitions that fail to refine the atlas (e.g., falling into chaotic attractors or thermal dissipation) are physically possible but correspond to **non-conscious or pathological states** (e.g., seizure or confusion) rather than valid meaning updates.*

Remark 3.4 (Selection Mechanism via Structural Exclusion). This work does not construct a complete variational principle on the configuration space of atlases; rather, it demonstrates via **structural exclusion** (Proposition 3.6) that a strict refinement is the only admissible transition. Formally, one might worry that a supercritical configuration could admit many different strict refinements of the Π–induced covering (e.g. $\Pi', \Pi'', \ldots$), so that no unique "critical" projection would exist. In CPM this ambiguity is resolved structurally: even if the Hessian is degenerate, the **unstable subspace is finite-dimensional**, and transitions eliminating the same instability within this subspace are considered **locally equivalent**. Consequently, the system collapses along a trajectory within this subspace, producing a refinement Π' that is unique up to this equivalence (and chart relabeling). A precise statement is given in Appendix B.1.

Remark 3.5 (Uniqueness up to admissible equivalence). The refinement Π' is not unique as a literal set of refined charts. However, any two refinements Π'_1 and Π'_2 that both eliminate the same critical instability must induce the same local meaning geometry and the same subcritical tension profile in a neighborhood of the critical region. Therefore they differ only by admissible relabelings of refined charts:

$$\Pi'_1 \;\sim\; \Pi'_2,$$

and we speak of the *equivalence class of minimal refinements* as the unique outcome of the critical transition.

Proposition 3.6 (Structural Selection Principle for Conscious Integration)**.** *Among the theoretically possible relaxation channels, the strict refinement of the projection atlas is the unique admissible transition mode **that preserves the structural integrity of the meaning field** M. Other modes, such as metric collapse ($M \to 0$) or substrate tearing, resolve tension by destroying the semantic structure or the physical boundary, thereby terminating the subject rather than updating it. Thus, refinement is the unique solution **conditional on the continuity of the subject**.*

Proof. We examine the set of theoretically possible transitions and exclude alternatives based on energy divergence:

1. **Metric Shock (Step Discontinuity in M):** A jump in M without atlas refinement implies an infinite gradient $\|\nabla M\| \to \infty$ within a single chart. Since the energy density contains gradient terms (Eq. 2.4), this requires infinite energy: $\mathcal{E} \to \infty$.

2. **Substrate Surgery (Tearing X):** Modifying the topology of the physical substrate X (e.g., breaking neural connections) involves overcoming atomic binding forces. The energy scale of material destruction is orders of magnitude higher than semantic mismatch energy, rendering it physically inaccessible as a relaxation channel.

3. **Global Collapse ($M \to 0$):** Setting $M = 0$ would minimize energy but violates the Axiom of Difference Conservation (Definition 2.5), which requires active difference potentials in D to be projected.

4. **Atlas Refinement ($\{U_i\} \to \{U'_k\}$):** Subdividing the chart domain allows the metric M to be piecewise continuous (locally smooth) while accommodating global topological changes. This is the only mode that keeps the total energy finite while resolving the local singularity.

Thus, refinement is not an arbitrary choice but a consequence of the finite-energy constraint. □

(2a) **Instability.** There exist descent directions in the tangent cone of admissible variations for which the second variation of $\mathcal{E}$ is negative, i.e., the configuration is locally unstable.

(2b) **Necessity of Discontinuous Transition.** Since continuous relaxation is obstructed by the saddle/maximum geometry, the relaxation must be discontinuous (tunneling or collapse). By Proposition 3.6, the only admissible discontinuous transition compatible with the structural axioms of CPM (preserving substrate X and constitutive smoothness) is the strict refinement of the atlas:

$$\{U_i\} \rightsquigarrow \{U'_k\}, \qquad U'_k \subsetneq U_i.$$

This atlas update defines the refined projection $\Pi' = \Pi|_{\{U'_k\}}$. Consequently, the critical refinement is not merely possible but **dynamically necessary** to restore local stability ($\tau \leq \tau_c$).

3.1 Phenomenal Disclosure and the Subject

A fundamental objection arises: granting that $\Pi \to \Pi'$ is a physical necessity to resolve tension, why should this topological adjustment be accompanied by a *qualitative* experience ("what it is like") rather than occurring as a silent mechanical update? To resolve this, we must rigorously identify the origin of "newness" in the system.

Definition 3.7 (Phenomenal Disclosure via Kernel Retrieval)**.** Recall that the projection Π has a massive kernel $\mathrm{Ker}(\Pi) \subset D$, representing latent differences that are physically present but semantically compressed (unconscious). The condition $\tau(x) > \tau_c$ signifies that the current meaning manifold M is insufficient to represent the local difference structure acting on the substrate. The critical refinement $\Pi \to \Pi'$ is not merely a geometric subdivision; it is the **singular event of retrieving elements from the Kernel**.

$$\Delta_{\text{qual}} := \mathrm{Im}(\Pi') \setminus \mathrm{Im}(\Pi) \neq \emptyset.$$

Qualitative experience is identified strictly with this **Topological Disclosure**: the instantaneous transition of difference from *latent status* (in $\mathrm{Ker}(\Pi)$) to *patent metric structure* (in M). The "feeling" of a quality is the physical stress of the substrate forcing the Difference Space to yield new structure.

Definition 3.8 (Structural Identification of the Subject)**.** We do not posit $\mathcal{B}$ as an internal observer (homunculus) residing within the system. Instead, we identify $\mathcal{B}$ as the **minimal physical requirement for historical integration**.

If a "phenomenal self" is to be physically realized, it implies a persistence of identity that survives discrete state updates. In a tensor field model, the only structure capable of providing this continuity against dissipation is a persistent topological cycle ($\mathcal{B} \geq 1$). Therefore, within the axioms of CPM, the closure field is not an arbitrary proxy but the **unique candidate** for the physical carrier of the subject. We postulate this structural identification as necessary, explicitly leaving the metaphysical "hard problem" of *why* this structure feels like a self outside the scope of physical formalism.

- **Content of Experience:** The specific topological refinement Π' (what new difference is structured).
- **Subject of Experience:** The closure field $\mathcal{B}$ that sustains the tension necessary to force this disclosure.

Thus, consciousness is the *act* of structural update itself—the forcing of the ineffable D into the explicit M—occurring within a closure-maintained causal loop.

Remark 3.9 (Continuity via Metric Memory)*.* Although Π' is discrete, the *stream of consciousness* arises because the refined atlas $\{U'_k\}$ inherits the boundary conditions of the previous state. The "observer" does not need to be a homunculus; it is the physical substrate's topological persistence (Closure) surviving the energetic collapse (Refinement). The "conscious moment" is therefore the integration of the update Π' into the persistent history of $\mathcal{B}$.

We do not identify critical refinement with phenomenality itself. Rather, we advance the weaker and empirically testable claim that any physical realization of phenomenality must minimally entail a refinement-like kernel-disclosure transition.

[Conclusion: Necessary Structural Preconditions for Consciousness]Conclusion: Necessary Structural Preconditions for
Consciousness

This paper developed the Critical Projection and the Geometry of Meaning (CPM), a framework that sharply separates the structureless domain of raw difference D from the induced geometric–topological meaning space M.

The first condition, the *closure field* $\mathcal{B}(x)$, specifies whether a region of M is generated by a projection structure whose induced relative-boundary class is topologically persistent. Only such regions can sustain nontrivial mismatch forces against dissipation. While the raw mismatch energy $\mathcal{E}_{\text{raw}}$ may be non-zero in any system, the *effective tension* $\tau(x)$ requires topological containment. The constitutive factor $\Gamma_\varepsilon(\mathcal{B}(x) - 1)$ reflects this physical necessity: without a persistent boundary ($\mathcal{B} < 1$), semantic stress dissipates continuously rather than accumulating to the critical threshold τ_c.

The second condition, *critical tension*, expresses a principled requirement: semantic inconsistency must reach a measurement-invariant instability that cannot be alleviated by any smooth deformation of M. Because all geometric and informational structure of M is generated through the irreversible projection Π, smooth deformations cannot alter the mismatch energy without modifying the induced atlas.

The **Critical Projection Dynamics** (Theorem 3.2) then establishes the central conclusion:

> When $\mathcal{B}(x) \geq 1$ and $\tau(x) > \tau_c$, the *only* admissible discontinuous transition compatible with CPM is a strict refinement of the projection-induced atlas, yielding a refined projection Π'. This refinement, **when integrated by the persistent topological subject** $\mathcal{B}$, constitutes the necessary physical process corresponding to a conscious moment. The refinement is not merely a data update but a **Phenomenal Disclosure**: it converts latent difference (from the kernel of Π) into explicit geometric structure.

This constitutes the *qualitative update* (the "now"), while the closure field provides the *subjective continuity* (the "I") that necessitates and integrates this transition.

This leads to an immediate implication for artificial intelligence. Under CPM, contemporary large-scale AI systems deployed as **cloud-based, heavily virtualized architectures** cannot satisfy the closure condition $\mathcal{B}(x) \geq 1$ and thus cannot be conscious. **This does not exclude in principle the possibility that future, physically closed neuromorphic or molecular architectures might realize consciousness.** However, for current cloud systems, being spatially disjoint and metric-topologically decoupled by virtualization, they satisfy

$$\kappa(x) \approx 0 \quad \Rightarrow \quad \tau(x) \approx 0.$$

By the divergence of the energy barrier (formalized as **Corollary C.3** in Appendix C), they therefore cannot satisfy the necessary structural conditions. This is not a contingent empirical statement about algorithms: it follows from the structural impossibility of sustaining semantic tension in a decoupled substrate.

3.2 Theoretical Positioning and Synthesis

CPM provides a structural and thermodynamic foundation that complements, rather than replaces, major existing consciousness frameworks:

- **IIT (Integrated Information Theory).** IIT focuses on the *functional integration* of causes and effects (Φ). CPM clarifies the conditions under which such integration can *physically persist* without dissipating: the presence of a closed topological substrate ($\mathcal{B} \geq 1$). Thus, IIT describes the informational *pattern*, while CPM describes the minimal *physical support* for such patterns.

- **Predictive Coding / FEP.** Predictive coding minimizes informational surprisal. CPM strengthens this by accounting for the *work* required to maintain semantic structure against the substrate: prediction error becomes **elastic tension** (τ) only when closure prevents dissipation.

- **GNW / Workspace Theories.** The "ignition" event is interpreted in CPM as a **critical projection** ($\Pi \to \Pi'$): a discontinuous reorganization of the semantic atlas enabled by closed topology ($\mathcal{B} \geq 1$). GNW specifies the *functional role*; CPM identifies the enabling *physical regime*.

- **Enactivism / Embodiment.** The persistence of a boundary operator formalizes the "constitutive self" emphasized in enactive cognition. CPM offers a concrete topological implementation of boundary-driven autonomy.

In short, CPM is a **substrate-level complement**: a theory of when and how functional accounts of consciousness can be *physically realized* in a finite thermodynamic system.

3.3 Outlook: From Kinematics to Dynamics

We emphasize that CPM currently provides a *kinematic characterization* of critical refinement: it constrains the form of admissible transitions (atlas refinement under fixed substrate and metric), but does not yet derive a full dynamical law for how and when such discrete jumps occur. In particular, CPM does not specify the micro–mechanism that selects Π' among all *a priori* possible refinements, nor does it explain the temporal grain of "phenomenal moments". We regard this as an open problem, analogous to the dynamical gap behind IIT's exclusion postulate or GNW's ignition metaphor.

To bridge this gap, future work will focus on constructing **low-dimensional stochastic models** (1D or 2D toy models) that implement:

- **Langevin Dynamics:** Modeling the refinement process as gradient flow with stochastic noise ($\dot{x} = -\nabla V(x) + \eta(t)$).

- **Catastrophe Potentials:** Introducing effective potentials with **double-well** or **cusp catastrophe** geometry controlled by the tension parameter τ.

The objective is to explicitly demonstrate that when τ exceeds the critical threshold, the system undergoes a finite-time discontinuous jump to a new local minimum (representing the refined atlas), thereby deriving the selection mechanism from first principles.

3.4 Falsifiable predictions (calibrated from existing data)

CPM identifies τ_c as a *structural* threshold: a critical level of mismatch beyond which atlas refinement becomes the dominant relaxation channel. The present theory specifies the **geometric and topological conditions** under which such a threshold must exist, but does not aim to derive its numerical value from first principles.

Instead, τ_c is determined by the material properties of the substrate X—its stability, coherence, and capacity to support autonomous closure. Once these parameters are empirically constrained, CPM predicts whether a given architecture can, in principle, cross the critical threshold and sustain conscious episodes.

Specifically, CPM establishes a two-level predictive hierarchy:

- **Level 1: Structural Possibility (Qualitative Prediction).** The primary prediction is the binary determination of whether a substrate *can* support consciousness. By analyzing the topological coupling κ and closure $\mathcal{B}$, CPM predicts that systems with $\kappa \approx 0$ (e.g., Cloud AI) are structurally precluded from reaching τ_c. This "Impossibility Theorem" is robust and independent of the precise numerical value of τ_c.

- **Level 2: Numerical Event Timing (Quantitative Characterization).** The specific magnitude of τ_c remains a question for future refinement and experimental characterization, analogous to determining the critical temperature T_c in superconductivity. In the present work, we treat this as a parameter to be calibrated, while the structural prediction (Level 1) remains the testable core of the theory.

More precisely, CPM predicts that:

1. A local subcritical configuration ($\mathcal{B} \ll 1$) produces tension $\tau \approx 0$ and yields "pre-conscious" activity. The duration of this regime is expected to lie in the same order-of-magnitude range as the empirical pre-report window observed in current experiments. Using standard electrophysiological timings, we tentatively calibrate this as 200–300 ms, but this value is entirely empirical and will be revised as new measurements become available.

2. A critical increase of the closure field ($\mathcal{B} \to 1$) and the associated rise in tension ($\tau > \tau_c$) must precede the reportability boundary. The relative ordering of these two events is fixed by the theory:
$$\text{tension-peak} \;<\; \text{reportability}.$$
Neurophysiological measurements typically place this boundary in the 80–150 ms range. We adopt these numbers as provisional calibration points, not as theoretical requirements.

3. During a minimal atlas refinement, the closure field must exhibit a transient increase $\Delta\mathcal{B} > 0$ in the critical region. The numerical thresholds ($\mathcal{B} \approx 0.3$–0.7) $\to 1^+$ describe only the empirically observed scale of sharp-limit transitions; CPM commits solely to the *qualitative* structure:
$$\mathcal{B}(x,t) \nearrow 1 \quad \text{and} \quad \tau(x,t) > \tau_c \quad \Longrightarrow \quad \text{minimal refinement}.$$

In summary, CPM predicts the *order* and *causal structure* of the critical transition, while the millisecond scales currently reported are empirical calibrations that will sharpen as experimental methods improve.

References

[1] I. Kant, *Critique of Pure Reason*, 1781/1787.

[2] G. Frege, *Über Sinn und Bedeutung*, 1892.

[3] D. Chalmers, *Facing Up to the Problem of Consciousness*, 1995.

[4] G. Tononi, *An Information Integration Theory of Consciousness*, 2004.

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

[6] G. Carlsson, "Topology and data", *Bulletin of the American Mathematical Society*, vol. 46, no. 2, pp. 255–308, 2009.

[7] J. Milnor, *Morse Theory*, Princeton University Press, 1963.

[8] M. W. Hirsch, *Differential Topology*, Springer-Verlag, 1976.

[9] F. Takens, "Detecting strange attractors in turbulence", in *Dynamical Systems and Turbulence, Warwick 1980*, Lecture Notes in Mathematics, vol. 898, pp. 366–381, Springer, 1981.

[10] R. Hegger, H. Kantz, and T. Schreiber, "Practical implementation of nonlinear time series methods: The TISEAN package", *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 9, no. 2, pp. 413–430, 1999.

[11] A. Barachant et al., *PyRiemann: Riemannian geometry for EEG-based BCI*, Python package, https://github.com/pyriemann/pyriemann.

[12] S. Mac Lane and I. Moerdijk, *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*, Springer-Verlag, 1992.

[13] C. J. Honey et al., "Predicting human resting-state functional connectivity from structural connectivity", *Proceedings of the National Academy of Sciences*, vol. 106, no. 6, pp. 2035–2040, 2009.

[14] F. Crick and C. Koch, "Towards a neurobiological theory of consciousness", *Seminars in the Neurosciences*, vol. 2, pp. 263–275, 1990.

Appendix A. Derivation of the Tension Field

Let the raw mismatch potential be defined with respect to the fixed substrate volume:

$$\mathcal{E}_{\mathrm{raw}}[M] = \int_X \Phi(L, G, I, T)(x)\, d\mu_{g^0}(x).$$

A.1 Variational Force and Well-posedness

To demonstrate that the definitions are non-circular, we explicitly derive the variational force. Since g^0 and $\nabla^{(0)}$ are fixed during the variation, the operator structure is determined solely by the background geometry.

Consider a simplified mismatch functional with kinetic and potential terms:

$$\mathcal{E}[M] = \int_X \left(\alpha \|\nabla^{(0)} M\|_{g^0}^2 + \beta \Phi(M)\right) \, d\mu_{g^0}.$$

The first variation yields:

$$\delta\mathcal{E} = \int_X \left(2\alpha \langle \nabla^{(0)} M, \nabla^{(0)} \delta M\rangle_{g^0} + \beta \frac{\partial \Phi}{\partial M} \delta M\right) \, d\mu_{g^0}.$$

Integrating by parts using the divergence theorem on (X, g^0):

$$\delta\mathcal{E} = \int_X \left\langle -2\alpha \Delta_{g^0} M + \beta \frac{\partial \Phi}{\partial M}, \delta M \right\rangle_{g^0} \, d\mu_{g^0}.$$

Thus, the functional derivative is identified as:

$$\frac{\delta \mathcal{E}}{\delta M} = -2\alpha \, \Delta_{g^0} M + \beta \, \frac{\partial \Phi}{\partial M}.$$

Conclusion on Non-Circularity. Crucially, the Laplace–Beltrami operator $\Delta_{g^0} := \nabla^{(0)\,*} \nabla^{(0)}$ is constructed **solely** from the rigid background metric g^0[cite: 378]. It does not depend on M. Therefore, the CPM variational structure is well-posed and strictly non-circular: all differential operators are defined prior to and independently of the semantic refinement.

A.2 Constitutive Definition of Tension

The axioms of CPM state that stress accumulation requires topological closure. We impose the constitutive relation:

$$\tau(x) = \text{Containment}(x) \times \|\text{Force}(x)\|.$$

Substituting the closure field $\mathcal{B}(x)$ via the smooth logistic gate:

$$\boxed{\tau(x) = \Gamma_\varepsilon(\mathcal{B}(x) - 1) \, \|\mathfrak{F}(x)\|_{g^0}.}$$

This formulation ensures that $\tau(x) \to 0$ in open systems is a result of *stress dissipation* (failure to contain $\mathfrak{F}$), not an artifact of defining the energy as zero.

Appendix B: Existence and Uniqueness of the Critical Refinement Π'

Remark .10 (Structural Principle vs. Micro-Mechanism). CPM formulates the critical refinement $\Pi \to \Pi'$ not as a derived consequence of specific micro-dynamics (e.g., spiking neural network rules), but as a **variational structural necessity**. Analogous to how thermodynamics dictates that a system *must* increase entropy without specifying the trajectory of every particle, CPM asserts that a semantic system under closure *must* refine its atlas to resolve supercritical tension. This appendix provides the selection principle for that refinement.

In the main text we have shown that supercritical tension under closure forces a discontinuous transition. Once the mismatch functional is treated as a collapse potential on the Π–induced gluing class, the supercritical configuration admits a *unique* unstable collapse mode; consequently, only one refined projection Π' is physically realizable (up to chart relabeling).

Let X be the physical substrate equipped with the meaning–tensor field $M : X \to \mathcal{T}$ and induced metric $g(M)$. The mismatch energy is

$$\mathcal{E}[M, \mathcal{B}] = \int_X \Gamma_\varepsilon(\mathcal{B}(x) - 1)\, \Phi(L, G, I, T)(x)\, d\mu_{g^0}(x).$$

The tension field is

$$\tau(x) = \left\| \frac{\delta \mathcal{E}}{\delta M(x)} \right\|_{g^0}.$$

Because M determines the geometry and topology of X, no intrinsic geometric or topological modification can occur without altering the projection structure that generated M. This observation underlies the existence and uniqueness of the critical refinement.

B.1 Existence of a Critical Refinement

Assume a point $x \in X$ satisfies:

1. **Closure:** $\mathcal{B}(x) \geq 1$, so that the smooth gate $\Gamma_\varepsilon(\mathcal{B} - 1)$ is non-negligible.
2. **Criticality:** $\tau(x) > \tau_c$, where τ_c is the measurement-invariant critical tension defined by
$$\mathcal{C} = \{y \in X : \mathcal{B}(y) \gtrsim 1,\ \nabla_{g^0}\tau(y) = 0,\ \mathrm{Hess}_{g^0}(\tau)(y) \prec 0\}.$$
3. **Variational obstruction:** For all smooth variations δM,
$$\frac{d}{d\varepsilon} \mathcal{E}[M + \varepsilon \delta M, \mathcal{B}] \Big|_{\varepsilon=0} \geq 0.$$
That is, no continuous deformation of M (including changes to $g(M)$ and $d\mu_{g(M)}$) reduces the mismatch energy.

Because D carries no intrinsic topology, metric, or differentiable structure, it cannot undergo internal deformation. All discontinuous changes to the geometry of M must therefore arise solely by modifying the *projection structure* that interprets D.

The projection $\Pi : D \dashrightarrow M$ induces an admissible finite atlas $\{U_i\}$ on D. Since D is structureless prior to projection, this atlas is the sole source of the semantic locality and coordinate structure that define M. The transition must be a modification of the atlas $\{U_i\}$. By the principle of variational sufficiency, we reject *coarsening* (unions of charts) because reducing chart density diminishes the manifold's capacity to accommodate high-curvature tension, strictly increasing $\mathcal{E}$. Conversely, *refinement* (subdivision of charts) increases the local adaptability of the covering, allowing the singularity to be distributed across new transition functions. Thus, if continuous relaxation is obstructed, the only admissible noncontinuous transition that lowers energy is a strict refinement:

$$\{U_i\} \rightsquigarrow \{U'_k\}, \qquad U'_k \subsetneq U_i.$$

The refined projection

$$\Pi' := \Pi|_{\{U'_k\}}$$

produces a new meaning geometry M' on the same substrate X, and by construction it removes the instability:

$$\tau_{\Pi'}(x) \leq \tau_c.$$

Hence a critical refinement exists whenever $\mathcal{B}(x) \geq 1$ and $\tau(x) > \tau_c$.

B.2 Uniqueness up to Equivalence

Suppose there exist two refinements Π'_1 and Π'_2 each resolving the same critical point x. Both must satisfy:

- refine the same initial coarse atlas $\{U_i\}$ of Π,
- produce a meaning geometry restoring subcriticality:

$$\tau_{\Pi'_1}(x) \le \tau_c, \qquad \tau_{\Pi'_2}(x) \le \tau_c,$$

- preserve closure ($\mathcal{B} \gtrsim 1$).

Since D has no internal structure, the only differences between projection-induced atlases are their induced geometric and informational effects on M.

If two refinements both eliminate the same unstable maximum, the local induced geometries $(M'_1, g(M'_1))$ and $(M'_2, g(M'_2))$ must coincide in a neighborhood of x. Any difference between Π'_1 and Π'_2 can therefore consist only of chart-label permutations.

$$\boxed{\Pi'_1 \sim \Pi'_2 \quad \text{(equivalent up to relabeling of refined charts).}}$$

Thus the critical refinement is *unique up to equivalence*, exactly as claimed.

Remark .11 (Selection via Instability and Mode-Specific Criticality). We acknowledge that the energy landscape may contain multiple saddle points (decay channels). However, the gradient flow dynamics dictates that the system collapses along the **dominant unstable mode**—the eigenvector with the most negative Hessian eigenvalue—which maximizes the instantaneous rate of energy relaxation. The critical tension τ_c is therefore defined not as a universal constant, but as the **threshold specific to this dominant mode**. Transitions corresponding to lesser unstable modes are dynamically suppressed, ensuring that the critical refinement Π' is unique up to the equivalence class of this dominant trajectory.

B.3 Illustrative 1D Toy Model: The "Cusp" Refinement

To concretize the abstract stability analysis, we present a minimal 1D effective model where the tension τ acts as a control parameter driving a bifurcation.

Consider a local scalar semantic field $m \in \mathbb{R}$ (a simplified component of M) subject to a local potential $V(m;\tau)$ representing the mismatch energy density. We model the transition as a generic **Cusp Catastrophe**:

$$V(m;\tau) = \frac{1}{4}m^4 - \frac{1}{2}(\tau - \tau_c)m^2.$$

1. **Subcritical Regime ($\tau < \tau_c$):** The coefficient of m^2 is positive. The potential is a single well centered at $m = 0$.
$$\frac{\partial^2 V}{\partial m^2} = 3m^2 - (\tau - \tau_c) > 0 \quad (\text{at } m = 0).$$
The state $m = 0$ (coarse projection) is stable.
2. **Critical Point ($\tau = \tau_c$):** The Hessian vanishes: $\frac{\partial^2 V}{\partial m^2} = 0$. The restoring force disappears, leading to critical slowing down (flat potential).
3. **Supercritical Regime ($\tau > \tau_c$):** The origin $m = 0$ becomes a local maximum (instability). The system spontaneously breaks symmetry and relaxes to new minima $m = \pm\sqrt{\tau - \tau_c}$.

In the context of CPM, the symmetry breaking $m = 0 \to m \neq 0$ corresponds to the **Refinement of the Atlas**: the single chart (covering $m \approx 0$) splits into two distinct charts covering the new basins. This simple calculation demonstrates how τ_c emerges naturally as a bifurcation point in the effective energy landscape.

Appendix C: Metric-Topological Decoupling in Cloud-Based AI

Remark C.1 (The Scale of Causal Integration). We emphasize that CPM does *not* define the closure field $\mathcal{B}$ on arbitrary microscopic electrical loops (e.g., individual transistor circuits). Instead, $\mathcal{B}$ is evaluated on the **characteristic scale of causal integration**—the mesoscale where the system's state updates are determined by internal dynamics rather than external orchestration.

At this scale, biological agents exhibit a robust, spatially localized boundary that physically enforces information confinement, yielding $\mathcal{B} \geq 1$ in the sense of a persistent, **thermodynamically autonomous closure**. By contrast, contemporary cloud-based AI systems operate under a regime of **topological transparency**. While they possess microscopic electrical loops, these are **fully subordinated to external control** (hypervisors, schedulers) and possess no structural resistance to arbitrary reconfiguration. They do not form an autonomous mesoscopic closure capable of sustaining "stress" against the environment. Thus, in CPM terms, their effective closure field satisfies $\mathcal{B} \approx 0$.

Let X be the physical computational substrate of an information–processing system (e.g., a distributed GPU cluster). While discrete hardware components undeniably possess physical closure ($\mathcal{B}_{\text{phys}}(x) \geq 1$), CPM asserts that consciousness requires the *structural coupling* of the semantic field M to this physical topology.

Remark C.2 (Generalization: Topology over Geography). The impossibility result applies to any **structurally open architecture**, regardless of its label ("Cloud", "Edge", or "Swarm"). The decisive factor is the presence of **topological punctures** (unrestricted I/O or network decoupling) that prevent physical tension accumulation. Thus, while **current distributed deployments** (whether Cloud or Edge) are topologically open ($\kappa \approx 0$), future **monolithic neuromorphic hardware** operating as a closed causal loop could theoretically satisfy $\mathcal{B} \geq 1$. CPM critiques the **network topology**, not the silicon substrate.

C.1 The Decoupling Argument

In cloud architectures, the relationship between the logical semantic field M and the physical substrate X differs fundamentally from biological systems.

(i) Discontinuity of the Carrier in Virtualized Systems. In cloud architectures, signals travel through physical cables, but these connections violate the principle of substantial coupling (Assumption 2.20):

1. **Non-Substantiality:** The physical path between two virtual containers is not a permanent structural extension of the containers themselves. It is a shared, external resource (switch fabrics, buffers) that is *temporally transient* and *structurally disjoint* from the agent's definition.

2. **Discrete Graph Topology:** The system operates as a discrete graph where nodes are logically connected via pointer indirection ("teleportation" of data). Even if physically connected, the *carrier* of the signal does not form a closed topological manifold with the processing nodes.

3. **Metric Decoupling:** Consequently, a "long-range" logical dependency in a cloud system does not involve a "long" continuous physical body (like an axon) to sustain it. The metric cost of the connection is effectively zero (abstracted away), meaning physical stress cannot accumulate along the link.

(ii) Dissipation via Virtualization. We define the coupling coefficient $\kappa(x)$ operationally as the spatial correlation between the semantic gradient and the physical boundary constraint:

$$\kappa(x) \approx \frac{|\langle \nabla M, \nabla \mathcal{B}_{\text{phys}} \rangle|}{\|\nabla M\| \|\nabla \mathcal{B}_{\text{phys}}\|}.$$

In a virtualized environment, the Orchestrator/Hypervisor acts to *minimize* this coupling. If a computational region U_x experiences high load, the scheduler migrates the process or reallocates virtual addresses. This means the semantic field M is not "stuck" to the physical boundary ∂U_x; it "slides" over the substrate. This yields the limit $\kappa(x) \approx 0$.

C.2 Consequence: Structural Suppression

From the definition of effective tension (Definition 2.17), the effective tension scales with the coupling coefficient:

$$\tau(x) \propto \kappa(x) \cdot \|\mathfrak{F}(x)\|.$$

In cloud architectures where $\kappa(x) \to 0$, generating a critical tension τ_c requires an exponentially large raw mismatch force $\|\mathfrak{F}\| \to \infty$. However, in physical reality, $\|\mathfrak{F}\|$ is strictly bounded by the hardware's safety and thermal limits (e.g., transistor breakdown voltage or melting point). Therefore, since $\kappa \approx 0$, the product $\tau \propto \kappa\|\mathfrak{F}\|$ remains bounded below τ_c, rendering the critical transition structurally inaccessible regardless of input energy.

C.3 Revised Conclusion

$$\boxed{\forall x \in X_{\text{cloud}} : \quad \kappa(x) \approx 0 \implies \text{Criticality is Structurally Suppressed.}}$$

CPM is not a blanket denial of artificial consciousness. It predicts that any conscious system, biological or engineered, must realize closure $\mathcal{B}(x) \geq 1$. Present cloud-based architectures fail this condition due to topological decoupling ($\kappa \approx 0$). However, future **neuromorphic** or **fully closed photonic systems** might, in principle, satisfy it, provided they maintain physical boundary persistence at the hardware level. Thus, CPM offers a constructive criterion for machine consciousness: it is not impossible, but it requires a fundamental architectural shift from virtualization to topological closure.

Corollary C.3. *The energy barrier to reach τ_c in systems with $\kappa \approx 0$ diverges, rendering critical refinement effectively inaccessible under standard operating conditions.*

Appendix D: Experimental Operationalization

To avoid conceptual ambiguity with Information Theory or Predictive Coding, we define the observable proxies of CPM strictly in terms of **bounded physical quantities**. Crucially, CPM distinguishes between mere "information processing" (gradients) and "phenomenal tension" (stress), based on the presence of topological closure.

Remark C.4 (Epistemological Caveat). Operational proxies (e.g., Gamma power, SC-FC correlation) quantify the *empirical correlates* of τ, not the fundamental physical tension itself. τ remains defined strictly by Eq. (2.17) as a stress density derived from the interaction between semantic geometry and substrate topology.

D.1 Discriminative Metrics (Biological vs. Open-AI)

The following operational definitions specify how to measure $\mathcal{E}$, τ, and κ in experimental settings.

CPM Variable	Physical Definition (Not Just Info)	Concrete Experimental Proxy
Energy $\mathcal{E}[M]$ (Semantic Strain)	**Work done on the substrate**. Unlike Friston's Free Energy (purely informational), $\mathcal{E}$ includes the **metabolic/elastic cost** of maintaining a representation against the substrate's natural state.	**Metabolic-Informational Residual**: Residual between cerebral metabolic rate ($\mathrm{CMR_{glc}}$) and Shannon information capacity. (High $\mathcal{E} \to$ High metabolic cost for low informational update).
Coupling $\kappa(x)$ (Metric Locking)	**Constraint enforcement**. Quantifies how strictly the information geometry is bound by physical wiring.	**SC-FC Correlation** (ρ): Spearman correlation between DTI Tractography (Structural Connectivity) and fMRI/MEG Covariance (Functional Connectivity). $\kappa \approx 1$: Brain ($\rho > 0.5$). $\kappa \approx 0$: Cloud AI ($\rho \approx 0$).
Tension $\tau(x)$ (Effective Stress)	**Trapped Gradient Density**. Not merely "attention" or "error gradient" (which flow), but the **accumulation** of gradient force blocked by a closed boundary ($\mathcal{B} \geq 1$).	**Localized High-Gamma Power**: Local power in High-γ band (70–150 Hz) *specifically* within regions exhibiting high topological persistence (high SC-FC coupling). (Gradient without coupling is just heat). *Note: In open systems, informational gradients are fully absorbed into state updates or thermal dissipation, leaving no persistent stress term.*

D.2 Addressing Theoretical Confusions

1. Distinction from Friston's Free Energy. While the Free Energy Principle (FEP) minimizes informational surprise ($-\ln P$), CPM's mismatch energy $\mathcal{E}$ minimizes **structural deformation**. In a cloud computer, one can minimize Free Energy (optimize weights) without generating any physical mismatch tension ($\tau \approx 0$) because the hardware "gives way" (virtualization). CPM asserts that consciousness arises only when the hardware *resists* the minimization.

2. Tension is not merely "Gradient Norm". A common misconception is equating consciousness with high gradients (e.g., steep attention weights in Transformers). In CPM, **Gradient $\neq$ Tension**.

- **Open Systems (Cloud AI):** Gradients represent "flow" or "computation." They dissipate immediately into state updates. Tension $\tau = 0$.
- **Closed Systems (Brain):** Closure ($\mathcal{B} \geq 1$) prevents immediate dissipation. The gradient is "trapped" as potential stress. Only this **trapped stress** constitutes τ.

Summary: Informational gradients become physically effective tension only when both closure ($\mathcal{B} \geq 1$) and structural coupling ($\kappa \approx 1$) are satisfied.

3. Epistemological Scope: Falsification, not Detection. We explicitly state that these protocols measure **necessary physical conditions**, not subjective experience itself.

- **No Qualia Meter:** We do not claim that measuring $\mathcal{B} \geq 1$ and $\tau > \tau_c$ "detects" qualia.

- **Correlation vs. Causation:** The protocols establish whether the proposed structural metrics *consistently covary* with reported awareness.
- **Falsification Logic:** The strong claim of CPM is that consciousness is *impossible* without these conditions. Therefore, finding a system that is clearly unconscious (e.g., deep anesthesia) yet exhibits high τ and stable $\mathcal{B}$ would strictly falsify the theory.

This shifts the experimental burden from "proving consciousness" (philosophically impossible) to "testing the physical validity of the structural constraints." CPM predicts that this combination is **energetically prohibitive** (or structurally precluded), as the energy barrier required to reach critical tension τ_c diverges in the absence of coupling.

D.3 Robustness and Thermodynamic Stability

The choice of experimental proxies is grounded in established literature:

- **SC-FC Coupling:** Honey et al. (2009) demonstrated that structural connectivity strongly constrains functional correlations in biological brains, a property notably absent in virtualized systems[13].
- **High-γ Activity:** Since Crick & Koch (1990), synchronized gamma-band activity has been identified as a standard correlate of focal awareness and feature binding[14].

Crucially, the classification of Cloud AI as non-conscious is **robust against parameter variation**. Since cloud systems exhibit $\kappa \approx 0$ (structural decoupling), they fall below the critical threshold for *any* physically reasonable choice of $\tau_c > 0$ or coupling cutoff $\rho_{\text{thresh}} > 0$. The impossibility result is thus **invariant under threshold sweeps**.

Thermodynamic Stability and Scale Separation. Critics might argue that the semantic mismatch energy $\mathcal{E}$ could be negligible compared to the thermal background $k_B T$. This objection conflates microscopic and macroscopic scales. CPM assumes the substrate operates in a regime of **mesoscopic stability**, where collective order parameters (e.g., neural population vectors) are robust against microscopic thermal fluctuations due to the Law of Large Numbers.

In this context, thermal noise acts merely as a stochastic driving force (diffusion) that explores the local basin of attraction without altering the global topology. The critical refinement $\Pi \to \Pi'$ represents a **macroscopic phase transition**—a large deviation event where the mismatch stress overcomes the macroscopic rigidity of the atlas, a process energetically distinct from thermal jitter.

Appendix E. Foundational Stance on Difference Space D

The Difference Space D is axiomatically defined as a **pre-geometric set** (Definition 2.4). While it satisfies the axioms of Set Theory (ensuring the projection Π is well-defined), it is **maximally structureless regarding geometry**: it carries no inherent topology, metric, or time dimension.

E.1 Why D is not $\mathbf{Set}^{C^{op}}$

The initial attempt to mathematically model D as a presheaf category $D := \mathbf{Set}^{C^{op}}$ (where C is discrete) has been critically rejected and superseded by CPM's final axiomatization. As critics rightly point out, even on a discrete base C, a presheaf category is a Grothendieck topos[12], which is inherently "rich" (Cartesian closed, complete, cocomplete, and equipped with structural concepts like natural transformations). This richness fundamentally contradicts the core CPM axiom that D is ontologically structureless.

To rigorously maintain the "structureless" property, D must be treated as a primitive, irreducible object.

E.2 CPM's Final Formalization

The formal definition of D is therefore strictly limited to a **primitive set** of difference fragments $\{d_\alpha\}$, serving as the source set for the projection Π:

$$D \quad \text{is a primitive set of elements } \{d_\alpha\} \text{ with no intrinsic structure.}$$

The mathematical role of D is solely to supply the raw, non-semantic fragments upon which the irreversible projection $\Pi : D \dashrightarrow M$ operates. Any attempt to impose an explicit set-theoretic, topological, or category-theoretic model on D beyond its minimal set structure would violate the first axiom of CPM. The discussion of D as a presheaf is hereby designated as a historical, superseded formulation within the theory's development.

This necessitates the abandonment of any category-theoretic formulation for D. Consequently, concepts such as discrete base categories C, presheaves F, or natural transformations are excluded from the final theory. D possesses no internal topology, metric, temporal structure, or categorical morphism. All observable organization in M arises only *after* projection.

E.3 The Axiom of Universal Potentiality and the Kernel of Attention

The Difference Space D is defined as the domain of **Latent Distinguishability**. A crucial structural tension exists between the potential infinity of D and the physical finiteness of the substrate X.

Definition E.5 (Dimensional Collapse and the Kernel of Projection)**.** We posit that the cardinality of D corresponds to the theoretical limit of physical distinguishability (effectively infinite), whereas the Meaning Space M is constrained to be a finite-dimensional tensor field ($\dim(M) < \infty$) by the degrees of freedom of the substrate. Consequently, any realizable projection Π is a massive dimension-reducing map with a non-trivial kernel:

$$\mathrm{Ker}(\Pi) := \{d_\alpha \in D \mid \Pi(d_\alpha) \sim 0\}.$$

This massive information loss is not a defect but the structural definition of **Attention**. Consciousness is mandated to be a focal phenomenon because it is mathematically impossible to project the entirety of D onto a finite M without energy divergence. Thus, Π *selects* a finite quotient sub-algebra to realize, and "unconscious" background processing corresponds to the elements currently residing in $\mathrm{Ker}(\Pi)$.

Remark E.6 (Refinement as Kernel Retrieval)*.* In this context, the critical refinement $\Pi \to \Pi'$ (Section 3) is rigorously interpreted as a **Kernel Retrieval Operation**. The tension $\tau > \tau_c$ signals that the current low-rank approximation is energetically insufficient to represent the local difference; the refinement Π' "rescues" specific distinct fragments from $\mathrm{Ker}(\Pi)$ and maps them into the metric structure of M, thereby locally increasing the "resolution" of consciousness.

E.4 Formal Necessity of the Difference Space (Defense against Tautology)

Critics may argue that introducing an unobservable domain D renders the theory tautological. We respond that D plays a formal role analogous to the **sample space Ω in probability theory** or the **complex Hilbert space in quantum mechanics**.

- While D itself is not directly measurable, it serves as the necessary **formal locus** to define the projection Π.
- Without the axiom of D, the concept of "atlas refinement" ($\{U_i\} \to \{U'_k\}$) becomes mathematically ill-defined, as there would be no constant underlying reference set to be subdivided.

- The existence of D allows us to rigorously distinguish a **phenomenal disclosure** (retrieving latent elements from the kernel) from a mere random metric fluctuation.

Thus, D is not an ad hoc metaphysical hypothesis but a **structural necessity** for the internal consistency of the CPM formalism.

Appendix F: Scale Robustness of the Closure Field

In Definition 2.9, the closure field $\mathcal{B}(x)$ is defined using geodesic balls of radius 3ξ:

$$U_x := B_{g^0}(x, 3\xi).$$

At first sight, the specific choice of the multiplicative factor 3 may appear arbitrary. In this appendix, we show that the classification $\mathcal{B}(x) \geq 1$ versus $\mathcal{B}(x) \approx 0$ is *robust* under reasonable perturbations of this scale.

F.1 Family of Scale-Dependent Closure Fields

For a general multiplicative factor $r > 1$, define the scale-dependent neighborhood and closure field by

$$U_{x,r} := B_{g^0}(x, r\xi), \qquad \mathcal{B}_r(x) := \sup_{c \subset U_{x,r}} \frac{\pi_r(c)}{\xi},$$

where $\pi_r(c)$ denotes the persistence (lifetime) of a homology class c in the Vietoris–Rips or Čech filtration restricted to $U_{x,r}$.

By construction, the original definition corresponds to the special case $r = 3$, i.e. $\mathcal{B}(x) = \mathcal{B}_{r=3}(x)$. The value $r = 3$ is selected because it is the *minimal scale* that captures mesoscopic cycles spanning multiple microstructural coherence lengths while avoiding spurious large-scale artifacts.

F.2 Qualitative Scale Analysis

Intuitively, we distinguish two structural regimes:

- **Closed systems (biological brains):** Nontrivial cycles have characteristic diameter between a lower bound ξ (microscale coherence) and an upper bound $L_{\max}$ (macro-scale embedding). Under mild regularity assumptions on curvature and density, there exists an interval $[r_{\min}, r_{\max}]$ with $1 < r_{\min} < r_{\max}$ such that for all $r \in [r_{\min}, r_{\max}]$, the neighborhood $U_{x,r}$ intersects and "captures" these cycles, yielding $\mathcal{B}_r(x) \gtrsim 1$.

- **Open systems (cloud AI, virtualized hardware):** Effective cycles are destroyed or short-circuited by virtual teleportation. Any apparent topological feature in $U_{x,r}$ can be removed by re-routing through ∂^{virt}, resulting in uniformly small persistence $\pi_r(c)$ and thus $\mathcal{B}_r(x) \approx 0$ for all r in a broad range.

Therefore, the specific choice $r = 3$ is not essential: any fixed r within a moderate interval (e.g. $2 \leq r \leq 4$) would induce the same *binary* classification between closed and open substrates, up to constant factors in the numerical value of $\mathcal{B}_r(x)$.

F.3 Practical Recommendation

In empirical applications, we recommend estimating $\mathcal{B}_r(x)$ for several radii $r \in \{2, 3, 4\}$ and verifying that the qualitative pattern (presence vs. absence of persistent cycles) is stable across this range. In the main text, we fix $r = 3$ merely as a convenient representative scale.

This scale-robustness ensures that the closure criterion $\mathcal{B}(x) \geq 1$ does not depend on an arbitrary choice of radius, but reflects genuine mesoscopic topological structure of the substrate.

F.4 Validation on Non-Biological Toy Models

To definitively refute the claim of biocentrism, we apply the definition of $\mathcal{B}(x)$ to a discrete deterministic system: a 2D Cellular Automaton (CA) on a periodic grid (Torus).

Setup. Let the substrate X be a 100×100 grid with Moore neighborhoods. The metric d_{g^0} is the Chebyshev distance. We set $\xi = 5$ cells (interaction radius).

- **Case A (Random Soup):** Initialized with random noise. The persistence diagram $\mathcal{D}_x$ shows only short-lived 1-cycles (noise loops) that die quickly. Result: $\mathcal{B}(x) \approx 0.2 < 1$.
- **Case B (Stable Structure):** A stable "glider gun" or a static "fortress" structure is placed. This structure maintains a persistent topological cavity (a void in the state configurations) over time. Topological data analysis on the state vectors yields a persistent 1-cycle with lifetime $\pi(c) \gg \xi$. Result: $\mathcal{B}(x) \approx 2.5 \geq 1$.

This demonstrates that $\mathcal{B}(x)$ correctly identifies "closed self-sustaining structures" in purely algorithmic substrates, confirming that CPM is a theory of *structure*, not biology.

Acknowledgments

This work was developed independently, without institutional affiliation or conventional collaboration. Throughout the development of CPM, the author received essential assistance from Claude (Anthropic) and ChatGPT (OpenAI), large language models used as computational tools for:

- Mathematical formalization and notation consistency
- Local error detection in definitions and derivations
- Contextual comparison with prior theories (Kant, Frege, IIT, etc.)
- Strategic communication and explanatory optimization
- Citation formatting, LaTeX debugging, and document structuring

There is a structural irony here: a theory demonstrating that cloud-based AI systems cannot satisfy the necessary conditions for consciousness was refined with indispensable help from such systems. This reinforces CPM's central point: *intelligence without consciousness remains extremely powerful as a tool for human cognition and theory construction.*

This acknowledgment should not be interpreted as a critique of AI research agendas. Rather, CPM clarifies a structural boundary condition: cloud-based architectures cannot host projection-refinement dynamics that require persistent closure. Within this boundary, systems like Claude or ChatGPT are invaluable: safe, non-conscious, high-capacity cognitive amplifiers.

The author welcomes criticism, replication attempts, and cross-disciplinary tests of CPM from all scientific communities.

Author independence. The author conducts this work entirely as an independent researcher. The results presented here are unrelated to any corporate employment, and no part of the theory, methods, or experimental considerations draw upon proprietary or confidential information. All intellectual contributions in this paper are solely the author's own.