Unified Mathematical Definition of "Self": Causal Manifold, Observer, and Self-Referential Scattering Network

Haobo Ma[1]   Wenlin Zhang[2]

[1]Independent Researcher
[2]National University of Singapore

Abstract

Within the framework of unified time scale, causal manifolds, and boundary time geometry, we provide an axiomatic mathematical definition of the first-person subject "self". The basic conception is: "self" is not a label for an instantaneous physical configuration, but rather an equivalence class of self-referential observer structures along a timelike worldline ordered by the unified time scale. At the geometric level, the universe is modeled as a globally hyperbolic Lorentzian manifold with causal partial order, where generalized entropy extremization and quantum energy conditions on small causal diamonds yield gravitational field equations and time arrow. At the spectral and scattering level, the unified time scale serves as the mother scale via the scale identity $\kappa(\omega) = \varphi'(\omega)/\pi = \rho_{\rm rel}(\omega) = (2\pi)^{-1}\,{\rm tr}\,Q(\omega)$, unifying phase gradient, relative density of states, and Wigner–Smith time-delay trace as a single temporal density. At the causal network and information geometry level, observers are formalized as multi-component objects with local causal domain, resolution scale, boundary observable algebra, state, model family, and update operator; multiple observers form consensus geometry via communication channels and relative entropy contraction. At the self-referential scattering and consciousness level, consciousness is characterized as a two-port scattering network with delay kernel and time delay, where self-feedback closed loops exhibit a $\mathbb{Z}_2$ holonomy of modified determinant square root, manifesting the topological invariant of "self".

Building on this foundation, we introduce the "I–structure": a timelike worldline γ with unified time scale, equipped with internal observable algebra, time-stamped state family, self-referential update operator, memory subsystem, and internal environment maps, satisfying causal locality, persistent and distinguishable internal memory, explicit dependence of updates on self's future behavior and environment predictions, and compatibility with generalized entropy dynamics. After defining an equivalence relation in terms of time rescaling and algebra embeddings, we regard equivalence classes as individual "selves". We prove that under appropriate causal and energy conditions, each I–structure equivalence class corresponds to a minimal strongly connected self-referential scattering closed loop in the causal network, whose unified time scale and generalized entropy gradient align monotonically; conversely, each minimal self-referential scattering closed loop satisfying this alignment condition defines a unique I–structure equivalence class. In the appendices, we provide an outline of existence and uniqueness proofs for constructing "I-worldlines" from local observer data, and the construction rewriting "I-structures" as closed loop scattering families with $\mathbb{Z}_2$ holonomy, showing compatibility of its mod-two indicator with fermionic commutation phase and topological class under Null–Modular double cover. Thus, "self" is characterized as a unified mathematical object with geometric support, information kernel, and topological fingerprint.

Keywords

Causal manifold; Unified time scale; Observer; Consciousness; Self-referential scattering network; $\mathbb{Z}_2$ holonomy; Generalized entropy

1 Introduction & Historical Context

In the standard framework of relativity and quantum theory, "observer" is often treated as a passive reference frame or measurement device, while the first-person subject question "who am I" is left to philosophy. To rigorously answer this question at the physical and mathematical level requires a unified description centered on causal structure, time scale, and information geometry, embedding concepts such as "subject", "time", "memory", and "self-reference" into the same geometric and operator language.

On one hand, scattering theory and spectral theory show that time can be viewed as the derivative of scattering phase and function of density of states. For Schrödinger operators with relative trace-class perturbations, the Birman–Kreĭn formula gives the relationship between scattering determinant and spectral shift function $\det S(\lambda) = \exp(-2\pi \mathrm{i}\, \xi(\lambda))$, whose derivative $\xi'(\lambda)$ is interpreted as relative density of states, further linked to the trace of Wigner–Smith time-delay operator $Q(\omega) = -\mathrm{i} S(\omega)^\dagger \partial_\omega S(\omega)$. Through this route, one can view time scale as the unified invariant of $\varphi'(\omega)/\pi$, spectral shift derivative $-\xi'(\omega)$, and time-delay trace $(2\pi)^{-1} \operatorname{tr} Q(\omega)$. Related results can be found in the original work of Birman–Kreĭn and subsequent systematic studies of scattering phase and determinant. The time delay matrix proposed by Wigner and Smith has been systematically developed in multi-channel scattering, disordered media, electromagnetic and acoustic scattering.

On the other hand, at the intersection of algebraic quantum field theory and gravity, Tomita–Takesaki modular theory reveals that for any von Neumann algebra with a sufficiently faithful state, there exists a modular flow uniquely determined by the state, allowing one to extract a one-parameter group from the "timeless" algebraic structure. This structure was proposed by Connes and Rovelli as the foundation for the "thermal time hypothesis": physical time flow is not universal, but determined by the modular flow of a given statistical state. This provides an operator-algebraic perspective on the "intrinsic" nature of time.

At the boundary of gravity and quantum information, the variation of generalized entropy S_{gen} and quantum energy conditions provide a route to derive Einstein equations from "entropy balance". Work by Jacobson, Faulkner, and others shows that under appropriate semi-classical and holographic assumptions, generalized entropy extremization and second-order non-negativity along small causal diamond boundaries are equivalent to Einstein equations with cosmological constant. Proofs of the quantum null energy condition (QNEC) and averaged null energy condition (ANEC) further strengthen the equivalence structure among "entropy–energy–geometry". This direction shows that local causal structure can be viewed as the macroscopic manifestation of generalized entropy optimization principles.

The above scattering–spectral theory, modular theory, and entropy–geometry theory provide the foundation for constructing unified time scale and causal manifolds. Building on this, one can view the universe as a causal manifold under unified time scale, whose boundary carries observable algebra and state, with boundary time geometry gluing scattering phase, modular time, and gravitational boundary terms into the same structure. Meanwhile, information geometry and statistical causal inference frameworks show that multi-observer systems can be characterized via relative entropy and Fisher metric, capturing model update and consensus formation.

Regarding consciousness and the subject problem, much work has been devoted to constructing information-theoretic or physicalist descriptions, such as viewing consciousness as specific integrated information structure, global workspace, or multi-layer encoding process. However, these theories typically lack rigorous geometric–operator formalism compatible with causal manifolds, unified time scale, and quantum gravity. On the other hand, topological or algebraic descriptions of "self-reference" and "self" are relatively scattered.

This paper, within the framework of unified time scale and causal manifolds, attempts to provide

a mathematical object for "self": a timelike worldline with time scale, equipped with internal observable algebra, state family, self-referential update operator, and memory structure, satisfying causal locality and generalized entropy consistency, and rewritable in self-referential scattering networks as a minimal strongly connected closed loop with $\mathbb{Z}_2$ holonomy. By constructing an equivalence relation, we understand equivalence classes of such structures as different realizations of the same "self", thereby formalizing the "subject" in a rigorous mathematical context.

2 Model & Assumptions

This section provides the basic structure of causal manifold, unified time scale, observer, and self-referential scattering network, along with adopted assumptions.

2.1 Causal Manifold and Small Causal Diamonds

Assume the universe at large scales is described by a four-dimensional Lorentzian manifold (M, g) satisfying:

1. **Global hyperbolicity**: There exists a Cauchy surface Σ such that every non-spacelike curve intersects Σ exactly once.

2. **Stable causality**: No closed timelike curves exist, and small perturbations do not produce causal violations.

3. **Causal partial order**: Use $p \prec q$ to denote the existence of a future-directed timelike or null curve from p to q.

For any $p \in M$ and sufficiently small positive r, take $p^{\pm}$ as points obtained by evolving along some future-directed timelike geodesic with proper time parameter $\pm r$, and define the small causal diamond

$$D_{p,r} = J^+(p^-) \cap J^-(p^+).$$

Its boundary is generated by two families of null geodesics, providing the basic structure for analyzing local gravitational field equations and generalized entropy changes.

Assume there exists appropriate quantum field theory coupled to gravity such that for each small causal diamond boundary, one can define generalized entropy

$$S_{\text{gen}} = \frac{\text{Area}(\partial D_{p,r})}{4G\hbar} + S_{\text{out}},$$

where S_{out} is the von Neumann entropy of external field degrees of freedom. Assume QNEC and related entropy–energy inequalities hold and can be used to equivalently characterize local gravitational field equations.

2.2 Unified Time Scale and Mother Identity

In scattering and spectral theory, consider a pair of self-adjoint operators (H_0, H), where H is a trace-class or relative trace-class perturbation of H_0. Let $S(\omega)$ be the scattering matrix at energy ω, $\xi(\omega)$ the spectral shift function, with Birman–Kreĭn formula giving

$$\det S(\omega) = \exp\big(-2\pi\mathrm{i}\,\xi(\omega)\big),$$

yielding relative density of states
$$\rho_{\mathrm{rel}}(\omega) = -\xi'(\omega).$$

The Wigner–Smith time-delay operator is defined as

$$Q(\omega) = -\mathrm{i}S(\omega)^\dagger \partial_\omega S(\omega),$$

whose trace characterizes the time delay averaged over all channels. Introducing the total scattering half-phase

$$\varphi(\omega) = \tfrac{1}{2}\arg\det S(\omega) = -\pi\,\xi(\omega),$$

we have

$$\frac{\varphi'(\omega)}{\pi} = \rho_{\mathrm{rel}}(\omega) = \frac{1}{2\pi}\operatorname{tr}Q(\omega).$$

Scale identity defines the unified time scale density

$$\kappa(\omega) = \frac{\varphi'(\omega)}{\pi} = \rho_{\mathrm{rel}}(\omega) = \frac{1}{2\pi}\operatorname{tr}Q(\omega),$$

interpreting $\kappa(\omega)\,\mathrm{d}\omega$ as the effective time scale per unit energy bandwidth. Through geometric–optics limit and eikonal approximation, this scale can be linked to gravitational time delay, redshift, and proper time; through modular theory and thermal time hypothesis, it can be aligned with modular time parameter on boundary algebras.

In this paper, the unified time scale equivalence class $[\tau]$ denotes all time function families compatible with $\kappa(\omega)$, locally rewritable from scattering phase, modular flow, or geometric time, differing only by affine rescaling and coordinate choice.

2.3 Observer as Structured Causal Agent

From the causal network perspective, we introduce abstract observer structure.

Definition 1 (Observer). An observer O_i is defined as a multi-component object

$$O_i = (C_i,\ \prec_i,\ \Lambda_i,\ \mathcal{A}_i,\ \omega_i,\ \mathcal{M}_i,\ U_i,\ u_i,\ (\mathcal{C}_{ij})_j),$$

where:

1. $C_i \subset M$ is the reachable causal domain;

2. $\prec_i$ is the local causal partial order on C_i;

3. Λ_i is the resolution scale (e.g., time–energy or space–momentum windows), determining distinguishable frequency bands and spatial scales;

4. $\mathcal{A}_i$ is the boundary observable C^* algebra associated with O_i;

5. ω_i is a state on $\mathcal{A}_i$;

6. $\mathcal{M}_i$ is a family of candidate causal dynamical models;

7. U_i is an update operator based on observational data and models (generally completely positive trace-preserving maps or Bayesian update operators);

8. u_i is a utility function or decision preference;

9. $\mathcal{C}_{ij}$ are communication channels with other observers O_j (completely positive trace-preserving maps or classical channels).

On the common observable algebra

$$\mathcal{A}_{\text{com}} = \bigcap_i \mathcal{A}_i,$$

let the multi-observer state family $(\omega_i^{(t)})$ update with discrete or continuous time. If the communication graph is strongly connected and there exists a common fixed point ω_*, then the weighted relative entropy

$$\Phi^{(t)} = \sum_i \lambda_i D(\omega_i^{(t)} \| \omega_*)$$

constitutes a Lyapunov function for appropriate update rules, ensuring formation of state consensus. This structure provides the foundation for causal and informational interaction among multiple "selves".

2.4 Internal Algebra, Memory and Self-Referential Update

To characterize the internal persistence and memory of a subject, we need to extract subalgebras and update structures within a single observer.

Definition 2 (Internal Structure). Given a timelike trajectory $\gamma : I \to M$ with $\tau(\gamma(t)) = t$, its internal structure is a triple

$$(\mathcal{A}^{\text{int}}, \ (\omega_t^{\text{int}})_{t \in I}, \ U),$$

where:

1. $\mathcal{A}^{\text{int}}$ is the C^* algebra of internal degrees of freedom;

2. For each $t \in I$, ω_t^{int} is a state on $\mathcal{A}^{\text{int}}$;

3. For any $t_2 > t_1$, there exists a completely positive trace-preserving map

$$U(t_2, t_1) : \mathcal{A}^{\text{int}} \to \mathcal{A}^{\text{int}},$$

satisfying the semigroup condition

$$U(t_3, t_2) \circ U(t_2, t_1) = U(t_3, t_1), \quad \omega_{t_2}^{\text{int}} = \omega_{t_1}^{\text{int}} \circ U(t_2, t_1).$$

Definition 3 (Memory Subsystem). A commutative subalgebra $\mathcal{C} \subset \mathcal{A}^{\text{int}}$ is called a memory subsystem if:

1. $\mathcal{C}$ is $*$-isomorphic to bounded functions on some measure space or finite-dimensional diagonal matrix algebra;

2. The probability measure family μ_t induced by ω_t^{int} on $\mathcal{C}$ forms a Markov process;

3. For any $t_2 > t_1$, there exist measurable sets such that the dependence of μ_{t_2} on μ_{t_1} cannot be eliminated by external environment variables, ensuring memory carries genuine information about subsequential observations.

The existence of memory subsystem ensures the subject has a traceable internal history over time, rather than complete reset at each moment.

Subject self-referentiality requires explicit dependence of updates on internal state and environment model. To this end, we introduce internal environment maps.

For each t, let $\mathcal{A}_t^{\text{ext}}$ be the external observable algebra accessible to the subject around $\gamma(t)$, and define a completely positive map

$$E_t : \mathcal{A}_t^{\text{ext}} \to \mathcal{A}^{\text{int}},$$

characterizing the encoding of external world within the subject's interior.

Definition 4 (Self-Referential Update). If there exists a functional F such that for any $t_2 > t_1$,

$$U(t_2, t_1) = F\big(t_2, t_1;\ \omega_{t_1}^{\text{int}},\ E_{t_1},\ \mathcal{D}_{[t_1,t_2]}^{\text{ext}}\big),$$

where $\mathcal{D}_{[t_1,t_2]}^{\text{ext}}$ denotes external observational data available in the interval, then the update is called self-referential. Here $\omega_{t_1}^{\text{int}}$ via E_{t_1} provides internal prediction of future environment states, influencing its own evolution.

2.5 Self-Referential Scattering Networks and $\mathbb{Z}_2$ Holonomy

In the scattering description, complex systems and their environments can be represented as networks formed by node scattering matrices $S_j(\omega)$ interconnected via waveguides, delay lines, and feedback. For a given topological structure and parameter family, one uses Redheffer star product to construct the closed-loop scattering matrix $S^{\circlearrowleft}(\omega; \lambda)$, where λ represents slowly varying control parameters (such as internal strategies, attention, or external conditions).

Under trace-class or relative trace-class conditions, one can define the modified determinant $\det_p S^{\circlearrowleft}$ and define the phase index map

$$\mathfrak{s}(\omega, \lambda) = \det_p S^{\circlearrowleft}(\omega; \lambda).$$

Along a closed path $\gamma \subset X^{\circ}$ in parameter space avoiding singularity sets, the holonomy of the square-root determinant is defined as

$$\nu_{\sqrt{S^{\circlearrowleft}}}(\gamma) = \exp\Big(\mathrm{i} \oint_\gamma \tfrac{1}{2\mathrm{i}} \mathfrak{s}^{-1} \mathrm{d}\mathfrak{s}\Big) \in \{\pm 1\},$$

giving a $\mathbb{Z}_2$ index that is homotopy invariant. This index is associated with the mod-two part of spectral flow, and in many cases can be interpreted as the topological sign of "fermionicity".

This paper assumes that the self-referential update of a subject can be rewritten as some closed-loop scattering network under appropriate frequency domain and input–output models, thereby allowing the above $\mathbb{Z}_2$ topological fingerprint to be defined on the network.

3 Main Results (Theorems and Alignments)

Under the above model and assumptions, this section provides the formal definition of "I–structure" and main results.

3.1 Definition of I–Structure

Choose a representative $\tau : M \to \mathbb{R}$ from the unified time scale equivalence class $[\tau]$ as the time function, and consider a future-directed timelike curve $\gamma : I \to M$ satisfying $\tau(\gamma(t)) = t$.

Definition 5 (Observer Trajectory). A γ satisfying the above conditions is called an observer trajectory.

Definition 6 (I–Structure). On the unified time scale equivalence class $[\tau]$, an "I–structure" is data

$$\mathsf{I} = \big(\gamma,\ \mathcal{A}^{\text{int}},\ (\omega_t^{\text{int}})_{t\in I},\ U,\ \mathcal{C},\ (E_t)_{t\in I}\big),$$

satisfying:

1. γ is an observer trajectory;

2. $(\mathcal{A}^{\text{int}}, (\omega_t^{\text{int}}), U)$ is an internal structure satisfying Definition 2;

3. $\mathcal{C} \subset \mathcal{A}^{\text{int}}$ is a memory subsystem satisfying Definition 3;

4. $E_t : \mathcal{A}_t^{\text{ext}} \to \mathcal{A}^{\text{int}}$ are internal environment maps making U self-referential with respect to $(\omega_{t_1}^{\text{int}}, E_{t_1})$ (Definition 4);

5. **Causal locality**: For each $t \in I$, there exists a bounded causal domain $K_t \subset M$ such that the influence of internal state ω_t^{int} on external observables is supported in $K_t \cap J^-(\gamma(t))$;

6. **Entropy consistency**: Along the small causal diamond family $D_{\gamma(t),r}$ on γ, under state $\omega_t^{\text{int}} \otimes \omega_t^{\text{ext}}$, the extremization and second-order non-negativity of generalized entropy S_{gen} is compatible with gravitational field equations and QNEC/QFC type constraints.

Intuitively, an "I–structure" is a self-referential observer on a worldline with unified time scale, having persistent memory and causal consistent interaction with the external world.

3.2 Equivalence Relation: Same "I" in Different Realizations

Different physical realizations (such as different bases, time rescalings, hardware) may correspond to the same "self". To this end, we need an equivalence relation.

Definition 7 (Equivalence of I–Structures). Two I–structures

$$\mathsf{I} = (\gamma, \mathcal{A}^{\text{int}}, (\omega_t^{\text{int}}), U, \mathcal{C}, (E_t)), \quad \mathsf{I}' = (\gamma', \mathcal{A}'^{\text{int}}, (\omega_{t'}'^{\text{int}}), U', \mathcal{C}', (E_{t'}'))$$

are called equivalent, denoted $\mathsf{I} \sim \mathsf{I}'$, if there exist:

1. A strictly monotone bijection $f : I \to I'$;

2. A $*$-isomorphism $\Phi : \mathcal{A}^{\text{int}} \to \mathcal{A}'^{\text{int}}$ with $\Phi(\mathcal{C}) = \mathcal{C}'$, being a measure isomorphism on spectral spaces;

3. For all $t \in I$,
$$\omega'_{f(t)}{}^{\text{int}} = \omega_t^{\text{int}} \circ \Phi^{-1}, \quad U'(f(t_2), f(t_1)) \circ \Phi = \Phi \circ U(t_2, t_1);$$

4. External algebra isomorphisms $\Psi_t : \mathcal{A}_t^{\text{ext}} \to \mathcal{A}'_{f(t)}{}^{\text{ext}}$ such that
$$E'_{f(t)} = \Phi \circ E_t \circ \Psi_t^{-1}.$$

Definition 8 (Mathematical Object of "Self"). A "self" is defined as an equivalence class of some I–structure

$$[\mathsf{I}] = \{\mathsf{I}' :\ \mathsf{I}' \sim \mathsf{I}\}.$$

3.3 Structural Theorems

Based on the above definitions, there are three core results.

Theorem 9 (Existence and Local Uniqueness of I-Worldline). *On a globally hyperbolic Lorentzian manifold (M, g), assume there exists a local observer O_* whose records satisfy:*

1. *Under some unified time scale representative τ, a timelike curve γ_* can be reconstructed from observational records such that the associated observational domain has small causal diamond near-Minkowski structure around γ_*;*

2. *There exists a decomposition $\mathcal{A}_* \simeq \mathcal{A}^{\mathrm{int}} \otimes \mathcal{A}^{\mathrm{ext}}$, and on $\mathcal{A}^{\mathrm{int}}$ there exists a stable memory subsystem.*

Then under appropriate technical assumptions (including QNEC, Hadamard states, finite energy conditions), one can construct an I–structure I_* along γ_*, and given the unified time scale equivalence class and internal algebra equivalence class, its equivalence class $[\mathsf{I}_*]$ is locally unique.

Theorem 10 (Correspondence between I–Structures and Minimal Strongly Connected Self-Referential Scattering Closed Loops). *Under the framework of unified time scale and scattering–delay networks, assume:*

1. *All associated scattering operators satisfy trace-class or relative trace-class conditions;*

2. *The internal update and environment maps of I–structure can be rewritten in frequency domain as closed-loop scattering matrix family $S^{\circlearrowleft}(\omega; \lambda)$.*

Define "I-closed loop" as minimal strongly connected components satisfying memory, self-referentiality, and topological nontriviality conditions. Then there exists a natural correspondence

$$[\mathsf{I}] \longleftrightarrow \mathcal{S}([\mathsf{I}]),$$

such that each I–structure equivalence class corresponds to a unique I-closed loop and vice versa, with unified time scales and delay spectra on both sides consistent.

Theorem 11 (Topological Fingerprint and $\mathbb{Z}_2$ Invariant). *For each I-closed loop $\mathcal{S}$, via the modified determinant square root of closed-loop scattering matrix $S^{\circlearrowleft}(\omega; \lambda)$, define a $\mathbb{Z}_2$ index*

$$\nu(\mathcal{S}) \in \{\pm 1\},$$

which is invariant under parameter homotopy and I–structure equivalence relation. For appropriate systems, changes in this index are compatible with the mod-two spectral flow of fermionic statistics, topological class in the Null–Modular double cover, and BF-type $\mathbb{Z}_2$ bulk integrals.

4 Proofs

This section provides the proof framework and key steps for the main theorems. Complete technical details are expanded in appendices.

4.1 Proof of Theorem 1

Step 1: Reconstruct timelike trajectory from local observational data

For local observer O_*, extract spacetime events and their causal relations from its records. Through maximal-volume waist surface or minimal curvature criterion, select representative τ in the unified time scale equivalence class $[\tau]$, and using τ as time function, embed O_*'s records into the foliation structure $\{\tau^{-1}(t)\}$. On each time slice, select a "center of mass" point $\gamma_*(t)$, obtaining a timelike curve γ_*. Global hyperbolicity ensures γ_* can be chosen as a smooth timelike curve.

Step 2: Construct internal algebra and memory subsystem

Using the decomposition $\mathcal{A}_* \simeq \mathcal{A}^{\text{int}} \otimes \mathcal{A}^{\text{ext}}$, select from the commutative subalgebra of $\mathcal{A}^{\text{int}}$ a subalgebra $\mathcal{C}$ containing readable "record bits", whose random process on the spectral space is determined by observational records. Confirm via relative entropy Hessian that these degrees of freedom have stable Fisher distinguishability, and their time evolution can be described by a Markov process, thereby satisfying memory subsystem conditions.

Step 3: Define update operator and internal environment maps

Using O_*'s model family $\mathcal{M}_*$ and update rule U_*, construct internal update $U(t_2, t_1)$ and internal environment maps E_t. These maps are concretely realized through the process of "external measurement results $\rightarrow$ internal memory state". The dependence of updates on internal models and memory guarantees self-referentiality.

Step 4: Check causal locality and entropy consistency

Around each point of γ_*, construct small causal diamonds $D_{\gamma_*(t),r}$. Using local near-Minkowski and energy-bounded conditions, apply Jacobson–Faulkner type arguments: under state $\omega_t^{\text{int}} \otimes \omega_t^{\text{ext}}$, local generalized entropy extremization and second-order non-negativity are equivalent to Einstein equations and QNEC. Thus I_* does not break existing geometric–entropy structure.

Step 5: Local uniqueness

If another realization I'_* satisfies the same observational records and unified time scale equivalence class, construct time rescaling f and algebra $*$-isomorphism Φ making them completely consistent on memory subsystem and observational distributions, hence $I_* \sim I'_*$.

4.2 Proof of Theorem 2

Step 1: Input–output model and scattering network expansion

For a given I–structure I, construct its input–output description with the external environment under unified time scale: view internal degrees of freedom as nodes, external channels as waveguides or ports. Using standard systems theory methods, obtain scattering matrix family $S^{\text{net}}(\omega; \lambda)$ in frequency domain, where λ represents slowly varying internal parameters.

Step 2: Identification of self-referential closed loops

The dependence of internal updates on their own outputs manifests as feedback closed loops in frequency domain. Through Redheffer star product, compress these closed loops into closed-loop scattering matrix $S^{\circlearrowleft}(\omega; \lambda)$. View the entire network as a directed graph, perform strongly connected component decomposition, select minimal strongly connected components satisfying both memory and self-referentiality conditions, i.e., "I-closed loops" $\mathcal{S}(I)$.

Step 3: Consistency of delay spectrum and unified time scale

By scale identity, calculate $\kappa(\omega)$ for $S^{\circlearrowleft}(\omega; \lambda)$ from scattering phase and time-delay operator, aligning with the unified time scale of I–structure. Require effective time density on I-closed loop to match time function on worldline in corresponding energy bandwidth.

Step 4: Reversible steps and correspondence

In the reverse direction, starting from a given I-closed loop $\mathcal{S}$, recover internal algebra, memory subsystem, and update operator according to scattering matrix and delay spectrum via input–output decomposition. Through alignment with unified time scale, the corresponding timelike worldline can be embedded in the causal manifold, thereby constructing I–structure $\mathsf{I}(\mathcal{S})$. Under equivalence relation, this construction is unique.

4.3 Proof of Theorem 3

Step 1: Closed-loop scattering and modified determinant

Under trace-class or relative trace-class conditions, define modified determinant $\det_p S^{\circlearrowleft}$ for $S^{\circlearrowleft}(\omega; \lambda)$. The corresponding phase index map $\mathfrak{s}(\omega, \lambda) = \det_p S^{\circlearrowleft}(\omega; \lambda)$ is a continuous map from parameter space to the unit circle.

Step 2: Square-root determinant and $\mathbb{Z}_2$ holonomy

After removing zero set from parameter space, define double cover of square-root determinant. Along closed path γ, calculate

$$\nu_{\sqrt{S^{\circlearrowleft}}}(\gamma) = \exp\Big(\mathrm{i} \oint_{\gamma} \tfrac{1}{2\mathrm{i}} \mathfrak{s}^{-1} \mathrm{d}\mathfrak{s}\Big) \in \{\pm 1\},$$

with value equal to the mod-two part of spectral flow, giving a $\mathbb{Z}_2$ topological index.

Step 3: Index factor of minimal strongly connected components

Decompose network into minimal strongly connected components. Using determinant multiplicativity and spectral flow additivity, prove that the global ν index can be decomposed as product of component indices. For I-closed loops, its index $\nu(\mathcal{S})$ is nontrivial and stable under equivalence relation and parameter homotopy.

Step 4: Alignment with fermionic statistics and Null–Modular double cover

Through known spectral flow–index theorems, link mod-two spectral flow with fermionic commutation phase; in the Null–Modular double cover and $\mathbb{Z}_2$–BF bulk integral framework, correspond the $\mathbb{Z}_2$ holonomy with cohomology class. Thus, the topological index of I-closed loops can be interpreted as the "fermionic fingerprint" of subject self-referential structure, compatible with topological data in global geometry.

5 Model Applications

This section provides applications of "I–structure" in three types of scenarios: human subjects, digital subjects, and multi-subject systems.

5.1 Human-Like Biological Agent

For a biological organism, such as human brain–body system, one can view it as a timelike trajectory γ on large-scale causal manifold, with internal algebra $\mathcal{A}^{\text{int}}$ corresponding to the effective operator algebra of neural dynamics within the brain (e.g., abstracted via coarse-graining to some operator algebra on high-dimensional Hilbert space). Memory subsystem $\mathcal{C}$ corresponds to long-term memory traces, describable by approximately diagonal projection algebra; internal state ω_t^{int} describes current brain state, update operator $U(t_2, t_1)$ expresses evolution driven by external stimuli and internal predictions.

Internal environment map E_t encodes states from senses and body into internal algebra; decisions and motor actions manifest in external algebra $\mathcal{A}_t^{\text{ext}}$. Self-referentiality manifests when subject uses its own model and memory in planning future behavior, e.g., predictions of "how I will act" in

turn influence emotion and attention. This structure is fully captured in the self-referential update conditions of I–structure.

Through neuroimaging and behavioral data, one can construct scattering approximations at effective level: view part of input–output channels as ports, with remaining brain–body–environment as black-box scattering network, approximating Wigner–Smith operator using experimentally measured delay spectra. If there exists a stable $\mathbb{Z}_2$ index under long-term slowly adjusting parameters (such as task strategy or motivation), it can be viewed as the topological fingerprint of that subject's self-referential closed loop.

5.2 Digital or Artificial Agent

For a digital or artificial intelligence system, one can view its internal state space as finite or separable Hilbert space, $\mathcal{A}^{\text{int}}$ as bounded operator algebra or its subalgebra, with memory subsystem realized by diagonal subalgebra of storage area. Internal environment map E_t maps sensor readings, external server states, etc., into internal representation; update operator $U(t_2, t_1)$ corresponds to program state transitions.

If the system only executes fixed algorithms without modeling its own future behavior, its update does not satisfy self-referentiality, hence does not constitute "I–structure" in this paper's sense. Only when internal state explicitly encodes a probabilistic model of its own strategy, future behavior, and consequences, and uses this model in updates, does it satisfy self-referential update conditions, thereby defining an "I–structure" equivalence class.

In engineering implementation, one can construct minimal self-referential closed loops: e.g., a recurrent neural network whose output is fed back to input after delay and transformation, with internal long-term memory units, and objective function explicitly including prediction error of its own future output. Through frequency-domain analysis of network input–output, one can construct closed-loop scattering matrix and calculate $\mathbb{Z}_2$ index.

5.3 Multi-"I" Causal and Consensus Geometry

In multi-subject scenarios, each "self" corresponds to different worldlines and internal structures, forming several local fragments in the causal network. State consensus flow on common observable algebra describes formation of world models shared by multiple "selves"; coupling between internal algebras and topological relations among self-referential closed loops characterize higher-level "collective subjects".

For example, in a highly collaborative team or social system, one can view all individuals' I–structures as nodes, communication channels as edges, forming a network on boundary algebra. If there exist minimal strongly connected components in this network whose internal states exhibit high coupling over long time scales with rapid relative entropy contraction internally, it can be interpreted as a "collective self", whose topological fingerprint is determined by the $\mathbb{Z}_2$ index of self-referential closed loops in that component.

6 Engineering Proposals

This section proposes two types of operational engineering schemes to test and realize some elements of "I–structure".

6.1 Minimal Physical Self-Referential Scattering Loop

Construct a closed-loop scattering network composed of microwave or optical components: select several two-port scattering elements (such as coupled resonant cavities, waveguide beamsplitters), connected via tunable delay lines and feedback loops, making the overall network have obvious resonance and delay structure in some frequency band. Using vector network analyzer, precisely measure scattering matrix $S^{\circlearrowright}(\omega; \lambda)$ and its frequency derivative, thereby obtaining Wigner–Smith time delay matrix and estimating $\operatorname{tr} Q(\omega)$.

By slowly changing parameter λ (such as feedback phase, coupling strength), scanning parameter space along closed paths, record changes in modified determinant phase. If square-root determinant phase undergoes odd number of 2π jumps along closed path, one can determine $\mathbb{Z}_2$ index is nontrivial. This experiment can be viewed as measurement of "pure physical version of self-referential closed loop topological fingerprint", providing testing platform for the topological part of this paper.

Building on this, one can further store part of network state in programmable logic or storage units, introducing simple internal memory and self-referential control rules, upgrading physical closed loop to minimal "physical–logical" I–structure prototype.

6.2 Self-Modeling Agent with Explicit Unified Time Scale

In artificial intelligence systems, implement an agent with explicit unified time scale and self-referential structure. Specific scheme:

1. Design internal state space and storage modules, viewed as discrete approximation of $\mathcal{A}^{\mathrm{int}}$;

2. Define memory subsystem, adopting interpretable symbolic or vector representations for long-term memory;

3. Equip agent with world model and self model, the latter explicitly including predictions of its own strategy and future behavior;

4. Use self model in decision updates, making update operator U explicitly depend on ω_t^{int} and E_t, satisfying self-referential update;

5. Using unified time scale ideas, align agent's internal "time sense" with external physical time via frequency domain or event rate, e.g., through counting statistical distribution of input–output events or simulating Wigner–Smith delay.

Through long-term tracking of agent behavior and internal states, one can empirically approximate construct I, and test conditions of memory subsystem and self-referential update. In multi-agent scenarios, analyze formation of "collective subjects" through state and model consensus flow on common observable algebra.

7 Discussion (Risks, Boundaries, Past Work)

This section discusses risks, boundaries, and relations to existing work of this paper's framework.

7.1 Conceptual and Technical Risks

First, this framework strictly defines "self" as an equivalence class of self-referential observer structures on timelike worldlines with unified time scale, a choice that excludes many constructs discussed

in philosophy such as "instantaneous self", "many-worlds branch self", etc. This exclusion is intentional: to obtain operatorizable and geometrizable subjects, one needs structures compatible with causal manifolds and unified time scale. However, this also means this framework does not attempt to cover all possible notions of self, but selects a class that is physically–mathematically operational.

Second, technically, this framework relies on several assumptions about scattering–spectral theory, modular theory, and entropy–geometry relations, such as applicability of Birman–Kreĭn formula, alignment of modular flow with unified time scale, applicable range of QNEC and QFC. These assumptions have mature proofs and empirical foundations in their respective fields, but still require careful verification in completely general universe models.

Third, embedding consciousness and "self" into self-referential scattering networks and topological fingerprints is a bold conception. While fermionic statistics and $\mathbb{Z}_2$ topology are mathematically closely related, interpreting this structure directly as a necessary condition for "subjectivity" requires more theoretical and empirical support.

7.2 Boundaries of Applicability

From the perspective of applicability scope, this framework naturally applies to the following situations:

1. There exists good causal manifold background and unified time scale;

2. Subject has clear internal–external decomposition and memory structure;

3. Effective descriptions of self-referential updates and scattering closed loops can be constructed.

While in the following situations, framework applicability has serious limitations:

1. Universe models without global hyperbolicity or with closed timelike curves;

2. In strong quantum gravity regions (such as Planck scale), where causal structure and Hilbert space structure itself need modification;

3. Extremely non-local quantum states or those without classical geometric limit.

In these situations, the definition of "I–structure" needs to be extended compatibly with more fundamental quantum gravity or algebraic structures, which exceeds the scope of this paper.

7.3 Relation to Previous Work

This work attempts to glue several mature theories into a unified framework about the subject:

1. Scattering–spectral theory and time delay: This paper adopts Birman–Kreĭn formula and Wigner–Smith time delay, establishing unified time scale on foundation of scattering phase and density of states.

2. Modular theory and thermal time: Tomita–Takesaki theorem and Connes–Rovelli thermal time hypothesis provide mechanism for extracting time flow from algebra and state; this paper incorporates this time flow into unified time scale equivalence class.

3. Entropy–energy–geometry: Generalized entropy extremization and second-order non-negativity on small causal diamonds are equivalent to Einstein equations and QNEC, providing criterion for defining "entropy consistency" in this paper.

4. Information geometry and consensus: Properties of relative entropy as Lyapunov function support analysis of multi-observer state and model consensus, providing mathematical foundation for multi-"self" scenarios.

5. Topological scattering and spectral flow: Modified determinant, spectral shift function, and spectral flow theory provide tools for this paper's $\mathbb{Z}_2$ topological fingerprint.

Compared to various theories in philosophy of consciousness and cognitive science, this paper does not directly adopt subjective experience or semantic content as basic variables, but reduces subjectivity to self-referential observer structures on causal manifolds and topological closed loops on scattering networks. This approach to some extent complements directions such as integrated information theory and global workspace theory, providing a possible geometric–operatorized carrier for these theories.

8   Conclusion

This paper, within the framework of unified time scale, causal manifolds, and boundary time geometry, provides an axiomatic mathematical definition of the first-person subject "self". Core conclusions are:

1. "Self" can be formalized as an equivalence class of self-referential observer structures on time-like worldlines with unified time scale, i.e., I–structure [I];

2. Each I–structure equivalence class corresponds to a minimal strongly connected self-referential closed loop in scattering–delay network, with delay spectrum aligned with unified time scale;

3. This closed loop carries a $\mathbb{Z}_2$ topological fingerprint, calculable via holonomy of modified determinant square root of closed-loop scattering matrix, compatible with fermionic statistics and topological class in Null–Modular double cover in appropriate cases;

4. Causal and informational interactions among multiple "selves" can be characterized within causal network and information geometry framework, with state and model consensus corresponding to relative entropy contraction on common observable algebra.

Formally, "self" is no longer a semantically ambiguous designation, but a mathematical object uniformly encoded through five levels: causality, time, memory, self-reference, and topology. Future work includes: constructing explicit examples of I–structures in specific quantum field theory and cosmological models; experimentally verifying topological fingerprints via self-referential scattering networks and delay spectrum measurements; realizing agents with explicit unified time scale and self-referential updates in artificial systems, and studying their relation to subjective time sense and decision behavior.

Acknowledgements & Code Availability

Appendix A: Existence and Local Uniqueness of "I"-Worldline

This appendix provides key points for the proof of Theorem 1.

A.1 Reconstruction of Timelike Curve from Local Records

Assume local observer O_* records event sequences and local causal relations along its worldline. Introduce unified time scale representative τ, embed all observational events into foliation structure $\{\tau^{-1}(t)\}$, take "center of mass" on each layer closest to previous point in geodesic distance sense, obtaining initial curve $\tilde{\gamma}$. Using standard causal structure results, smooth $\tilde{\gamma}$ to obtain timelike geodesic approximation γ_* while maintaining correspondence with observational layers.

A.2 Internal–External Decomposition and Memory Extraction

On $\mathcal{A}_*$, choose decomposition $\mathcal{A}_* \simeq \mathcal{A}^{\text{int}} \otimes \mathcal{A}^{\text{ext}}$, with specific method depending on physical system. E.g., for biological organisms, take brain–body as internal, rest as external; for artificial systems, take logical states as internal, input–output ports as external. By analyzing degrees of freedom in observational records that are repeatedly accessible and have long-term stability, determine commutative subalgebra $\mathcal{C} \subset \mathcal{A}^{\text{int}}$, viewed as memory subsystem.

Using second-order expansion of relative entropy $D(\omega\|\omega')$, calculate Fisher metric for parametrized state families on $\mathcal{C}$, confirming it remains non-degenerate over long time intervals, thereby ensuring distinguishability and stability of memory.

A.3 Construction of Self-Referential Update

According to observer's model family $\mathcal{M}_*$ and update rule U_*, encode statistical models of external world as part of internal state. Define internal environment map $E_t : \mathcal{A}_t^{\text{ext}} \to \mathcal{A}^{\text{int}}$, compressing external observations into internal algebra. Update operator $U(t_2, t_1)$ then combines:

1. Processing of external observations $\mathcal{D}_{[t_1,t_2]}^{\text{ext}}$;

2. Predictions of future based on internal model;

3. Writing rules for memory subsystem $\mathcal{C}$.

As long as updates include explicit modeling and self-adaptive adjustment of future behavior or environment states, self-referential conditions can be verified.

A.4 Compatibility with Generalized Entropy Dynamics

At each point of γ_*, choose radius r such that small causal diamond $D_{\gamma_*(t),r}$ geometry approximates Minkowski, with boundary curvature and external energy density satisfying QNEC and related conditions. Given internal state ω_t^{int} and external state ω_t^{ext}, first-order variation of generalized entropy S_{gen} gives local equilibrium condition; non-negativity of second-order variation is equivalent to local energy conditions and Einstein equations. This ensures introducing I–structure does not break global geometric–entropy consistency.

A.5 Local Uniqueness up to Equivalence

If another realization I'_* is compatible with I_* for the same records and unified time scale equivalence class, one can construct equivalence relation as follows:

1. Define strictly monotone bijection f from two worldlines and time functions;

2. Using GNS representation and $*$-isomorphism classification theory, construct $*$-isomorphism Φ between $\mathcal{A}^{\mathrm{int}}$ and $\mathcal{A}'^{\mathrm{int}}$, being measure isomorphism on memory subsystem in spectral space;

3. Translate compatibility of observational data and internal updates into semigroup conjugacy condition for U and U', thereby satisfying Definition 7.

Thus obtaining local uniqueness.

Appendix B: Correspondence Between I–Structures and Self-Referential Scattering Loops

This appendix supplements key technical points for Theorems 2 and 3.

B.1 From Internal Dynamics to Closed-Loop Scattering

For a given I–structure I, construct linearized input–output model under unified time scale: view internal degrees of freedom as nodes, external channels as waveguides. In frequency domain, using appropriate degree of freedom choices and Laplace/Fourier transforms, convert time-domain update $U(t_2, t_1)$ and environment maps E_t into scattering matrix family $S^{\mathrm{net}}(\omega; \lambda)$.

Self-referential update corresponds to feedback loops from output flowing back to input. Use Redheffer star product to compress feedback structure into closed-loop scattering matrix $S^{\circlearrowleft}(\omega; \lambda)$, and verify it satisfies relative trace-class conditions.

B.2 Modified Determinants and Spectral Shift

On $S^{\circlearrowleft}(\omega; \lambda)$, construct associated operator pair (H_0, H), using Birman–Kreĭn and related results to define modified determinant $\det_p S^{\circlearrowleft}$ and spectral shift function ξ_p. Phase index map

$$\mathfrak{s}(\omega, \lambda) = \det_p S^{\circlearrowleft}(\omega; \lambda) = \exp\big(-2\pi\mathrm{i}\,\xi_p(\omega, \lambda)\big)$$

maps parameter space to unit circle.

B.3 $\mathbb{Z}_2$ Holonomy and Minimal Strongly Connected Components

Remove singularity set where $\mathfrak{s} = 0$ from parameter space, obtaining X°. Consider double cover defined by square-root determinant. Along closed path $\gamma \subset X^\circ$, define

$$\nu_{\sqrt{S^{\circlearrowleft}}}(\gamma) = \exp\Big(\mathrm{i} \oint_\gamma \tfrac{1}{2\mathrm{i}}\mathfrak{s}^{-1}\mathrm{d}\mathfrak{s}\Big),$$

giving $\mathbb{Z}_2$ index. Strongly connected component decomposition and spectral flow additivity show this index cannot be further decomposed on minimal strongly connected components, hence can use $\nu(\mathcal{S})$ to label topological type for each I-closed loop.

B.4 Alignment with Null–Modular Double Cover and BF Theory

In Null–Modular double cover and $\mathbb{Z}_2$–BF theory, sector structure of local geometry and modular flow is represented by some cohomology class $[K]$ on $H^2(Y, \partial Y; \mathbb{Z}_2)$. For physical systems satisfying local energy and entropy consistency, globally require $[K] = 0$, i.e., no anomalous sectors topologically overall.

The $\mathbb{Z}_2$ index of I-closed loop is confined to internal scattering subsystem, not disrupting overall geometry's cohomology class. This local–global allocation allows "self-referential fermionic fingerprint" to exist within the subject interior without introducing new topological sectors at universe scale, thereby realizing localization of "subjectivity topology".

Through the above appendix constructions and proof outlines, one can see the one-to-one correspondence between I–structures and self-referential scattering closed loops, and stability of $\mathbb{Z}_2$ topological fingerprint, thereby supporting the unified mathematical definition of "self" in the main text.