From Coherence to Consequent Nature

A Formal Approach to Process-Relational Theology

Mathias Leonhardt*

May 2026 (v8.0)

Working paper — Pre-conference draft, May 2026.
Accepted for presentation at the European Academy of Religion (EuARe) 2026, Philosophy of Religion panel, Rome, 2 July 2026 (Session 155, paper ID 1395). *Not yet peer-reviewed; comments welcome at* `mathias-leonhardt@gmx.de`.

Version 8 (May 2026). This paper is the revised and substantially expanded version of the author's abstract submitted to EuARe 2026. The theological framing of the title reflects the panel's focus; the technical core has been deepened since submission to include formal convergence, Lyapunov stability, NP-hardness, and replicator dynamics. Version 8 corrects an attribution gap regarding Thagard & Verbeurgt (1998), makes the idealising character of the gradient-ascent dynamics explicit, and refines the relation between NP-hardness and Bedau's (1997) weak-emergence criterion.

Abstract

We present a formal axiom system $\mathcal{C}_5$ for agent-relative value theory, grounded in the concept of *coherence* as constraint satisfaction over relational structures. Modeling coherence as a quadratic form, we establish six main results: (1) coherence-seeking agents under projected gradient dynamics converge monotonically to locally optimal states; (2) a continuous-time gradient flow formulation with Lyapunov stability, establishing that coherence attractors are asymptotically stable equilibria; (3) global coherence maximization is NP-hard under the binary restriction (restating Thagard and Verbeurgt's 1998 result in our quadratic-form framework), which we explicitly connect to Bedau's (1997) weak-emergence criterion; (4) the coherence landscape generically supports multiple stable attractors with path-dependent convergence; (5) spectral criteria for attractor stability derived from the coherence matrix; and (6) a replicator dynamics for aggregation weights, replacing the stipulated weights of prior versions with emergent, coherence-driven selection.

The aggregation mechanism can either stipulate weights or derive them dynamically via replicator dynamics, where average coherence increases monotonically. As an application, we develop a structural mapping to process philosophy (Whitehead, 1929), demonstrating that certain non-personal "God-concepts"—in particular Whitehead's consequent nature and Spinoza's *Deus sive Natura*—admit formal reconstruction as emergent coherence attractors, doubly emergent in both structure and weighting. This mapping is compatible with recent work on emergent moral properties (Baysan, 2025).

Keywords: coherence theory, weak emergence, computational complexity, value aggregation, process philosophy, process ontology, structural theology, non-personal God-concepts, Whitehead, Lyapunov stability, replicator dynamics

*Independent researcher, Hamburg. Web: ki-mathias.de. Feedback welcome at `mathias-leonhardt@gmx.de`.
AI assistance: Drafting and LaTeX typesetting were supported by Claude (Anthropic) via Claude Code; all axioms, proofs, and interpretive choices are the author's.

1 Introduction

1.1 Motivation

How do individual value orientations give rise to collective evaluative structures? This question connects formal value theory, social choice, and the philosophy of emergence. We approach it through a minimal axiom system that models *coherence-capable agents*—entities that orient toward structural consistency in their evaluations.

Our central insight is that aggregated value structures are *weakly emergent* in the technical sense of Bedau (1997): they are ontologically reducible but epistemically non-trivial, resisting decomposition into sums of individual contributions.

1.2 Contributions

This paper does not introduce a fundamentally new mathematical theorem on coherence complexity: the NP-hardness of coherence maximisation was already established by Thagard and Verbeurgt (1998). It also does not introduce a new theological doctrine: Whitehead's consequent nature and Spinoza's *Deus sive Natura* are centuries-old. The contribution of this paper is a *structural synthesis* of existing results into a single dynamical and complexity-theoretic framework, with a process-ontological reading of the resulting attractor structure. Concretely, eight contributions:

(C1) Axiom System $\mathcal{C}_5$: A formal system for coherence-capable agents that models orientation, distortion, and positive evaluative structure with explicit domain restrictions.

(C2) Coherence Functional with Dynamics: A constraint-satisfaction model of coherence inspired by Thagard (1989) and Thagard and Verbeurgt (1998), formalised as a quadratic form over relational structures, with both a discrete update rule and convergence guarantee (Theorem 5.6) and a continuous-time gradient flow with Lyapunov stability (Theorem 3.15). The dynamical assumptions are explicitly framed as modelling idealisations (Remark 3.10).

(C3) Restating NP-Hardness in Quadratic-Form Framework: A restatement of the Thagard and Verbeurgt (1998) NP-hardness result in the quadratic-form formulation (Theorem 5.9; cf. Remark 5.8), needed for the subsequent Lyapunov and spectral analyses. The mathematical content is not new; the embedding is.

(C4) Multiple Attractors: A demonstration that the coherence landscape generically supports multiple stable configurations (Theorem 5.11), with path-dependent convergence (Corollary 5.13).

(C5) Spectral Characterisation: A spectral analysis of the coherence matrix (Proposition 5.17 and Remark 5.18), characterising attractor stability via local optimality conditions and bounding the number of stable coherent regimes through the eigenvalue structure.

(C6) Dynamic Aggregation Weights: A replicator dynamics (Theorem 6.15) coupling individual coherence to aggregate selection, yielding monotonically increasing average coherence. This is positioned as the empirically substantive coherence-seeking claim of the framework (cf. Section 6.4); the individual gradient flow is the idealisation, the aggregate selection is the structural observation.

(C7) Bridge to Bedau's Weak Emergence: An explicit connection between the complexity result and Bedau's 1997 weak-emergence criterion (Definition 6.2), with the qualification that NP-hardness is a strict subcase of Bedau's broader criterion (Remark 6.3). This bridge was not drawn by Thagard & Verbeurgt.

(C8) Process-Ontological Application: A structural mapping to process philosophy, showing that Whitehead's "consequent nature" and Spinoza's *Deus sive Natura* admit formal reconstruction as *doubly emergent* coherence attractors at the aggregate level.

The novelty is the bridge, not the bricks. The bricks — coherence-as-constraint-satisfaction (Thagard 1989, Thagard & Verbeurgt 1998), weak emergence (Bedau 1997), replicator dynamics (Taylor & Jonker 1978), consequent nature (Whitehead 1929), *Deus sive Natura* (Spinoza 1677) — are taken from the existing literature with attribution. The contribution is the integration: embedding the Thagard–Verbeurgt complexity result in a Lyapunov-dynamical framework with spectral characterisation, coupling it to a replicator dynamics at the aggregate level, connecting both to Bedau's weak-emergence criterion, and reading the resulting positive attractor as a structural reconstruction of non-personal God-concepts in the Whitehead–Spinoza tradition.

1.3 What This Paper Is and Is Not

This paper IS:	This paper is NOT:
A formal axiom system with explicit assumptions	A proof of God's existence
A structural mapping (formal $\leftrightarrow$ process ontology)	An argument for theism
An exploration of weak emergence in value theory	A claim about ontological emergence
Axioms (A1–A5) that primarily fix vocabulary and scope	A derivation of moral truths from axioms alone
Methodologically modest and conditional	A moral excuse for harmful behavior

1.4 Structure of the Paper

Section 2 reviews related work. Section 3 presents formal foundations, including a continuous gradient flow with Lyapunov stability. Section 4 introduces $\mathcal{C}_5$. Section 5 derives main results, including spectral analysis. Section 6 treats emergence, aggregation, and replicator dynamics. Section 7 develops the process-ontological application. Section 8 addresses responsibility concerns. Section 9 provides meta-analysis. Section 10 concludes.

2 Related Work

2.1 Emergence

The emergence literature distinguishes *weak* from *strong* emergence. Bedau (1997) defines weak emergence as ontologically reducible but epistemically intractable without simulation: "the only kind of real emergence." Chalmers (2006) formalizes strong emergence as involving properties not deducible even in principle from physical facts.

Our framework explicitly adopts weak emergence. We avoid Kim's (1999) causal exclusion argument by not claiming that emergent structures have autonomous causal powers—they are structural properties of aggregated systems.

> *Our notion of emergence aligns with Bedau's (1997) "weak emergence": ontologically reducible but epistemically non-trivial.*

2.2 Coherence Theory

Thagard (1989) models coherence as constraint satisfaction over relational structures, implemented in the ECHO system. Thagard and Verbeurgt (1998) subsequently provided a complete formalisation of the coherence problem and proved that its maximisation is NP-hard. Thagard (1998) extends the framework to ethics. Our coherence functional C_h is a formal abstraction of this tradition; Theorem 5.9 below restates the Thagard–Verbeurgt complexity result in the quadratic-form formulation used here.

> *Following Thagard (1989), we model coherence as constraint satisfaction over relational structures.*

The epistemological foundations of coherentism are developed by BonJour (1985); Quine famously remarked that “a coherence theory is evidently the lot of ethics.”

2.3 Value Aggregation and Social Choice

Arrow (1951) proves that no aggregation procedure satisfies all reasonable axioms simultaneously. Sen (1982) argues that ordinal preferences are insufficient—interpersonal comparisons of welfare are necessary.

Our aggregation operator V^* sidesteps Arrow's impossibility by *stipulating* weights α_h rather than deriving them from individual orderings:

> *Our aggregation operator V^* sidesteps Arrow's impossibility by stipulating weights rather than deriving them from individual orderings.*

This is a modeling choice, not a solution to the impossibility problem. In Section 6.4, we complement this with a dynamic alternative: replicator dynamics from evolutionary game theory.

2.4 Evolutionary Game Theory

The replicator equation, introduced by Taylor and Jonker (1978), describes how the frequency of strategies in a population evolves based on relative fitness. Hofbauer and Sigmund (1998) provide a comprehensive treatment, including the fundamental result that average fitness increases monotonically under replicator dynamics (the “Fisher fundamental theorem” analogue). We apply this framework to aggregation weights in Section 6.4, with individual coherence playing the role of fitness.

2.5 Process Philosophy and Structural Theology

Whitehead (1929) develops a dipolar conception of God: the *primordial nature* (realm of pure potentials) and the *consequent nature* (actualized through interaction with the world). Whitehead writes: “It requires converse with the immanent world for God to emerge in all actuality.”

Our structural God $\mathfrak{G} = \mathcal{P}^*$ exhibits formal parallels to Whitehead's consequent nature: an emergent structure arising from the relational dynamics of actual entities.

Hartshorne (1967) develops this into process theism, emphasizing God as relational and affected by the world. Cobb (1965) extends Whitehead's theology systematically.

2.6 Spinoza and Deus sive Natura

Spinoza's *Ethics* (1677) identifies God with Nature (*Deus sive Natura*), distinguishing *natura naturans* (active, naturing nature) from *natura naturata* (passive, natured nature). Spinoza's formal method—definitions, axioms, propositions—anticipates our approach.

> *Spinoza's “Deus sive Natura” anticipates our structural identification: God as the totality of positively-valued relational structure, not a transcendent agent.*

2.7 Moral Emergence

Baysan (2025) defends emergent moral non-naturalism: moral properties depend on descriptive properties plus normative bridge principles, yielding emergent but non-causal powers. Our positive evaluative structure $B(h)$ exhibits a similar dependence pattern:

> *Baysan (2025) defends emergent moral properties with noncausal powers; our positive evaluative structure $B(h)$ exhibits a similar dependence pattern.*

3 Formal Foundations

3.1 Ontological Primitives

Definition 3.1 (Entity Space)**.** Let $\mathcal{E}$ be a set of *entities*, partitioned as:

$$\mathcal{E} = \mathcal{E}_{\text{physical}} \cup \mathcal{E}_{\text{mental}} \cup \mathcal{E}_{\text{abstract}} \cup \mathcal{E}_{\text{relational}} \tag{1}$$

Definition 3.2 (Relational Structure)**.** Define the *relational graph* $\mathcal{G} := (\mathcal{E}, \mathcal{R}, \omega)$ where $\mathcal{R} \subseteq \mathcal{E} \times \mathcal{E}$ and $\omega : \mathcal{R} \to \mathbb{R}$ assigns valence to relations.

3.2 Coherence-Capable Agents

Not all entities are agents, and not all agents are coherence-capable. We introduce:

Definition 3.3 (Coherence-Capable Agent)**.** An entity $h \in \mathcal{E}$ is a *coherence-capable agent* if:

(i) h possesses a value functional $V_h : \mathcal{E} \to \mathbb{R}$

(ii) h exhibits *responsiveness* to coherence gradients: changes in $C_h(\mathcal{G})$ tend to influence h's behavior

(iii) h is capable of *error*: h can act contrary to coherence maximization

The set of all coherence-capable agents is denoted $\mathcal{H} \subset \mathcal{E}$.

Remark 3.4 (Scope Limitation)*.* The axiom system $\mathcal{C}_5$ applies only to coherence-capable agents. Entities that lack value functionals, responsiveness, or error-capability are outside the scope. This is a *domain restriction*, not a universal claim about all beings.

3.3 Coherence: Formal Definition

Following Thagard (1989), we model coherence as constraint satisfaction:

Definition 3.5 (Coherence Functional)**.** For a coherence-capable agent $h \in \mathcal{H}$, the *coherence functional* is:

$$C_h(\mathcal{G}) := \sum_{(x,y) \in \mathcal{R}} V_h(x) \cdot \omega(x, y) \cdot V_h(y) \tag{2}$$

This measures structural consistency between the agent's value assignments and the relational graph.

Remark 3.6*.* This bilinear form is analogous to energy in Hopfield networks and consistency measures in belief revision theory (BonJour, 1985).

3.4 The Agent-Relative Value Functional

Assumption 1 (Value Functional)**.** For each $h \in \mathcal{H}$, the value functional $V_h : \mathcal{E} \to \mathbb{R}$ satisfies:

$$V_h(x \mid \mathcal{G}) = \alpha_h(x) + \sum_{y \in \mathcal{E}} \beta_h\big(\omega(x, y)\big) \tag{3}$$

where α_h captures intrinsic contribution and β_h captures relational contribution.

3.5 Predicate Formalization

Symbol	Interpretation
$O(h)$	h exhibits persistent orientation toward coherence maximization
$H(h)$	h performs actions that reduce aggregate coherence
$\mathcal{D}(h)$	h operates under a distortion of the coherence landscape
$E(h)$	h exhibits systematic error in coherence-seeking behavior
$B(h)$	h possesses a positive evaluative structure
$\text{MaxCoh}(h)$	h has attained maximal coherence state

3.6 Dynamics: State Vectors and Coherence Optimization

The coherence functional C_h (Definition 3.5) is a bilinear form over agent-relative values. To derive non-trivial results, we now make the optimization structure explicit.

Definition 3.7 (State Vector and Domain)**.** For each agent $h \in \mathcal{H}$, define the *state vector* $\boldsymbol{v}_h \in \mathbb{R}^{|\mathcal{E}|}$ with components $v_h(x) := V_h(x)$. Let $W \in \mathbb{R}^{|\mathcal{E}| \times |\mathcal{E}|}$ denote the *coherence matrix* with entries $W_{xy} := \omega(x, y)$ for $(x, y) \in \mathcal{R}$ and $W_{xy} := 0$ otherwise. Then:

$$C_h(\boldsymbol{v}_h) = \boldsymbol{v}_h^\top W \boldsymbol{v}_h \tag{4}$$

which is the matrix form of Definition 3.5.

Definition 3.8 (Box Constraint)**.** The *feasible domain* for value assignments is:

$$\mathcal{B} := [-1, 1]^{|\mathcal{E}|} \tag{5}$$

This reflects a natural boundedness assumption: no entity is valued beyond ± 1 in normalized units. The coherence-maximization problem for agent h is then $\max_{\boldsymbol{v}_h \in \mathcal{B}} C_h(\boldsymbol{v}_h)$.

Definition 3.9 (Projected Gradient Ascent)**.** Given step size $\eta > 0$, the *coherence update rule* is:

$$\boldsymbol{v}_h^{t+1} = \Pi_{\mathcal{B}}\big(\boldsymbol{v}_h^t + \eta \nabla C_h(\boldsymbol{v}_h^t)\big) \tag{6}$$

where $\Pi_{\mathcal{B}}$ denotes component-wise projection onto $[-1, 1]$ and $\nabla C_h(\boldsymbol{v}) = (W + W^\top)\boldsymbol{v}$.

This operationalizes the "responsiveness to coherence gradients" from Definition 3.3(ii): agents adjust their value assignments in the direction of increasing coherence, subject to bounded rationality (finite step size η) and domain constraints.

Remark 3.10 (Gradient Ascent as Modelling Idealisation). The choice of projected gradient ascent is a *modelling idealisation*, not an empirical derivation. It follows the convention established by Festinger (1957)'s cognitive dissonance theory, Hopfield networks (Hopfield, 1982), and Thagard's ECHO architecture (Thagard, 1989): cognitive systems are modelled as energy-minimising / coherence-maximising dynamical systems. Real cognitive trajectories are noisier than the deterministic flow (6) suggests; individuals can experience coherence-decreasing transitions (trauma, disconfirming evidence, manipulation), captured in our framework via the Distortion Operator (Axiom 3). The deterministic gradient flow captures the *asymptotic tendency* under the idealisation, not the moment-by-moment empirical trajectory of any particular agent. A stochastic extension (Langevin dynamics) would replace strict monotonicity with concentration of the stationary distribution on coherence attractors; we leave this to future work. The empirically substantive coherence-seeking claim of this framework operates not at the individual but at the aggregate level via the replicator dynamics of Section 6.4.

Definition 3.11 (Binary Restriction)**.** For *sharp* value assignments (full commitment or full rejection), define:

$$\mathcal{B}_{\pm} \coloneqq \{-1, +1\}^{|\mathcal{E}|} \tag{7}$$

The binary coherence-maximization problem is $\max_{\boldsymbol{v}_h \in \mathcal{B}_{\pm}} C_h(\boldsymbol{v}_h)$.

Remark 3.12 (Connection to Hopfield Networks). Under $\mathcal{B}_{\pm}$, the coherence functional $\boldsymbol{v}^{\top} W \boldsymbol{v}$ is formally identical to the energy of a Hopfield network *(with sign reversal)*. This is not a coincidence: Hopfield networks model constraint satisfaction in associative memory, and our framework models constraint satisfaction in value space. The analogy extends to the existence of multiple local optima (stored patterns/coherence attractors) and path-dependent convergence. This strengthens the connection noted in Remark following Definition 3.5.

3.7 Continuous Gradient Flow and Lyapunov Stability

The discrete projected gradient ascent (Definition 3.9) has a natural continuous-time counterpart. This continuous formulation provides the mathematical foundation for the "emergent attractors" language: coherence attractors are not merely fixed points of an iterative scheme but *asymptotically stable equilibria* of a dynamical system.

Definition 3.13 (Projected Gradient Flow)**.** The *coherence gradient flow* on $\mathcal{B}$ is the differential inclusion:

$$\frac{d\boldsymbol{v}_h}{dt} = \Pi_{T_{\mathcal{B}}(\boldsymbol{v}_h)}\big[(W + W^{\top})\boldsymbol{v}_h\big] \tag{8}$$

where $\Pi_{T_{\mathcal{B}}(\boldsymbol{v}_h)}$ denotes orthogonal projection onto the *tangent cone* of $\mathcal{B}$ at $\boldsymbol{v}_h$:

$$T_{\mathcal{B}}(\boldsymbol{v}) = \big\{d \in \mathbb{R}^{|\mathcal{E}|} \;\big|\; \forall i: \;(v_i = -1 \Rightarrow d_i \geq 0) \wedge (v_i = 1 \Rightarrow d_i \leq 0)\big\}.$$

Remark 3.14. At interior points of $\mathcal{B}$ (where $|v_i| < 1$ for all i), the tangent cone is $\mathbb{R}^{|\mathcal{E}|}$ and the flow reduces to $\dot{\boldsymbol{v}}_h = (W + W^{\top})\boldsymbol{v}_h$. The projection matters only at boundary points, where it prevents trajectories from leaving the feasible domain.

Theorem 3.15 (Lyapunov Stability of Coherence Flow)**.** *Along trajectories of the projected gradient flow* (8)*:*

(i) ***Monotonicity:*** $\dfrac{dC_h}{dt} \geq 0$*. That is, C_h is a Lyapunov function for the flow.*

(ii) ***Existence of ω-limits:*** *Every trajectory has a non-empty ω-limit set, and every ω-limit point satisfies the KKT conditions for* $\max_{\boldsymbol{v} \in \mathcal{B}} C_h(\boldsymbol{v})$*.*

(iii) ***Asymptotic stability:*** *Every strict local maximum of C_h on $\mathcal{B}$ is an asymptotically stable equilibrium of* (8)*.*

The projected dynamical system (8) is well-posed in the sense of Dupuis and Nagurney (1993): since $\mathcal{B}$ is convex and compact and the right-hand side $(W + W^{\top})\boldsymbol{v}$ is Lipschitz continuous, existence and uniqueness of absolutely continuous solutions follow from standard results on projected dynamical systems (see also Nagurney and Zhang, 1996, Ch. 2). The LaSalle invariance principle applies in this setting because solutions are confined to the compact set $\mathcal{B}$.

Intuition. The coherence functional is non-decreasing along the flow because the projected gradient is, by construction, a feasible ascent direction. Trajectories cannot escape the compact box, so limit sets exist; the LaSalle invariance principle then forces these limits onto KKT-stationary points. Strict local maxima are isolated and thus asymptotically stable. *Full proof: Appendix A.1.*

Remark 3.16 (Connection to Discrete Dynamics). The projected gradient ascent of Definition 3.9 is the forward Euler discretization of the gradient flow (8) with step size η. Theorem 5.6 (discrete monotonicity) and Theorem 3.15 (continuous Lyapunov stability) are thus complementary perspectives on the same underlying dynamics. The continuous formulation substantiates the "emergent attractors" language: coherence attractors are genuine dynamical attractors, not merely fixed points.

4 The Axiom System $\mathcal{C}_5$

We now present the axiom system $\mathcal{C}_5$ (Coherence-5). The name reflects the five axioms and the central role of coherence; it is unrelated to the modal logic S5.

Axiom 1 (Orientation of Coherence-Capable Agents)**.**

$$\forall h \in \mathcal{H} : O(h) \tag{9}$$

All coherence-capable agents exhibit persistent orientation toward coherence maximization.

Remark 4.1 (On Axiom 1). This is partially analytic: it restricts the domain rather than making a universal empirical claim. Entities that do not orient toward coherence are, by definition, not in $\mathcal{H}$.

Axiom 2 (Coherence-Harm Incompatibility)**.**

$$\forall h \in \mathcal{H} : \mathrm{MaxCoh}(h) \Rightarrow \neg H(h) \tag{10}$$

Maximal coherence states are incompatible with coherence-reducing actions.

Axiom 3 (Distortion Induces Error)**.**

$$\forall h \in \mathcal{H} : \mathcal{D}(h) \Rightarrow E(h) \tag{11}$$

Distorted coherence landscapes lead to systematic error in coherence-seeking.

Axiom 4 (Harm Under Orientation Implies Distortion)**.**

$$\forall h \in \mathcal{H} : \big(O(h) \wedge H(h)\big) \Rightarrow \mathcal{D}(h) \tag{12}$$

A coherence-oriented agent who causes harm must be operating under distortion.

Remark 4.2 (On Axiom 4: Distortion Is Not Exculpation). Distortion ($\mathcal{D}$) denotes a *model–world mismatch*: the agent's internal coherence landscape diverges from the relational structure $\mathcal{G}$. This is a descriptive diagnosis, not a judgment about diminished responsibility. Typical instances include self-serving bias, information asymmetry, and ideological capture—all of which are compatible with full moral accountability. See Section 8 for a detailed discussion.

Axiom 5 (Orientation Implies Positive Structure)**.**

$$\forall h \in \mathcal{H} : O(h) \Rightarrow B(h) \tag{13}$$

Persistent orientation toward coherence implies the presence of a positive evaluative structure.

Remark 4.3 (On Axiom 5: Definitional Character). Axiom 5 has a *definitional* character: it stipulates that we *call* coherence-orientation "positive evaluative structure." This is a terminological choice. We are explicit about this to avoid pseudo-depth.

5 Main Results

Theorem 5.1 (Universal Positive Structure)**.** *Under $\mathcal{C}_5$:*

$$\mathcal{C}_5 \vdash \forall h \in \mathcal{H} : B(h) \tag{14}$$

Proof. Let $h \in \mathcal{H}$ be arbitrary.

(1) By Axiom 1: $O(h)$.

(2) By Axiom 5: $O(h) \Rightarrow B(h)$.

(3) By *modus ponens*: $B(h)$.

(4) Since h was arbitrary: $\forall h \in \mathcal{H} : B(h)$.

□ □

Remark 5.2. This proof is short because the conclusion is largely built into the axioms. The theorem's value is *organizational*: it names a property holding for all members of $\mathcal{H}$.

Corollary 5.3 (Harm as Distortion-Induced Error)**.** *For all $h \in \mathcal{H}$:*

$$H(h) \Rightarrow \big(\mathcal{D}(h) \wedge E(h)\big) \tag{15}$$

Proof. Suppose $H(h)$.

(1) By Axiom 1: $O(h)$.

(2) From (1) and hypothesis: $O(h) \wedge H(h)$.

(3) By Axiom 4: $\mathcal{D}(h)$.

(4) By Axiom 3: $E(h)$.

(5) Thus: $\mathcal{D}(h) \wedge E(h)$.

□ □

Theorem 5.4 (Distortion Removal)**.** *Let $h' \in \mathcal{H}$ denote a temporal successor state. Then:*

$$\big(\mathcal{D}(h) \wedge \neg\mathcal{D}(h')\big) \Rightarrow \neg H(h') \tag{16}$$

Proof. Suppose $\mathcal{D}(h)$ and $\neg\mathcal{D}(h')$. Assume for contradiction $H(h')$. By Axiom 4: $\mathcal{D}(h')$. Contradiction. □ □

The preceding results follow directly from the axioms. We now derive *substantive* results that exploit the mathematical structure of the coherence functional (Section 3.6) and are *not* built into the axioms.

5.1 Monotone Coherence Ascent and Convergence

Lemma 5.5 (Lipschitz Gradient)**.** *The gradient $\nabla C_h(\boldsymbol{v}) = (W + W^\top)\boldsymbol{v}$ is Lipschitz continuous with constant $L = \|W + W^\top\|_2$ (spectral norm).*

Proof. For any $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^{|\mathcal{E}|}$:

$$\|\nabla C_h(\boldsymbol{u}) - \nabla C_h(\boldsymbol{v})\|_2 = \|(W + W^\top)(\boldsymbol{u} - \boldsymbol{v})\|_2 \leq \|W + W^\top\|_2 \cdot \|\boldsymbol{u} - \boldsymbol{v}\|_2.$$

□ □

Theorem 5.6 (Monotone Coherence Ascent)**.** *Let $\boldsymbol{v}_h^t$ evolve according to the projected gradient ascent* (6) *with step size $\eta \leq 1/L$ where $L = \|W + W^\top\|_2$. Then:*

(i) ***Monotonicity:*** *$C_h(\boldsymbol{v}_h^{t+1}) \geq C_h(\boldsymbol{v}_h^t)$ for all t.*

(ii) ***Convergence:*** *The sequence $(C_h(\boldsymbol{v}_h^t))_{t\geq 0}$ converges. Every accumulation point of $(\boldsymbol{v}_h^t)$ satisfies the Karush–Kuhn–Tucker (KKT) conditions for $\max_{\boldsymbol{v}\in\mathcal{B}} C_h(\boldsymbol{v})$.*

Intuition. The quadratic ascent lemma for L-smooth functions guarantees that one step in the gradient direction increases the objective by at least $\frac{\eta}{2}\|\nabla C\|^2$; projection onto the convex box $\mathcal{B}$ preserves this ascent property. The sequence is monotone and bounded, hence convergent, with accumulation points satisfying KKT stationarity. *Full proof: Appendix A.2.*

Remark 5.7 (Philosophical Interpretation)*.* Theorem 5.6 provides the formal basis for the "orientation toward coherence" postulated in Axiom 1. Under bounded rationality (finite η), coherence-capable agents move monotonically toward higher coherence *in the idealised model.* The KKT convergence shows they reach locally optimal states—but not necessarily the global optimum. This gap between local and global coherence is precisely the space where distortion (Axiom 3) operates.

Two qualifications are in order. First, the monotonicity statement is a property of the postulated dynamics, not an empirical claim about all human cognition: real cognitive trajectories can exhibit coherence-decreasing transitions, as discussed in Remark 3.10. The Lyapunov result (Theorem 3.15) describes the noise-free idealisation. Second, the empirically substantive coherence-seeking claim of this framework lies at the aggregate level (Section 6.4), not at the level of any individual trajectory.

5.2 Computational Intractability of Global Coherence

We now establish that global coherence maximization is, in general, computationally intractable. This is the formal core of our emergence claim and gives mathematical substance to Bedau's intuition that weakly emergent properties are those for which no general analytic shortcut is known.

Remark 5.8 (Prior Result)*.* The NP-hardness of coherence-as-constraint-satisfaction was established by Thagard and Verbeurgt (1998), using a reduction from 3-SAT in a slightly different formalisation. The theorem below restates the result in the quadratic-form formulation employed throughout this paper, which is needed for the subsequent Lyapunov and spectral analyses (Sections 3 and 5) and for the integration with replicator dynamics (Section 6.4). The mathematical content is not new; the embedding in a single dynamical framework is.

Theorem 5.9 (NP-Hardness of Coherence Maximization)**.** *Under the binary restriction $\mathcal{B}_\pm = \{-1,+1\}^{|\mathcal{E}|}$ (Definition 3.11), the problem*

$$\max_{\boldsymbol{v}\in\mathcal{B}_\pm} \boldsymbol{v}^\top W \boldsymbol{v} \tag{17}$$

is NP-hard.

Intuition. Setting $W := -A_H$ for the adjacency matrix of an arbitrary graph H turns binary coherence maximisation into Max-Cut on H. Since Max-Cut is NP-hard (Karp, 1972), so is binary coherence maximisation. *Full reduction: Appendix A.3.*

Remark 5.10 (Why This Matters for Emergence)*.* Theorem 5.9 establishes that *finding the globally optimal coherence structure is computationally intractable* under standard complexity-theoretic assumptions (P $\neq$ NP). This goes significantly beyond Definition 6.1: the global optimum is not merely non-separable, but *epistemically inaccessible* without exhaustive computation.

This provides formal support for Bedau's (1997) weak emergence criterion (Definition 6.2): NP-hardness gives a precise sense in which no general polynomial-time shortcut exists (under standard assumptions). Note that NP-hardness does not strictly entail simulation-only derivability—heuristic methods may perform well in structured instances—but it rules out a *general* efficient shortcut.

5.3 Multiple Coherence Attractors

Agents converge to *local* optima (Theorem 5.6), but these need not be unique. We now show that the coherence landscape generically supports multiple attractors.

Theorem 5.11 (Existence of Multiple Attractors)**.** *For any $n \geq 4$, there exists a coherence matrix $W \in \mathbb{R}^{n\times n}$ such that $C(\boldsymbol{v}) = \boldsymbol{v}^\top W \boldsymbol{v}$ under $\mathcal{B}_\pm$ has at least two distinct local maxima $\boldsymbol{v}^{(1)} \neq \boldsymbol{v}^{(2)}$ (where a local maximum is a point such that flipping any single component does not increase C).*

Intuition. For $n = 4$, partition the entities into two communities with strong positive intra-community coupling and weak negative inter-community coupling. Both $(+1, +1, -1, -1)$ and $(-1, -1, +1, +1)$ become local maxima by symmetry: each represents one consistent way of aligning the communities. The construction generalises to $n \geq 4$. *Full construction: Appendix A.4.*

Remark 5.12 (Philosophical Interpretation: Plurality of Coherent Regimes). Theorem 5.11 establishes that the coherence landscape admits *multiple stable configurations*—different ways of being internally coherent. This has significant philosophical implications:

(i) **Value pluralism:** Different coherent value systems can coexist without one being objectively superior (both are local optima).

(ii) **Community structure:** The proof construction shows that multiple attractors arise naturally from community structure—tightly coupled subgroups that disagree with each other.

(iii) **Process ontology:** In Whitehead's framework, this corresponds to the possibility of multiple coherent "societies" with different dominant characteristics. The consequent nature of God ($\mathcal{P}^*$) would then depend on which attractor the system has settled into—making it genuinely *path-dependent*.

Corollary 5.13 (Path Dependence)**.** *Under asynchronous updates (agents update sequentially rather than simultaneously), there exist coherence matrices W and initial states $\boldsymbol{v}^0$ such that different update schedules S_1, S_2 lead to different limit states:*

$$\lim_{t\to\infty} \boldsymbol{v}^t_{S_1} \neq \lim_{t\to\infty} \boldsymbol{v}^t_{S_2}.$$

Proof. This follows directly from Theorem 5.11: given multiple local maxima with non-trivial basins of attraction, the update order determines which basin the trajectory enters. A concrete example with $n = 3$ and frustrated coupling (analogous to the Hopfield network with multiple stored patterns) demonstrates this; details are deferred to future work. □ □

5.4 Spectral Analysis of the Coherence Landscape

The coherence matrix W has been treated so far as a "black box" encoding relational structure. We now analyze its spectrum, which reveals structural constraints on the number and stability of coherence attractors.

Definition 5.14 (Symmetrized Coherence Matrix)**.** Let $W_s := \frac{1}{2}(W + W^\top)$ denote the symmetrization of W. Its eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ (where $n = |\mathcal{E}|$) are real since W_s is symmetric.

Remark 5.15. The coherence functional can be written $C_h(\boldsymbol{v}) = \boldsymbol{v}^\top W \boldsymbol{v} = \boldsymbol{v}^\top W_s \boldsymbol{v}$, since $\boldsymbol{v}^\top A \boldsymbol{v} = \boldsymbol{v}^\top A_s \boldsymbol{v}$ for any matrix A and its symmetrization A_s. Hence the spectrum of W_s (not W) governs the coherence landscape.

Definition 5.16 (Spectral Coherence Gap)**.** The *spectral coherence gap* is:

$$\Delta(W) := \lambda_1(W_s) - \lambda_2(W_s). \tag{18}$$

Proposition 5.17 (Spectral Stability Criterion)**.** *Let $\boldsymbol{v}^*$ be a local maximum of C_h on $\mathcal{B}_\pm$ (the binary domain). Then:*

(i) ***Local optimality condition:*** *For every component i with $v_i^* \in \{-1, +1\}$:*

$$v_i^* \cdot [(W + W^\top)\boldsymbol{v}^*]_i \geq 0. \tag{19}$$

That is, the gradient component agrees in sign with the current assignment.

(ii) ***Stability under single-flip perturbation:*** *The coherence decrease from flipping component i is:*

$$C_h(\boldsymbol{v}^*) - C_h(\boldsymbol{v}^{*,\neg i}) = 4v_i^* \cdot [(W_s)\boldsymbol{v}^*]_i \tag{20}$$

where $\boldsymbol{v}^{*,\neg i}$ denotes $\boldsymbol{v}^*$ with the i-th component flipped. This is positive iff the local optimality condition holds strictly.

Proof. (i) At a local maximum on $\mathcal{B}_\pm$, no single-flip improves the objective. The change from flipping v_i^* to $-v_i^*$ is $\Delta_i = -4v_i^*[(W_s)\boldsymbol{v}^*]_i$. Requiring $\Delta_i \leq 0$ gives $v_i^*[(W_s)\boldsymbol{v}^*]_i \geq 0$.

(ii) Direct computation: $C_h(\boldsymbol{v}^{*,\neg i}) = (\boldsymbol{v}^*)^\top W_s \boldsymbol{v}^* - 4v_i^*[(W_s)\boldsymbol{v}^*]_i$. The difference follows. □

□

Remark 5.18 (Spectral Attractor Heuristic). Let k^+ denote the number of strictly positive eigenvalues of W_s. Then:

(i) On $\mathcal{B}$ (continuous domain), the coherence functional $C_h(\boldsymbol{v}) = \boldsymbol{v}^\top W_s \boldsymbol{v}$ achieves its maximum at a vertex of $\mathcal{B}$ that lies in the span of the top eigenvectors of W_s.

(ii) On $\mathcal{B}_\pm$ (binary domain), the number of local maxima is at most 2^n (trivially), but the number of *dynamically stable* attractors—those attracting a positive-measure basin under gradient flow—is expected to be bounded by $O(2^{k^+})$.

This can be motivated as follows.

(i) On the continuous domain, C_h is a quadratic form and thus maximized on the boundary of $\mathcal{B}$ (in fact at vertices, since it is bilinear). The maximum value is $\sum_i \lambda_i u_i^2$ where u_i are the coordinates of the maximizer in the eigenbasis. Only positive eigenvalues contribute positively.

(ii) The bound is motivated by the Hopfield network analogy (Remark following Definition 3.11). In the Hopfield setting, Amit et al. (1985) show that the network can store approximately $0.14n$ stable patterns before recall degrades. More precisely, stable attractors correspond to eigenvectors associated with the largest eigenvalues of the coupling matrix. By analogy, only the k^+ positive eigenvalues of W_s contribute to stable attractors on $\mathcal{B}_\pm$, suggesting the $O(2^{k^+})$ bound. A rigorous proof for the general coherence setting remains open.

Remark 5.19 (Spectral Gap and Attractor Robustness). A large spectral coherence gap $\Delta(W) \gg 0$ implies that the *dominant* coherence attractor (associated with λ_1) is well-separated from secondary attractors. This makes the dominant regime robust under perturbations of the relational structure. Conversely, when $\Delta(W) \approx 0$, multiple coherent regimes are nearly iso-energetic, and small changes in the relational graph $\mathcal{G}$ can shift the system between attractors—a formal analogue of "paradigm shifts" in value systems.

6 Emergence and Aggregation

6.1 Emergence: Definitions

We distinguish two layers of emergence that are often conflated:

Definition 6.1 (Non-Separability)**.** A global functional $\Phi : \mathcal{G} \to \mathbb{R}$ is *non-separable* if:

$$\nexists \phi : \mathcal{E} \to \mathbb{R} \text{ such that } \Phi(\mathcal{G}) = \sum_{x \in \mathcal{E}} \phi(x) \tag{21}$$

That is, Φ cannot be expressed as a sum of local (single-entity) functions.

Definition 6.2 (Weak Emergence after Bedau)**.** A macro-property is *weakly emergent* in the sense of Bedau (1997) if it is derivable from the system's micro-level specification but only by simulation—that is, no general analytic shortcut exists.

Remark 6.3 (Bedau's Criterion is Broader than NP-Hardness). Bedau's weak emergence covers a family of computational obstructions: it encompasses *undecidability* (Bedau's own paradigmatic example is the universal Turing machine embeddable in Conway's Game of Life, where the halting problem becomes inaccessible) as well as polynomial-time-hardness (NP-hardness and beyond). NP-hardness is therefore a *strict subcase* of Bedau's criterion, not a re-statement of it. Our framework provides a *computational-complexity instance* of weak emergence: coherence maximisation under the binary restriction is NP-hard, hence epistemically intractable in the polynomial sense. This is sufficient for the structural mapping pursued in Section 7, but it does not exhaust Bedau's criterion. The relation is one of *sufficient asymmetry*: NP-hardness rules out a general analytic shortcut, but does not preclude that specific instances admit one; full Bedau emergence in the simulation-only sense is a stronger and potentially separable claim.

Remark 6.4 (Scope of Emergence Claims). Non-separability (Definition 6.1) is a minimal structural criterion: it captures irreducibility to single-entity properties. It does *not* entail ontological emergence, causal autonomy, or strong supervenience failure. Bedau's weak emergence (Definition 6.2) is strictly stronger: it adds an epistemic/computational dimension. Our NP-hardness result (Theorem 5.9) provides formal support for this stronger criterion under standard complexity assumptions, with the qualification of Remark 6.3.

Definition 6.1 captures the *minimal* notion of emergence (non-separability). Our main results (Section 5) establish a substantially stronger property:

Definition 6.5 (Computational Emergence)**.** A global functional $\Phi : \mathcal{G} \to \mathbb{R}$ is *computationally emergent* if determining $\max \Phi$ (or any optimizer) is NP-hard in the size of $\mathcal{G}$.

Remark 6.6. Computational emergence implies non-separability (Definition 6.1), but is strictly stronger: it asserts not merely structural irreducibility but *epistemic intractability*. By Theorem 5.9, the coherence functional C is computationally emergent under binary restrictions. This provides formal support for Bedau's (1997) weak emergence criterion (Definition 6.2): NP-hardness gives a precise sense in which no general polynomial-time shortcut is known, though it does not strictly entail that simulation is the *only* route.

6.2 Aggregation

Definition 6.7 (Aggregated Value Functional)**.**

$$V^*(x) := \sum_{h \in \mathcal{H}} \alpha_h \cdot V_h(x) \quad \text{where } \alpha_h \geq 0, \sum_h \alpha_h = 1 \tag{22}$$

Remark 6.8 (Relation to Arrow's Impossibility). By stipulating weights α_h rather than deriving them from preference orderings, we sidestep Arrow's impossibility theorem. This is a modeling choice; we do not claim to solve the aggregation problem. However, the stipulation can be replaced by a *dynamical* alternative—see Section 6.4 below.

Definition 6.9 (Global Positive Attractor)**.**

$$\mathcal{P}^* := \{x \in \mathcal{E} \mid V^*(x) > 0\} \tag{23}$$

Proposition 6.10 (Weak Emergence of Aggregated Structures)**.** *Under Assumption 1 with $\beta_h \not\equiv 0$, both V^* and $\mathcal{P}^*$ are weakly emergent.*

Proof. By (3), V_h depends on relational weights $\omega(x, y)$. These encode pairwise information not reducible to single-entity properties. Hence, $V^* = \sum_h \alpha_h V_h$ inherits this non-separability. □
□

Theorem 6.11 (Non-Reducibility of the Global Attractor)**.** *In general:*

$$\mathcal{P}^* \neq \bigcup_{h\in\mathcal{H}} \mathcal{P}_h \quad \textit{and} \quad \mathcal{P}^* \neq \bigcap_{h\in\mathcal{H}} \mathcal{P}_h \tag{24}$$

where $\mathcal{P}_h := \{x \in \mathcal{E} \mid V_h(x) > 0\}$ denotes the individual positive attractor of agent h.

Intuition. Two counterexamples suffice: agents can disagree on a value with equal weights so the aggregated value is zero while one agent rates it positively (cancellation), and one strongly positive but lightly weighted agent can be outweighed by another agent's mild negative rating (asymmetric amplification). These interference effects are the mechanism by which aggregation produces genuinely emergent structures. *Full counterexamples: Appendix A.5.*

6.3 A Substantive Model: The Cooperative Community

Example 6.12 (Cooperative Community Model). Let $\mathcal{E} = \{a, b, c, r_1, r_2\}$ where a, b, c are agents and r_1, r_2 are relational goods ("trust", "mutual aid").

Define $\mathcal{H} = \{a, b, c\}$ with value functionals:

$$V_a : r_1 \mapsto +1,\ r_2 \mapsto +1,\ b \mapsto +0.5,\ c \mapsto +0.5 \tag{25}$$
$$V_b : r_1 \mapsto +1,\ r_2 \mapsto +0.8,\ a \mapsto +0.5,\ c \mapsto +0.5 \tag{26}$$
$$V_c : r_1 \mapsto +0.9,\ r_2 \mapsto +1,\ a \mapsto +0.5,\ b \mapsto +0.5 \tag{27}$$

This model satisfies all axioms non-vacuously with $|\mathcal{H}| = 3$.

6.4 Dynamic Aggregation: Replicator Dynamics

While Theorem 3.15 (Lyapunov stability) and Theorem 5.6 (monotone ascent) describe the *idealised individual* dynamics, the *empirically substantive* coherence-seeking claim of this framework lies at the aggregate level. Individual gradient trajectories are a modelling convention (Remark 3.10); the dynamics introduced in this section operates on the population of agents and selects coherence-bearing perspectives over time. Cultural traditions, scientific paradigms, and religious structures all exhibit selection processes consistent with replicator dynamics: ideas that yield higher coherence with the prevailing structure reproduce more reliably than those that do not. The individual gradient trajectory is an idealisation; the aggregate coherence selection is a structural observation about cultural epistemology.

The weights α_h in Definition 6.9 were stipulated. We now introduce a dynamical mechanism that allows these weights to *emerge* from the system's coherence structure, drawing on evolutionary game theory (Taylor and Jonker, 1978; Hofbauer and Sigmund, 1998).

Definition 6.13 (Replicator Dynamics for Aggregation Weights)**.** Let $\boldsymbol{\alpha} = (\alpha_h)_{h\in\mathcal{H}}$ with $\alpha_h \geq 0$ and $\sum_h \alpha_h = 1$. The *coherence-driven replicator dynamics* is:

$$\frac{d\alpha_h}{dt} = \alpha_h\big(C_h - \bar{C}\big) \quad \text{where} \quad \bar{C} := \sum_{h\in\mathcal{H}} \alpha_h C_h \tag{28}$$

denotes the weighted average coherence.

Remark 6.14. This is the standard replicator equation from evolutionary game theory, with individual coherence C_h playing the role of fitness. Agents whose coherence exceeds the average gain weight; those below average lose weight. The simplex $\{\boldsymbol{\alpha} \geq 0, \sum_h \alpha_h = 1\}$ is invariant under this dynamics.

Theorem 6.15 (Monotone Average Coherence)**.** *Under the replicator dynamics* (28) *with fixed individual coherence values C_h:*

$$\frac{d\bar{C}}{dt} = \mathrm{Var}_{\boldsymbol{\alpha}}(C) \geq 0 \tag{29}$$

where $\mathrm{Var}_{\boldsymbol{\alpha}}(C) := \sum_{h\in\mathcal{H}} \alpha_h (C_h - \bar{C})^2$ *is the $\boldsymbol{\alpha}$-weighted variance of individual coherence values.*

Remark 6.16 (Timescale Separation). Theorem 6.15 assumes fixed individual coherence values C_h. In the full system, agents simultaneously optimize their state vectors $\boldsymbol{v}_h$ (Section 3.6) while aggregation weights α_h evolve under replicator dynamics. The decoupled analysis is justified under a *two-timescale assumption*: if states $\boldsymbol{v}_h$ equilibrate fast relative to weight adaptation ($\eta_{\text{state}} \gg \eta_{\text{weight}}$), the fixed-fitness regime approximates the slow manifold. A fully coupled analysis of the joint dynamics $(\boldsymbol{v}_h(t), \alpha_h(t))$ is left for future work.

Intuition. This is the discrete analogue of Fisher's fundamental theorem: differential reproduction by fitness causes the mean fitness to rise at a rate equal to the variance of fitness across the population. Direct computation confirms $d\bar{C}/dt = \mathrm{Var}_{\boldsymbol{\alpha}}(C) \geq 0$. *Full derivation: Appendix A.6.*

Corollary 6.17 (Equilibrium Characterization)**.** *Under replicator dynamics, $\bar{C}$ converges. At equilibrium, all agents with $\alpha_h > 0$ have identical coherence values. Agents with below-average coherence satisfy $\alpha_h \to 0$: they are "selected out" by the dynamics.*

Proof. Since $\bar{C}$ is non-decreasing and bounded above (coherence values are bounded on $\mathcal{B}$), it converges. At any limit point, $d\bar{C}/dt = 0$ requires $\mathrm{Var}_{\boldsymbol{\alpha}}(C) = 0$, which holds iff all agents with $\alpha_h > 0$ share the same coherence value. For any agent h with $C_h < \bar{C}$, we have $d\alpha_h/dt < 0$ persistently, so $\alpha_h \to 0$. □ □

Remark 6.18 (Arrow Revisited). The stipulated weights (Remark 6.8) are now revealed as the *static* special case of a richer dynamics. The replicator equation provides a *structural* answer to "which weights?": the dynamically stable weights are those in which all surviving agents are equally coherent. This does not "solve" Arrow's impossibility—the theorem concerns ordinal preferences, not coherence-weighted aggregation—but it shifts the question from *which weights are fair?* to *which weights are dynamically stable?* The answer emerges from the system rather than being imposed externally.

Remark 6.19 (Process-Theological Significance). Whitehead's "creative advance" gains a precise formalization. Under replicator dynamics, the weighting of perspectives (α_h) is not externally given but arises through the process itself. The structural God $\mathfrak{G} = \mathcal{P}^*$ becomes *doubly emergent*: emergent in structure (Theorem 5.9: determining the optimal $\mathcal{P}^*$ is NP-hard) and emergent in weighting (Theorem 6.15: the aggregation weights emerge from coherence-driven selection). This double emergence strengthens the parallel with process theology, where the consequent nature of God is shaped by the actual occasions of experience—both in what those occasions are and in how they contribute to the divine life.

7 Application: Process-Ontological Reconstruction

7.1 The Bridge Thesis

> **Bridge Thesis:** If one adopts (i) the axiom system $\mathcal{C}_5$, (ii) Definition $D_{\mathfrak{G}}$, and (iii) a non-personal interpretation of theological language, then certain theological concepts become *formally derivable* as theorems about emergent coherence attractors.

This is an *interpretive reconstruction*: a demonstration that the formal framework's structural properties can be mapped onto process-ontological concepts. It is not a derivation of theological claims from mathematics. The formal framework stands independently.

It bears emphasis that the mapping operates at the *collective*, not individual, level. Whitehead's consequent nature is itself a collective entity: the unification of all actual occasions in the divine experience. The corresponding structural role in our framework is played by the aggregate coherence selection of Section 6.4, not by the idealised individual gradient flow. The Lyapunov result of Theorem 3.15 characterises the idealised individual coherence-seeking; the replicator result of Theorem 6.15 characterises the cultural-level coherence-selection that sustains the collective attractor $\mathcal{P}^*$. Both are needed to make the Whitehead mapping precise: individual coherence-seeking generates the local dynamics, aggregate selection generates the stable collective structure.

7.2 Definition $D_{\mathfrak{G}}$ (Structural God)

Definition 7.1 (Structural God)**.** Under $\mathcal{C}_5$ and the aggregation framework:

$$\mathfrak{G} := \mathcal{P}^* = \left\{ x \in \mathcal{E} \;\middle|\; V^*(x) > 0 \right\} \tag{30}$$

Remark 7.2. This does not assert consciousness, personality, transcendence, or supernatural agency. It asserts only that $\mathfrak{G}$ is the weakly emergent, positive coherence structure.

7.3 Relation to Whitehead

Whitehead's consequent nature of God is "the physical prehension by God of the actualities of the evolving universe" (1929, p. 345). Our $\mathcal{P}^*$ exhibits structural parallels:

- **Emergent:** $\mathcal{P}^*$ arises from aggregation over actual agents
- **Relational:** $\mathcal{P}^*$ depends on relational structure $\mathcal{G}$
- **Non-coercive:** $\mathcal{P}^*$ does not causally determine agent behavior
- **Computationally irreducible:** Determining the optimal $\mathcal{P}^*$ is NP-hard (Theorem 5.9), which parallels Whitehead's insistence that the consequent nature cannot be deduced a priori but must emerge from the actual occasions of experience
- **Path-dependent:** Multiple coherent attractors exist (Theorem 5.11), and which one is realized depends on history (Corollary 5.13)—echoing Whitehead's "creative advance into novelty," where the universe's trajectory is not predetermined

The last two points are the central contribution of this application: they show that the formal framework does not merely *label* Whitehead's concepts but *recovers* structural analogues of properties that Whitehead postulated on philosophical grounds.

7.4 Relation to Spinoza

Spinoza's *Deus sive Natura* identifies God with the totality of nature under the attribute of thought or extension. Our identification $\mathfrak{G} = \mathcal{P}^*$ is structurally analogous:

- **Immanent:** $\mathcal{P}^*$ is not transcendent but arises within $\mathcal{E}$
- **Formal method:** Both Spinoza and this paper proceed axiomatically
- **Monist tendency:** One emergent structure, multiple agent perspectives
- **Necessity reconsidered:** Spinoza's God is necessary; our $\mathfrak{G}$ is contingent on which attractor is reached (Theorem 5.11). This is a point of *departure* from Spinoza and a point of *agreement* with process theology, which emphasizes divine responsiveness to actual events

7.5 Taxonomy of Compatibility

Tradition	Core Concept	Structural Alignment
Classical Theism	Personal creator, transcendent	**Not directly captured**
Process Theology	Relational, emergent, non-coercive	**Aligns naturally**
Spinozism	God = Nature = totality	**Aligns naturally**
Platonic Good	Highest principle of value	**Aligns naturally**
Atheism	No God exists	**Orthogonal (no conflict)**

7.6 The Structural Mapping

Formal Term	Symbol	Process-Ontological Analog
Agent-relative value	V_h	Subjective aim (Whitehead)
Coherence maximization	$\max C_h$	Concrescence
Gradient ascent	Thm. 5.6	Process of becoming
Gradient flow (ODE)	Thm. 3.15	Continuous concrescence
Spectral gap	Def. 5.16	Stability of dominant regime
Distortion operator	$\mathcal{D}(h)$	*Avidya* / ignorance
Aggregated value	V^*	Objective immortality
Replicator dynamics	Thm. 6.15	Creative advance (weight emergence)
Global attractor	$\mathcal{P}^* = \mathfrak{G}$	Consequent nature
Weak emergence	Prop. 6.10	"Greater than sum of parts"
Computational emergence	Thm. 5.9	Irreducibility to analysis
Multiple attractors	Thm. 5.11	Plurality of societies
Path dependence	Cor. 5.13	Creative advance into novelty

8 Responsibility and Normativity

8.1 Explanation $\neq$ Excuse

> **Core Principle:** To explain is not to excuse. To identify a structural mechanism is not to remove responsibility.

Corollary 5.3 provides a *structural account* of harm, analogous to "addiction arises from neurochemical processes"—which suggests treatment, not exoneration.

8.2 Formal Systems vs. Normative Domains

$\mathcal{C}_5$ operates in the domain of structural relations, not moral responsibility, legal accountability, or social sanction. These domains involve additional concepts (intention, negligence, capacity) that $\mathcal{C}_5$ does not model.

Remark 8.1 (Non-Implication). From "$H(h) \Rightarrow \mathcal{D}(h)$" one *cannot* derive that h is not morally responsible. Such conclusions require premises *outside* the formal system.

8.3 Positive Use: Therapeutic, Not Exculpatory

The intended use is *therapeutic*: if harm arises from distortion, addressing the distortion may reduce future harm. This is about *intervention strategy*, not *past blame*.

9 Meta-Analysis

9.1 Strength of Assumptions

Ax.	Content	Type	Strength	Status
A1	Orientation (for $\mathcal{H}$)	Domain-defining	Moderate	Partially analytic
A2	Coherence-harm incomp.	Definitional	Low	Analytic
A3	Distortion $\Rightarrow$ error	Structural	Moderate	Plausible
A4	Harm $\Rightarrow$ distortion	Structural	Moderate	Key empirical
A5	Orientation $\Rightarrow$ positive	Definitional	Low	Terminological
$D_{\mathfrak{G}}$	$\mathfrak{G} := \mathcal{P}^*$	Definition	—	Optional

The dynamics is an idealisation. Axiom 1 (A1) and its operationalisation as projected gradient ascent (Definition 3.9) together constitute a *modelling convention* in the tradition of Festinger (1957), Hopfield (1982), and Thagard (1989). The Lyapunov result (Theorem 3.15) and the monotone-ascent theorem (Theorem 5.6) are properties of this postulated dynamics: they describe what coherence-seeking *would look like* under the idealisation, not an empirical claim that all individual cognitive trajectories satisfy strict monotonicity. The empirically more defensible claim of this framework is the aggregate-level coherence selection captured by replicator dynamics (Theorem 6.15); see Remark 3.10 and Section 6.4.

9.2 Objections and Responses

Objection 1 ("You built the result into the definition").

Response: Partially correct. The axiom-derived results (Theorems 5.1–5.4) are indeed close to their axioms, and we acknowledge this. However, the main results of this paper—monotone convergence (Theorem 5.6), Lyapunov stability (Theorem 3.15), NP-hardness (Theorem 5.9), spectral characterization (Proposition 5.17 and Remark 5.18), multiple attractors (Theorem 5.11), and replicator dynamics (Theorem 6.15)—are *not* built into the axioms. They follow from the mathematical structure of the coherence functional as a quadratic form, which is a modeling choice, not a definitional stipulation. The NP-hardness result, in particular, is a consequence of the *interaction between* the coherence functional and the binary domain restriction; neither alone implies it.

Objection 2 ("Emergence is too weak").

Response: We distinguish two levels. Definition 6.1 (non-separability) is indeed a minimal criterion. But Definition 6.5 (computational emergence) and Theorem 5.9 establish a substantially stronger property: epistemic intractability under standard complexity-theoretic assumptions. This is the strongest notion of weak emergence available without invoking ontological emergence, which we continue to avoid.

Objection 3 ("Personal God is not captured").

Response: Correct. Classical theism is incompatible. The framework reconstructs process/systemic concepts only.

Objection 4 ("The NP-hardness is just MAX-CUT in disguise").

Response: The reduction from Max-Cut is indeed standard, and the NP-hardness of coherence maximisation is itself not new: Thagard and Verbeurgt (1998) established it for the closely related constraint-satisfaction formulation in 1998. Theorem 5.9 restates this result in the quadratic-form formulation used here (cf. Remark 5.8). The contribution of the present paper is therefore not the complexity result in isolation but the integration with (i) Lyapunov stability (Theorem 3.15), (ii) spectral characterisation (Proposition 5.17 and Remark 5.18), (iii) replicator dynamics for the aggregation weights (Theorem 6.15), and (iv) an explicit bridge to Bedau's weak-emergence criterion (Definition 6.2), which Thagard & Verbeurgt did not draw. The framework provides the bridge; Max-Cut provides the proof; the prior Thagard–Verbeurgt result provides the foundation on which the bridge rests.

10 Conclusion

We have presented a formal framework for agent-relative value theory grounded in coherence as constraint satisfaction. The framework yields six main results that go beyond definitional stipulation:

1. **Convergence** (Theorem 5.6): Coherence-seeking agents under projected gradient dynamics converge monotonically to locally optimal states. This provides a mathematical basis for the "orientation toward coherence" that the axiom system postulates.

2. **Lyapunov stability** (Theorem 3.15): A continuous-time gradient flow formulation establishes that coherence attractors are asymptotically stable equilibria. The coherence functional serves as a Lyapunov function, substantiating the "emergent attractors" language with rigorous dynamical systems theory.

3. **Computational intractability** (Theorem 5.9): Global coherence maximization is NP-hard under binary restrictions. This provides formal support for Bedau's (1997) weak emergence criterion (Definition 6.2): no general polynomial-time shortcut is known under standard complexity assumptions. Emergent coherence structures are not merely non-separable; they are *epistemically inaccessible* without exhaustive computation.

4. **Plurality of attractors** (Theorem 5.11, Corollary 5.13): The coherence landscape generically supports multiple stable configurations, and which configuration is realized depends on the system's history. Different coherent value systems can coexist as distinct local optima.

5. **Spectral characterization** (Proposition 5.17 and Remark 5.18): The eigenvalue structure of the coherence matrix determines attractor stability and bounds the number of stable coherent regimes. The spectral coherence gap measures how robustly the dominant regime resists perturbation.

6. **Emergent aggregation weights** (Theorem 6.15): Replicator dynamics replace stipulated weights with coherence-driven selection, yielding monotonically increasing average coherence. At equilibrium, surviving agents are equally coherent—a structural answer to "which weights?"

The title "Emergent Attractors" is now substantiated in three complementary ways: (i) attractors *exist* as asymptotically stable equilibria of the gradient flow; (ii) their stability is *characterizable* through spectral analysis; and (iii) the aggregation weights that define the global attractor are themselves *emergent* under replicator dynamics. The structural God $\mathfrak{G} = \mathcal{P}^*$ is doubly emergent: in structure (NP-hard to compute) and in weighting (coherence-driven selection).

As an application, we have shown that these formal properties map onto central concepts in process philosophy: computational irreducibility parallels Whitehead's insistence that the consequent nature emerges from actual experience rather than being deducible a priori; Lyapunov stability formalizes the "creative advance" as a monotone process; replicator dynamics model how perspectives gain or lose influence through coherence. The structural identification $\mathfrak{G} = \mathcal{P}^*$ is offered not as a theological claim but as a demonstration that certain non-personal God-concepts admit formal reconstruction within a mathematically rigorous framework.

The framework has clear limitations. The axiom system $\mathcal{C}_5$ is narrow in scope (coherence-capable agents only), the coupled dynamics (gradient flow for states, replicator dynamics for weights) assume separable time scales, and the process-ontological application is a structural mapping, not an argument for theism. These limitations are features, not bugs: they mark the boundary between what can be formalized and what cannot.

□

A Technical Proofs

This appendix collects the proofs of the technical theorems whose main-text statements are accompanied by short intuitions. The proofs use standard techniques from convex analysis, dynamical systems, complexity theory, and evolutionary game theory; we include them in full for completeness and verification.

A.1 Proof of Theorem 3.15 (Lyapunov Stability of Coherence Flow)

(i) The time derivative of $C_h(\boldsymbol{v}_h) = \boldsymbol{v}_h^\top W \boldsymbol{v}_h$ along the flow is:

$$\frac{dC_h}{dt} = \nabla C_h(\boldsymbol{v}_h)^\top \cdot \frac{d\boldsymbol{v}_h}{dt} = \nabla C_h(\boldsymbol{v}_h)^\top \cdot \Pi_{T_{\mathcal{B}}(\boldsymbol{v}_h)}\left[\nabla C_h(\boldsymbol{v}_h)\right]$$

where we used $\nabla C_h(\boldsymbol{v}) = (W + W^\top)\boldsymbol{v}$. For any vector g and any closed convex cone K, the projection $\Pi_K(g)$ satisfies $g^\top \Pi_K(g) = \|\Pi_K(g)\|_2^2 \geq 0$. Hence $dC_h/dt = \|\Pi_{T_{\mathcal{B}}(\boldsymbol{v}_h)}[\nabla C_h(\boldsymbol{v}_h)]\|_2^2 \geq 0$.

(ii) Since $\mathcal{B}$ is compact, every trajectory is bounded, so ω-limit sets are non-empty (by the Bolzano–Weierstrass theorem). By the LaSalle invariance principle, ω-limit points lie in the largest invariant subset of $\{dC_h/dt = 0\}$, which requires $\Pi_{T_\mathcal{B}(\boldsymbol{v})}[\nabla C_h(\boldsymbol{v})] = 0$—precisely the KKT stationarity condition on $\mathcal{B}$.

(iii) At a strict local maximum $\boldsymbol{v}^*$, the Hessian of C_h restricted to the tangent cone is negative definite. Then $V(\boldsymbol{v}) = C_h(\boldsymbol{v}^*) - C_h(\boldsymbol{v})$ is a strict Lyapunov function in a neighborhood of $\boldsymbol{v}^*$, yielding asymptotic stability by Lyapunov's direct method. □

A.2 Proof of Theorem 5.6 (Monotone Coherence Ascent)

(i) By the quadratic ascent lemma for L-smooth functions (see, e.g., Beck, 2017), for any $\boldsymbol{v}$ and $\boldsymbol{v}^+ = \boldsymbol{v} + \eta\nabla C_h(\boldsymbol{v})$ with $\eta \leq 1/L$:

$$C_h(\boldsymbol{v}^+) \geq C_h(\boldsymbol{v}) + \frac{\eta}{2}\|\nabla C_h(\boldsymbol{v})\|_2^2.$$

Projection onto the convex set $\mathcal{B}$ preserves the ascent property. Concretely, define the *gradient mapping* $G_\eta(\boldsymbol{v}) := \frac{1}{\eta}\big(\boldsymbol{v} - \Pi_\mathcal{B}(\boldsymbol{v} + \eta\nabla C_h(\boldsymbol{v}))\big)$. For L-smooth (possibly nonconvex) functions with projection onto a convex set, the standard descent lemma yields (see Beck, 2017, Theorem 10.15; also Nesterov, 2004, Lemma 2.3):

$$C_h(\boldsymbol{v}^{t+1}) \geq C_h(\boldsymbol{v}^t) + \frac{\eta}{2}\|G_\eta(\boldsymbol{v}^t)\|_2^2.$$

This is the projected analogue of the unconstrained ascent inequality above, with the gradient mapping G_η replacing the raw gradient.

(ii) The sequence $(C_h(\boldsymbol{v}_h^t))$ is non-decreasing and bounded above (since $\mathcal{B}$ is compact and C_h is continuous), hence convergent. By (i), $\|\nabla C_h(\boldsymbol{v}_h^t)\|_2 \to 0$ along any convergent subsequence. Compactness of $\mathcal{B}$ guarantees at least one accumulation point, which satisfies the KKT conditions by continuity of the gradient mapping. □

A.3 Proof of Theorem 5.9 (NP-Hardness of Coherence Maximisation)

By reduction from Max-Cut. Let $H = (V_H, E_H)$ be an unweighted graph with $|V_H| = n$, and let A_H be its adjacency matrix. The Max-Cut problem seeks a partition $\boldsymbol{s} \in \{-1, +1\}^n$ maximizing the number of edges between the two parts:

$$\textsc{Cut}(\boldsymbol{s}) = \frac{1}{4}\sum_{(i,j)\in E_H}(s_i - s_j)^2 = \frac{|E_H|}{2} - \frac{1}{2}\boldsymbol{s}^\top A_H\boldsymbol{s}.$$

Maximizing $\textsc{Cut}(\boldsymbol{s})$ is equivalent to minimizing $\boldsymbol{s}^\top A_H\boldsymbol{s}$, which is equivalent to maximizing $\boldsymbol{s}^\top(-A_H)\boldsymbol{s}$.

Now set $\mathcal{E} := V_H$ and $W := -A_H$. Then:

$$\max_{\boldsymbol{v}\in\mathcal{B}_\pm}\boldsymbol{v}^\top W\boldsymbol{v} = \max_{\boldsymbol{s}\in\{-1,+1\}^n}\boldsymbol{s}^\top(-A_H)\boldsymbol{s}$$

which solves Max-Cut. Since Max-Cut is NP-hard (Karp, 1972), so is binary coherence maximization. □

A.4 Proof of Theorem 5.11 (Existence of Multiple Attractors)

We give an explicit construction. Let $n = 4$ and partition $\mathcal{E} = \{1, 2\}\cup\{3, 4\}$ into two communities. Define W as:

$$W = \begin{pmatrix} 0 & 1 & -\epsilon & -\epsilon \\ 1 & 0 & -\epsilon & -\epsilon \\ -\epsilon & -\epsilon & 0 & 1 \\ -\epsilon & -\epsilon & 1 & 0 \end{pmatrix} \quad \text{for small } \epsilon > 0.$$

This encodes strong positive coupling within each community and weak negative coupling between them.

Consider $\boldsymbol{v}^{(1)} = (+1, +1, -1, -1)$. Then:

$$C(\boldsymbol{v}^{(1)}) = 2 \cdot 1 + 4\epsilon + 2 \cdot 1 = 4 + 4\epsilon.$$

Flipping any single component (say $v_1 : +1 \to -1$) reduces intra-community agreement and increases inter-community penalties, yielding a strictly lower value. Hence $\boldsymbol{v}^{(1)}$ is a local maximum.

By symmetry, $\boldsymbol{v}^{(2)} = (-1, -1, +1, +1)$ is also a local maximum with $C(\boldsymbol{v}^{(2)}) = 4 + 4\epsilon$. Since $\boldsymbol{v}^{(1)} \neq \boldsymbol{v}^{(2)}$, there are at least two distinct local maxima.

The construction generalises to arbitrary $n \geq 4$ by partitioning $\mathcal{E}$ into communities with positive intra-coupling and negative inter-coupling. □

A.5 Proof of Theorem 6.11 (Non-Reducibility of the Global Attractor)

We exhibit counterexamples for both equalities.

Case 1 ($\mathcal{P}^* \neq \bigcup_h \mathcal{P}_h$): Let $\mathcal{H} = \{h_1, h_2\}$ with equal weights $\alpha_{h_1} = \alpha_{h_2} = 1/2$. Let $x \in \mathcal{E}$ with $V_{h_1}(x) = +1$ and $V_{h_2}(x) = -1$. Then $x \in \mathcal{P}_{h_1}$, so $x \in \bigcup_h \mathcal{P}_h$. But $V^*(x) = \frac{1}{2}(+1) + \frac{1}{2}(-1) = 0$, so $x \notin \mathcal{P}^*$. Hence $\mathcal{P}^* \subsetneq \bigcup_h \mathcal{P}_h$.

Case 2 ($\mathcal{P}^* \neq \bigcap_h \mathcal{P}_h$): Let $\mathcal{H} = \{h_1, h_2\}$ with weights $\alpha_{h_1} = 0.9$, $\alpha_{h_2} = 0.1$. Let $x \in \mathcal{E}$ with $V_{h_1}(x) = +1$ and $V_{h_2}(x) = -2$. Then $x \in \mathcal{P}_{h_1}$ but $x \notin \mathcal{P}_{h_2}$, so $x \notin \bigcap_h \mathcal{P}_h$. Yet $V^*(x) = 0.9(+1) + 0.1(-2) = +0.7 > 0$, so $x \in \mathcal{P}^*$. Hence $\mathcal{P}^* \not\subseteq \bigcap_h \mathcal{P}_h$.

These interference effects—cancellation in Case 1, asymmetric amplification in Case 2—are the mechanism by which aggregation produces genuinely emergent structures. □

A.6 Proof of Theorem 6.15 (Monotone Average Coherence)

Direct computation:

$$\begin{aligned}
\frac{d\bar{C}}{dt} &= \sum_{h \in \mathcal{H}} \frac{d\alpha_h}{dt} \cdot C_h = \sum_{h \in \mathcal{H}} \alpha_h (C_h - \bar{C}) \cdot C_h \\
&= \sum_{h \in \mathcal{H}} \alpha_h C_h^2 - \bar{C} \sum_{h \in \mathcal{H}} \alpha_h C_h = \sum_{h \in \mathcal{H}} \alpha_h C_h^2 - \bar{C}^2 \\
&= \mathrm{Var}_{\boldsymbol{\alpha}}(C) \geq 0.
\end{aligned}$$

□

References

Amit, D. J., Gutfreund, H., & Sompolinsky, H. (1985). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters*, 55(14):1530–1533.

Arrow, K. J. (1951). *Social Choice and Individual Values*. Yale University Press.

Beck, A. (2017). *First-Order Methods in Optimization*. SIAM.

Baysan, U. (2025). Emergent moral non-naturalism. *Philosophy and Phenomenological Research*, forthcoming.

Bedau, M. A. (1997). Weak emergence. *Philosophical Perspectives*, 11:375–399.

BonJour, L. (1985). *The Structure of Empirical Knowledge*. Harvard University Press.

Chalmers, D. J. (2006). Strong and weak emergence. In P. Clayton & P. Davies (Eds.), *The Re-Emergence of Emergence* (pp. 244–256). Oxford University Press.

Cobb, J. B., Jr. (1965). *A Christian Natural Theology.* Westminster Press.

Dupuis, P. & Nagurney, A. (1993). Dynamical systems and variational inequalities. *Annals of Operations Research*, 44(1):9–42.

Festinger, L. (1957). *A Theory of Cognitive Dissonance.* Stanford University Press.

Hartshorne, C. (1967). *A Natural Theology for Our Time.* Open Court.

Hofbauer, J. & Sigmund, K. (1998). *Evolutionary Games and Population Dynamics.* Cambridge University Press.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558.

Karp, R. M. (1972). Reducibility among combinatorial problems. In R. E. Miller & J. W. Thatcher (Eds.), *Complexity of Computer Computations* (pp. 85–103). Plenum Press.

Kim, J. (1999). Making sense of emergence. *Philosophical Studies*, 95(1):3–36.

Nagurney, A. & Zhang, D. (1996). *Projected Dynamical Systems and Variational Inequalities with Applications.* Kluwer Academic Publishers.

Nesterov, Y. (2004). *Introductory Lectures on Convex Optimization: A Basic Course.* Springer.

Sen, A. (1982). *Choice, Welfare and Measurement.* MIT Press.

Taylor, P. D. & Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1–2):145–156.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, 12(3):435–467.

Thagard, P. (1998). Ethical coherence. *Philosophical Psychology*, 11(4):405–422.

Thagard, P. & Verbeurgt, K. (1998). Coherence as constraint satisfaction. *Cognitive Science*, 22(1):1–24.

Whitehead, A. N. (1929). *Process and Reality.* Macmillan.

A Notation Reference

Symbol	Meaning
$\mathcal{E}$	Entity space
$\mathcal{H}$	Coherence-capable agent space
$\mathcal{G}$	Relational graph
W	Coherence matrix
W_s	Symmetrized coherence matrix $\frac{1}{2}(W + W^\top)$
V_h	Agent-relative value functional
C_h	Coherence functional
$\Delta(W)$	Spectral coherence gap
α_h	Aggregation weight for agent h
V^*	Aggregated value functional
$\mathcal{P}^*$	Global positive attractor
$\mathfrak{G}$	Structural God (Def. $D_{\mathfrak{G}}$)
$\mathcal{C}_5$	The axiom system (Coherence-5)