

Toward a Mathematical Formalization of the SB-HC4A Cosmological Model: A Recommended Approach

Matthias Gruber

Independent researcher

ORCID: 0009-0005-9697-1665

Correspondence: matthias@matthiasgruber.com

Abstract

The Singularity-Bounded Holographic Class 4 Automaton (SB-HC4A; Gruber, 2026a) proposes that the universe is a self-referential holographic Class 4 automaton bounded at every scale by singularity surfaces — information-impermeable boundaries of maximum density — where the observable interior is a decompressed projection of boundary-encoded information. The model was derived from a convergence of the Five-Class computational taxonomy, the Four-Model Theory of consciousness (FMT; Gruber, 2015, 2026b), the FMT formalization’s continuous model-space framework (Gruber, 2026c), and ’t Hooft’s (1993, 2016) holographic automaton interpretation. The model currently operates in natural language supplemented by structural arguments. This paper outlines a recommended mathematical formalization strategy. Eight formalization modules are proposed: (1) a measure-theoretic definition of the five computational classes replacing Wolfram’s visual classification, (2) a topological characterization of singularity boundaries as equivalence classes of Bekenstein-saturated surfaces, (3) a formal definition of holographic rule sets via dimensional compression operators, (4) the self-referential fixed point $\Phi(U) = U$ grounded in Lawvere’s fixed-point theorem and the theory of coalgebras, (5) the consciousness-cosmology structural identity as a functor between categories, (6) the necessity argument in modal logic, (7) energy-information equivalence as a duality principle, and (8) the cognitive ceiling problem formalized as a computability-theoretic bound on self-knowledge. A phased build order prioritizes components that interface with existing mathematical physics. The formalization is offered as a research program specification for mathematically trained collaborators; verification of the formal apparatus is explicitly deferred to domain experts.

Keywords: cosmology, cellular automata, holographic principle, formalization, criticality, singularity, self-referential closure, fixed-point theory, category theory, modal logic, Bekenstein bound

1. Introduction

1.1 The Formalization Gap

The SB-HC4A cosmological model (Gruber, 2026a) makes five structural claims: (a) the universe operates at Wolfram Class 4 — the edge of chaos; (b) singularities at every physical scale (Planck regime, particle interiors, event horizons, cosmological horizons, temporal endpoints) are structurally identical instances of the same information boundary; (c) the universe is a holographic Class 4 automaton whose rules, dynamics, and output are all holographic; (d) this architecture is self-referentially closed: $\Phi(U) = U$; and (e) consciousness, as described by the Four-Model Theory (FMT; Gruber, 2015, 2026b), is a local, scale-reduced instance of the same computational pattern.

These claims are currently expressed in natural language with structural arguments. The elimination argument for Class 4 is informal. The singularity unification is asserted by shared properties, not by a formal equivalence relation. The holographic rule set is defined verbally. The self-referential fixed point is stated but not derived. The consciousness-cosmology mapping is presented as a table of correspondences, not as a mathematical morphism.

Without formalization, the model’s flexibility is a liability. A verbally specified “structural identity” between consciousness and cosmology could potentially accommodate any observation post hoc. The necessity argument — that the SB-HC4A is the unique structure consistent with the five axioms — requires formal demonstration, not verbal persuasion. And the cognitive ceiling problem (Weak Point 5 of Gruber, 2026a) — that Class 4 observers may be constitutionally incapable of determining whether they are describing the universe or the limits of their own cognition — needs formal characterization before it can be analyzed.

1.2 The Three-Paper Formalization Program

This paper is the third in a series of formalization specifications. The first (Gruber, 2026c) specified the formalization of the Four-Model Theory of consciousness, proposing a continuous model-space framework with a model density $\rho(s, \nu, t)$, transfer entropy for permeability, criticality operationalization, ESM redirection dynamics, self-referential closure as a fixed point, and category-theoretic architecture. The second (Gruber, 2026d) specified the formalization of the Recursive Intelligence Model, proposing a domain-structured knowledge manifold, motivation as a consciousness functional, coupled stochastic differential equations, the ignition threshold as a stochastic bifurcation, and social coupling dynamics.

The present paper addresses the cosmological level. The three formalizations are not independent — the FMT formalization provides the consciousness-side apparatus that the cosmology formalization must map onto, and the RIM formalization provides the intermediate-scale dynamics that connect phenomenal experience to the learning processes that produce cosmological models. The full theoretical architecture, once formalized, forms a three-level system: cosmological substrate → consciousness → intelligence, with mathematical morphisms connecting each level.

1.3 Scope and Limitations

This paper specifies a formalization research program. It proposes mathematical frameworks, defines quantities, and outlines a build order. It does *not* verify the mathematical apparatus — the author is not a mathematician, and formal verification is explicitly deferred to domain experts. The equations and formal structures presented here are intended as precise specifications of *what needs to be formalized* and *which mathematical tools are appropriate*, not as proven theorems.

The challenge is acute. The cosmological model touches mathematical physics (general relativity, quantum field theory, holographic duality), theoretical computer science (computational complexity, automata theory, Kolmogorov complexity), mathematical logic (fixed-point theorems, incompleteness, modal logic), and category theory. No single mathematician commands all these domains. The formalization is therefore designed as a modular program where each module can be developed independently by domain specialists, with category theory providing the integration language.

The paper is structured as follows. Section 2 proposes measure-theoretic definitions of the five computational classes. Section 3 formalizes singularity boundaries. Section 4 defines holographic rule sets. Section 5 addresses self-referential closure. Section 6 constructs the consciousness-cosmology

functor. Section 7 formalizes the necessity argument. Section 8 addresses energy-information equivalence. Section 9 formalizes the cognitive ceiling. Section 10 proposes a phased build order. Section 11 identifies what formalization would buy and what it cannot deliver.

2. Formalizing the Five Computational Classes

2.1 The Problem with Wolfram’s Classification

Wolfram’s (2002) four-class taxonomy of cellular automaton behavior — static, periodic, random, complex — is empirical: classes are assigned by visual inspection of space-time diagrams. This is scientifically useful for classification but mathematically unsatisfying. The classes lack formal definitions, the boundaries between classes are imprecise, and the extension to continuous, noisy, high-dimensional systems (like the physical universe) is undefined.

The five-class refinement (Gruber, 2015, 2026a) — splitting Wolfram’s Class 3 into fractal (computationally reducible) and random (computationally irreducible, maximal Kolmogorov complexity) — sharpens the taxonomy but inherits the informality. For the SB-HC4A to be a mathematical object, each class needs a formal definition.

2.2 Proposed Measure-Theoretic Definitions

Consider a discrete dynamical system $\Sigma = (S, f, s_0)$ where S is the state space, $f: S \rightarrow S$ is the transition function, and s_0 is the initial state. The orbit of s_0 under f is the sequence $O = \{s_0, f(s_0), f^2(s_0), \dots\}$. Define the following quantities:

Topological entropy $h_{\text{top}}(f)$: The exponential growth rate of the number of distinguishable orbits as resolution increases (Adler, Konheim, & McAndrew, 1965). Measures the “complexity” of the dynamics.

Kolmogorov complexity rate $\kappa(O)$: The asymptotic Kolmogorov complexity per symbol of the orbit sequence. Measures the incompressibility of the output.

Lyapunov exponent λ_{max} : The maximum exponential divergence rate of nearby trajectories. Measures sensitivity to initial conditions.

Computational reducibility $r(\Sigma)$: Whether the state at time t can be computed in fewer than t steps. Formally: $r(\Sigma) = 1$ if there exists an algorithm A such that $A(s_0, t) = f^t(s_0)$ in time $o(t)$; $r(\Sigma) = 0$ otherwise.

Computational universality $u(\Sigma)$: Whether the system can simulate any Turing machine. $u(\Sigma) = 1$ if Σ is Turing-complete; $u(\Sigma) = 0$ otherwise.

The five classes are then defined by their signatures on these quantities:

Class	h_{top}	$\kappa(O)$	λ_{max}	$r(\Sigma)$	$u(\Sigma)$	Informal
1	0	0	< 0	1	0	Static: con- verges to fixed point

Class	h_{top}	$\kappa(O)$	λ_{max}	$r(\Sigma)$	$u(\Sigma)$	Informal
2	0	0	≤ 0	1	0	Periodic: finite-period orbits
3	> 0	$< h_{\text{top}}$	≤ 0	1	0	Fractal: self-similar, reducible
4	> 0	> 0	≈ 0	0	1	Complex: irreducible, universal
5	max	max	> 0	0	0	Random: maximal complexity, no structure

2.3 Key Distinctions

Class 3 vs. Class 4: Both have positive topological entropy, but Class 3 is computationally reducible ($r = 1$) while Class 4 is not. This captures the intuition that fractal systems, however visually complex, are “shortcuttable” — you can compute the Sierpinski triangle’s state at time 10^{100} without running 10^{100} steps, because the structure is self-similar and the rule is algebraically tractable (Wolfram, 2002, p. 870). Class 4 systems resist this: Rule 110 cannot be shortcut (Cook, 2004).

Class 4 vs. Class 5: Both are computationally irreducible, but Class 4 is computationally universal ($u = 1$) while Class 5 is not. A truly random system cannot simulate a specific Turing machine because simulation requires deterministic control over the output — which contradicts maximal Kolmogorov complexity. Class 4 achieves maximum *structured* complexity; Class 5 achieves maximum *unstructured* complexity.

The Lyapunov signature: Class 4 has $\lambda_{\text{max}} \approx 0$ (edge of chaos), which is the formal expression of Langton’s (1990) observation that Class 4 lives at the boundary between order ($\lambda_{\text{max}} < 0$) and chaos ($\lambda_{\text{max}} > 0$). Class 5 has $\lambda_{\text{max}} > 0$ — sensitivity to initial conditions without the structured information processing that characterizes Class 4.

2.4 The Class 4 Expressibility Theorem

The cosmological model’s first premise — that Class 4 is the maximum complexity achievable by expressible rules — requires a formal statement and proof sketch.

Claim (Expressibility Ceiling). Let $\Sigma = (S, f, s_0)$ be a dynamical system whose transition function f is specified by a finite description of length L (a “rule” in the Wolfram sense). Then the Kolmogorov complexity rate of the orbit satisfies:

$$\kappa(O) \leq L / t \rightarrow 0 \text{ as } t \rightarrow \infty$$

That is: the long-run information rate of any finitely-specified system is zero. The orbit is always compressible to “apply rule f from initial condition s_0 .”

Corollary. A system with $\kappa(O) > 0$ in the infinite limit (Class 5) cannot have a finite rule specification. Class 5 requires rules of infinite Kolmogorov complexity — rules that cannot be written down.

Implication. Class 4 — computationally universal, irreducible, with structured positive-entropy dynamics — is the most complex behavior achievable by finitely specifiable systems. It is the expressibility ceiling.

This argument needs rigorous formulation by a computability theorist. The key technical challenge is making the transition from Kolmogorov complexity (defined for finite strings) to the asymptotic complexity rate of infinite orbits rigorous. The relevant tools are algorithmic information theory (Downey & Hirschfeldt, 2010) and the theory of computable measures (Zvonkin & Levin, 1970).

2.5 Extension to Continuous Systems

The physical universe is not a discrete cellular automaton — it is (at least approximately) a continuous dynamical system on a field-theoretic state space. The five-class definitions must extend to this setting.

The key observation: the five measures (h_{top} , κ , λ_{max} , r , u) are all defined for continuous dynamical systems, not just discrete automata:

- Topological entropy extends to flows (Bowen, 1971).
- Kolmogorov complexity rate extends via the theory of algorithmic randomness for real-valued sequences (Hertling & Weihrauch, 2003).
- Lyapunov exponents are standard for ODEs and PDEs (Eckmann & Ruelle, 1985).
- Computational reducibility extends to the computability of solutions: can the state at time t be computed without integrating through all intermediate times? (Pour-El & Richards, 1989).
- Computational universality for continuous systems: certain PDEs are Turing-complete (Moore, 1991; Cardona et al., 2021).

The classification table (Section 2.2) therefore applies directly to continuous systems, with the caveat that the measures are harder to compute and the class boundaries become less sharp (continuous systems may interpolate between classes in parameter space, with Class 4 occupying a critical manifold between the ordered and chaotic phases).

2.6 The Containment Hierarchy

The verbal model claims that Class 4 contains all lower classes as subprocesses. Formally:

Claim (Containment). Let Σ_4 be a Class 4 system. For each $k \in \{1, 2, 3\}$, there exists a subsystem $\Sigma_k \subset \Sigma_4$ and a projection π_k such that $\pi_k(\Sigma_4)$ is Class k .

This follows from computational universality: a Turing-complete system can simulate any computable dynamics, including Class 1 (constant output), Class 2 (periodic output), and Class 3 (fractal/reducible output). The Game of Life demonstrates this concretely — it contains still lifes (Class 1), oscillators (Class 2), and self-similar growth patterns (Class 3) as embedded phenomena.

The containment is strict: no Class k system for $k < 4$ can contain a Class 4 subsystem (a non-universal system cannot simulate a universal one).

3. Singularity Boundaries as Topological Objects

3.1 The Equivalence Relation

The cosmological model claims that singularities at all scales — Planck regime, particle interiors, event horizons, cosmological horizons, temporal endpoints — are instances of the same information boundary. This claim requires a formal equivalence relation.

Definition (Information Boundary). An *information boundary* is a codimension-1 surface B in a spacetime manifold M such that:

(IB1) **Information impermeability:** No signal crosses B . Formally, B is an apparent horizon: for every future-directed causal curve γ originating in the interior of B , γ does not exit the region bounded by B . (For cosmological horizons, “interior” and “exterior” reverse — but the impermeability condition holds from the observer’s side.)

(IB2) **Bekenstein saturation:** The information content of B satisfies $I(B) = A(B) / (4 l_P^2)$, where $A(B)$ is the surface area and l_P is the Planck length. The boundary stores the maximum information permitted by the Bekenstein-Hawking formula.

(IB3) **Computational domain bound:** B defines the boundary of a computational domain — the set of spacetime events whose states can influence and be influenced by each other through causal processes.

Definition (Boundary Equivalence). Two information boundaries B_1 and B_2 are *equivalent* ($B_1 \sim B_2$) if they satisfy the same three conditions (IB1–IB3) with identical information-theoretic properties up to scale:

$$I(B_1) / A(B_1) = I(B_2) / A(B_2) = 1 / (4 l_P^2)$$

This is an equivalence relation (reflexive, symmetric, transitive).

3.2 The Singularity Inventory as an Equivalence Class

The cosmological model’s claim becomes: all entries in the singularity inventory (Gruber, 2026a, Section 5.1) belong to the same equivalence class under \sim . Specifically:

Singularity	IB1	IB2	IB3
Planck regime	No sub-Planck measurement possible	Maximum information density at Planck scale	Bounds the computational domain from below
Particle “interiors”	Interiors inaccessible	Claimed: Planck-scale Bekenstein saturation	Particles are atomic computational units

Singularity	IB1	IB2	IB3
Event horizons	No signal escapes	Bekenstein-Hawking entropy = $A/4l_P^2$	Bounds the causally connected region
Cosmological horizon	No signal from beyond	Gibbons-Hawking entropy = $A/4l_P^2$	Bounds the observable universe
Big Bang	No “before” accessible	Maximum density/temperature → maximum information density	Bounds the temporal computational domain (past)
Heat death / Big Crunch	Infinite dilution or reconvergence	Boundary of the computational domain (future)	Bounds the temporal computational domain (future)

The strongest entries are event horizons (all three conditions established) and the cosmological horizon (Gibbons & Hawking, 1977). The weakest is the particle interior claim, which is a prediction of the model rather than an established result.

3.3 The Topological Characterization

Define the **boundary space** B as the set of all information boundaries in the spacetime manifold M . The equivalence relation \sim partitions B into equivalence classes. The model claims there is a single class: $[B]_{\sim} = B$ — all information boundaries are equivalent.

The boundary space inherits a topology from M . The scale-invariance claim becomes:

Claim (Scale Invariance of Boundary Structure). The boundary space B , equipped with the topology inherited from M , is connected. Moreover, there exists a continuous scaling map $\sigma_\lambda: B \rightarrow B$ for each $\lambda > 0$ such that $\sigma_\lambda(B) \sim B$ for all $B \in B$ and all λ .

This states that the boundary structure is self-similar: zooming in or out on a boundary produces another boundary in the same equivalence class. This is the formal expression of the model’s claim that singularities at all scales share the same information-theoretic character.

3.4 Connection to Existing Mathematics

The formalization connects to several established programs:

Black hole thermodynamics (Bekenstein, 1973; Hawking, 1975): The Bekenstein-Hawking entropy formula $S = A / (4l_P^2)$ is the paradigm case of IB2. The generalized second law — that the total entropy of matter plus the black hole boundary never decreases — is a conservation law across the boundary.

The holographic principle ('t Hooft, 1993; Susskind, 1995; Bousso, 2002): The covariant entropy bound $S(L) \leq A(B(L)) / (4l_P^2)$ for any light sheet L bounds the entropy of a region by its boundary area. This is the general statement of IB2.

Causal structure in general relativity (Penrose, 1979; Hawking & Ellis, 1973): The theory of causal boundaries, event horizons, and apparent horizons provides the mathematical framework for IB1. The key tools are Penrose-Carter diagrams, the notion of a trapped surface, and the theory of causal sets.

Loop quantum gravity (Rovelli, 2004; Ashtekar & Lewandowski, 2004): Area quantization in LQG — the result that area comes in discrete quanta proportional to l_P^2 — provides a natural interpretation of the Planck-scale boundary as a discrete information-storage surface. Each quantum of area stores approximately one bit.

The technical challenge is proving that the equivalence relation \sim is nontrivial — that the Planck-scale boundary, event horizons, and cosmological horizons demonstrably satisfy the same formal conditions, not merely analogous ones. This requires tools from quantum gravity that do not yet exist in complete form.

4. Holographic Rule Sets

4.1 The Informal Definition

The cosmological model distinguishes three relationships between holographic systems and Class 4 automata (Gruber, 2026a, Section 6.1):

1. A holographic substrate produces Class 4 dynamics.
2. A Class 4 automaton produces holographic output.
3. A Class 4 automaton whose rule structure is itself holographic.

Relationship 3 — the holographic rule set — is the most novel and the least well-defined. What does it mean for a dynamical system’s rules to be “holographic”?

4.2 Dimensional Compression Operators

Definition (Holographic Rule Set). A rule set R for a dynamical system on a d -dimensional manifold M^d is *holographic* if there exists a dimensional compression operator $C: R^{(d)} \rightarrow R^{(d-1)}$ such that:

(HR1) **Compression:** The full d -dimensional rule $R^{(d)}$ can be faithfully encoded in a $(d-1)$ -dimensional specification $R^{(d-1)} = C(R^{(d)})$, with no information loss.

(HR2) **Local decompression:** The d -dimensional dynamics at any point $x \in M^d$ can be recovered from $R^{(d-1)}$ restricted to the $(d-1)$ -dimensional boundary ∂M^d in a neighborhood of x . Local application of $R^{(d-1)}$ at each boundary site implicitly computes global (non-local) relationships in the interior.

(HR3) **Universality preservation:** If $R^{(d)}$ defines Class 4 dynamics on M^d , then $R^{(d-1)}$ on ∂M^d also defines Class 4 dynamics (the compression preserves computational universality).

4.3 Connection to AdS/CFT

The definition of a holographic rule set is directly modeled on the AdS/CFT correspondence (Maldacena, 1998), which provides the most concrete example:

- A gravitational theory in $(d+1)$ -dimensional anti-de Sitter space (the “bulk” — $R^{(d+1)}$)
- is fully equivalent to a conformal field theory on the d -dimensional boundary (the boundary theory — $R^{(d)}$)
- with the boundary theory encoding all bulk physics (HR1)

- and local boundary operators reconstructing bulk observables in their causal wedge (HR2; Ryu & Takayanagi, 2006; Dong, 2016)

The SB-HC4A model generalizes this: the universe’s rule structure is holographic not because of a specific AdS/CFT duality but because any Class 4 system at every scale has its information encoded on its boundary (the singularity surface) with the interior dynamics being a decompression of the boundary encoding.

4.4 The Rule Set as a Coalgebra

A more abstract characterization uses coalgebra theory (Rutten, 2000), which provides the natural mathematical language for systems whose behavior unfolds over time from a specification:

Define the state space X and a functor $F: \text{Set} \rightarrow \text{Set}$ that describes the system’s “type of behavior” (what can be observed at each step). A coalgebra for F is a pair $(X, \alpha: X \rightarrow F(X))$. The rule set R is the coalgebra structure map α — it specifies how the system’s state determines its next observable behavior.

A **holographic coalgebra** would then be a coalgebra (X, α) together with a boundary functor $B: \text{Set} \rightarrow \text{Set}$ and a natural transformation $\eta: F \Rightarrow B \circ G$ such that:

- G is a “restriction to boundary” functor
- η encodes the full coalgebra dynamics in the boundary restriction
- The coalgebra can be recovered from the boundary restriction

This is highly abstract and likely requires substantial development by a category theorist. The value of the coalgebraic formulation is that it makes the holographic property a *structural* property of the system’s specification — not a property of any particular state or trajectory, but of the rule structure itself.

4.5 Holographic Automata in the Literature

't Hooft (2016) proposed the Cellular Automaton Interpretation of quantum mechanics, in which quantum mechanics is emergent from deterministic automaton dynamics at the Planck scale. The holographic rule set, in 't Hooft’s framework, is the automaton rule at the Planck scale that produces quantum behavior as an effective description at larger scales. The dimensional compression (HR1) corresponds to the renormalization group flow from Planck-scale to observable-scale physics — each step in the RG flow “compresses” the full dynamics into an effective description at a coarser scale.

Verlinde’s (2011) entropic gravity proposal — that gravity is not a fundamental force but an entropic phenomenon arising from the holographic distribution of information on screens — provides another concrete example of a holographic rule set: the gravitational “rule” (Newton’s law, or its relativistic generalization) is derived from the information content of holographic screens.

5. Self-Referential Closure

5.1 The Fixed-Point Condition

The model’s central formal claim is self-referential closure:

$$\Phi(U) = U$$

where U is the SB-HC4A (the universe) and Φ is the “compute the output” operator. The holographic rules encode the system; the Class 4 dynamics decompress the encoding; the holographic output re-encodes the result; the computation and the system are the same thing.

This is a fixed-point statement. The mathematical program is to define Φ rigorously and prove (or specify the conditions under which) a fixed point exists.

5.2 Lawvere’s Fixed-Point Theorem

Lawvere’s (1969) fixed-point theorem generalizes the diagonal arguments of Cantor, Gödel, and Turing into a single category-theoretic statement: In a Cartesian closed category C , if there exists a point-surjective morphism $e: A \rightarrow B^A$ (where B^A is the exponential object), then every endomorphism $f: B \rightarrow B$ has a fixed point.

The cosmological application: Let C be the category of computable dynamical systems. Let A = the set of specifications (rule sets) and B = the set of dynamical behaviors (orbits). A point-surjective morphism $e: A \rightarrow B^A$ means “every map from specifications to behaviors can be specified” — i.e., the system is expressive enough to encode the relationship between rules and dynamics.

If the SB-HC4A is computationally universal (Class 4, $u = 1$), it can simulate any computable dynamics — including its own. The point-surjectivity condition is therefore satisfied. By Lawvere’s theorem, every endomorphism on the behavior space has a fixed point. In particular, the “compute and re-encode” operator Φ has a fixed point: $\Phi(U) = U$.

5.3 Coalgebraic Fixed Points

An alternative formalization uses the theory of final coalgebras (Aczel & Mendler, 1989; Barr, 1993). For a functor F , the final coalgebra — if it exists — is the terminal object in the category of F -coalgebras. It has the property that it is the unique (up to isomorphism) coalgebra that is “maximally consistent”: it contains all possible behaviors and is a fixed point of the coalgebra-building process.

Define F as the functor capturing the SB-HC4A’s behavior type: $F(X) = (\text{holographic boundary information}) \times (\text{Class 4 interior dynamics})$. The final F -coalgebra, if it exists, would be the unique self-consistent SB-HC4A: the system that is its own final coalgebra — whose behavior IS its specification.

The fixed-point condition $\Phi(U) = U$ becomes: U is the final coalgebra for the SB-HC4A behavior functor. This is a stronger claim than merely “a fixed point exists” — it claims that U is the *unique* self-consistent system of this type.

5.4 Connection to the FMT Fixed Point

The FMT formalization (Gruber, 2026c, Section 6.3) defines the consciousness fixed point as: $\Phi_c(m) = m$, where Φ_c is the “model of” operator and m^* is the ESM state at which the model and the modeled coincide. The cosmological model claims this is the same formal structure at a different scale.

To make this precise, define:

- $\Phi_{\text{cosmo}}: U_{\text{cosmo}} \rightarrow U_{\text{cosmo}}$ (the cosmological “compute the output” operator)
- $\Phi_{\text{consciousness}}: M_{\text{ESM}} \rightarrow M_{\text{ESM}}$ (the consciousness “model of” operator)

The structural identity claim is that there exists a structure-preserving map (a functor; see Section 6) that maps Φ_{cosmo} to $\Phi_{\text{consciousness}}$. Both operators are self-referential endomorphisms on their respective domains, and both have fixed points. The claim is that the fixed-point structure is isomorphic — not merely analogous.

5.5 Inexpressibility as a Formal Consequence

Godel's first incompleteness theorem (Godel, 1931): Any consistent formal system F that is sufficiently expressive (can represent arithmetic) contains statements that are true but not provable in F .

The SB-HC4A is computationally universal (Class 4), hence sufficiently expressive. If the SB-HC4A is self-referentially closed ($\Phi(U) = U$), then U is a formal system that contains itself as a subsystem. By Godel's theorem, there exist truths about U that cannot be proven from within U .

More specifically, the **Weltformel** (world equation) — a complete specification of U — cannot be a statement within U , because U 's own self-referential structure guarantees the existence of truths about U that U cannot prove. The “world equation” is therefore not an equation but the process U itself — it can only be expressed by running it.

This connects to Chaitin's (1966) extension of Godel: no formal system of complexity K can prove theorems about systems of complexity greater than $K + c$ (for a constant c). Since U is at least as complex as any internal formal system, no internal specification can fully capture U .

6. The Consciousness-Cosmology Functor

6.1 Two Categories

The consciousness-cosmology structural identity (Gruber, 2026a, Section 7) is currently expressed as a table of correspondences. To make this a mathematical claim, we need a morphism between mathematical objects.

Following the FMT formalization's category-theoretic approach (Gruber, 2026c, Section 7), define two categories:

- **Cosmo** (Cosmological): Objects are physical states of the universe (field configurations, metric tensors, matter distributions). Morphisms are physical dynamics — time evolution governed by the laws of physics.
- **Consc** (Consciousness): Objects are the FMT model-density states $\rho(s, \nu, t)$ (from the FMT formalization). Morphisms are the dynamics of the model density — the Fokker-Planck equation of the FMT formalization, substrate dynamics, permeability changes.

6.2 The Structural Identity Functor

The structural identity claim is that there exists a functor $I: \text{Cosmo} \rightarrow \text{Consc}$ that preserves the relevant structure:

Cosmo	I	Consc
Singularity boundary	\mapsto	Implicit-explicit boundary

Cosmo	I	Consc
Observable interior	\mapsto	Explicit models
Holographic rule structure	\mapsto	Distributed implicit knowledge
Class 4 dynamics	\mapsto	Cortical criticality
$\Phi_{\text{cosmo}}(U) = U$	\mapsto	$\Phi_c(m) = m$
Information conservation across boundary	\mapsto	Information conservation across implicit-explicit split
Godelian inexpressibility	\mapsto	Meta-Problem

For I to be a proper functor, it must:

(F1) **Map objects to objects:** Each physical state maps to a model-density state.

(F2) **Map morphisms to morphisms:** Physical dynamics map to model-density dynamics, preserving composition (if state A evolves to B and B to C in Cosmo, the mapped states evolve correspondingly in Consc).

(F3) **Preserve identity:** The identity morphism in Cosmo (no change) maps to the identity morphism in Consc.

6.3 What the Functor Preserves

The functor I is not expected to preserve all structure — the cosmological scale and the consciousness scale differ in their physical content, their energy ranges, their characteristic time scales, and their spatial extent. What I preserves is the *computational architecture*:

1. **The boundary structure:** The functor maps information-impermeable boundaries to information-opaque boundaries.
2. **The compression-decompression relationship:** The functor maps holographic encoding/decoding to implicit/explicit information processing.
3. **The criticality condition:** The functor maps Class 4 dynamics to Class 4 dynamics (both systems operate at the edge of chaos).
4. **The fixed-point structure:** The functor maps self-referential closure to self-referential closure.

More precisely, the functor should be a **forgetful functor** that discards scale-specific content while preserving computational-architectural structure. The category-theoretic framework for this is the theory of **abstract model theory** (Barwise, 1974) or **institutions** (Goguen & Burstall, 1992), which formalize the notion of “same structure, different content.”

6.4 The Scale Functor

The fractal nature of the SB-HC4A suggests not just a single functor but a family of functors parameterized by scale:

$$I_\lambda: \text{Cosmo}_\lambda \rightarrow \text{Cosmo}_{\{\lambda'\}}$$

where Cosmo_λ is the category of cosmological dynamics at scale λ . The self-similarity of Class 4 systems means that each scale hosts the same computational architecture. The consciousness-cosmology functor is then $I_{\text{brain}}: \text{Cosmo}_{\{l_{\text{brain}}\}} \rightarrow \text{Consc}$ — the specialization of the scale functor at the brain-relevant scale ($\sim 10^{-2}$ m).

This formulation makes the fractal claim precise: the universe's computational architecture is scale-invariant, and consciousness is the instance at one particular scale. The scale functor I_λ implements the self-similar nesting that Class 4 dynamics produce (Class 4 contains Class 3/fractal as a subprocess — the scale invariance IS a Class 4 subprocess).

7. The Necessity Argument

7.1 The Five Axioms in Formal Language

The cosmological model's necessity argument (Gruber, 2026a, Section 10) claims that the SB-HC4A is the unique structure consistent with five axioms. Formalizing this requires expressing the axioms and the derivation in a formal logical framework.

The natural tool is modal logic — specifically, the S5 system for metaphysical necessity (Kripke, 1963), extended with predicates for computational and information-theoretic properties.

Formal axiom system:

Let the domain of discourse be the class of all possible dynamical systems. Define predicates:

- $E(x)$: “ x exists” (is physically instantiated)
- $C_k(x)$: “ x has computational class k ” for $k \in \{1, 2, 3, 4, 5\}$
- $U(x)$: “ x is computationally universal”
- $SOC(x)$: “ x self-maintains criticality via self-organized criticality”
- $Contains(x, k)$: “ x contains a Class k subsystem”
- $IB(x)$: “ x is bounded by information-impermeable boundaries at every scale”
- $Hol(x)$: “ x has holographic structure”
- $SRC(x)$: “ x is self-referentially closed”

Axiom A1 (Ontological Necessity): $\square \exists x E(x)$ (Necessarily, something exists.)

Axiom A2 (Computational Character): $\forall x [E(x) \rightarrow \exists k C_k(x)]$ (Everything that exists has a computational class.)

Axiom A3 (Criticality Stability): $\forall x [C_4(x) \leftrightarrow (U(x) \wedge SOC(x) \wedge \forall k < 4 Contains(x, k))]$ (Class 4 is uniquely characterized by universality + self-maintenance + containment of lower classes.)

Axiom A4 (Information Bound): $\forall x [E(x) \rightarrow IB(x)]$ (Everything that exists is bounded by information horizons.)

Axiom A5 (Holographic Encoding): $\forall x [IB(x) \rightarrow Hol(x)]$ (Information boundaries encode interior information holographically.)

7.2 The Derivation

Theorem (SB-HC4A Necessity). From A1–A5:

$$\square \exists x [E(x) \wedge C_4(x) \wedge IB(x) \wedge Hol(x) \wedge SRC(x)]$$

Proof sketch:

1. By A1: $\square \exists x E(x)$. Something necessarily exists.
2. By A2: This thing has some computational class k .

3. The elimination argument (Section 3.2 of Gruber, 2026a):
 - $k = 1$ or $k = 2$: The universe contains consciousness and universal computation. By A3, only Class 4 supports these. A Class 1 or 2 system cannot contain a Class 4 subsystem (strict containment hierarchy). Eliminated.
 - $k = 3$: Class 3 is computationally reducible. Universal computation requires irreducibility. Eliminated.
 - $k = 5$: If $k = 5$, the system has no expressible rules (Section 2.4). Physics — the expressible-rule description of the universe — would be impossible. By A2, the system has computational character; Class 5 makes this character inexpressible. Eliminated (by coherence with A2, or by abduction: physics is not impossible).
 - Therefore $k = 4$.
4. By A4: The Class 4 system is bounded by information horizons.
5. By A5: The boundaries encode the interior holographically.
6. By A3 (universality) + step 5 (holographic encoding): A universal system with holographic encoding can encode its own rules on its boundary. The holographic rules produce Class 4 dynamics that produce holographic output — a self-referential loop. By Lawvere's theorem (Section 5.2), this loop has a fixed point: $\text{SRC}(x)$. \square

7.3 The Uniqueness Claim

Theorem (Uniqueness). The SB-HC4A is the unique (up to isomorphism) structure satisfying all five axioms simultaneously.

Proof strategy: Show that removing any axiom admits alternatives, and that the conjunction of all five constrains the system to a single architectural class:

- Without A1: Nothing need exist — vacuous satisfaction.
- Without A2: The existent thing need not have computational character — no classification possible.
- Without A3: Any computational class is permissible — no specific dynamics predicted.
- Without A4: No information boundaries — no singularity structure.
- Without A5: No holographic encoding — no dimensional compression.

The conjunction forces: Class 4 dynamics + information boundaries at every scale + holographic encoding + self-referential closure. This is precisely the SB-HC4A definition.

7.4 Strengths and Weaknesses of the Formal Argument

Strengths: The argument is valid if the axioms and the elimination are accepted. The conclusion follows logically. The use of modal logic makes the necessity claim precise — it is not “the universe happens to be SB-HC4A” but “any possible physical universe must be SB-HC4A.”

Weaknesses:

1. **A1 is metaphysical**, not mathematical. “Something exists” is not a theorem of any formal system — it is an axiom about reality. The formalization makes this transparent but cannot justify it.
2. **The elimination of Class 5** relies on abduction (physics would be impossible), not deduction. A Class 5 universe that locally appears Class 4 to Class 4 observers is logically consistent — just explanatorily catastrophic. The modal-logical formalization can express this as: the

necessity holds *within the scope of A2* (that the system's dynamics are expressible), which may be question-begging if the universe is Class 5 with inexpressible dynamics.

3. **A3 encodes a substantive claim** (that Class 4 is uniquely self-maintaining) as an axiom. A skeptic could reject A3 — perhaps multiple classes are self-maintaining under different conditions. The axiom's justification is empirical (self-organized criticality is observed in Class 4 systems) and theoretical (universality + containment are unique to Class 4), but it is not a logical necessity.
 4. **A5 (holographic encoding)** is well-supported for gravitational systems (AdS/CFT, black hole thermodynamics) but is not proven as a universal principle for all physical systems. The formalization makes this dependence explicit.
-

8. Energy-Information Equivalence

8.1 The Duality

The cosmological model's Weak Point 1 (Gruber, 2026a, Section 9.1) is the unproven energy-information identity: $E = I$. The formalization must specify what this means and what it would take to prove.

Definition (Energy-Information Duality). Energy and information are dual descriptions of the same physical quantity if there exists a bijection $D: E \rightarrow I$ (where E is the set of energy configurations and I is the set of information configurations) such that:

(EI1) **Conservation equivalence:** Energy conservation (first law of thermodynamics) is equivalent to information conservation (unitarity of quantum mechanics, no-cloning theorem).

(EI2) **Bound equivalence:** The Bekenstein bound $I_{\max} = 2\pi RE / (\hbar c \ln 2)$ is a definitional relationship, not a constraint between independent quantities.

(EI3) **Transformation equivalence:** Every energy transformation corresponds to an information transformation and vice versa, with Landauer's principle ($\Delta E \geq kT \ln 2$ per bit erased) as the marginal exchange rate.

8.2 Existing Support

Several results in physics point toward $E = I$:

- **Landauer's principle** (Landauer, 1961; experimentally confirmed by Berut et al., 2012): Erasing information has a minimum energy cost. This establishes a lower bound on the energy-information exchange rate.
- **The Bekenstein bound** (Bekenstein, 1981): Maximum information content of a region is proportional to its energy content and its radius. If E and I were independent, this bound would be a remarkable coincidence. If $E = I$, it is a tautology.
- **Black hole thermodynamics** (Bekenstein, 1973; Hawking, 1975): Black holes have entropy $S = A / (4l_P^2)$, temperature $T = \hbar c^3 / (8\pi GMk)$, and satisfy the laws of thermodynamics. The information content of a black hole is its energy content, up to dimensional constants.

- **The holographic principle** ('t Hooft, 1993; Susskind, 1995): The maximum entropy (= information) of a region scales with surface area, not volume. This is a statement about information density that is simultaneously a statement about energy density.
- **The ER = EPR conjecture** (Maldacena & Susskind, 2013): Einstein-Rosen bridges (wormholes) are equivalent to Einstein-Podolsky-Rosen entanglement (quantum information). Geometry (energy) and entanglement (information) are the same thing.

8.3 The Formalization Strategy

A full derivation of $E = I$ would require a fundamental theory of quantum gravity. Short of that, the formalization can specify:

1. **The equivalence conditions** (EI1–EI3) as testable predictions.
 2. **The dimensional analysis:** E and I have different units (joules vs. bits). The conversion factor involves fundamental constants: 1 bit = $kT \ln 2$ joules (at temperature T). At the Planck temperature ($T_P = \sqrt{(\hbar c^5 / G k^2)} \approx 1.4 \times 10^{32}$ K), 1 bit = $E_P = \sqrt{(\hbar c^5 / G)} \approx 1.22 \times 10^9$ J. This is the Planck energy — the natural unit where E and I become the same number.
 3. **The category-theoretic formulation:** E and I are two representations of the same underlying object, related by a natural isomorphism. The category of energy configurations and the category of information configurations are equivalent categories.
-

9. The Cognitive Ceiling as a Formal Constraint

9.1 The Problem Restated

The deepest objection to the SB-HC4A model (Gruber, 2026a, Section 9.5) is that Class 4 observers may be constitutionally incapable of determining whether the universe is Class 4 or merely appears Class 4 to Class 4 cognition. The formalization must make this objection precise.

9.2 Computability-Theoretic Formulation

Definition (Cognitive Ceiling). A Class k observer can only detect dynamics up to Class k . Formally: let O_k be an observer (measurement/modeling apparatus) of computational class k . Let U be the universe. The observer's model of U is:

$$M(O_k, U) = \pi_k(U)$$

where π_k is a projection operator that maps U 's dynamics onto the subspace of Class $\leq k$ behavior. If U is Class $j > k$, then $M(O_k, U) = \pi_k(U) \neq U$ — the observer's model is a strict lower-class projection of reality.

Theorem (Indistinguishability). A Class 4 observer cannot distinguish between: - (a) U is Class 4 (and the model is accurate), and - (b) U is Class 5, but $\pi_4(U)$ appears Class 4 (and the model is a projection artifact).

Proof sketch: By definition, $M(O_4, U) = \pi_4(U)$. If U is Class 4, then $\pi_4(U) = U$ and M is faithful. If U is Class 5, then $\pi_4(U) \neq U$ but $M(O_4, U) = \pi_4(U)$ is still Class 4. The observer sees Class 4 in both cases. \square

9.3 The Self-Referential Trap

This connects to Gödel's incompleteness: the SB-HC4A, if self-referentially closed, contains truths about itself that it cannot prove from within (Section 5.5). Whether U is Class 4 or Class 5 may be one such undecidable truth — a fact about the universe that no internal formal system can determine.

Formally: define the predicate $\text{True_Class}(U, k) = \text{"the universe's true computational class is } k\text{"}$. The cognitive ceiling theorem states:

$$\neg \exists \text{ proof } P \text{ within } U: [P \text{ proves } \text{True_Class}(U, 4)] \vee [P \text{ proves } \text{True_Class}(U, 5)]$$

if U is self-referentially closed and sufficiently expressive (by Gödel's first incompleteness theorem applied to the self-referential system).

9.4 The Model's Self-Prediction

This is the most remarkable formal feature: the SB-HC4A model *predicts its own potential unfalsifiability*. The cognitive ceiling is not an external objection brought against the model — it is a consequence of the model's own self-referential structure. If the model is correct, then:

1. The universe is a self-referentially closed Class 4 system.
2. Class 4 observers cannot determine whether the universe is Class 4 or Class 5.
3. Therefore, Class 4 observers cannot verify the model — which is exactly what a self-referentially closed system predicts about subsystems' ability to fully characterize the whole.

This is either the deepest confirmation (the model's epistemological predictions are consistent with its ontological claims) or the deepest flaw (the model immunizes itself against falsification). The formalization makes this dilemma precise but cannot resolve it — resolution requires tools from outside the model's own computational class, which, by the cognitive ceiling, are unavailable.

9.5 Connection to the Meta-Problem

The FMT identifies the Meta-Problem of consciousness (Chalmers, 2018) as structurally identical: consciousness cannot fully model its own substrate because the ESM's self-model is a projection (analogous to π_4), not a faithful representation.

Formally: the ESM's model of the substrate is $M(\text{ESM}, \text{substrate}) = \pi_{\text{explicit}}(\text{substrate}) \neq \text{substrate}$. The substrate has implicit structure that the explicit model cannot access — just as a Class 5 universe would have structure that a Class 4 observer cannot detect.

The consciousness-cosmology functor (Section 6) maps the cosmological cognitive ceiling to the consciousness Meta-Problem:

$$I(\text{cognitive ceiling in Cosmo}) = \text{Meta-Problem in Consc}$$

This is the formal expression of the paper's central symmetry claim: the same epistemological limitation operates at both scales, generated by the same self-referential architecture.

10. Phased Build Order

The formalization project is substantial. A pragmatic build sequence, ordered by the maturity of the relevant mathematical tools and the proximity to existing physics:

Phase 1: Highest Priority (Connects to Existing Mathematical Physics)

Module 3 (Sections 3.1–3.4) — Singularity boundaries as equivalence classes. The mathematical tools exist: causal structure theory (Hawking & Ellis, 1973), the holographic principle ('t Hooft, 1993; Bousso, 2002), black hole thermodynamics (Bekenstein, 1973; Hawking, 1975), and loop quantum gravity's area quantization (Rovelli, 2004). The task is to define the equivalence relation precisely and verify that the known singularity types satisfy it. This is a paper-length project for a mathematical relativist.

Module 5.2–5.3 (Self-referential closure via Lawvere) — The fixed-point argument. Lawvere's theorem is proven. The application to the SB-HC4A requires specifying the category, the exponential object, and the point-surjection. This is a paper-length project for a category theorist familiar with Lawvere's work and with background in mathematical physics.

Phase 2: Core Formalism (Requires Dedicated Development)

Module 2 (Section 2) — Measure-theoretic class definitions. The individual measures (topological entropy, Kolmogorov complexity rate, Lyapunov exponents, computational reducibility) are well-defined. The challenge is proving that they jointly characterize the five classes as claimed — particularly the Class 3/4 boundary (reducibility vs. irreducibility) and the Class 4/5 boundary (universality vs. non-universality). The Expressibility Ceiling claim (Section 2.4) needs a rigorous proof by an algorithmic information theorist.

Module 4 (Section 4) — Holographic rule sets. The AdS/CFT correspondence provides the paradigm case, but the general definition (HR1–HR3) needs development. The coalgebraic formulation (Section 4.4) is the most promising approach but requires original category-theoretic work. Collaboration with researchers working on the interface of quantum gravity and category theory (cf. Baez & Stay, 2011; Abramsky & Coecke, 2004) is recommended.

Phase 3: Deep Formalism (Hardest, Highest Potential Impact)

Module 6 (Section 6) — The consciousness-cosmology functor. This requires both the FMT formalization (Gruber, 2026c) and the cosmological formalization to be sufficiently developed. The functor construction is the mathematical expression of the model's strongest claim — that consciousness and cosmology implement the same architecture. Demonstrating that this functor exists, is non-trivial, and preserves the claimed structural properties would be a substantial mathematical achievement.

Module 7 (Section 7) — The necessity argument in modal logic. The formal argument depends on the formalization of the axioms and the elimination. The hardest part is the elimination of Class 5 (abductive, not deductive) and the justification of A3 (that Class 4 is uniquely self-maintaining). A modal logician working with a computability theorist could assess whether the argument is formally valid and what assumptions it requires.

Phase 4: Speculative Extensions

Module 8 (Section 8) — Energy-information equivalence. This depends on progress in quantum gravity. The formalization can specify the equivalence conditions and the testable predictions, but a proof of $E = I$ is beyond current physics.

Module 9 (Section 9) — Cognitive ceiling formalization. The computability-theoretic formulation is straightforward; the profound question (whether the model’s self-predicted unfalsifiability is a strength or a weakness) is philosophical, not mathematical. The formalization’s contribution is to make the question precise.

11. What Formalization Buys — And What It Cannot

11.1 What It Buys

Precision on the central claims. The verbal model says “singularities are structurally identical at all scales.” The formalization says: “singularities belong to the same equivalence class under the relation \sim , defined by conditions IB1–IB3.” The verbal model says “the consciousness-cosmology mapping is not analogy but structural identity.” The formalization says: “there exists a functor $I: \text{Cosmo} \rightarrow \text{Consc}$ preserving computational-architectural structure.” These are not interchangeable statements — the formal versions are testable and falsifiable in ways the verbal versions are not.

Constraint on the model. A verbally specified “self-referential closure” can mean almost anything. The Lawvere fixed-point formalization constrains it: self-referential closure requires a specific category-theoretic structure (a Cartesian closed category with a point-surjective morphism). If the SB-HC4A’s dynamics do not satisfy this structure, the fixed-point argument fails. The formalization turns a vague philosophical claim into a checkable mathematical condition.

Interoperability with physics. The singularity equivalence relation connects to black hole thermodynamics, the holographic principle, and loop quantum gravity. The holographic rule set definition connects to AdS/CFT. The Class 4 characterization connects to the theory of dynamical systems at criticality. The formalization positions the SB-HC4A model within the existing landscape of mathematical physics, rather than floating free as a verbal speculation.

Connection to the FMT and RIM formalizations. The consciousness-cosmology functor (Section 6) and the cognitive ceiling / Meta-Problem connection (Section 9.5) formally link the three-level theoretical architecture: cosmological substrate \rightarrow consciousness \rightarrow intelligence. The three formalization papers, taken together, specify a single mathematical framework that connects cosmology, consciousness, and intelligence through category-theoretic morphisms.

Sharpened objections. The cognitive ceiling formalization (Section 9) makes the model’s deepest vulnerability precise: the indistinguishability theorem (Section 9.2) states exactly what a Class 4 observer cannot determine. This is more useful than the verbal formulation — it specifies the exact conditions under which the model is unfalsifiable and the exact sense in which this is a structural prediction rather than an evasion.

11.2 What It Cannot

The formalization cannot prove the axioms. A1 (something exists) is metaphysical. A3 (Class 4 is uniquely self-maintaining) is empirical. A5 (holographic encoding is universal) is conjectural.

The formalization derives consequences from axioms; it does not justify the axioms themselves. The model is only as strong as its weakest axiom.

The formalization cannot resolve the cognitive ceiling. If the indistinguishability theorem (Section 9.2) is correct, no amount of formal work conducted within the universe can determine whether the universe is truly Class 4 or merely appears so. The formalization makes this limitation precise but cannot overcome it — overcoming it would require computational resources exceeding Class 4, which are by definition unavailable.

The formalization cannot prove $E = I$. Energy-information equivalence depends on a theory of quantum gravity. The formalization specifies what $E = I$ means and what its consequences would be, but the proof (if one is possible) belongs to fundamental physics, not to the formalization program.

The formalization cannot close the gap between structure and ontology. Demonstrating that a formal architecture is self-consistent does not demonstrate that the universe instantiates that architecture. Mathematics can show that the SB-HC4A is a well-defined, self-consistent object. Physics must determine whether the universe is an SB-HC4A.

Even the best formalization cannot prove that the SB-HC4A model describes reality rather than the cognitive ceiling of its human author’s Class 4 brain. Its value is to make the model’s claims precise enough to be *clearly right or clearly wrong* — which is, in the end, the only honest standard a theory can meet.

12. Conclusion

The SB-HC4A cosmological model requires mathematical formalization to constrain its claims, connect to existing mathematical physics, and sharpen both its predictions and its acknowledged vulnerabilities. The formalization strategy proposed here addresses the model’s core open questions: What does “Class 4” mean precisely? What is the equivalence relation on singularity boundaries? What does it mean for a rule set to be holographic? What category-theoretic structure makes self-referential closure rigorous? In what formal sense is the consciousness-cosmology mapping a structural identity rather than an analogy?

Eight modules are proposed, spanning measure theory (class definitions), topology (singularity boundaries), category theory (holographic rule sets, self-referential closure, the consciousness-cosmology functor), modal logic (the necessity argument), thermodynamic information theory (energy-information equivalence), and computability theory (the cognitive ceiling). The modular design allows domain specialists to contribute independently, with category theory providing the integration language.

The formalization also achieves a specific meta-theoretical result: it makes precise the model’s prediction of its own potential unfalsifiability. The indistinguishability theorem (Section 9.2) — that Class 4 observers cannot distinguish between a Class 4 universe and a Class 5 universe that locally appears Class 4 — is either the model’s deepest structural prediction or its deepest flaw. The formalization cannot resolve this dilemma, but it can state it with mathematical precision, which is the prerequisite for any future resolution.

This paper specifies the program. Its execution requires collaboration across mathematical physics, category theory, computability theory, and modal logic — and the intellectual honesty to discover

that some of these formalizations may reveal the model to be wrong in specific, identifiable ways. That would be a feature, not a bug.

References

- Abramsky, S., & Coecke, B. (2004). A categorical semantics of quantum protocols. *Proceedings of the 19th Annual IEEE Symposium on Logic in Computer Science*, 415–425.
- Aczel, P., & Mendler, N. (1989). A final coalgebra theorem. *Lecture Notes in Computer Science*, 389, 357–365.
- Adler, R. L., Konheim, A. G., & McAndrew, M. H. (1965). Topological entropy. *Transactions of the American Mathematical Society*, 114(2), 309–319.
- Ashtekar, A., & Lewandowski, J. (2004). Background independent quantum gravity: A status report. *Classical and Quantum Gravity*, 21(15), R53–R152.
- Baez, J. C., & Stay, M. (2011). Physics, topology, logic and computation: A Rosetta Stone. *Lecture Notes in Physics*, 813, 95–172.
- Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. Springer.
- Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, 59(4), 381–384.
- Barr, M. (1993). Terminal coalgebras in well-founded set theory. *Theoretical Computer Science*, 114(2), 299–315.
- Barwise, J. (1974). Axioms for abstract model theory. *Annals of Mathematical Logic*, 7(2–3), 221–265.
- Beggs, J. M., & Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *Journal of Neuroscience*, 23(35), 11167–11177.
- Bekenstein, J. D. (1973). Black holes and entropy. *Physical Review D*, 7(8), 2333–2346.
- Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D*, 23(2), 287–298.
- Berut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R., & Lutz, E. (2012). Experimental verification of Landauer’s principle linking information and thermodynamics. *Nature*, 483, 187–189.
- Bousso, R. (2002). The holographic principle. *Reviews of Modern Physics*, 74(3), 825–874.
- Bowen, R. (1971). Entropy for group endomorphisms and homogeneous spaces. *Transactions of the American Mathematical Society*, 153, 401–414.
- Cardona, R., Miranda, E., Peralta-Salas, D., & Presas, F. (2021). Constructing Turing complete Euler flows in dimension 3. *Proceedings of the National Academy of Sciences*, 118(19), e2026818118.
- Chaitin, G. J. (1966). On the length of programs for computing finite binary sequences. *Journal of the ACM*, 13(4), 547–569.

- Chalmers, D. J. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies*, 25(9-10), 6–61.
- Cook, M. (2004). Universality in elementary cellular automata. *Complex Systems*, 15(1), 1–40.
- Dong, X. (2016). The gravity dual of Renyi entropy. *Nature Communications*, 7, 12472.
- Downey, R. G., & Hirschfeldt, D. R. (2010). *Algorithmic Randomness and Complexity*. Springer.
- Eckmann, J. P., & Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, 57(3), 617–656.
- Gibbons, G. W., & Hawking, S. W. (1977). Cosmological event horizons, thermodynamics, and particle creation. *Physical Review D*, 15(10), 2738–2751.
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38, 173–198.
- Goguen, J. A., & Burstall, R. M. (1992). Institutions: Abstract model theory for specification and programming. *Journal of the ACM*, 39(1), 95–146.
- Gruber, M. (2015). *Die Emergenz des Bewusstseins*. Self-published.
- Gruber, M. (2026a). The Singularity-Bounded Holographic Class 4 Automaton: A cosmological model from consciousness theory. Manuscript.
- Gruber, M. (2026b). The Four-Model Theory of Consciousness: A simulation-based framework unifying the Hard Problem, binding, and altered states. *Zenodo* preprint. <https://doi.org/10.5281/zenodo.18669891>
- Gruber, M. (2026c). Toward a mathematical formalization of the Four-Model Theory: A recommended approach. Manuscript.
- Gruber, M. (2026d). Toward a mathematical formalization of the Recursive Intelligence Model: A recommended approach. Manuscript.
- Hawking, S. W. (1975). Particle creation by black holes. *Communications in Mathematical Physics*, 43, 199–220.
- Hawking, S. W., & Ellis, G. F. R. (1973). *The Large Scale Structure of Space-Time*. Cambridge University Press.
- Hertling, P., & Weihrauch, K. (2003). Random elements in effective topological spaces with measure. *Information and Computation*, 181(1), 32–56.
- 't Hooft, G. (1993). Dimensional reduction in quantum gravity. In *Salamfestschrift* (pp. 284–296). World Scientific.
- 't Hooft, G. (2016). *The Cellular Automaton Interpretation of Quantum Mechanics*. Springer.
- Kripke, S. A. (1963). Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16, 83–94.
- Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 183–191.
- Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D*, 42(1–3), 12–37.

- Lawvere, F. W. (1969). Diagonal arguments and Cartesian closed categories. *Lecture Notes in Mathematics*, 92, 134–145.
- Maldacena, J. (1998). The large-N limit of superconformal field theories and supergravity. *Advances in Theoretical and Mathematical Physics*, 2(2), 231–252.
- Maldacena, J., & Susskind, L. (2013). Cool horizons for entangled black holes. *Fortschritte der Physik*, 61(9), 781–811.
- Moore, C. (1991). Generalized shifts: Unpredictability and undecidability in dynamical systems. *Nonlinearity*, 4(2), 199–230.
- Penrose, R. (1979). Singularities and time-asymmetry. In *General Relativity: An Einstein Centenary Survey* (pp. 581–638). Cambridge University Press.
- Pour-El, M. B., & Richards, J. I. (1989). *Computability in Analysis and Physics*. Springer.
- Rovelli, C. (2004). *Quantum Gravity*. Cambridge University Press.
- Rutten, J. J. M. M. (2000). Universal coalgebra: A theory of systems. *Theoretical Computer Science*, 249(1), 3–80.
- Ryu, S., & Takayanagi, T. (2006). Holographic derivation of entanglement entropy from the anti-de Sitter space/conformal field theory correspondence. *Physical Review Letters*, 96(18), 181602.
- Shew, W. L., & Plenz, D. (2013). The functional benefits of criticality in the cortex. *The Neuroscientist*, 19(1), 88–100.
- Susskind, L. (1995). The world as a hologram. *Journal of Mathematical Physics*, 36(11), 6377–6396.
- Verlinde, E. (2011). On the origin of gravity and the laws of Newton. *Journal of High Energy Physics*, 2011(4), 29.
- Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media.
- Zvonkin, A. K., & Levin, L. A. (1970). The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms. *Russian Mathematical Surveys*, 25(6), 83–124.