Cohomological and Thermodynamic Guardrails for Causal World Models: A Neuro-Symbolic Architecture for Verifiable Scientific Intelligence

Peter De Ceuster

November 15, 2026

Abstract

Contemporary foundation models operate predominantly as stateless, dissipative autoregressive token predictors lacking intrinsic representations of physical conservation laws, global causal consistency, or formal mathematical truth. In our recent perspective on the two-hundred-year trajectory of artificial intelligence to 2226, we argued that stateless token prediction will yield to continuously learning, causal world models tightly coupled with formal theorem provers, rendering mathematical and physical hallucination structurally unrepresentable within the verified subspace. Here, we construct the explicit mathematical and algorithmic architecture required to realize this synthesis. First, we constrain continuous neural latent trajectories on a state manifold $\mathcal{M}$ using the non-equilibrium thermodynamic GENERIC (General Equation for Non-Equilibrium Reversible–Irreversible Coupling) formalism, pairing skew-symmetric Poisson operators $L(z) = -L(z)^{\mathsf{T}}$ (encoding Casimir and topological gauge invariants such as $\mathrm{d}(*F - \theta F - J_s) = 0$) with symmetric positive-semidefinite friction operators $M(z) \succeq 0$ parameterized via Weil–Li quadratic positivity functionals. This operator structure guarantees exact conservation laws alongside strict compliance with the Second Law of Thermodynamics ($\dot{S} \geq 0$). Second, generalizing the de Rham exactness condition $[J_s]_{\mathrm{dR}} = 0$ and cohomology obstruction classes $\omega(\eta)$ from our extended electrodynamics framework, we organize distributed local causal models as sections of a cellular sheaf $\mathcal{F}$ over a relational state complex X. Multi-scale contradictions and contextual hallucinations manifest strictly as non-vanishing first sheaf cohomology classes $[\delta^0 s] \in H^1(X; \mathcal{F})$, quantifiable via the sheaf Laplacian $\Delta_{\mathcal{F}} = (\delta^0)^\star \delta^0$. Third, drawing on Tannakian realization functors in Grothendieck's theory of motives and our epistemological demarcation into proved mathematics [**P**], empirical evidence [**E**], and structural theses [**T**], we define a functorial verification bridge into dependent type theory (Lean 4) governed by an active "upstream" epistemic control loop that halts passive statistical drift whenever cohomological obstruction energy exceeds tolerance.

Keywords: Neuro-symbolic AGI, causal world models, GENERIC non-equilibrium thermodynamics, cellular sheaf cohomology, dependent type theory, Weil–Li positivity, formal verification, motives.

Contents

1 Introduction: From Autoregressive Drift to Structural Invariance **3**

2 The Typed Epistemic Stratification: [P], [E], and [T] **4**

3 Thermodynamic and Gauge Guardrails via GENERIC Brackets and Weil–Li Positivity **5**
3.1 Manifold Parameterization via the GENERIC Formalism 5
3.2 Weil–Li Positivity Parameterization and Topological Gauge Casimirs 5

4 Cellular Sheaf Cohomology and Obstruction Classes as Hallucination Detectors **6**
4.1 The Relational Context Complex and Cellular Sheaf 6
4.2 Sheaf Laplacian, Cohomology Obstruction $\omega(\eta)$, and Hallucination Theorem . . . 7

5 Functorial Realization into Dependent Type Theory and the "Upstream" Epistemic Loop **8**
5.1 Motivic Realization Inspiration and the Type-Realization Functor 8
5.2 The "Upstream" Epistemic Control Loop (The Salmon's Reflex) 9

6 Hard Boundaries: Chaos, Irreducibility, and Undecidability **10**

7 Architectural Comparison and Synthesis with Paper 1 **10**

8 Conclusion: The Complete Two-Pillar Foundation **10**

1 Introduction: From Autoregressive Drift to Structural Invariance

Forecasting the evolution of machine intelligence across a multi-century horizon requires separating transient software engineering heuristics from fundamental physical and mathematical invariants [1]. In our companion hardware study [2], we addressed the first two invariants of twenty-third-century intelligence—non-equilibrium thermodynamics and relativistic synchronization latency—by formulating a sub-Landauer optomagnetic computing substrate based on non-thermal Floquet–Chern switching [7] and Electromagnetically Induced Transparency (EIT) in Bose–Einstein condensates [3, 5].

In the present paper, we address the third foundational invariant identified in *Intelligence in 2226* [1]: **the structural elimination of mathematical and physical hallucination via neuro-symbolic causal world models coupled to formal verification kernels**.

Modern frontier language models minimize cross-entropy loss over sequential token distributions $P(w_t \mid w_{<t})$. Despite impressive empirical fluency, unconstrained autoregressive architectures suffer from three deep structural pathologies when deployed in scientific discovery, theoretical physics, and high-consequence engineering:

1. **Thermodynamic and Symplectic Blindness:** Neural latent representations evolve through unconstrained affine layers and activation functions that neither conserve symplectic phase-space volume, respect gauge continuity laws, nor enforce non-negative entropy production. When rolled out over long horizons, continuous world models routinely predict trajectories that violate energy-momentum conservation or the Second Law of Thermodynamics.

2. **Local-to-Global Cohomological Inconsistency:** Attention mechanisms aggregate context via soft inner-product weights without verifying whether locally plausible inferences glue together into a globally consistent causal structure. A model can assert $A \implies B$ in one local context window, $B \implies C$ in a second, and $A \implies \neg C$ in a third without detecting the global topological obstruction.

3. **Epistemological Conflation and Passive Drift:** Autoregressive decoders conflate deductive mathematical proofs, finite-sample empirical correlations, and speculative heuristic analogies into a single undifferentiated stream of token probabilities. In the language of our philosophical inquiry *The Salmon's Revolt Against a Mechanical World* [4], a purely statistical predictor drifts passively downstream with the thermodynamic current of training correlations, lacking the intentional, "upstream" agency required to halt generation, recognize an internal contradiction, and demand rigorous proof.

Remarkably, the mathematical machinery required to resolve all three pathologies already exists across our recent work in extended topological electrodynamics [3, 5], spectral number theory [8], and arithmetic geometry [6, 9]:

- In Ref. [8], we introduced a strict epistemological stratification partitioning all scientific propositions into **Proved Mathematics** [**P**], **Empirical and Statistical Evidence** [**E**], and **Interpretive Structural Theses** [**T**], and showed how infinite-dimensional positivity conditions (Weil's quadratic functional and Li's criterion) confine spectral zeros to a critical geometric locus.

- In Refs. [3,5], we demonstrated that extending Maxwell's equations via a higher-dimensional three-form current $J_s \in \Omega^3(X)$ ($\mathrm{d} * F = J_s$) requires J_s to be globally exact ($[J_s]_{\mathrm{dR}} = 0 \in H^3_{\mathrm{dR}}(X)$), and that non-trivial cohomology obstruction classes $\omega(\eta)$ produce measurable phase and visibility defects.

- In Refs. [6, 9], we examined how functorial realizations on Voevodsky's triangulated category of mixed motives $DM(k)$ linearize algebraic varieties into verifiable cohomological invariants, and outlined how hybrid human–AGI–formal proof systems can navigate multi-tiered mathematical landscapes.

Here, we synthesize these tools into a unified **GENERIC–Sheaf–Type (GST) Neuro-Symbolic Architecture**. Section 2 formalizes the typed epistemic lattice $([\mathbf{P}], [\mathbf{E}], [\mathbf{T}])$. Section 3 constructs the thermodynamic world-model dynamics using Poisson–GENERIC brackets and Weil–Li positivity operators. Section 4 formulates cellular sheaf cohomology over relational state complexes and proves that logical and multi-scale hallucinations correspond strictly to non-vanishing cohomology classes $H^1(X; \mathcal{F}) \neq 0$. Section 5 defines the functorial realization into dependent type theory (Lean 4) and the upstream epistemic control loop. Section 6 analyzes fundamental boundaries—computational irreducibility, Lyapunov chaos, and Gödelian undecidability—and Section 7 provides a complete system specification and comparative evaluation.

2 The Typed Epistemic Stratification: [P], [E], and [T]

In *Convergent Structural Evidence for the Riemann Hypothesis* [8], we demonstrated that centuries of confusion in foundational mathematics and theoretical physics stem from failing to demarcate proved theorems from asymptotic numerical evidence and structural conjectures. To prevent a neuro-symbolic intelligence from laundering a heuristic conjecture into an assumed theorem during multi-step reasoning, every internal representation, latent trajectory, and symbolic proposition in our architecture carries an explicit **epistemic grade** drawn from a bounded meet-semilattice.

Definition 2.1 (Epistemic Grading Semilattice)**.** Let $(\mathfrak{E}, \preceq, \wedge)$ be the three-element totally ordered epistemic grading lattice

$$\mathfrak{E} = \{[\mathbf{T}] \prec [\mathbf{E}] \prec [\mathbf{P}]\}, \tag{1}$$

where the grades carry the following operational semantics:

1. [**P**] (**Formally Verified / Invariant**): Propositions possessing a machine-checked proof term $\pi : \tau$ in a sound dependent type theory (e.g., Lean 4 kernel) or trajectories identically satisfying an exact operator identity (such as $\mathrm{d}^2 = 0$ or skew-symmetric Casimir conservation).

2. [**E**] (**Empirically Bounded / Statistically Validated**): Claims supported by finite-resolution experimental measurements, bounded numerical integration, or statistical confidence intervals $(\mu, \sigma, \delta_{\text{conf}})$ with vanishing local sheaf coboundary error below threshold ϵ_{tol}.

3. [**T**] (**Structural Thesis / Abductive Hypothesis**): Candidate mathematical conjectures, counterfactual world-model rollouts, or post-Grothendieckian structural analogies [6] generated by the neural proposal engine prior to formal or empirical discharge.

In any composite inference step combining premises $\phi_1 : g_1$ and $\phi_2 : g_2$ with $g_1, g_2 \in \mathfrak{E}$, the epistemic grade of the derived conclusion ψ is governed by the lattice meet operator:

$$\frac{\Gamma \vdash \phi_1 : g_1 \qquad \Gamma \vdash (\phi_1 \to \psi) : g_2}{\Gamma \vdash \psi : g_1 \wedge g_2}, \qquad g_1 \wedge g_2 = \min_{\preceq}(g_1, g_2). \tag{2}$$

By Rule (2), an abductive thesis [**T**] or empirical approximation [**E**] can never silently contaminate the [**P**]-verified kernel; elevating a claim from [**T**] or [**E**] to [**P**] requires an explicit **functorial verification discharge** (Section 5).

3 Thermodynamic and Gauge Guardrails via GENERIC Brackets and Weil–Li Positivity

We first construct the continuous causal world model so that violations of energy conservation, gauge continuity, and the Second Law of Thermodynamics are structurally unrepresentable at the differential operator level.

3.1 Manifold Parameterization via the GENERIC Formalism

Let $\mathcal{M}$ be a smooth N-dimensional latent state manifold representing the coarse-grained physical macrostate $z \in \mathcal{M}$ of a modeled dynamical system. Rather than learning an unconstrained neural ODE $\dot{z} = f_\theta(z)$, we parameterize the vector field $\dot{z} \in T_z\mathcal{M}$ using the **General Equation for Non-Equilibrium Reversible–Irreversible Coupling (GENERIC)** [10, 11]:

$$\frac{\mathrm{d}z}{\mathrm{d}t} = L(z)\,\mathrm{d}E(z) + M(z)\,\mathrm{d}S(z), \tag{3}$$

where:

- $E : \mathcal{M} \to \mathbb{R}$ and $S : \mathcal{M} \to \mathbb{R}$ are smooth scalar functionals representing total system energy and thermodynamic entropy, respectively, with cotangent differentials $\mathrm{d}E(z), \mathrm{d}S(z) \in T_z^*\mathcal{M}$.
- $L(z) : T_z^*\mathcal{M} \to T_z\mathcal{M}$ is a skew-symmetric Poisson bivector operator satisfying the Jacobi identity $\{F, \{G, H\}\}_L + \text{cyclic} = 0$ for $\{F, G\}_L \equiv \langle \mathrm{d}F,\, L(z)\,\mathrm{d}G\rangle$.
- $M(z) : T_z^*\mathcal{M} \to T_z\mathcal{M}$ is a symmetric, positive-semidefinite irreversible friction operator encoding dissipative relaxation.

To decouple reversible symplectic transport from irreversible entropy production, $L(z)$ and $M(z)$ must satisfy the two mutual degeneracy (orthogonality) conditions:

$$L(z)\,\mathrm{d}S(z) = 0, \qquad M(z)\,\mathrm{d}E(z) = 0 \qquad \forall z \in \mathcal{M}. \tag{4}$$

3.2 Weil–Li Positivity Parameterization and Topological Gauge Casimirs

In both our study of the Riemann Hypothesis [8] and our work on Grothendieck's Standard Conjectures [9], a central theme is that geometric confinement and stability are governed by **positivity conditions**—specifically Weil's quadratic positivity functional $W(f * \tilde{f}) \geq 0$, Li's sequence $\lambda_n \geq 0$, and the Hodge–Lefschetz polarization form. We apply this exact structural principle to construct $L(z)$ and $M(z)$ in the neural world model so that Eqs. (3)–(4) hold identically for all network weights $\theta \in \Theta$.

Proposition 3.1 (Structural Parameterization of Conservation and Second-Law Invariance)**.** *Let $A_\theta(z) \in \mathbb{R}^{N\times N}$ and $B_\theta(z) \in \mathbb{R}^{N\times r}$ be arbitrary smooth neural matrix fields parameterized by weights $\theta \in \Theta$, and define the projection operators onto the orthogonal complements of $\mathrm{d}S(z)$ and $\mathrm{d}E(z)$:*

$$P_{\mathrm{d}S}^{\perp}(z) \equiv \mathbb{I}_N - \frac{\mathrm{d}S(z) \otimes \mathrm{d}S(z)^{\mathsf{T}}}{\|\mathrm{d}S(z)\|^2 + \varepsilon_0}, \qquad P_{\mathrm{d}E}^{\perp}(z) \equiv \mathbb{I}_N - \frac{\mathrm{d}E(z) \otimes \mathrm{d}E(z)^{\mathsf{T}}}{\|\mathrm{d}E(z)\|^2 + \varepsilon_0}. \tag{5}$$

Construct the operators $L_\theta(z)$ and $M_\theta(z)$ via the sandwich forms:

$$L_\theta(z) = P_{\mathrm{d}S}^{\perp}(z)\left(A_\theta(z) - A_\theta(z)^{\mathsf{T}}\right)P_{\mathrm{d}S}^{\perp}(z)^{\mathsf{T}}, \tag{6}$$

$$M_\theta(z) = P_{\mathrm{d}E}^{\perp}(z)\,B_\theta(z)B_\theta(z)^{\mathsf{T}}\,P_{\mathrm{d}E}^{\perp}(z)^{\mathsf{T}}. \tag{7}$$

Then in the exact limit $\varepsilon_0 \to 0^+$ on the regular locus $\{\mathrm{d}E \neq 0, \mathrm{d}S \neq 0\}$, every trajectory $z(t)$ generated by Eq. (3) under any choice of network weights $\theta \in \Theta$ satisfies:

$$\begin{aligned} &\frac{\mathrm{d}E(z(t))}{\mathrm{d}t} = 0 \quad (\textit{First Law}), \\ &\frac{\mathrm{d}S(z(t))}{\mathrm{d}t} = \left\| B_\theta(z)^\mathsf{T} P^\perp_{\mathrm{d}E}(z)^\mathsf{T} \mathrm{d}S(z) \right\|_2^2 \geq 0 \quad (\textit{Second Law}). \end{aligned} \tag{8}$$

Proof. By construction, $L_\theta(z)^\mathsf{T} = -L_\theta(z)$ and $L_\theta(z)\mathrm{d}S(z) = 0$, while $M_\theta(z)^\mathsf{T} = M_\theta(z) \succeq 0$ and $M_\theta(z)\mathrm{d}E(z) = 0$. Applying the chain rule along Eq. (3) gives

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \langle \mathrm{d}E,\, L_\theta \mathrm{d}E \rangle + \langle \mathrm{d}E,\, M_\theta \mathrm{d}S \rangle = 0 + \langle M_\theta \mathrm{d}E,\, \mathrm{d}S \rangle = 0,$$

where the first term vanishes by skew-symmetry of L_θ and the second by degeneracy. Similarly,

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \langle \mathrm{d}S,\, L_\theta \mathrm{d}E \rangle + \langle \mathrm{d}S,\, M_\theta \mathrm{d}S \rangle = -\langle L_\theta \mathrm{d}S,\, \mathrm{d}E \rangle + \left\langle \mathrm{d}S,\, P^\perp_{\mathrm{d}E} B_\theta B_\theta^\mathsf{T} (P^\perp_{\mathrm{d}E})^\mathsf{T} \mathrm{d}S \right\rangle = \left\| B_\theta^\mathsf{T} (P^\perp_{\mathrm{d}E})^\mathsf{T} \mathrm{d}S \right\|_2^2 \geq 0.$$

□

Furthermore, when modeling electrodynamic, magnetohydrodynamic, or quantum-optical fields, we incorporate our generalized topological continuity law from Refs. [3, 5]:

$$\mathcal{C}_{\mathrm{gauge}}[F, \theta, J_s] \equiv \mathrm{d}\left(*F - \theta F - J_s\right) = 0, \tag{9}$$

as a **Casimir functional** $C_k(z)$ of the Poisson bracket ($L_\theta(z)\,\mathrm{d}C_k(z) = 0$ and $M_\theta(z)\,\mathrm{d}C_k(z) = 0$) by extending the projection operators in Eqs. (6)–(7) to the orthogonal complement of $\mathrm{span}\{\mathrm{d}E, \mathrm{d}C_1, \ldots, \mathrm{d}C_m\}$. Consequently, the world model is constrained to the physical gauge-invariant submanifold at epistemic grade [**P**].

4 Cellular Sheaf Cohomology and Obstruction Classes as Hallucination Detectors

While Section 3 enforces local differential conservation laws on a single state patch, a scientific intelligence must synthesize overlapping local contexts, multi-scale physical regimes (e.g., microscopic quantum optics vs. macroscopic continuum electrodynamics [3]), and distributed causal graphs. In *The Salmon's Revolt Against a Mechanical World* [4], we invoked the image of the **beads and the necklace**: individual contexts (beads) are distinct, yet their coherence depends on the relational thread (the necklace) binding them into an unbroken whole.

Mathematically, the exact structure that formalizes local data glued across a relational network is a **cellular sheaf** [12, 13]. Moreover, our work on extended electrodynamics [3, 5] already established the paradigm for detecting global inconsistency via cohomology: the modified Maxwell equation $\mathrm{d} * F = J_s$ admits a globally defined gauge field F on a manifold X if and only if the de Rham cohomology class $[J_s]_{\mathrm{dR}} \in H^3_{\mathrm{dR}}(X)$ vanishes identically ($[J_s]_{\mathrm{dR}} = 0$), and any topological mismatch manifests as a non-zero cohomology obstruction class $\omega(\eta)$.

4.1 The Relational Context Complex and Cellular Sheaf

Definition 4.1 (Causal World-Model Sheaf)**.** Let $X = (V, E, F_{\mathrm{cells}})$ be a regular cell complex (or hypergraph) whose vertices $v \in V$ index local causal sub-models, physical coordinate patches, or reasoning steps, and whose edges $e = (u, v) \in E$ (with incidence relation $u \trianglelefteq e$, $v \trianglelefteq e$) represent overlapping variables, boundary conditions, or shared logical predicates. A **cellular sheaf of Hilbert spaces** $\mathcal{F}$ on X assigns:

1. To each vertex $v \in V$, a local latent state space (stalk) $\mathcal{F}(v) \cong \mathbb{R}^{d_v}$ carrying local world-model predictions $s_v \in \mathcal{F}(v)$.

2. To each edge $e \in E$, an interface comparison space $\mathcal{F}(e) \cong \mathbb{R}^{d_e}$ encoding shared physical observables or logical propositions.

3. To each incident pair $v \trianglelefteq e$, a bounded linear (or locally linearized smooth) restriction map $\mathcal{F}_{v \trianglelefteq e} : \mathcal{F}(v) \to \mathcal{F}(e)$ projecting the local state s_v onto the shared interface e.

The space of 0-cochains $C^0(X;\mathcal{F}) = \bigoplus_{v \in V} \mathcal{F}(v)$ represents the aggregate collection of all local predictions $s = (s_v)_{v \in V}$ across the architecture, while the space of 1-cochains $C^1(X;\mathcal{F}) = \bigoplus_{e \in E} \mathcal{F}(e)$ lives on the overlaps. Given an orientation on each edge $e = (u \to v)$, the **sheaf coboundary operator** $\delta^0 : C^0(X;\mathcal{F}) \to C^1(X;\mathcal{F})$ measures the disagreement across every interface:

$$(\delta^0 s)_e \equiv \mathcal{F}_{v \trianglelefteq e}(s_v) - \mathcal{F}_{u \trianglelefteq e}(s_u) \in \mathcal{F}(e). \tag{10}$$

4.2 Sheaf Laplacian, Cohomology Obstruction $\omega(\eta)$, and Hallucination Theorem

Equipping $C^0(X;\mathcal{F})$ and $C^1(X;\mathcal{F})$ with inner products induces the adjoint operator $(\delta^0)^\star : C^1(X;\mathcal{F}) \to C^0(X;\mathcal{F})$ and the positive-semidefinite **sheaf Laplacian**:

$$\Delta_{\mathcal{F}} \equiv (\delta^0)^\star \delta^0 : C^0(X;\mathcal{F}) \to C^0(X;\mathcal{F}). \tag{11}$$

By the discrete Hodge decomposition theorem, the space of global sections $H^0(X;\mathcal{F})$ is the kernel of the sheaf Laplacian:

$$H^0(X;\mathcal{F}) = \ker(\delta^0) = \ker(\Delta_{\mathcal{F}}) = \left\{ s \in C^0(X;\mathcal{F}) \;\middle|\; \mathcal{F}_{v \trianglelefteq e}(s_v) = \mathcal{F}_{u \trianglelefteq e}(s_u) \;\; \forall e = (u,v) \in E \right\}. \tag{12}$$

Definition 4.2 (Cohomological Hallucination Functional and Obstruction Class)**.** For any candidate world-model state assignment $s \in C^0(X;\mathcal{F})$ and edge gauge connection $\eta \in C^1(X;\mathcal{F})$ representing relational translations between sub-models, define:

(i) The **Sheaf Disagreement (Hallucination) Energy**:

$$\mathcal{E}_{\text{obs}}(s) \equiv \frac{1}{2} \langle s,\, \Delta_{\mathcal{F}} s \rangle_{C^0} = \frac{1}{2} \left\| \delta^0 s \right\|_{C^1}^2 = \frac{1}{2} \sum_{e=(u,v) \in E} \| \mathcal{F}_{v \trianglelefteq e}(s_v) - \mathcal{F}_{u \trianglelefteq e}(s_u) \|_{\mathcal{F}(e)}^2 . \tag{13}$$

(ii) The **First Cohomology Obstruction Class** $\omega(\eta) \in H^1(X;\mathcal{F}) = \ker(\delta^1)/\operatorname{im}(\delta^0)$, which measures irreducible loop inconsistencies (frustrated cycles $v_0 \to v_1 \to \cdots \to v_k = v_0$) that cannot be removed by any local coordinate recalibration $\delta^0 s$.

Theorem 4.3 (Cohomological Criterion for Multi-Scale and Logical Consistency)**.** *Let $(X, \mathcal{F})$ be a causal world-model sheaf coupling M local inference patches. Then:*

(i) *A candidate inference state $s \in C^0(X;\mathcal{F})$ is globally self-consistent across all overlapping causal, physical, and logical boundaries if and only if $\mathcal{E}_{\text{obs}}(s) = 0$, i.e., $s \in \ker(\Delta_{\mathcal{F}}) = H^0(X;\mathcal{F})$.*

(ii) *Suppose the local sub-models enforce pairwise compatibility constraints $\mathcal{F}_{v \trianglelefteq e}(s_v) - \mathcal{F}_{u \trianglelefteq e}(s_u) = \eta_e$ for an interface offset 1-cocycle $\eta \in \ker(\delta^1) \subset C^1(X;\mathcal{F})$. A globally consistent state $s \in C^0(X;\mathcal{F})$ satisfying $\delta^0 s = \eta$ exists if and only if the cohomology obstruction class vanishes identically:*

$$[\omega(\eta)] = 0 \in H^1(X;\mathcal{F}). \tag{14}$$

Whenever $[\omega(\eta)] \neq 0$, the local assumptions contain a structural contradiction (a logical paradox or incompatible multi-scale physical closure), and the minimum achievable hallucination energy is bounded below by the norm of the harmonic Hodge projection $\mathcal{H}(\eta) \in \ker(\Delta_{\mathcal{F}}^{(1)})$:

$$\min_{s \in C^0(X;\mathcal{F})} \frac{1}{2} \left\| \delta^0 s - \eta \right\|_{C^1}^2 = \frac{1}{2} \left\| \mathcal{H}(\eta) \right\|_{C^1}^2 > 0. \tag{15}$$

Proof. Part (i) follows immediately from the positive-semidefiniteness of $\Delta_{\mathcal{F}} = (\delta^0)^\star \delta^0$ and Eq. (13). For Part (ii), by the orthogonal Hodge decomposition on 1-cochains,

$$C^1(X;\mathcal{F}) = \operatorname{im}(\delta^0) \oplus^\perp \ker(\Delta_{\mathcal{F}}^{(1)}) \oplus^\perp \operatorname{im}((\delta^1)^\star), \qquad \ker(\delta^1) = \operatorname{im}(\delta^0) \oplus^\perp \ker(\Delta_{\mathcal{F}}^{(1)}),$$

where $\Delta_{\mathcal{F}}^{(1)} = \delta^0(\delta^0)^\star + (\delta^1)^\star \delta^1$ is the 1-form Hodge Laplacian and $\ker(\Delta_{\mathcal{F}}^{(1)}) \cong H^1(X;\mathcal{F})$. Decomposing any cocycle $\eta \in \ker(\delta^1)$ as $\eta = \delta^0 s_0 + \mathcal{H}(\eta)$ with $\mathcal{H}(\eta) \perp \operatorname{im}(\delta^0)$, the residual norm is $\left\|\delta^0 s - \eta\right\|^2 = \left\|\delta^0(s - s_0)\right\|^2 + \left\|\mathcal{H}(\eta)\right\|^2$, which achieves zero if and only if $\mathcal{H}(\eta) = 0 \iff [\omega(\eta)] = 0 \in H^1(X;\mathcal{F})$. □

Remark 4.4 (Bridge to Extended Electrodynamics and Hardware Implementation). Theorem 4.3 is the exact algebraic-topological counterpart of our electrodynamic exactness constraint $\mathrm{d} * F = J_s \implies [J_s]_{\mathrm{dR}} = 0$ in Ref. [5] and our interferometric visibility degradation $1 - \mathcal{V} \sim \omega(\eta)^2 \Lambda^{-2k}$ in Ref. [3]. Even more remarkably, when executed on the reversible optomagnetic photonic hardware of Paper 1 [2], the harmonic projection $\mathcal{H}(\eta)$ can be evaluated physically via optical wave interference in a coupled Mach–Zehnder waveguide mesh: destructive interference visibility directly measures the sheaf coboundary residual $\left\|\delta^0 s - \eta\right\|^2$ at the speed of light.

5 Functorial Realization into Dependent Type Theory and the "Upstream" Epistemic Loop

Enforcing GENERIC thermodynamic invariance (Section 3) and vanishing sheaf cohomology $H^1(X;\mathcal{F}) = 0$ (Section 4) ensures that continuous world-model states are physically conservative and mutually coherent at epistemic grade [**E**]. To elevate mathematical and structural derivations to the absolute verification grade [**P**], we couple the sheaf space $H^0(X;\mathcal{F})$ to a formal dependent type-theoretic kernel (such as the Calculus of Inductive Constructions implemented in Lean 4 [14]).

5.1 Motivic Realization Inspiration and the Type-Realization Functor

In Grothendieck's program and our recent investigation into how future mathematicians and mathematical AGI could unlock the theory of motives [6, 9], the category of motives acts as a universal arithmetic-geometric hub equipped with faithful **realization functors** $(\omega_{\mathrm{Betti}}, \omega_{\mathrm{dR}}, \omega_\ell)$ that map geometric objects into concrete linear categories equipped with Galois and Hodge structures.

Analogously, let $\mathbf{Harm}(X;\mathcal{F})$ denote the category of harmonic global sections $s \in \ker(\Delta_{\mathcal{F}})$ equipped with GENERIC vector fields (L_θ, M_θ), and let $\mathcal{T}_{\mathrm{Lean}}$ denote the syntactic category of contexts, types, and proof terms in a dependent type theory with universe hierarchy $\mathrm{Type}_0 : \mathrm{Type}_1 : \cdots$.

Definition 5.1 (Functorial Neuro-Symbolic Realization)**.** The **Type-Realization Functor** $\mathcal{R}_{\mathrm{type}} : \mathbf{Harm}(X;\mathcal{F}) \to \mathcal{T}_{\mathrm{Lean}}$ assigns:

1. To each vertex stalk $\mathcal{F}(v)$ and its local coordinate algebra, a formal context $\Gamma_v \vdash \tau_v : \mathrm{Type}_u$ declaring the state spaces, differential operators, and physical constants in $\mathcal{T}_{\mathrm{Lean}}$.

2. To each edge restriction constraint $\mathcal{F}_{v \trianglelefteq e}(s_v) = \mathcal{F}_{u \trianglelefteq e}(s_u)$, a propositional identity type

$$\mathrm{Id}_{\mathcal{F}(e)}(\mathcal{F}_{v \trianglelefteq e}(s_v),\ \mathcal{F}_{u \trianglelefteq e}(s_u)) : \mathrm{Prop}. \tag{16}$$

3. To a candidate proof or symbolic derivation trace generated by the neural policy over X, a formal proof term π_e inhabiting Eq. (16). A global section $s \in H^0(X; \mathcal{F})$ is promoted from grade [**E**] or [**T**] to grade [**P**] if and only if the type-checker kernel verifies

$$\Gamma_X \vdash \Pi_{e \in E}\, \pi_e : \prod_{e=(u,v) \in E} \mathrm{Id}_{\mathcal{F}(e)}(\mathcal{F}_{v \trianglelefteq e}(s_v),\ \mathcal{F}_{u \trianglelefteq e}(s_u)) = \texttt{true}. \tag{17}$$

5.2 The "Upstream" Epistemic Control Loop (The Salmon's Reflex)

In a standard autoregressive transformer, token generation proceeds monotonically forward in time: even when internal activations become contradictory, the model continues sampling tokens, compounding hallucinations exponentially.

In contrast, our architecture implements an **Upstream Epistemic Reflex**—mathematically operationalizing the final image of *The Salmon's Revolt Against a Mechanical World* [4], in which an intentional agent refuses to drift passively downstream with mechanical inertia and instead turns upstream toward the sources of order and verification.

At every reasoning step t, before emitting any symbolic proposition or executing a physical control action, the system evaluates the coupled **Lyapunov–Cohomology Obstruction Functional**:

$$\begin{aligned} \mathcal{V}_{\mathrm{total}}(z, s; \pi) \equiv & \underbrace{\frac{1}{2} \langle s,\ \Delta_{\mathcal{F}} s \rangle_{C^0}}_{\text{Sheaf Cohomology Error } \mathcal{E}_{\mathrm{obs}}} + \underbrace{\lambda_{\mathrm{G}} \left\| L_\theta(z) \mathrm{d}S(z) \right\|_2^2 + \lambda_{\mathrm{G}} \left\| M_\theta(z) \mathrm{d}E(z) \right\|_2^2}_{\text{GENERIC Numerical Integration Residual}} \\ & + \underbrace{\lambda_{\mathrm{T}}\, \mathbb{I}_{\{\mathrm{TypeCheck}(\pi) = \texttt{fail}\}}}_{\text{Formal Kernel Rejection}}. \end{aligned} \tag{18}$$

The inference engine operates across three strictly gated regimes:

1. **Regime I — Downstream Laminar Flow ($\mathcal{V}_{\mathrm{total}} = 0$, Grade [P]):** When the sheaf coboundary vanishes, GENERIC invariants hold exactly, and the Lean 4 kernel accepts the proof term π, the derivation is committed to the permanent verified knowledge graph at epistemic grade [**P**].

2. **Regime II — Bounded Empirical Diffusion ($0 < \mathcal{E}_{\mathrm{obs}}(s) \leq \epsilon_{\mathrm{tol}}$, Grade [E]):** When the formal proof term is absent or intractable due to continuous empirical noise, but $\mathcal{E}_{\mathrm{obs}}(s) \leq \epsilon_{\mathrm{tol}}$, the system evolves s via **sheaf heat diffusion**

$$\frac{\mathrm{d}s}{\mathrm{d}\tau} = -\gamma_{\mathrm{diff}}\, \Delta_{\mathcal{F}} s \tag{19}$$

   to project s orthogonally onto the harmonic subspace $H^0(X; \mathcal{F})$, outputting the result with an explicit [**E**] confidence bound.

3. **Regime III — Upstream Epistemic Reversal ($\mathcal{E}_{\mathrm{obs}}(s) > \epsilon_{\mathrm{tol}}$ or $[\omega(\eta)] \neq 0$):** When the Hodge projection detects a non-vanishing cohomology obstruction $\|\mathcal{H}(\eta)\| > 0$ or kernel rejection, forward decoding is **structurally halted**. The system computes the maximal frustration cycle $C^* = \arg\max_{C \subset X} \sum_{e \in C} |(\mathcal{H}(\eta))_e|^2$ in $H^1(X; \mathcal{F})$ and initiates an upstream backward search: either (a) pruning the defective [**T**] hypothesis on cycle C^*, (b) invoking a hybrid quantum-binary / tactic search over the five-tier motivic hierarchy [6] to synthesize a missing lemma, or (c) requesting a targeted physical experiment to resolve empirical underdetermination.

6 Hard Boundaries: Chaos, Irreducibility, and Undecidability

A central thesis of *Intelligence in 2226* [1] is that stripping away science-fiction tropes requires respecting the hard boundaries imposed by nature and mathematical logic. No neuro-symbolic architecture—even one operating on sub-Landauer optomagnetic hardware with formal type-checking—becomes an omniscient oracle. Instead, the GST architecture explicitly quantifies and respects three irreducible ceilings:

Proposition 6.1 (Fundamental Horizon Bounds of the GST Architecture)**.** *Let a GST neuro-symbolic system operate with physical measurement precision $\delta z_0 > 0$, thermodynamic energy budget E_{comp}, and formal proof search horizon N_{steps}. Then:*

(i) ***Deterministic Chaos and Lyapunov Horizon:*** *For any GENERIC dynamical system on $\mathcal{M}$ exhibiting a positive maximal Lyapunov exponent $\lambda_{\max} > 0$, pointwise state predictions $z(t)$ degrade from grade $[\mathbf{E}]$ to uninformative beyond the Lyapunov time horizon*

$$t_{\text{Lyap}} = \frac{1}{\lambda_{\max}} \ln\left(\frac{\epsilon_{\text{tol}}}{\delta z_0}\right), \tag{20}$$

even though global thermodynamic invariants ($E(z(t)) = E_0$, $\dot{S}(z(t)) \geq 0$) and topological Casimirs ($\mathrm{d}(*F - \theta F - J_s) = 0$) remain exact at grade $[\mathbf{P}]$ for all $t \in [0, \infty)$.

(ii) ***Computational Irreducibility and Complexity Barriers:*** *For cellular or many-body quantum systems whose time evolution is computationally irreducible (or* PSPACE*/*BQP*-hard), no shortcut functor exists to predict the state at step T in $o(T)$ operations; only coarse-grained hydrodynamic and cohomological invariants can be shortcut at grade $[\mathbf{P}]$.*

(iii) ***Gödel–Tarski Undecidability and Empirical Underdetermination:*** *By Gödel's First Incompleteness Theorem, any consistent recursively enumerable dependent type theory $\mathcal{T}_{\text{Lean}}$ capable of expressing Peano arithmetic contains true propositions ϕ that possess no proof term π of length $\leq N_{\text{steps}}$. Furthermore, when two distinct physical models $\mathcal{M}_1, \mathcal{M}_2$ yield identical restrictions on all accessible experimental edges $E_{\text{exp}} \subset E$, the sheaf coboundary cannot distinguish them without active experimental intervention expanding E_{exp}.*

Proposition 6.1 highlights why the $[[\mathbf{P}], [\mathbf{E}], [\mathbf{T}]]$ epistemic lattice is indispensable: beyond the Lyapunov horizon t_{Lyap} or in undecidable arithmetic regimes, a conventional AI fabricates confident pointwise hallucinations, whereas the GST architecture automatically restricts its $[\mathbf{P}]$-claims to conserved structural invariants while tagging pointwise forecasts beyond t_{Lyap} as $[\mathbf{T}]$.

7 Architectural Comparison and Synthesis with Paper 1

Table 1 contrasts the GENERIC–Sheaf–Type (GST) Neuro-Symbolic Architecture against 2026 autoregressive foundation models, heuristic Retrieval-Augmented / Tool-Use wrappers, and standard Physics-Informed Neural Networks (PINNs).

8 Conclusion: The Complete Two-Pillar Foundation

Together with our companion hardware manuscript [2], this work completes a concrete, self-contained bridge between our 2026 theoretical physics and pure mathematics research corpus and the long-horizon architecture of twenty-third-century intelligence [1]:

Table 1: Comparison of AI paradigms across thermodynamic invariance, global consistency, formal verification, and epistemic typing (structural capabilities by architectural design).

Architectural Property	Autoregressive LLM (2026)	Heuristic RAG / Agent Wrapper	Standard PINNs	**GENERIC–Sheaf–Type Architecture (This Work)**
Conservation ($\dot{E} = 0$, $\mathrm{d} * F = J_s$)	None (soft token correlation)	Post-hoc external simulator call	Soft penalty loss $\mathcal{L}_{\mathrm{PDE}} > 0$	Hard operator projection ($L = -L^{\mathsf{T}}$, $M\mathrm{d}E = 0$)
Second Law ($\dot{S} \geq 0$)	Violated	Unchecked	Violated	Exact PSD parameterization ($M = P_{\mathrm{d}E}^{\perp} B B^{\mathsf{T}} (P_{\mathrm{d}E}^{\perp})^{\mathsf{T}} \succeq 0$)
Multi-Scale / Logical Consistency	Soft attention (no loop check)	Prompt-based self-critique	Single domain mesh only	Cellular Sheaf Laplacian $\Delta_{\mathcal{F}} = (\delta^0)^{\star}\delta^0$, $H^1(X;\mathcal{F})$
Mathematical Proof Soundness	Unverified text generation	External call without sheaf alignment	N/A	Functorial realization $\mathcal{R}_{\mathrm{type}}$ into Lean 4 ([**P**]-gated)
Epistemic Typing	Conflated	Conflated	Empirical only	Strict Lattice $[\mathbf{T}] \prec [\mathbf{E}] \prec [\mathbf{P}]$
Response to Contradiction	Passive drift	Heuristic retry	Gradient stall	Upstream Reflex on frustrated H^1 cycle
Target Hardware Substrate	Dissipative CMOS	Dissipative CMOS	Dissipative CMOS	Sub-Landauer Optomagnetic [2]

1. **Paper 1 (The Physical Substrate):** By uniting high-frequency Zeeman–Floquet optomagnetism [7] with pseudoscalar electrodynamics ($\mathbf{E}_{\mathrm{opt}} \cdot \mathbf{B}_{\mathrm{opt}} \neq 0$) and EIT slow-light BEC waveguides [3, 5], we established a physical blueprint for non-thermal, reversible in-memory computing operating below the conventional Landauer dissipation floor.

2. **Paper 2 (The Formal & Causal Architecture):** By uniting GENERIC non-equilibrium thermodynamics with Weil–Li positivity [8, 9], cellular sheaf cohomology obstructions [3, 5], motivic realization functors [6], and the active epistemic agency of *The Salmon's Revolt* [4], we constructed a neuro-symbolic world-model architecture in which conservation violations and logical contradictions within the verified subspace are structurally intercepted before output.

As low-wattage, zero-cloud intelligence becomes quietly embedded into critical physical infrastructure, scientific discovery, medicine, and governance over the coming decades, cryptographic provenance and formal mathematical verification will form the bedrock of cognitive autonomy and democratic legitimacy. The GENERIC–Sheaf–Type framework provides an immediately implementable mathematical foundation for that transition.

References

[1] P. De Ceuster, *Intelligence in 2226: Physical Bounds, Neuro-Symbolic Architecture, and the Restructuring of Human Civilization*, Preprint, Zenodo / `https://peterdeceuster.uk/papers/` (Uploaded November 14, 2026).

[2] P. De Ceuster, *Sub-Landauer Optomagnetic Logic: Non-Thermal Floquet–Chern Switching and Pseudoscalar Photonic Gates for Reversible Neuromorphic Computing*, Companion Preprint, Zenodo / `https://peterdeceuster.uk/papers/` (November 15, 2026).

[3] P. De Ceuster, *Topological Unification of Axion Electrodynamics and Higher-Dimensional Photon–Higgs Continuity*, Preprint, Zenodo / `https://peterdeceuster.uk/papers/` (Uploaded November 7, 2026).

[4] P. De Ceuster, *The Salmon's Revolt Against a Mechanical World*, Publication, Zenodo / `https://peterdeceuster.uk/papers/` (Uploaded November 3, 2026).

[5] P. De Ceuster, *Topological Currents in Extended Electrodynamics, Negative Effective Dispersion, and Supersymmetric Operator Protection*, Publication, Zenodo / `https://peterdeceuster.uk/papers/` (Uploaded October 29, 2026).

[6] P. De Ceuster, *How Future Mathematicians, Mathematical AGI, and Quantum-Binary Systems Could Unlock the Theory of Motives*, Publication, Zenodo / `https://peterdeceuster.uk/papers/` (Uploaded September 11, 2026).

[7] P. De Ceuster, *Beyond Electric-Dipole Coupling: Optical Zeeman Torque, Floquet Gauge Dynamics, and Topological Switching in Moiré Heterostructures*, Journal Article / Preprint, Zenodo / `https://peterdeceuster.uk/papers/` (Uploaded September 6, 2026).

[8] P. De Ceuster, *Convergent Structural Evidence for the Riemann Hypothesis: Or Why the RH Will Eventually Be Proven*, Publication, Zenodo / `https://peterdeceuster.uk/papers/` (Uploaded September 3, 2026).

[9] P. De Ceuster, *Beyond Grothendieck's Standard Conjectures: Derived Stacks, Automorphic Correspondences, and the Motivic t-Structure*, Publication, Zenodo / `https://peterdeceuster.uk/papers/` (Uploaded September 2, 2026).

[10] M. Grmela and H. C. Öttinger, *Dynamics and Thermodynamics of Complex Fluids. I. Development of a General Formalism*, Physical Review E **56**(6), 6620–6632 (1997).

[11] H. C. Öttinger, *Beyond Equilibrium Thermodynamics*, Wiley-Interscience, Hoboken, NJ (2005).

[12] J. M. Curry, *Sheaves, Cosheaves and Applications*, Ph.D. Thesis, University of Pennsylvania, arXiv:1303.3255 (2014).

[13] J. Hansen and R. Ghrist, *Toward a Spectral Theory of Cellular Sheaves*, Journal of Applied and Computational Topology **3**(4), 315–358 (2019).

[14] L. de Moura and S. Ullrich, *The Lean 4 Theorem Prover and Programming Language*, in *Automated Deduction – CADE 28*, Lecture Notes in Computer Science **12699**, 625–635, Springer (2021).