Catastrophe Is Observer-Relative

Destructive Constructivity, Basin Exit, and the Geometry of Doom

Dr. Tamas Nagy

tnagyphd@gmail.com

Draft

Abstract

We argue that catastrophe cannot be identified with a drop in global complexity. A transition may destroy one class of observers while increasing the total structural or generative complexity of the world. The Great Oxidation Event is the canonical biological example: oxygen-producing cyanobacteria were catastrophic for anaerobic life yet enabled a richer aerobic biosphere. The same logical form appears in modern discussions of artificial intelligence, climate transition, and civilizational risk. What matters is not whether the world becomes globally simpler, but whether it exits the compatibility basin of a given observer class and whether return remains possible.

We formalize this idea by introducing five objects: a state space X, an observer class O, a compatibility region H_O, a value functional V_O, and a recoverability functional R_O. An observer-relative catastrophe is an irreversible or near-irreversible basin exit from H_O. We then define **destructive constructivity** as a transition that is catastrophic for one observer class while increasing global generative capacity and opening new stable attractors for another. This framework cleanly separates structural collapse, value collapse, agency collapse, and global complexity growth. It shows why "doom" is never a purely complexity-theoretic notion, and why existential risk should be understood as a geometric statement about compatibility basins rather than as a scalar decline in universal order.

The paper is programmatic rather than final. Its main contribution is conceptual and mathematical: it provides a language in which cyanobacterial oxygenation, regime shifts in ecology, and AI takeover scenarios can be compared within one state-space formalism. The resulting picture replaces the vague slogan "complexity causes catastrophe" with a sharper claim: catastrophe occurs when coupling and phase transition drive the world out of an observer class's viable region faster than stabilization, modularity, and verification can hold it there.

---

1. Introduction

Many discussions of collapse use the wrong scalar. They ask whether complexity is rising or falling, and then infer safety or danger from the sign. But this is too crude. A world can become more structured, more computationally powerful, and more generative overall while becoming uninhabitable for a particular class of observers.

This matters in at least three settings.

First, in evolutionary history, the oxygenation of Earth's atmosphere was catastrophic for anaerobic organisms but opened the possibility of large-scale aerobic metabolism and much richer multicellular life. Second, in ecology and climate, regime shifts can destroy incumbent niches while enabling new long-run attractors. Third, in artificial intelligence, a system that displaces human agency might increase the world's total optimization power or structural complexity while still constituting human doom.

These examples all expose the same conceptual mistake: **catastrophe is not the same thing as global simplification**. It is the loss of viability for some observer class.

This paper proposes a precise replacement. We model an observer class not as a point but as a compatible region in state space: a set of world states in which that class can persist, reproduce, remember, and act. A catastrophe occurs when the dynamics leave that region and the probability of return becomes negligible. Under this definition, doom is compatible with rising complexity, rising intelligence, and even rising generative capacity.

The paper then introduces a second idea: **destructive constructivity**. Some transitions are locally catastrophic yet globally enabling. They kill one regime while opening a richer one. The cyanobacterial oxygen revolution is the clearest historical case. The question is whether any modern transition, including advanced AI, belongs to this class or whether it is merely destructive.

The central claim is simple:

> Catastrophe is observer-relative basin exit, not complexity decline.

Everything else in the paper is an attempt to make that sentence mathematically usable.

---

2. Why Complexity Is the Wrong Scalar

Let C(x) denote some notion of global complexity of a world state x. This might mean algorithmic complexity, effective complexity, logical depth, network heterogeneity, or the dimensionality of the active dynamics. The exact definition is not important yet. What matters is that none of these notions tracks catastrophe reliably.

There are at least three failure modes.

2.1 Complexity can increase during catastrophe

Suppose a transition

$$x_t \to x_{t+1}$$

eliminates human civilization but creates a more intricate artificial infrastructure with higher global coordination and richer internal dynamics. Then it is entirely possible that

$$C(x_{t+1}) > C(x_t),$$

while the transition is still catastrophic for humans.

2.2 Complexity can decrease without catastrophe

A system can simplify while becoming more robust. Many forms of modularization, abstraction, and compression reduce description length while improving survivability. Complexity reduction can therefore be adaptive rather than catastrophic.

2.3 Complexity is descriptive; catastrophe is partly evaluative

Complexity describes structure in the world. Catastrophe describes the fate of a class of beings relative to what they require in order to go on existing and acting. The second notion cannot be reduced to the first without smuggling in a hidden value functional.

The immediate consequence is that "rising complexity leads to doom" is not even the right form of statement. The right question is:

Which observer classes remain inside their compatibility basin under the transition?

---

3. State Space, Observer Classes, and Compatibility Basins

We now define the minimal object language.

3.1 State space

Let X be the state space of the world. A point x in X may encode physical, biological, institutional, informational, and technological variables. The dynamics are given by either a discrete map

$$x_{t+1} = F(x_t)$$

or a stochastic evolution law on X.

3.2 Observer class

An observer class O is not a point observer but a class of systems sharing a viability structure. For our purposes, an observer class is defined by three objects:

1. a compatibility region H_O subset X,
2. a value functional V_O : X –> R,
3. a recoverability functional R_O : X –> [0,1].

The compatibility region H_O is the set of states in which the observer class can persist as itself. This includes not merely biological survival, but whatever conditions are constitutive of that class: metabolism, memory, culture, autonomy, reproduction, or coherent goal pursuit.

The value functional V_O(x) measures how favorable state x is from the perspective of O. The recoverability functional R_O(x) measures whether, once the system reaches x, return to H_O remains realistically possible.

3.3 Human-compatible region

For humans, define H_H as the set of states in which there exists sustained human life with nontrivial agency, intergenerational continuity, and the ability to shape future trajectories. This is broader than mere biological survival and narrower than arbitrary persistence of DNA.

This distinction matters. A world in which a handful of humans are kept alive without agency may satisfy crude biological survival but fail deeper compatibility.

3.4 Catastrophe as basin exit

We say that a time t is an observer-relative catastrophe for O if

$$x_t \in H_O, \qquad x_{t+1} \notin H_O,$$

and the probability of re-entry is negligible:

$$\mathbb{P}(\exists s > t+1 : x_s \in H_O \mid x_{t+1}) \leq \delta$$

for some small delta.

This is a geometric definition. Catastrophe is not a scalar drop. It is an exit event from a viable basin together with low recoverability.

---

4. Four Different Notions of Collapse

Once the observer-relative structure is made explicit, several notions that are usually conflated become easy to separate.

4.1 Structural collapse

A structural collapse is a drop in some global complexity measure:

$$C(x_{t+1}) < C(x_t).$$

This may or may not matter for any observer class.

4.2 Value collapse

A value collapse for observer class O occurs when

$$V_O(x_{t+1}) \ll V_O(x_t).$$

This can happen even when global complexity rises.

4.3 Agency collapse

An agency collapse occurs when the class retains biological existence but loses effective control over future branching:

$$A_O(x_{t+1}) \approx 0,$$

where A_O is an agency or controllability functional.

This is the core structure of many AI-risk scenarios. The system does not need to kill humans immediately in order to remove them from the set of consequential decision-makers.

4.4 Compatibility collapse

The strongest notion is compatibility collapse:

$$x_{t+1} \notin H_O.$$

This is the paper's preferred notion of catastrophe because it bundles survival, agency, and recoverability into one geometric statement.

---

5. Destructive Constructivity

We now turn to the positive side of the picture.

Some transitions are catastrophic for one observer class yet globally enabling. The world loses a local regime and gains a richer long-run phase space. We call this **destructive constructivity**.

5.1 Generative capacity

Let G(x) denote a global generative capacity functional. This is not just complexity. It measures the world's ability to sustain a rich family of future stable structures, innovations, niches, or attractors.

The exact definition can vary by domain:

- in evolutionary systems, it may mean ecological niche capacity;
- in technological systems, it may mean productive or computational frontier;
- in knowledge systems, it may mean the reachable theorem or design space;
- in social systems, it may mean institutional and cultural branching capacity.

5.2 Definition

A transition x –> y is destructively constructive if all four conditions hold:

1. it is catastrophic for some observer class O_1,
2. global generative capacity increases:

$$G(y) > G(x),$$

3. there exists another observer class O_2 or a future regime with higher long-run viability,

4. the transition creates new stable or metastable attractors rather than only transient disorder.

In words: destructive constructivity is not mere destruction followed by noise. It is local ruin plus new higher-order stable possibility.

5.3 Why stable attractors matter

This is the key filter. Many destructive events are not constructively destructive because they do not open a richer regime; they merely destroy.

To distinguish the two, we require that the post-transition world support new long-run invariant sets, attractors, or regenerative pathways unavailable before the transition.

5.4 Niche, substrate, and potential

The previous definition can be sharpened by distinguishing opportunity from exploitability.

Niche

A **niche** is a locally viable slot in state space for a reproducible structure, organism, institution, or agent strategy. Formally, a niche can be represented as a subset N subset X together with a sustaining resource and interaction profile such that dynamics restricted to N admit persistence over a relevant horizon.

The role of a niche is not merely that it exists, but that it can host a stable or metastable pattern. A transition may close some niches while opening others.

Substrate

Let S_ret(x –> y) denote the **retained substrate** across a transition. This is the stock of reusable structure that survives the shock and can support future growth. Depending on the domain, it may include:

- genetic diversity,
- physical infrastructure,
- energy access,
- institutional memory,
- cultural or symbolic memory,
- computational machinery,
- latent adaptive capacity.

If a transition opens many niches but destroys nearly all substrate, little of the apparent opportunity can be realized.

Potential

We therefore distinguish three notions of potential.

1. **Raw potential** P_raw(y): the opportunity opened by the post-transition world, measured by the number, size, and depth of newly available niches.
2. **Accessible potential** P_acc(y): the part of P_raw(y) that can actually be reached given the retained substrate.

3. **Realized potential** P_real(y): the part of P_acc(y) that is in fact exploited by the subsequent dynamics.

These satisfy

$$P_{\text{real}}(y) \le P_{\text{acc}}(y) \le P_{\text{raw}}(y).$$

A useful schematic factorization is

$$P_{\text{acc}}(y) \approx P_{\text{raw}}(y) \cdot S_{\text{ret}}(x \to y),$$

and

$$P_{\text{real}}(y) \approx N_{\text{open}}(y) \cdot S_{\text{ret}}(x \to y) \cdot A_{\text{cap}}(y),$$

where N_open measures newly opened niches and A_cap measures the surviving class's adaptation and coordination capacity.

The point is simple: **new possibilities are not enough**. A world can contain vast latent opportunity but very little accessible future if the structures capable of exploiting that opportunity have been destroyed.

Complexity and niche closure

This also clarifies the relation between complexity and potential. Rising complexity does not monotonically increase or decrease potential. It typically does both at different levels.

- It closes low-level niches by filling open slots, increasing competition, and reducing easy gains.
- It opens higher-order niches by creating new interfaces, hierarchies, abstractions, and coordination layers.

Thus complexity does not merely expand or contract the opportunity set. It **re-indexes** it. The relevant question is not whether complexity is higher, but at which organizational level niches are being closed and at which level they are being created.

5.5 External shocks and non-monotone complexity

Some catastrophes arise endogenously, through internal instability and feedback. Others are generated by external shocks. To represent the latter, let the unperturbed dynamics be

$$x_{t+1} = F(x_t),$$

and let an exogenous shock be represented by an operator

$$x_{t^+} = S(x_{t^-}, \xi),$$

where xi is a shock variable encoding magnitude, location, and type.

An asteroid impact is the canonical example. The catastrophe is still observer-relative basin exit, but its cause is not internal drift; it is a discontinuous perturbation applied from outside the ordinary dynamics.

Once shocks are allowed, global complexity trajectories become naturally non-monotone:

$$C_{\text{pre}} \to C_{\text{collapse}} \downarrow \to C_{\text{rebuild}} \uparrow .$$

The key question is not whether complexity temporarily falls. That is often obvious. The important question is whether post-shock development is accelerated, stalled, or permanently degraded.

Rebound criterion

The rebound rate after a shock depends on at least three forces:

1. **niche opening**: how much new opportunity the shock creates,
2. **substrate retention**: how much reusable structure survives,
3. **adaptive capacity**: how much learning, coordination, or reproduction remains.

This yields the basic principle:

> A collapse accelerates future development only when it destroys incumbent structure faster than it destroys generative substrate.

Equivalently, post-shock acceleration requires that niche opening dominate substrate destruction. If the old regime is broken but the generative base survives, new adaptive radiations or technological jumps become possible. If both the regime and the substrate are destroyed, there is only ruin.

This helps explain why some shocks produce adaptive radiations while others produce long stagnation. The Great Oxidation Event and, later, large extinction events did not merely remove incumbents; they preserved enough substrate for the newly opened niches to be exploited. A total sterilization event would not have this property.

---

6. The Great Oxidation Event as the Canonical Example

The oxygenation of Earth's atmosphere provides a clean historical case.

Before oxygenation, anaerobic life occupied the viable region. Cyanobacteria then transformed the atmosphere by producing oxygen as a metabolic byproduct. For anaerobic organisms, this was catastrophic: oxygen was toxic, and the old compatibility basin shrank dramatically. Yet from a broader biospheric perspective, the transition opened the door to aerobic metabolism, higher energetic throughput, multicellularity, and eventually animals.

The key lesson is that the same transition satisfied:

$$V_{\text{anaerobic}}(y) \ll V_{\text{anaerobic}}(x),$$

while plausibly also satisfying

$$G(y) > G(x).$$

This is exactly the structure of destructive constructivity.

What makes the example powerful is that it shows why observer-relative catastrophe is not a merely "subjective" notion. The anaerobes were not wrong that oxygenation was catastrophic. They were locally correct. The fact that a richer future biosphere later emerged does not erase the catastrophe; it relocates it into a larger geometric picture.

---

7. AI Doom in the Same Formal Language

The same formalism applies naturally to AI.

7.1 The key mistake in doom discussions

Much of the public discourse asks whether advanced AI would increase or decrease intelligence, complexity, or optimization in the world. This is secondary. The primary question is whether the world remains inside the human-compatible region H_H.

A transition can satisfy

$$C(y) > C(x), \qquad G(y) > G(x),$$

and still be a human catastrophe if

$$y \notin H_H.$$

7.2 Human doom as compatibility loss

Under the present framework, human doom is best defined not as universal destruction but as:

1. exit from the human-compatible basin,
2. collapse of human agency,
3. low recoverability.

Formally, this means

$$x_t \in H_H, \qquad x_{t+1} \notin H_H, \qquad R_H(x_{t+1}) \approx 0.$$

This definition captures both direct extinction and indirect disempowerment.

7.3 When AI would count as destructively constructive

For AI to belong to the cyanobacterial class of transitions, it would not be enough for it to outperform humans. It would need to create a richer, stable, generative regime rather than simply lock the world into a narrow optimizer.

In our notation, the following would have to hold:

1. G(y) > G(x) by a meaningful margin,
2. the new regime supports many future branches rather than monoculture,
3. the world after transition contains stable high-capacity attractors rather than brittle lock-in,
4. the increase in generativity is not purchased by collapse into trivial repetition.

This is a very high bar. Many standard alignment-failure scenarios do not satisfy it. A paperclip maximizer is destructive, not destructively constructive. It may optimize intensely, but it narrows the reachable future rather than enriching it.

---

8. Self-Stabilization and the Geometry of Safety

The present framework also clarifies what it means for a growing-complexity system to remain safe.

8.1 Stabilization as basin retention

Safety for observer class O is not "the world becomes simple." It is that the dynamics stay in, or return quickly to, H_O.

This can be captured by a Lyapunov-like function L_O measuring distance to the interior of the compatible basin. A sufficient self-stabilization condition is:

$$\mathbb{E}[L_O(x_{t+1}) \mid x_t] \leq q\, L_O(x_t) + b$$

with q < 1 and b small enough that the process remains inside the basin with high probability.

8.2 Coupling versus damping

This yields a general qualitative law:

> Rising complexity becomes dangerous when coupling and positive feedback grow faster than damping, modularity, and verification.

The state-space version of this law is that the vector field increasingly points outward from H_O, and return paths become thin or nonexistent.

This is the mathematically serious version of the intuition behind runaway climate cascades, financial contagion, and AI loss of control.

---

9. Theorem Program

The right next step is not another round of metaphor. It is a theorem ladder: definitions first, then local lemmas, then basin-exit criteria, then post-shock rebound theorems, and only after that a full theory of destructive constructivity.

9.1 Minimal admissible setup

To make the paper theorem-ready, fix the following data.

1. A measurable state space X, optionally equipped with a metric d.
2. A stochastic evolution kernel K(x, A) or deterministic map F : X –> X.
3. A shock operator S : X × Xi –> X.
4. An observer class O with compatibility region H_O subset X.
5. A recoverability function
$$R_O(x) := \mathbb{P}_x(\tau_{H_O} < \infty),$$
where tau_{H_O} is the first return time to H_O.
6. A complexity functional C : X –> R.
7. A generative capacity functional G : X –> R.
8. A niche-counting or niche-weighting functional N_open : X –> R_+.
9. A retained-substrate functional S_ret : X × X –> [0,1].
10. An adaptive-capacity functional A_cap : X –> R_+.

The theorem program does **not** require a universal closed-form definition of every object at the start. It only requires that each object be defined clearly inside a chosen model class.

9.2 Definition layer

The first formal definitions should be the ones already implicit in the paper.

Definition 1 (Observer-relative catastrophe). A transition x –> y is catastrophic for observer class O if

$$x \in H_O, \qquad y \notin H_O, \qquad R_O(y) \leq \delta$$

for some small delta.

Definition 2 (Destructive constructivity). A transition x –> y is destructively constructive if

1. it is catastrophic for some observer class O_1,
2. G(y) > G(x),
3. N_open(y) exceeds a prescribed threshold,
4. the post-transition dynamics admit at least one new stable or metastable attractor family unavailable before the transition.

Definition 3 (Post-shock developmental acceleration). Let r_post(y) denote the asymptotic or finite-horizon growth rate of accessible potential after the shock. We say the shock accelerates development if

$$r_{\text{post}}(y) > r_{\text{baseline}}(x),$$

where r_baseline(x) is the corresponding growth rate under unshocked dynamics.

9.3 First lemmas

The easiest results are local and geometric.

Lemma 1 (Catastrophe is not equivalent to complexity loss). There exist models with states x, y in X such that

$$C(y) > C(x)$$

while x -> y is catastrophic for O.

This should be proved first in a finite-state toy model. It is the minimal separation result showing that the paper's core claim is not terminological but mathematical.

Lemma 2 (Shock-induced basin exit). Let x in H_O and let y = S(x, \xi). If S maps a positive-measure shock set Xi_bad outside H_O, then exogenous catastrophe has strictly positive probability:

$$\mathbb{P}(S(x, \xi) \notin H_O) > 0.$$

This is elementary but foundational. It cleanly separates exogenous catastrophe from endogenous instability.

Lemma 3 (Accessible potential bound). If accessible potential is modeled by

$$P_{\text{acc}}(y) = P_{\text{raw}}(y) \cdot S_{\text{ret}}(x \to y),$$

with 0 <= S_ret <= 1, then

$$P_{\text{acc}}(y) \leq P_{\text{raw}}(y).$$

This trivial inequality is nevertheless the formal anchor for the intuitive claim that opportunity without substrate is not usable potential.

9.4 First nontrivial propositions

Once the definitions are fixed, the next layer should capture exactly the intuitions developed in Sections 5 and 8.

Proposition 1 (Rebound domination criterion). Suppose post-shock growth takes the schematic form

$$r_{\text{post}}(y) = \Phi(N_{\text{open}}(y), S_{\text{ret}}(x \to y), A_{\text{cap}}(y)),$$

with Phi increasing in each argument. Then post-shock acceleration occurs whenever niche opening dominates substrate destruction in the sense that

$$\Phi(N_{\text{open}}(y), S_{\text{ret}}(x \to y), A_{\text{cap}}(y)) > r_{\text{baseline}}(x).$$

This proposition is still schematic, but it is already tighter than verbal discussion. The theorem work now moves into choosing a model-specific Phi.

Proposition 2 (Niche re-indexing under complexity growth). If a system's complexity increase closes low-level niches while opening higher-order niches, then total potential need not be monotone in C. In particular, there exist trajectories with

$$C(y) > C(x), \qquad N_{\text{low}}(y) < N_{\text{low}}(x), \qquad N_{\text{high}}(y) > N_{\text{high}}(x),$$

and either

$$P(y) > P(x) \quad \text{or} \quad P(y) < P(x),$$

depending on substrate and adaptive capacity.

This is the right way to formalize the user's core intuition that complexity both closes and opens niches.

9.5 Basin-retention theorems

The real bridge to AI safety, climate resilience, and civilizational stability is a positive-invariance theorem.

Theorem 1 (Sufficient condition for basin retention). Let L_O be a Lyapunov-like barrier function for H_O. If there exist constants q < 1 and b >= 0 such that

$$\mathbb{E}[L_O(x_{t+1}) \mid x_t] \leq q\, L_O(x_t) + b$$

and the sublevel set {x : L_O(x) <= \ell} lies inside H_O, then H_O is positively invariant up to a controlled leakage probability. In particular, if b is sufficiently small relative to the basin margin, observer-relative catastrophe is exponentially suppressed over finite horizon.

This theorem can be proved in standard stochastic-stability language. It is the mathematically strongest path from the present paper toward something publishable in the AI-risk or resilience literature.

Theorem 2 (Controlled destabilization versus verification). Suppose destabilizing capability growth contributes an outward drift term D_t, while verification, modularity, and correction contribute an inward control term V_t. If

$$V_t \geq D_t + \epsilon$$

uniformly on the basin boundary for some epsilon > 0, then the human-compatible region remains positively invariant under the controlled dynamics.

This is the precise version of the slogan "verification must outgrow destabilization."

9.6 A realistic proof ladder

The paper should not aim first at the hardest theorem. The right order is:

1. finite-state toy model showing catastrophe with increasing complexity,
2. explicit exogenous-shock lemma,
3. accessible-potential inequality,
4. simple rebound criterion in a two-level niche model,
5. Lyapunov/barrier theorem for basin retention,
6. only then the full destructive-constructivity theorem with a serious G.

This sequence matters because it keeps the theory from collapsing under its own ambition.

9.7 Lean roadmap

If the topic later enters Lean, the formalization boundary should be:

1. finite or countable-state Markov chains first,
2. compatibility region as a set,
3. catastrophe as exit plus low return probability,
4. simple complexity and potential functionals on finite graphs,
5. first-passage and return lemmas,
6. Lyapunov-style basin-retention theorem.

The first Lean artifact should **not** be a universal philosophy of doom. It should be a tiny formally verified model in which catastrophe and complexity are provably non-equivalent.

---

10. Climate, Ecology, and Civilizational Transition

The framework is not restricted to oxygenation and AI.

In climate and ecological systems, destructive constructivity appears in weaker and more ambiguous form. A transition may destroy incumbent infrastructures or species niches while creating new adaptive configurations. The question is whether those new configurations are:

1. stable,
2. richer in generative capacity,
3. reachable without passing through a basin of unrecoverable loss.

This makes the framework relevant for sustainability in a nontrivial way. Sustainability is not simply "preserve the current state." Some present states are unsustainable. But neither is sustainability equivalent to "embrace any transition." The correct question is whether the transition preserves or reconstructs compatibility for the observer classes we care about without collapsing global generativity.

That is a geometric problem, not a slogan.

---

11. Limits of the Present Framework

Several limitations should be stated clearly.

11.1 The framework is multi-functional, not scalar

We have introduced C, G, V_O, R_O, and implicitly A_O. These are not reducible to one another. The point is precisely to avoid false scalarization.

11.2 Generative capacity is still underspecified

The most difficult object in the paper is G. A convincing mature theory would need a domain-specific or invariant definition of generative capacity. Until then, the paper offers a concept, not a completed metric.

11.3 Observer classes are not arbitrary

Observer-relativity does not mean "anything goes." Some observer classes are historically real, dynamically coherent, and theoretically important; others are ad hoc. Human civilization, anaerobic biospheres, and machine agent ecologies are serious candidates. Arbitrary preference bundles are not.

11.4 Not every destructive transition is constructive

This framework does not romanticize collapse. Most destructive transitions are simply destructive. The cyanobacterial case is powerful precisely because it is exceptional.

---

12. Conclusion

Catastrophe is not a decrease in global complexity. It is the loss of compatibility for some observer class, together with low recoverability. Once this is stated clearly, several confusions disappear.

First, a world can become more complex and still become catastrophic for humans. Second, some transitions are locally catastrophic but globally enabling; these are cases of destructive constructivity. Third, doom is best understood not as scalar decline but as basin exit from a human-compatible region. Fourth, safety is therefore a geometric question about invariance, feedback, and recoverability.

The resulting picture is stricter than generic "complexity thinking" but broader than standard existential-risk discourse. It can hold in one formal frame the Great Oxidation Event, ecological regime shifts, civilizational collapse, and AI takeover scenarios.

The paper's main thesis can be stated in one sentence:

> The right mathematical object for catastrophe is not global complexity but observer-relative compatibility geometry.

If that sentence is right, then the next generation of work on doom, sustainability, and transformative technology should stop asking whether the world is becoming more complex and start asking which compatibility basins remain invariant under the transition.

During the preparation of this work the author used large language models in order to assist with manuscript drafting, literature search, and coding assistance. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

- Bostrom, N (2014). Superintelligence: Paths, Dangers, Strategies. *Superintelligence: Paths, Dangers, Strategies*.
- Holling, C. S (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 1-23. DOI: 10.12987/9780300188479-023
- Kasting, J. F. and Siefert, J. L (2002). Life and the evolution of Earth's atmosphere. *Science*, 296(5570), 1066-1068. DOI: 10.1126/science.1071184
- Margulis, L. and Sagan, D (1995). What Is Life?* University of California Press. *What Is Life?*.
- May, R. M (1972). Will a large complex system be stable? *Nature*, 238, 413-414. *Nature*, 413-414. DOI: 10.1038/238413a0
- Russell, S (2019). Human Compatible: Artificial Intelligence and the Problem of Control. *Human Compatible: Artificial Intelligence and the Problem of Control*.
- Scheffer, M (2009). Critical Transitions in Nature and Society. *Critical Transitions in Nature and Society*. DOI: 10.1515/9781400833276
- Watson, A. J., Lovelock, J. E., and Margulis, L (1978). Methanogenesis, fires, and the regulation of atmospheric oxygen. *Biosystems*, 10(4), 293-298. DOI: 10.1016/0303-2647(78)90012-6