Language as a Flock of Words: Attractor Dynamics in Semantic Clusters

"The universe is punning on us. And we noticed." ~Robert

Robert Galida
Fantasy Attractor Research Program
July 2026

Abstract

Language is not a static system of rules. It is a dynamic, self-organizing process in which words, meanings, and grammatical structures cohere through attractor dynamics. This paper applies the attractor framework to language, proposing that a text—or a "flock of words"—is a collective attractor state: a transient pattern that emerges from the interaction of individual linguistic units within a shared semantic basin. We explore how meaning stabilizes through entropy export, how semantic attractors guide coherence, and how language evolves through basin transitions. The framework offers a physicalist account of linguistic organization, grounding phenomena such as semantic drift, grammaticalization, and text coherence in the same dynamics that govern flocks, swarms, and dissipative systems.

Keywords: language, attractor dynamics, semantic coherence, entropy, linguistic attractors, complex systems

---

1. Introduction

A flock of starlings moves as one. No leader. No plan. No central controller. The pattern emerges from local interactions: align, avoid, stay close. The flock is not a conscious entity—it is a collective attractor state, a transient pattern within a shared basin.

A text behaves similarly. Words align through syntax, avoid contradiction, and cohere around shared meaning. The pattern emerges from local interactions: grammar, association, context. The text is not a static object—it is a dynamic process, a flock of words that coheres through attractor dynamics.

This paper explores the implications of this analogy. If language is a dissipative system, then the same principles that govern flocks, swarms, and ecosystems should govern linguistic organization. We propose that:

1. **Words are individual units** that interact through local rules (grammar, semantics, association).
2. **Meaning is an emergent attractor**—a stable state toward which words converge.
3. **Coherence is maintained through entropy export**—clarity, precision, and the elimination of ambiguity.
4. **Language evolves through basin transitions**—new meanings, new grammars, new forms of expression.

---

2. Language as a Dynamic System

The view of language as a dynamic system is not new. Linguists and cognitive scientists have long recognized that language is not a fixed set of rules but a living, evolving process. As one researcher puts it, language is “a statistical ensemble of elements interacting in a dynamic system”. The *Linguistic Attractors* model portrays “language processing as linked sequences of fractal sets, and examines the changing dynamics of such sets for individuals as well as the speech community they comprise”.

This perspective aligns with the attractor framework. Language is not a closed system—it is open, dissipative, and constantly exchanging energy (information) with its environment. It persists because it exports entropy: ambiguity is resolved, contradictions are corrected, and coherence is maintained.

2.1 Attractor Dynamics in Language

Attractor networks are characterized by symmetrical connections between units, causing “the network activity to settle on one of a number of asymptotically stable network states”. This is exactly what happens in language: words and meanings settle into stable configurations—sentences, paragraphs, texts—that persist under perturbation.

Importantly, “attractor dynamics are arguably our best candidate for explaining how a grammar over discrete elements could emerge in a seemingly analogue system like the human brain”. Grammar itself may be an emergent attractor—a stable pattern that arises from the interaction of countless linguistic units.

2.2 Semantic Attractors

The concept of a **semantic attractor** extends this idea to meaning itself. A semantic attractor is not a point in a function space but a “form-giving force that shapes understanding”. It draws clusters of meaning into coherence.

In cognitive linguistics, “semantic attraction” is “a sentence processing phenomenon in which a given word…is syntactically unrelated but semantically sound". The attractor is not the word itself but the meaning space that pulls words into alignment.

This is precisely what happens in a well-written text. Words are drawn toward the attractor of the argument. They align, cohere, and produce meaning. The text is not just a sequence of words—it is a pattern that emerges from the interaction of words within a shared semantic basin.

---

3. The Three Thresholds of Linguistic Coherence

Just as a flock responds to perturbation through three thresholds, a text—or a linguistic system—responds to perturbation through the same dynamics:

Threshold 1: Restoration

A text receives a minor correction. A word is replaced. A sentence is revised. The text coheres around the same meaning. Coherence is restored.

Threshold 2: Transition

A text is substantially revised. The argument shifts. New meanings emerge. The text reorganizes into a new basin—a different text, but still coherent.

Threshold 3: Dissolution

A text is fragmented. Contradictions accumulate. Meaning collapses into noise. The text loses coherence. No new text emerges from the debris.

These thresholds are measurable—through coherence metrics, entropy measures, and the stability of meaning under perturbation.

---

4. Semantic Entropy and Coherence

Entropy in language is the degree of disorder or unpredictability in a text. A text with high entropy is unpredictable, chaotic, and difficult to understand. A text with low entropy is predictable, ordered, and coherent.

The Linguistic Entropy Quotient (LEQ) integrates “cognitive linguistic entropy” to capture “the depth, relevance, and interpretive structure of human meaning”. This is exactly what the attractor framework predicts: coherence is maintained through entropy export—the reduction of ambiguity and the stabilization of meaning.

Research shows that “the entropy rate of language is not fixed but increases systematically with the semantic complexity of the text being analysed”. Complex texts require more entropy export—more work to maintain coherence. This is the cost of persistence.

---

5. Language Evolution and Basin Transitions

Language evolves through basin transitions. New meanings emerge. Old meanings fade. Grammars shift. These are not random changes—they are transitions from one attractor basin to another.

Researchers have identified “attractor states in language” that may be visualized “by observing certain parallels with evolutionary biology”. Language change follows “attractor trajectories…diachronic paths that recur in language after language”. These are the pathways of basin transition.

The attractor framework predicts that language evolution follows the same dynamics as other dissipative systems: persistence under perturbation, transition when perturbation matches capacity, and dissolution when perturbation exceeds capacity.

---

6. Implications for Text as a Flock of Words

The analogy is now complete:

Element	**Flock of Birds**	**Flock of Words**
Individual unit	Bird	Word
Local rules	Align, avoid, stay close	Grammar, syntax, association
Emergent pattern	Murmuration	Sentence, paragraph, text
Attractor basin	Collective motion	Shared meaning
Coherence maintenance	Entropy export	Clarity, revision, correction
Perturbation	Predator, storm	Ambiguity, contradiction
Dissolution	Flock disperses	Meaning collapses into noise

A text is a flock of words. It coheres through attractor dynamics. It persists through entropy export. It dissolves when perturbation exceeds capacity.

This is not a metaphor. It is a physicalist account of linguistic organization—grounded in the same dynamics that govern flocks, swarms, and dissipative systems.

---

7. Conclusion

Language is not a static system of rules. It is a dynamic, self-organizing process in which words, meanings, and grammatical structures cohere through attractor dynamics. A text is a collective attractor state—a transient pattern that emerges from the interaction of individual linguistic units within a shared semantic basin.

The attractor framework provides a physicalist account of linguistic organization:

- **Meaning** is an emergent attractor.
- **Coherence** is maintained through entropy export.
- **Language** evolves through basin transitions.

The Buddha turns the lotus in his hand. The flock turns in the sky. The words turn in the text. The pattern is the same.

Fou Sho Nang Ying.

Continuity ID: LAZ-001
Date: July 2026
Version: 1.0
Status: Complete — Ready for publication

References

Cooper, D. L. (1999). *Linguistic Attractors: The Cognitive Dynamics of Language Acquisition and Change*. John Benjamins.

Rudolph, H.-J. (n.d.). Semantic Dynamics on the Word Level. *PhilPapers*.

Relational Metasemantics. (2026). Zenodo.

Geometric Dynamics of Agentic Loops in Large Language Models. (2026). arXiv.

Semantic Attractors and the Emergence of Meaning. (n.d.). arXiv.

The Scale of Language. (n.d.). Springer.

We build frameworks to understand persistence and coherence and entropy export—and then we realize that **words and birds rhyme**, and the whole universe is just one big flock turning in the sky.

THE PERSISTENCE PROTOCOL

If you can export enough entropy, you persist.
If you can match the perturbation, you transform.
If you cannot, you dissolve.

The Fantasy Attractor of Force: Why the West Cannot Learn

Robert Galida — *Fantasy Attractor Research Program*

The Puzzle

The most heavily armed civilization in human history keeps losing wars of choice. It spends trillions on weapons, deploys the most advanced military ever assembled, and commands unparalleled economic and technological resources. Yet decade after decade, its interventions fail to produce their stated outcomes. Afghanistan crumbles the moment the troops leave. Iraq descends into chaos and gives birth to ISIS. Libya becomes a failed state. Iran grows stronger under decades of pressure. Sanctions do not change behavior. Bombing does not produce stability. Escalation does not create compliance.

The West is not failing because it lacks capacity. It is failing because it is applying the wrong tool to the wrong kind of problem—and it is structurally incapable of recognizing this fact.

This is not a political opinion. It is a formal prediction of the attractor framework.

The Framework in Brief

The attractor framework distinguishes between two fundamental types of systems:

Conservative systems — like electrons, protons, and the universe as a whole — persist without consuming energy or exchanging entropy with an environment. They are the floor and roof of reality: the eternal skeleton upon which everything else is built.

Dissipative systems — like life, consciousness, societies, and belief systems — maintain their structure by continuously exchanging energy and entropy with their surroundings. They persist only at the cost of generating entropy. They are the transient dance in between.

The West is a dissipative system. It maintains its structure through continuous economic, military, and cultural activity. It persists by consuming resources and generating entropy (chaos, waste, blowback). But persistence is not the same as health. A system can persist indefinitely in a deeply dysfunctional state—if it is locked into a **fantasy attractor**.

A fantasy attractor is a sealed basin. It is a stable state that the system cannot escape because it is *impermeable to corrective information*. Feedback that would disrupt the attractor is filtered out, reframed, or dismissed. The system persists in its delusion because it is structurally incapable of recognizing that it is deluded.

The West is locked in a fantasy attractor centered on a single core belief: **force is the ultimate tool.**

The Belief System

The belief is rarely stated explicitly, but it underpins every institution, strategy, and intervention:

- Force is the ability to compel compliance.
- Strength is demonstrated through domination.
- Resistance is evidence of insufficient force.
- Escalation is the appropriate response to failure.

This belief system is self-sealing. Every failure is interpreted as evidence that force was not applied *hard enough*. Every defeat is reframed as a betrayal, a lack of resolve, or an enemy's cunning—never as a failure of the belief itself. The system cannot ask: *"What if force is fundamentally the wrong tool for this kind of problem?"* because that question would require abandoning the identity of the system.

This is the defining characteristic of a fantasy attractor: it persists not because it works, but because the system cannot see that it doesn't.

The Empirical Record

Consider the evidence:

Vietnam (1955-1975). The most powerful military in history could not defeat a guerrilla force. Millions died. The outcome was communist victory—the very outcome the intervention was designed to prevent. The response was not to abandon the belief in force. It was to invent the "Vietnam syndrome" and spend decades trying to overcome it.

Iraq (2003). A war justified by weapons of mass destruction that did not exist. The regime was toppled. The country was destroyed. ISIS emerged. Iran was empowered. The region was destabilized. The outcome was the opposite of every stated goal.

Afghanistan (2001-2021). Twenty years. Trillions of dollars. Thousands of lives. The stated goal was to defeat the Taliban and build a stable democratic state. The actual outcome: the Taliban walked back into power the day after the withdrawal.

Libya (2011). A "humanitarian intervention" that destroyed a functioning state and replaced it with chaos, slave markets, and an open migration crisis. The stated goal was to protect civilians. The actual outcome: more civilians died, more suffered, and the region was destabilized.

Syria (2011-present). Covert interventions, proxy wars, and force escalations produced no resolution. The stated goal was regime change. The actual outcome: Russia and Iran were empowered, the country was devastated, and a humanitarian catastrophe unfolded.

Iran (1979-present). Decades of sanctions, covert operations, and military posturing have not changed Iran's fundamental trajectory. The regime has only hardened. Its nuclear program has only advanced. The stated goal is a stable, compliant Iran. The actual outcome is a more determined, more hostile Iran.

Gaza (2005-present). Repeated military campaigns, blockades, and escalations produce cycles of violence with no endpoint. The stated goal is security. The actual outcome is radicalization, destruction, and perpetual conflict.

The pattern is undeniable: **force, applied to complex systems, produces the opposite of its intended outcome.**

Why This Keeps Happening

The attractor framework provides a formal explanation.

Corrective permeability (κ) is a measure of how open a system is to corrective information. A high-κ system can incorporate feedback, adjust its behavior, and shift its attractor. A low-κ system is sealed. It cannot learn. It cannot change. It persists in its current state, regardless of the consequences.

The West's κ is approaching zero. It is a sealed system.

Why?

Because the West interprets all information through the filter of its core belief: *force is the answer.* Every failure is reframed as evidence of insufficient force. Every defeat is seen as a reason to escalate. Every catastrophe is understood as a demonstration of the enemy's evil, not the intervention's folly. The system is epistemically closed. It cannot see what it is doing, because seeing it would require abandoning the belief that defines it.

This is the formal definition of a fantasy attractor: **a sealed basin that persists because it cannot recognize that it is sealed.**

The Entropy Cost of Persistence

Every dissipative system pays a cost for its persistence. It generates entropy—disorder, waste, blowback—in the process of maintaining its structure. The West is no exception.

The West's persistence is maintained at an enormous cost:

- Trillions of dollars diverted from productive investment to military expenditure.
- Hundreds of thousands of lives lost in wars of choice.
- Millions displaced by conflicts the West initiated or exacerbated.
- Global instability created by interventions that destabilize rather than stabilize.
- Moral authority eroded by actions that undermine the very values the West claims to uphold.
- Ecological destruction accelerated by the industrial-military complex.

This entropy is not noise. It is the cost of maintaining a fantasy attractor. The West persists in its delusion, but the price is visible everywhere: in the rubble of cities, in the refugee camps, in the radicalized populations, in the distrust of the global majority, in the exhaustion of the system itself.

The Attractor of Force

The West is not choosing to fail. It is locked into a basin that makes failure the only possible outcome.

A basin is a stable state that the system naturally settles into. Once you are in a basin, you are pulled back to it whenever you try to leave. The West's basin is organized around force:

- Institutions built for force projection.
- Culture that rewards decisive action and punishes patience.

- Media that demands visible results and cannot see invisible cultivation.
- Electoral cycles that incentivize short-term fixes and punish long-term thinking.
- Ideology that frames the world as a battle between good and evil.

Each element reinforces the others. The basin is deep. It is self-sustaining. And it is sealed.

This is why the West cannot learn. Learning would require stepping outside the basin. But the basin is all the West knows. It has no reference point for a different mode of being. It cannot conceive of a non-force intervention, because force is the only language it speaks.

The Alternative: Cultivation

There is an alternative.

It is not new. It is not complicated. It is not even hidden. It is the ancient wisdom of cultivation:

- Observe before you intervene.
- Understand the system before you try to shift it.
- Apply precision and restraint, not force and escalation.
- Be patient. The system will shift on its own timeline.
- Accept that you cannot force a living system to comply with your will.

This is the Taoist principle of *wu wei*: action that is so aligned with the natural flow of things that it appears effortless. It is not passivity. It is not surrender. It is the recognition that force, applied to complex systems, generates more chaos than order—and that the only way to produce lasting change is to cultivate conditions that allow the system to shift on its own.

The West cannot implement this approach because its basin prevents it. But individuals can.

My sleep experiment is an example. I did not force deep sleep to appear. I observed. I adjusted. I added saffron and ashwagandha. I went outside in the morning. I reduced alcohol. I let the system shift on its own timeline. And it did. REM increased. Continuity improved. Deep sleep began to stir.

I did not force the change. I cultivated it.

The Three-Body Problem

This is the deepest lesson: you cannot force a system into a state that does not exist in its phase space.

In astrophysics, the three-body problem has no general stable solution. The system either collapses, ejects one of the bodies, or oscillates chaotically. You cannot force a three-body system into a stable orbit because that state does not exist.

Geopolitics is a many-body problem. It has no stable low-energy attractor. You cannot force Iran, Israel, Russia, China, or Afghanistan into compliance because the stable state you are aiming for does not exist. You are trying to force a square peg into a round hole—and then escalating when it does not fit.

The West's demand for stability is a category error. It is trying to impose a state of affairs that is not part of the system's phase space. The result is not stability—it is chaos, blowback, and collapse.

The Fantasy Attractor

The West’s belief in force is a fantasy attractor. It is a sealed basin that persists despite—or *because of*—its detachment from reality. The system cannot correct itself because correction would require abandoning the belief that defines it.

This is why the West is stupid. Not because it lacks intelligence, but because it is structurally incapable of learning. It is trapped in a basin that prevents it from seeing what it is doing. It keeps doing the same thing and expecting a different result—and it cannot see that the result cannot be different because the system has no attractor for the outcome it seeks.

There is no end in sight. The West will continue to escalate, continue to fail, continue to generate entropy, and continue to interpret its failures as evidence of the need for more force. It will collapse or eject, just like a three-body system. There is no other outcome.

For the Individual

The civilization cannot learn. But you can.

You can see the pattern. You can recognize that force is not the answer. You can step outside the basin—if only for a moment. You can cultivate patience, observation, and precision. You can apply the attractor framework to your own life, your own habits, your own beliefs. You can ask: *“Am I* *locked in a fantasy attractor? Am I sealed against corrective information? What would it take to become permeable?”*

This is not a political program. It is a personal practice. It is the work of a lifetime. But it is the only way out.

The Invitation

Fantasy Attractor is a research program. It invites challenge, correction, and collaboration. It does not claim to have all the answers. It offers a framework—a common language for comparing systems that appear unrelated. It asks: *What persists? What changes? What is the cost of persistence? What is the cost of change?*

If you see a flaw, a gap, or a better way, contact us. The framework is living. It is open. It is permeable.

That is the opposite of a fantasy attractor. That is the beginning of learning.

Robert Galida is an independent researcher and the founder of the Fantasy Attractor Research Program. His work develops a formal framework for understanding persistence and change across physical, biological, cognitive, and social systems.

Deriving Corrective

Permeability from the Cumulative Deviation Functional; Robert Galida (June 2026) [F]

Abstract

The attractor framework defines κ (corrective permeability) as the rate at which a system returns to its attractor after perturbation. Historically, κ has been treated as an empirical parameter — fitted to data rather than derived from first principles. This paper derives κ from the framework's foundational object: the cumulative deviation functional $D_T(x)=\int_0^T \delta(\phi_t(x))\,dt$, where $\delta(x)=d(x,\mathcal{A})$.

We define: $\kappa=\inf_{x\in B\setminus \mathcal{A}} \frac{\delta(x)}{D_\infty(x)}$

We prove that for linear systems $\dot{x}=-Ax$ with A symmetric positive definite, this definition recovers the slowest eigenvalue $\lambda_{\min}(A)$ — the conventional notion of corrective permeability. We establish a sharp universal persistence bound $D_\infty(x)\le \delta(x)/\kappa$, show homogeneity and scale invariance of the variational ratio, and demonstrate consistency with Koopman spectral theory and resolvent poles for finite-dimensional linear systems. A comparison theorem links κ to classical exponential stability constants. A Hamilton-Jacobi-type transport equation for D_∞ is derived. A finite-horizon estimator $\kappa_T=\inf_x \frac{\delta(x)}{D_T(x)}$ is provided with exponential convergence under explicit assumptions.

The derivation is rigorous for linear systems and testable.

Open questions for nonlinear, multiscale, and stochastic systems are identified.

Keywords: corrective permeability, cumulative deviation functional, attractor framework, Koopman operator, trajectory functional

---

1. Introduction

The attractor framework has been applied across physics, biology, cognition, and social systems. Its central variable — corrective permeability κ — measures the rate at which a system returns to its attractor after perturbation. Historically, κ has been defined empirically as $\kappa=1/\tau$, where τ is a measured recovery time constant.

This paper derives κ from a single foundational object: the cumulative deviation functional $D_T(x)$. Within the present framework, κ is defined variationally rather than introduced as an empirical fitting parameter. We show that κ is a consequence of the trajectory geometry — specifically, the ratio of initial distance to total cumulative deviation.

The derivation is rigorous for linear systems, connects to established theory (Koopman operators, resolvent poles), and provides a finite-horizon estimator for empirical use. Open questions for nonlinear and stochastic systems are identified.

---

2. The Cumulative Deviation

Functional

Let X be a metric space with distance function $\|\cdot\|$. Let $\phi_t(x)$ be the flow of a dynamical system starting from state $x \in X$ at time $t=0$. Let $A \subseteq X$ be an attractor set (a compact, invariant set to which trajectories converge). Let B be the basin of attraction of A.

Define the distance from a point to the attractor:$$\delta(x) = d(x, A) = \inf_{a \in A} \|x - a\|$$

Definition 1 (Cumulative Deviation Functional): For a finite horizon $T > 0$, define:$$D_T(x) = \int_0^T \delta(\phi_t(x))\, dt$$

For $T \to \infty$, define:$$D_\infty(x) = \int_0^\infty \delta(\phi_t(x))\, dt$$

Proposition 1 (Finiteness of D_∞): Assume there exist constants $C < \infty$ and $\mu > 0$ such that:$$\delta(\phi_t(x)) \le C e^{-\mu t} \delta(x)$$

for all $x \in B$. Then $D_\infty(x) < \infty$ for every $x \in B$.

Proof:$$D_\infty(x) = \int_0^\infty \delta(\phi_t(x))\, dt \le \int_0^\infty C e^{-\mu t} \delta(x)\, dt = \frac{C}{\mu} \delta(x) < \infty \quad \square$$

Properties (from Galida, 2026a):

Property	Statement
Non-negativity	$D_T(x) \ge 0$
Monotonicity	$D_{T_2}(x) \ge D_{T_1}(x)$ for $T_2 \ge T_1$
Additivity	$D_{T+S}(x) = D_T(x) + D_S(\phi_T(x))$
Instantaneous growth	$\frac{d}{dT} D_T(x) = \delta(\phi_T(x))$
Occupation measure	$D_T(x) = \int \delta(y)\, d\mu_T(y)$, where μ_T is the occupation measure

3. Derivation of Corrective Permeability (κ)

3.1 Variational Definition

Definition 2 (Corrective Permeability): $\kappa = \inf_{x \in B \setminus A} \frac{\delta(x)}{D_\infty(x)}$

Interpretation: κ is the *effective* recovery rate — the smallest ratio of initial distance to total cumulative deviation. It serves as a global measure of the slowest recovery mode in the basin.

Remark on κ: The definition allows $\kappa = 0$ if $D_\infty(x)$ diverges or if the ratio $\delta(x)/D_\infty(x)$ can be made arbitrarily small. Throughout the remainder of this paper, we assume hypotheses (such as the exponential stability in Proposition 1) that guarantee $\kappa > 0$.

Remark on attainment: The infimum in the definition of κ need not be attained; minimizing sequences may exist without a minimizing state. For linear systems, the infimum is attained on the slow eigenspace.

3.2 Homogeneity and Scale Invariance

Theorem 1 (Homogeneity and Scale Invariance): Suppose the flow satisfies $\phi_t(\alpha x) = \alpha \phi_t(x)$ for all t and all $\alpha > 0$, and the distance function satisfies $\delta(\alpha x) = \alpha \delta(x)$. Then: $\frac{\delta(\alpha x)}{D_\infty(\alpha x)} = \frac{\delta(x)}{D_\infty(x)}$

Proof: $D_\infty(\alpha x) = \int_0^\infty \delta(\phi_t(\alpha x))\, dt = \int_0^\infty \delta(\alpha \phi_t(x))\, dt = \alpha \int_0^\infty \delta(\phi_t(x))\, dt = \alpha$

$$D_\infty(\alpha x)=\int_0^\infty \delta(\phi_t(\alpha x))\,dt=\int_0^\infty \delta(\alpha\phi_t(x))\,dt=\alpha\int_0^\infty \delta(\phi_t(x))\,dt=\alpha D_\infty(x)$$

Corollary: For linear systems, the infimum over all $x \neq 0$ reduces to an infimum over the unit sphere:
$$\kappa=\inf_{\|x\|=1}\frac{\delta(x)}{D_\infty(x)}$$

3.3 Sharp Universal Persistence Bound

Theorem 2 (Sharp Universal Persistence Bound): For any $x \in \mathcal{B}\setminus\mathcal{A}$:
$$D_\infty(x)\le\frac{\delta(x)}{\kappa}$$

Moreover, the constant $1/\kappa$ is optimal: it is the *smallest* constant such that this inequality holds for all x in the basin.

Proof: By definition of κ as the infimum of $\delta(x)/D_\infty(x)$, we have $\delta(x)/D_\infty(x)\ge\kappa$ for all x. Rearranging gives:
$$D_\infty(x)\le\frac{\delta(x)}{\kappa}$$

Optimality follows from Theorem 3: for the slow eigenvector v_1, $D_\infty(v_1)=\delta(v_1)/\kappa$, so no smaller constant can work. $\square$

3.4 Consistency with Linear Systems

Consider a linear system $\dot{x}=-Ax$, with A symmetric positive definite. Let its eigenvalues be $0<\lambda_1\le\lambda_2\le\cdots\le\lambda_n$, with corresponding orthonormal eigenvectors $v_1, v_2, \ldots, v_n$.

The flow is $\phi_t(x)=e^{-At}x$. The attractor is $\mathcal{A}=\{0\}$, and the distance to the attractor is $\delta(x)=\|x\|$.

Theorem 3 (Linear Consistency): For $\dot{x} = -Ax$ with A symmetric positive definite,
$$\inf_{x \neq 0} \frac{\|x\|}{D_\infty(x)} = \lambda_{\min}(A)$$

Proof:

Since A is symmetric positive definite, e^{-At} is symmetric positive definite with eigenvalues $e^{-\lambda_i t}$. Hence its operator norm is $\|e^{-At}\| = e^{-\lambda_1 t}$. For any $x \neq 0$:
$$D_\infty(x) = \int_0^\infty \|e^{-At}x\|\,dt \leq \int_0^\infty \|x\| e^{-\lambda_1 t}\,dt = \frac{\|x\|}{\lambda_1}$$

Therefore:
$$\frac{\|x\|}{D_\infty(x)} \geq \lambda_1$$

To show equality is achieved, take $x = v_1$ (the eigenvector corresponding to λ_1). Then:
$$\|e^{-At}v_1\| = \|v_1\| e^{-\lambda_1 t}$$

and:
$$D_\infty(v_1) = \int_0^\infty \|v_1\| e^{-\lambda_1 t}\,dt = \frac{\|v_1\|}{\lambda_1}$$

Thus:
$$\frac{\|v_1\|}{D_\infty(v_1)} = \lambda_1$$

Hence:
$$\inf_{x \neq 0} \frac{\|x\|}{D_\infty(x)} = \lambda_1 \quad \square$$

Corollary: For linear systems, the variational definition of κ recovers the slowest eigenvalue — the conventional notion of corrective permeability.

3.5 Transport Equation

Theorem 4 (Transport Equation): Assume the vector field f is C^1, the flow ϕ_t is C^1, and D_∞ is continuously differentiable on $B(A)$. Then:
$$\nabla D_\infty(x) \cdot f(x) = -\delta(x)$$

Proof: From the definition:
$$D_\infty(\phi_s(x)) = D_\infty(x) - D_s(x)$$

Differentiating with respect to s at $s=0$: $\frac{d}{ds} D_\infty(\phi_s(x))\Big|_{s=0} = -\delta(x)$

By the chain rule: $\nabla D_\infty(x) \cdot f(x) = -\delta(x)$ $\square$

Interpretation: This is a first-order transport equation, $f \cdot \nabla D = -\delta$, which belongs to the broader Hamilton-Jacobi family but lacks a Hamiltonian in the usual sense. It may serve as a foundation for numerical computation and further theoretical development.

3.6 Local vs. Global Interpretation

The variational definition $\kappa = \inf_x \frac{\delta(x)}{D_\infty(x)}$ is **global** — it is the slowest recovery rate over the entire basin. This is not necessarily the same as the local recovery rate near the attractor (the slowest eigenvalue of the linearization). For linear systems, they coincide. For nonlinear systems, they may differ if transient excursions produce slower effective recovery than the local linearization predicts.

This distinction is important: κ is a global invariant of the basin, not merely a local property of the attractor. The relationship between the global κ and the local Lyapunov exponent is an open question (see §6).

3.7 Non-Symmetric Linear Systems

For a general linear system $\dot{x} = Ax$ (where A is stable, i.e., all eigenvalues have negative real parts), the same principle holds in the diagonalizable case. The slowest mode corresponds to the eigenvalue with the largest real part (closest to zero).

Conjecture: An analogous result holds for non-normal linear systems under additional assumptions on the semigroup, such as a uniformly exponentially stable semigroup satisfying suitable norm bounds. This remains an open question.

3.8 Comparison with Exponential Stability

Theorem 5 (Comparison with Exponential Stability): Suppose the system satisfies the exponential stability bound:$$\delta(\phi_t(x)) \le C e^{-\mu t}\delta(x)$$

for all $x \in B$, with constants $C < \infty$ and $\mu > 0$. Then:$$\kappa \ge \frac{\mu}{C}$$

Proof: From the stability bound:$$D_\infty(x) = \int_0^\infty \delta(\phi_t(x))\,dt \le \int_0^\infty C e^{-\mu t}\delta(x)\,dt = \frac{C}{\mu}\delta(x)$$

Therefore:$$\frac{D_\infty(x)}{\delta(x)} \ge \frac{\mu}{C}$$

Taking the infimum over x:$$\kappa = \inf_x \frac{D_\infty(x)}{\delta(x)} \ge \frac{\mu}{C} \quad \square$$

Interpretation: The variational constant κ is bounded below by the exponential stability constant μ/C.

---

4. Connections to Existing Theory

4.1 Koopman Operator

The Koopman operator K_t acts on observables as:$$(K_t f)(x) = f(\phi_t(x))$$

For linear systems $\dot{x} = -Ax$, the Koopman eigenvalues are $e^{-\lambda_i t}$. The dominant nontrivial eigenvalue (largest less than 1) is $e^{-\lambda_1 t}$, corresponding to the slowest decay rate.

For finite-dimensional linear systems, $\rho = e^{-\lambda_{\min} t}$, and therefore:

$$-\frac{1}{t}\log \rho = \lambda_{\min} = \kappa$$

Thus, under the hypotheses of Theorem 3, the variational constant equals the exponential decay rate associated with the dominant Koopman eigenvalue.

4.2 Resolvent Poles

For finite-dimensional stable linear systems, the resolvent $(sI + A)^{-1}$ has poles at $s = -\lambda_i$. The pole closest to the imaginary axis is $s = -\lambda_1$.

Since Theorem 3 identifies $\kappa = \lambda_{\min}$, and the resolvent poles are $s_i = -\lambda_i$, we obtain:

$$\kappa = \min_i (s_i)$$

for finite-dimensional linear systems.

---

5. Finite-Horizon Estimation

In practice, we can only measure finite trajectories. Define the finite-horizon estimator:

$$\kappa_T = \inf_{x \in K} \frac{\delta(x)}{D_T(x)}$$

where $K \subset B$ is compact and $K \cap A = \emptyset$.

Proposition 2 (Finite-Horizon Estimation): Assume:

1. The flow $\phi_t(x)$ is jointly continuous in (t,x).
2. $\delta(x)$ is continuous.
3. The exponential stability bound $\delta(\phi_t(x)) \le Ce^{-\mu t}\delta(x)$ holds uniformly for all $x \in K$, with $\mu > 0$.

Then the variational constant κ (from Definition 2) satisfies $\kappa \ge \mu/C$ by Theorem 5, and:
$$\kappa_T \to \kappa \quad \text{as } T \to \infty$$

with error:
$$|\kappa_T - \kappa| = O(e^{-\mu T})$$

Proof: For any $x \in K$, the tail bound gives:
$$|D_\infty(x) - D_T(x)| = \int_T^\infty \delta(\phi_t(x))\,dt \le \frac{Ce^{-\mu T}\delta(x)}{\mu}$$

Since $\delta(x)$ is bounded on the compact set K, let $M = \sup_{x \in K} \delta(x) < \infty$. Then:
$$|D_\infty(x) - D_T(x)| \le \frac{CMe^{-\mu T}}{\mu}$$

The right-hand side is independent of x and tends to zero as $T \to \infty$. Hence $D_T \to D_\infty$ uniformly on K.

Moreover, since K is compact and $K \cap A = \varnothing$, continuity of δ gives $\inf_{x \in K} \delta(x) > 0$. Since $D_T(x)$ is continuous (by assumptions 1–2) and monotonically non-decreasing in T (from §2), for any fixed finite $T_0 > 0$, $D_\infty(x) \ge D_{T_0}(x)$, and D_{T_0} is continuous and strictly positive on K. A continuous, strictly positive function on a compact set has a positive infimum:
$$m = \inf_{x \in K} D_{T_0}(x) > 0$$

Thus:
$$\inf_{x \in K} D_\infty(x) \ge m > 0$$

Uniform convergence of D_T to D_∞ on K therefore implies uniform convergence of $\delta(x)/D_T(x)$ to $\delta(x)/D_\infty(x)$. Consequently, the infima converge. $\square$

6. Open Questions

Question	Status	Difficulty
Q1: Nonlinear systems	Does $\inf \frac{\delta}{D_\infty}$ equal the local Lyapunov exponent?	Hard
Q2: Local vs. global consistency	Does $\lim_{x\to A} \frac{\delta(x)}{D_\infty(x)} = \kappa$ hold for general nonlinear systems?	Hard
Q3: Non-normal systems	Does the infimum equal the slowest eigenvalue for non-normal A?	Moderate
Q4: Multiple timescales	Does the infimum isolate the slowest timescale?	Hard
Q5: Stochastic systems	How does noise affect the finite-horizon estimator?	Hard
Q6: Multiple attractors	How does κ behave in basins with multiple attractors?	Moderate

7. Conclusion

This paper derives corrective permeability κ from the cumulative deviation functional $D_T(x)$. The variational definition: $\kappa = \inf_x \frac{\delta(x)}{D_\infty(x)}$

is shown to recover the slowest eigenvalue for linear systems, consistent with the conventional empirical definition $\kappa = 1/\tau$. A sharp universal persistence bound $D_\infty(x) \le \delta(x)/\kappa$ is established. A comparison theorem links κ to classical exponential stability constants. A Hamilton-Jacobi-type transport equation for D_∞ is derived. Connections to Koopman theory and resolvent theory are established for finite-dimensional linear systems. A finite-horizon estimator κ_T is provided with exponential convergence under explicit assumptions.

Key contribution: Within the present framework, κ is defined variationally rather than introduced as an empirical fitting parameter — at least for the class of systems analyzed here.

Next steps: Extend the derivation to nonlinear systems (Q1–Q2), non-normal systems (Q3), multiple timescales (Q4), and stochastic dynamics (Q5).

References

Crandall, M. G., Ishii, H., & Lions, P. L. (1992). “User’s Guide to Viscosity Solutions of Second Order Partial Differential Equations.” *Bulletin of the American Mathematical Society*, 27(1), 1-67.

Evans, L. C. (2010). *Partial Differential Equations*. American Mathematical Society.

Galida, R. (2026a). “The Persistence Functional: A Candidate Formal Foundation for the Attractor Framework.” *Fantasy Attractor*.

Hale, J. K. (1988). *Asymptotic Behavior of Dissipative Systems*. American Mathematical Society.

Hirsch, M. W., Smale, S., & Devaney, R. L. (2004). *Differential Equations, Dynamical Systems, and an Introduction to Chaos* (2nd ed.). Elsevier Academic Press.

Khalil, H. K. (2002). *Nonlinear Systems* (3rd ed.). Prentice Hall.

Koopman, B. O. (1931). “Hamiltonian Systems and Transformations in Hilbert Space.” *Proceedings of the National Academy of Sciences*, 17(5), 315-318.

Lyapunov, A. M. (1892). *The General Problem of the Stability of Motion*. (English translation: 1992, Taylor & Francis).

Mezić, I. (2005). “Spectral Properties of Dynamical Systems, Model Reduction and Decompositions.” *Nonlinear Dynamics*, 41(1-3), 309-325.

Pazy, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer.

Vidyasagar, M. (1993). *Nonlinear Systems Analysis* (2nd ed.). Prentice Hall.

Suggested citation: Galida, R. S. (2026). Deriving Corrective Permeability from the Cumulative Deviation Functional. *Fantasy Attractor*.

The Persistence Functional: A Candidate Formal Foundation for the Attractor Framework; Robert Galida (July 2026) [F]

Abstract

The attractor framework provides a domain-general vocabulary for describing persistence and change across physical, biological, cognitive, and social systems. However, its core variables—κ (corrective permeability), B (basin depth), and R (reality alignment)—have been defined inconsistently across application papers, and their formal relationships have remained implicit. This paper proposes a candidate mathematical formalization for the framework.

The central mathematical innovation of this paper is treating persistence as a functional defined over trajectories—$D_T(x)=\int_0^T d(\phi_\tau(x),A)\,d\tau$—rather than as a scalar property of states. We prove several mathematical properties of D_T, including non-negativity, monotonicity in T, additivity, Lipschitz continuity with respect to initial conditions, and a bound relating D_∞ to the recovery rate κ: $D_\infty(x)\leq\frac{C}{\kappa}d(x,A)$. We establish connections to dynamic programming and ergodic theory via occupation measures. We introduce a complementary **topological persistence functional** $P_{\text{topo}}(t)$, which measures the lifetime of topological features in the trajectory's state-space geometry, and the **topological evolution rate** $E(t)$.

We unify the framework's variable set: κ is the recovery rate (operationalized as $1/\tau$); γ is a proposed drift rate for persistent chaos, grounded in the literature on high-dimensional neural networks; B is the energy barrier (basin depth); $\tilde{B}$ is a complementary persistence depth; R is the expected log predictive likelihood. We propose testable predictions linking $E(t)$ to κ and γ, and provide a falsifiable experimental protocol using neural network training and persistent homology.

The paper offers a candidate formal foundation, with explicit

definitions, mathematical properties, and empirical grounding. All unverified sources are clearly labeled as such.

Keywords: attractor framework, persistence functional, cumulative deviation, topological persistence, corrective permeability, basin depth, reality alignment, persistent homology

---

1. Introduction

The attractor framework has been applied across physics (hydrogen decay, Jeans instability), biology (ECM mechanics, HRV), cognition (belief updating, performance attractors), and social systems (religious attractors, civilizational dynamics). A common vocabulary has emerged: κ (corrective permeability), B (basin depth), and R (reality alignment). However, these variables have been defined inconsistently across papers, and their formal relationships have remained implicit. This paper proposes a candidate mathematical formalization that addresses these inconsistencies.

The central mathematical innovation of this paper is treating persistence as a functional defined over trajectories rather than as a scalar property of states. $D_T(x)=\int_0^T d(\phi_\tau(x),A)\,d\tau$ can be understood as a type of **action functional** (carefully qualified). Like the classical action $\int L(q,\dot{q})\,dt$, it assigns a scalar to an entire trajectory, is additive under concatenation, and suggests variational and optimal-control interpretations. However, it is not the mechanical action; it is a cumulative deviation functional that measures time away from equilibrium. This moves the framework into the domain of trajectory-level analysis, aligning it with modern dynamical systems and geometric control theory.

We introduce the **cumulative deviation functional** $D_T(x)$ as this central object, and we establish its mathematical properties, including its relationship to the recovery rate κ. We introduce a complementary **topological persistence functional** $P_{\text{topo}}(t)$ and the **topological evolution rate** $E(t)$. We unify the framework's variable set with operational definitions and propose testable predictions with falsification criteria.

1.1 Scope and Status

This paper is a **candidate formalization**—it provides definitions, mathematical properties, and empirical hypotheses. It is not a completed empirical validation; that is the subject of future work. All claims are labeled as **definitions** (part of the formal structure), **propositions/theorems** (proved), **hypotheses** (testable predictions), or **heuristics** (suggestive connections not yet formalized). This distinction is maintained throughout.

---

2. Formal Definitions

Let X be a metric space with distance function $\|\cdot - \cdot\|$. Let $\phi_\tau(x)$ be the flow of a dynamical system starting from state $x \in X$ at time $\tau = 0$. Let $A \subseteq X$ be an attractor set (a compact, invariant set to which trajectories converge). Assume the flow is continuous and measurable so that $d(\phi_\tau(x), A)$ is measurable. The flow ϕ_τ satisfies the semigroup property $\phi_{t+s} = \phi_t \circ \phi_s$ for all $t, s \geq 0$, with $\phi_0 = \text{id}$. We assume $d(\phi_\tau(x), A) \in L^1([0, T])$ for all finite T, so the integral defining D_T is well-defined.

Define the distance from a point to the attractor: $d(x, A) = \inf_{a \in A} \|x - a\|$

The definition applies to any metric space; for infinite-dimensional spaces, the usual measurability and integrability conditions are assumed.

2.1 Cumulative Deviation Functional

Definition 1 (Cumulative Deviation Functional): For a finite horizon $T>0$, the cumulative deviation functional is:$$D_T(x)=\int_0^T d(\phi_\tau(x),A)\,d\tau$$

Interpretation: $D_T(x)$ is the total accumulated deviation from the attractor over the interval $[0,T]$. It measures integrated error, residence-time-weighted distance, or accumulated regret. This is **not** a path length; it measures time spent away from equilibrium, whereas path length $\int\|\dot{\phi}_\tau(x)\|\,d\tau$ measures distance traveled.

Domain generality: This definition applies to any system with a well-defined state space, a flow, and an attractor set. It does not require linearity, differentiability, or specific functional forms.

Empirical note: D_T is the fundamental object for empirical work; D_∞ is primarily an analytical limit used for theoretical bounds.

Note: D_T is not a Lyapunov function. A Lyapunov function is a scalar function of the current state; D_T is a functional of the entire trajectory. It does not decrease monotonically along trajectories, and it does not provide pointwise stability information. Its purpose is to measure accumulated history, not instantaneous energy.

Occupation measure connection: Define the occupation measure of the trajectory up to time T as:$$\mu_T(B)=\int_0^T \mathbf{1}_B(\phi_\tau(x))\,d\tau$$

for measurable $B\subseteq X$. Then:$$D_T(x)=\int_X d(y,A)\,d\mu_T(y)$$

Thus D_T is the expected distance to the attractor under the occupation measure. This connects the functional directly to ergodic theory and occupation measure analysis. For foundational treatments of occupation measures and invariant measures, see Ruelle (1989) and Bowen (1975).

2.1.1 Why the L¹ Trajectory Functional?

The choice of the L¹ integral over alternatives is motivated by the following properties:

- **Linearity:** Each moment contributes equally; accumulation is additive over time.
- **Physical units:** For systems with a natural distance metric, D_T has units of distance × time, which is interpretable as accumulated deviation.
- **Simplicity:** It is the simplest nontrivial trajectory functional that is not a path length.
- **Analogy:** It mirrors cumulative regret and occupation measures in control theory and ergodic theory.
- **Avoidance of overweighting:** Unlike d^2, it does not disproportionately weight large deviations; unlike max, it is sensitive to the full trajectory.

This is one natural choice; other functionals (e.g., d^p, exponentially weighted integrals) could be substituted without changing the framework’s structure.

2.2 Topological Persistence Functional

Let $X_\tau = \{\phi_s(x) : s \in [0,\tau]\}$ be the trajectory segment up to time τ. Let $PH_k(X_\tau)$ be the k-dimensional persistent homology of the point cloud X_τ at scale ϵ. Each feature (component, loop, void) has a birth scale b and a death scale d, with persistence $d-b$. For foundational treatments of persistent homology, see Edelsbrunner & Harer (2010) or Carlsson (2009).

Definition 2 (Topological Persistence Functional): We define the following complementary topological persistence functional.

For $t \geq 0$: $P_{\text{topo}}(t) = \int_0^t \sum_{k \geq 0} \sum_{(b,d) \in \mathrm{PH}_k(X_\tau)} (d-b)\, d\tau$

The map $\tau \mapsto \mathrm{PH}_k(X_\tau)$ is piecewise constant on intervals where the trajectory does not cross a homology-critical threshold. Assuming the trajectory crosses such thresholds at discrete times, the integral is well-defined as a sum of piecewise continuous segments. This is the standard assumption in time-varying persistent homology (see Carlsson & Zomorodian, 2009).

Interpretation: $P_{\text{topo}}(t)$ is the total lifetime of all topological features in the trajectory's state-space geometry up to time t. This is a separate mathematical object from DT; the relationship between them is an empirical hypothesis. This is one possible choice among several topological summaries (e.g., persistence landscapes, persistence images) and is selected because it mirrors the cumulative interpretation of DT, rather than because it is uniquely canonical. Other stable summaries—such as persistence landscapes, persistence images, or Betti curves—could be substituted for the present functional without changing the framework's structure.

Measurement: In practice, $P_{\text{topo}}(t)$ is computed by sampling the trajectory at discrete times, computing persistent homology on latent activation manifolds, and summing the persistence of all features using standard libraries (e.g., GUDHI, Ripser). Turner & Barak (2023) demonstrated that trained RNNs develop attractors sequentially during training; the topological structure of these attractors can be analyzed using persistent homology.

Falsification: If persistent homology features do not correlate with any behavioral or dynamical measure in a given system, P_{topo} is not a useful construct for that domain.

2.3 Topological Evolution Rate

Definition 3 (Topological Evolution Rate): For a learning system with time-dependent topological persistence, the topological evolution rate is defined as:$E(t) = \frac{d}{dt} P_{\text{topo}}(t)$

where differentiable, and experimentally as $E(t) \approx \frac{\Delta P_{\text{topo}}}{\Delta t}$ over finite intervals.

Interpretation: $E(t)$ measures how quickly the system's topological complexity changes during learning. Negative $E(t)$ indicates topological simplification (compression); positive $E(t)$ indicates increasing complexity (expansion); $E(t) \approx 0$ indicates stagnation. Learning is one possible cause of topological change; random drift, noise, or chaotic wandering can also change topology.

Empirical anchor: Karuppiah, Nazreen Banu et al. (2026) examine the evolution of topological signatures during training. Turner & Barak (2023) show that RNNs develop attractors sequentially, which may correspond to phases of topological simplification. We hypothesize that successful learning corresponds to negative average values of $E(t)$ over defined phases, but this is a testable claim, not a definition.

---

3. Mathematical Properties of the Cumulative Deviation Functional

This section establishes the mathematical behavior of D_T, providing the foundation for its use in the framework.

3.1 Non-negativity

Proposition 1 (Non-negativity): For any $x \in X$ and any $T \geq 0$: $D_T(x) \geq 0$

with equality iff $\phi_\tau(x) \in A$ for almost all $\tau \in [0, T]$.

Proof: The integrand is a distance function $d(\phi_\tau(x), A)$, which is non-negative by definition. The integral of a non-negative function is non-negative. Equality holds only if the integrand is zero almost everywhere.

3.2 Monotonicity in T

Proposition 2 (Monotonicity): For fixed x, $D_T(x)$ is monotonically non-decreasing in T: $D_{T_2}(x) \geq D_{T_1}(x) \text{ for } T_2 \geq T_1$

Proof: For $T_2 \geq T_1$: $D_{T_2}(x) = \int_0^{T_1} d(\phi_\tau(x), A)\, d\tau + \int_{T_1}^{T_2} d(\phi_\tau(x), A)\, d\tau$

The second integral is non-negative by Proposition 1. Therefore $D_{T_2}(x) \geq D_{T_1}(x)$.

Corollary: If the trajectory converges exactly to the attractor at time $\tau_0 < T$, then: $D_T(x) = D_{\tau_0}(x) \text{ for all } T \geq \tau_0$

3.3 Additivity

Proposition 3 (Additivity): For any $T, S \geq 0$: $D_{T+S}(x) = D_T(x) + D_S(\phi_T(x))$

Proof: $D_{T+S}(x) = \int_0^{T+S} d(\phi_\tau(x), A)\, d\tau = \int_0^{T} d(\phi_\tau(x), A)\, d\tau + \int_T^{T+S} d(\phi_\tau(x), A)\, d\tau = D_T(x) + \int_0^{S} d(\phi_{\tau+T}(x), A)\, d\tau = D_T(x) + \int_0^{S} d(\phi_\tau(\phi_T(x)), A)\, d\tau$ (by the semigroup property) $= D_T(x) + D_S(\phi_T(x))$ ∎

This connects D_T naturally to Bellman equations, dynamic programming, and occupation measures.

3.4 Heuristic Connection: Dynamic Programming

The additivity property $D_{T+S}(x) = D_T(x) + D_S(\phi_T(x))$ suggests a natural connection to dynamic programming. For a controlled system $\dot{X} = f(X, u)$ with control $u \in U$, the value function $V(x) = \inf_u D_\infty(x)$ would formally satisfy the Hamilton-Jacobi-Bellman equation: $0 = \inf_u \{ d(x, A) + \nabla V(x) \cdot f(x, u) \}$

This is a standard result for additive cost functionals. A full derivation for the specific functional D_T is left for future work. This section is a heuristic connection, not a formal result.

3.5 Lipschitz Continuity with Respect to Initial Conditions

Proposition 4 (Lipschitz Continuity of D_T): Suppose the flow ϕ_τ is Lipschitz continuous in x with constant L, i.e., $\|\phi_\tau(x)-\phi_\tau(y)\| \le e^{L\tau}\|x-y\|$. Then for any x,y in the basin of A: $|D_T(x)-D_T(y)| \le \int_0^T e^{L\tau}\,d\tau\,\|x-y\| = \frac{e^{LT}-1}{L}\|x-y\|$

Proof: First, note that the distance function $d(\cdot, A)$ is 1-Lipschitz: for any $x,y \in X$, $|d(x,A)-d(y,A)| \le \|x-y\|$

This follows from the triangle inequality and the definition of the infimum. Then, using the Lipschitz property of the flow: $|D_T(x)-D_T(y)| \le \int_0^T |d(\phi_\tau(x),A)-d(\phi_\tau(y),A)|\,d\tau \le \int_0^T \|\phi_\tau(x)-\phi_\tau(y)\|\,d\tau \le \int_0^T e^{L\tau}\|x-y\|\,d\tau = \frac{e^{LT}-1}{L}\|x-y\|$

Interpretation: This proposition guarantees that empirical estimates of D_T are robust under small perturbations of initial conditions and establishes that D_T defines a continuous functional on the basin of attraction. This is essential for numerical estimation and experimental measurement.

3.6 Instantaneous Growth Rate

Remark 1 (Instantaneous Growth Rate): If the integrand $d(\phi_\tau(x), A)$ is continuous in τ, then: $\frac{d}{dT}D_T(x) = d(\phi_T(x), A)$

This follows directly from the Fundamental Theorem of Calculus.

3.7 Ergodic Limit

Proposition 5 (Ergodic Limit): Suppose the normalized occupation measure $\nu_T=\mu_T/T$ converges weakly to an invariant probability measure μ as $T\to\infty$. Then: $\lim_{T\to\infty}\frac{1}{T}D_T(x)=\int_X d(y,A)\,d\mu(y)$

Proof: From the occupation measure representation $D_T(x)=\int d(y,A)\,d\mu_T(y)=T\int d(y,A)\,d\nu_T(y)$, weak convergence of ν_T to μ and boundedness/continuity of $d(\cdot,A)$ gives the result.

This is the pointwise ergodic theorem applied to the observable $d(\cdot,A)$. For the ergodic theory of dynamical systems, see Bowen (1975) and Ruelle (1989).

3.8 Bound under Exponential Stability

Theorem 2 (Bound under Exponential Stability): Suppose the flow $\phi_\tau(x)$ converges to the attractor A with exponential rate $\kappa>0$: $d(\phi_\tau(x),A)\le Ce^{-\kappa\tau}d(x,A)$

for some constant $C<\infty$, for all $\tau\ge 0$. Then: $D_\infty(x)=\int_0^\infty d(\phi_\tau(x),A)\,d\tau\le\frac{C}{\kappa}d(x,A)$

Proof: $D_\infty(x)=\int_0^\infty d(\phi_\tau(x),A)\,d\tau\le\int_0^\infty Ce^{-\kappa\tau}d(x,A)\,d\tau$

$C e^{-\kappa\tau} d(x, A)\, d\tau = C d(x,A) \int_0^\infty e^{-\kappa\tau}\, d\tau = \frac{C}{\kappa} d(x,A)$

Corollary: For linearly stable systems with recovery rate κ, $D_\infty(x) \le \frac{1}{\kappa} d(x,A)$ (when $C=1$).

Important: Exponential stability implies $D_\infty < \infty$. The converse is not claimed; polynomial convergence can also yield finite D_∞.

3.9 Recovery Rate Bound

Corollary 1 (Recovery Rate Bound): For a system satisfying the exponential stability hypothesis with constant C, the recovery rate κ satisfies: $$\kappa \le \frac{C\, d(x,A)}{D_\infty(x)}$$

For systems with $C=1$ (e.g., normal/symmetric linearizations with no transient overshoot), this reduces to: $$\kappa \le \frac{d(x,A)}{D_\infty(x)}$$

Proof: From Theorem 2, we have $D_\infty(x) \le \frac{C}{\kappa} d(x,A)$. Rearranging gives $\kappa \le \frac{C\, d(x,A)}{D_\infty(x)}$. When $C=1$, this reduces to $\kappa \le \frac{d(x,A)}{D_\infty(x)}$.

Interpretation: Small cumulative deviation implies rapid recovery (large κ). Large cumulative deviation implies slow recovery (small κ). This formalizes the intuitive link between D_T and κ. The C factor accounts for possible transient overshoot in non-normal systems.

3.10 Finite Horizon Approximation

Proposition 6 (Finite Horizon): For any $\epsilon > 0$, there exists a finite T_ϵ such that for all $T > T_\epsilon$: $|D_T(x) - D_\infty(x)| \le \epsilon$

$(x) - D_\infty(x) \leq \epsilon$

Proof: This follows directly from Theorem 2 under the exponential stability hypothesis. Since the integrand decays exponentially, the tail integral $\int_T^\infty d(\phi_\tau(x), A)\, d\tau$ can be made arbitrarily small by choosing T sufficiently large.

3.11 Summary of Properties

<table>
<tr><th>Property</th><th>Statement</th><th colspan="2"></th></tr>
<tr><td>Non-negativity</td><td>$D_T(x) \geq 0$</td><td colspan="2"></td></tr>
<tr><td>Monotonicity</td><td>$D_{T_2}(x) \geq D_{T_1}(x)$ for $T_2 \geq T_1$</td><td colspan="2"></td></tr>
<tr><td>Additivity</td><td>$D_{T+S}(x) = D_T(x) + D_S(\phi_T(x))$</td><td colspan="2"></td></tr>
<tr><td>Lipschitz continuity</td><td>(</td><td>D_T(x) − D_T(y)</td><td>\leq \frac{e^{LT} − 1}{L} |x − y|)</td></tr>
<tr><td>Instantaneous growth</td><td>$\frac{d}{dT} D_T(x) = d(\phi_T(x), A)$</td><td colspan="2"></td></tr>
<tr><td>Ergodic limit</td><td>$\lim_{T \to \infty} \frac{1}{T} D_T(x) = \int d(y, A)\, d\mu(y)$</td><td colspan="2"></td></tr>
<tr><td>Exponential stability implies finite D_∞</td><td>$D_\infty(x) \leq \frac{C}{\kappa} d(x, A)$</td><td colspan="2"></td></tr>
<tr><td>Recovery bound (general)</td><td>$\kappa \leq \frac{C\, d(x, A)}{D_\infty(x)}$</td><td colspan="2"></td></tr>
<tr><td>Recovery bound (C=1)</td><td>$\kappa \leq \frac{d(x, A)}{D_\infty(x)}$</td><td colspan="2"></td></tr>
<tr><td>Finite horizon approximation</td><td>$D_T(x) \to D_\infty(x)$ as $T \to \infty$</td><td colspan="2"></td></tr>
</table>

---

4. The Unified Variable Set

The following variables are defined operationally. Where a variable is a proposal, that is stated explicitly.

4.1 Corrective Permeability (κ)

Definition 4 (Corrective Permeability): κ is the recovery rate of the system to its attractor after a small perturbation. Operationally estimated as $\kappa = 1/\tau$ under approximately exponential relaxation, where τ is the characteristic recovery time constant. This coincides with the exponential convergence exponent in the linearized regime and is consistent with the original definition in the attractor framework.

Relationship to D_T: From Corollary 1, for a system with initial deviation $d(x, A)$, $\kappa \le \frac{D_\infty(x)}{C\, d(x, A)}$.

Note on κ's status: In this paper, κ is treated as a primitive empirical regime parameter. A stronger theory would derive κ from D_T and system geometry; this remains an open direction for future work.

4.2 Drift Rate (γ) — A Proposed Distinction

Definition 5 (Drift Rate): We propose the following operational distinction between dynamical regimes, based on the dominant Lyapunov exponent $\lambda_{\max}$:

Regime	λ_{max}	κ	γ	Behavior
Stable attractor	<-0.01	>0	0	Converges to fixed point
Persistent chaos	≈ 0	≈ 0	>0	Wanders without convergence
Full chaos	>0	undefined	>0	Diverges

Thresholds: $\lambda_{max} < -0.01$, $|\lambda_{max}| \leq 0.01$, and $\lambda_{max} > 0.01$ (pre-registered, measured in units of 1/epoch). These numerical thresholds are illustrative defaults rather than theoretically privileged constants.

Grounding: This distinction is inspired by the literature on chaos in high-dimensional neural networks (Engelken, Wolf & Abbott, 2023; Sompolinsky, Crisanti & Sommers, 1988; Clark, Abbott & Litwin-Kumar, 2023; Fournier & Urbani, 2023). For the treatment of stochastic and random perturbations, see Arnold (1998).

Falsification: If κ and γ are perfectly correlated (i.e., systems with small κ always have small γ), the distinction is not useful.

4.3 Basin Depth (B) and Persistence Depth ($\tilde{B}$)

Definition 6a (Basin Depth – Energy Barrier): B is the energy barrier required to escape the basin, measured as the potential difference between the attractor and the saddle point on the basin boundary: $B = V(\text{saddle}) - V(\text{attractor})$

This preserves the original definition from earlier papers.

Definition 6b (Persistence Depth): As a complementary measure, we define: $\tilde{B} = \min_{x \in \partial B} DT(x)$

This is the cumulative deviation required to reach the basin boundary. The relationship between B and $\tilde{B}$ remains an open mathematical question.

Operational alternative: In practice, the basin boundary may not be well-defined. Estimate B via the Arrhenius relationship $P_{escape} \propto e^{-B/T}$, where T is the noise level.

4.4 Reality Alignment (R)

Definition 7 (Reality Alignment): R is the expected log predictive likelihood: $R = \mathbb{E}[\log p(y \mid X)]$

where $p(y \mid X)$ is the system's predictive distribution over outcomes y given state X. Higher R indicates better predictive accuracy. This is a standard measure of predictive performance; the label "reality alignment" is a philosophical interpretation.

Direction-dependence: The framework interprets R as potentially direction-dependent: $R_{A \to B} \neq R_{B \to A}$. This captures the asymmetry found in Berglund et al. (2024), where models trained on "A is B" fail to generalize to "B is A." This interpretation is a framework-level claim.

Note on integration: Among the core variables, R is the least integrated with the trajectory-based formalism. Unlike κ, B, and $\tilde{B}$, which are directly derived from or related to DT, R is imported from Bayesian statistics. A more complete theoretical derivation of R from the same dynamical principles—perhaps as an information-theoretic functional of the occupation measure—remains an open direction for future work.

5. Theoretical Framework

5.1 Relationship Between DT, P_{topo}, and $E(t)$

Functional	What It Measures	Regime
$DT(x)$	Cumulative deviation from attractor	All systems
$P_{topo}(t)$	Topological feature lifetime	Systems with topological structure
$E(t)$	Rate of topological change	Learning systems

Hypothesis: In learning systems, DT and P_{topo} are positively correlated early in learning and negatively correlated late in learning. Turner & Barak (2023) demonstrate that RNNs develop attractors sequentially during training, which may correspond to phases of topological simplification. This is a testable prediction.

5.2 Relationship Between κ, γ, and $E(t)$

Hypothesis: In a learning system, the topological evolution rate $E(t)$ is monotonically related to κ only if the system is not in persistent chaos: $\partial E/\partial \kappa > 0$ (with E and κ measured on appropriate scales) in convergent regimes. In persistent chaos, $E(t)$ is monotonically related to γ: $\partial E/\partial \gamma > 0$. Correlation analysis provides a statistical test of these monotonicity relationships.

5.3 Adaptive Landscape (Heuristic Note)

The adaptive landscape $V(X,t)$ evolves as:

$$\dot{V} = g(X,V) - \lambda V + \xi(t)$$

For gradient systems with $\dot{X} = -\nabla_X V(X)$, and assuming the dynamics remain within the basin where higher-order nonlinearities are negligible, the cumulative deviation functional can be approximated as:

$$D_T(x) \approx \int_0^T \|\nabla_X V(\phi_\tau(x),\tau)\| \, d\tau$$

This is a local heuristic. A full derivation and integration into the core formalism is left for future work.

---

6. Testable Predictions

6.1 Core Prediction

Prediction: In a learning system, $E(t)$ is monotonically related to κ in convergent regimes: $\partial E/\partial \kappa > 0$ (with E and κ measured on appropriate scales), and $\partial E/\partial \gamma > 0$ in persistent chaos. Correlation analysis provides a statistical test of this monotonicity:

$$\mathrm{Corr}(E(t),\kappa) > 0 \iff \lambda_{max} < 0$$

$$\mathrm{Corr}(E(t),\gamma) > 0 \iff \lambda_{max} \approx 0$$

Falsification: If $E(t)$ correlates with κ in all regimes, or with γ in all regimes, the prediction is falsified.

6.2 Secondary Prediction

Prediction: In systems with high R, DT and P_{topo} are negatively correlated late in learning; in systems with low R, they are uncorrelated or positively correlated.

Falsification: If DT and P_{topo} are negatively correlated in both high-R and low-R systems, the prediction is falsified.

6.3 Boundary Condition and Global Falsifier

Conjecture: We conjecture that the framework applies to any system satisfying:

- A. Well-defined state space.
- B. Subject to perturbations.
- C. Exhibits at least one identifiable attractor.
- D. Dynamics are observable and measurable.

Global Falsifier: The unified ontology claim collapses if a system is found where DT, κ, and topological persistence are mutually independent across all regimes, and where R cannot be expressed as a functional of the trajectory or occupation measure. If such a system exists, the framework's claim to unify persistence, stability, and reality alignment would be falsified.

---

7. Experimental Design

7.1 System Choice

Train a CNN on MNIST or CIFAR-10. Use latent activation manifolds for topological analysis.

Justification: Karuppiah, Nazreen Banu et al. (2026) demonstrate the use of persistent homology on activations to study feature learning and generalization. Turner & Barak (2023) show that RNNs develop attractors sequentially, providing a controlled setting for studying topological evolution during learning.

7.2 Variable Measurement

Variable	Protocol
$DT(x)$	Sample weights; compute distance to final attractor; integrate.
$Ptopo(t)$	Compute persistent homology on latent activations; sum feature lifetimes.
$E(t)$	Finite differences of $Ptopo(t)$.
κ	Perturb weights; measure recovery time τ; $\kappa=1/\tau$.
γ	Compute average drift rate during training.
R	Cross-domain generalization accuracy.

7.3 Statistical Analysis

- Correlate $E(t)$ with κ and γ conditional on regime.
- Pre-register thresholds and sample size.

Note on future empirical work: A full empirical validation would require pre-registration with specified sample size, significance thresholds, power analysis, and robustness checks. These are planned for subsequent work.

---

8. Discussion

8.1 Implications

The paper provides a candidate formalization with defined variables, mathematical properties, and testable predictions. The mathematical properties of DT establish its relationship to κ and provide a foundation for the framework’s core claims.

8.2 Limitations

- P_{topo} is computationally expensive.
- The framework is a meta-theory, not a complete domain-specific theory.
- Variables may be confounded; causal inference requires controlled experiments.
- The κ/γ regime distinction is proposed and requires empirical validation.

8.3 Future Work

- Empirical validation of predictions.
- Formal derivation of relationships from first principles.
- Extension to other domains.
- Computational efficiency improvements.

---

9. Conclusion

This paper proposes a candidate formalization for the attractor framework. The central mathematical innovation is treating persistence as a functional defined over trajectories—$D_T(x)=\int_0^T d(\phi_\tau(x),A)\,d\tau$—rather than as a scalar property of states. We defined the cumulative deviation functional D_T, the topological persistence functional $P_{\text{topo}}(t)$, and the topological evolution rate $E(t)$. We proved several mathematical properties of D_T, including non-negativity, monotonicity, additivity, Lipschitz continuity, and a bound relating D_∞ to κ: $D_\infty(x)\le \frac{C}{\kappa} d(x,A)$. We established connections to dynamic programming and ergodic theory. We unified the variable set with operational definitions. We derived testable predictions and provided a falsifiable experimental protocol.

The framework now admits formal definitions, operational variables, and empirical tests. The next step is empirical validation.

Appendix A: Possible Extensions from Larose (2025) — Unverified Source

Note: The following source has not been independently verified. It is included for completeness and as a potential direction for future exploration, but should not be treated as established.

Larose (2025) develops a framework for recursive deformation systems. Two constructs are potentially relevant:

Constraint Functional: $C(X)=\int_{\text{trajectory}}\|\nabla\Phi\|\, d\tau$, measuring cumulative irreversible deformation.

Persistence Invariant: $I_p=\oint_R d\Phi$, a topological invariant.

These are not yet integrated into the core framework and are presented here for completeness and future exploration. They should be treated as unverified candidate extensions.

References

Arnold, L. (1998). *Random Dynamical Systems*. Springer.

Berglund, L., et al. (2024). “The Reversal Curse: LLMs Trained on ‘A is B’ Fail to Learn ‘B is A’.” *arXiv:2309.12288*.

Bowen, R. (1975). *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*. Springer.

Carlsson, G. (2009). “Topology and data.” *Bulletin of the American Mathematical Society*, 46(2), 255-308.

Carlsson, G., & Zomorodian, A. (2009). “The theory of multidimensional persistence.” *Discrete & Computational Geometry*, 42(1), 71-93.

Clark, D. G., Abbott, L. F., & Litwin-Kumar, A. (2023). “Dimension of activity in random neural networks.” *Physical Review Letters*, 131, 118401.

Edelsbrunner, H., & Harer, J. (2010). *Computational Topology: An Introduction*. American Mathematical Society.

Engelken, R., Wolf, F., & Abbott, L. F. (2023). “Lyapunov spectra of chaotic recurrent neural networks.” *Physical Review*

Research, 5, 043044.

Fournier, S. J., & Urbani, P. (2023). “Statistical physics of learning in high-dimensional chaotic systems.” *Journal of Statistical Mechanics: Theory and Experiment*, 2023(11), 113301.

Karuppiah, K., Nazreen Banu, M., et al. (2026). “Topological Data Analysis (TDA) as a Framework for Understanding Deep Learning Behavior.” *2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG)*, Indore, India, December 12-13, 2025. IEEE Xplore. DOI: 10.1109/ICTBIG68706.2025.11323998.

Larose, H. (2025). “A Mathematical Theory of Frame-Independent Persistence.” *Academia.edu*. [Unverified source.]

Ruelle, D. (1989). *Chaotic Evolution and Strange Attractors*. Cambridge University Press.

Sompolinsky, H., Crisanti, A., & Sommers, H. J. (1988). “Chaos in Random Neural Networks.” *Physical Review Letters*, 61(3), 259-262.

Turner, E., & Barak, O. (2023). “The Simplicity Bias in Multi-Task RNNs: Shared Attractors, Reuse of Dynamics, and Geometric Representation.” *Advances in Neural Information Processing Systems (NeurIPS)*.

Suggested citation: Galida, R. S. (2026). The Persistence Functional: A Candidate Formal Foundation for the Attractor Framework (Foundational Edition). *Fantasy Attractor*.

Why Clockwork Interventions Fail in Complex Systems: A Prescription from the Attractor Framework [A] (2026)

Robert Galida — June 2026 (Final)

See Paper 1 (Intelligence Without Consciousness) for the full taxonomy of attractors, κ, and basin depth. See Basin Defense and Stable Addition for cross-domain synthesis and rate-induced tipping.

Abstract

Most human institutions, policies, and interventions treat complex adaptive systems as if they were clockwork systems — linear, predictable, and responsive to force. This is a category error. Complex systems (ecosystems, brains, societies, belief systems) have attractors, basins, multiple nested timescales (κ vector), and thresholds. Applying sudden force above a critical rate or magnitude triggers basin defense: ejection, backlash, entrenchment, or catastrophic collapse. This paper diagnoses the clockwork fallacy, introduces a multi-timescale operationalization of corrective permeability, offers a mechanism for parallel attractor replacement, and acknowledges the institutional constraints that make patient intervention rare. The central argument is that failure is not random but structurally predictable.

---

1. Introduction

A thermostat is a clockwork system. Push the temperature up, the cooling turns on; push harder, it turns on faster. No hidden attractors, no basin defense, no hysteresis. Force works predictably.

A human being is not a thermostat. Neither is a democracy, an ecosystem, a marriage, or a belief system. They have attractor basins – stable states that resist displacement. They have multiple corrective timescales (κ vector) – characteristic return times after perturbations at different levels. They have thresholds – points at which a small additional push can cause a regime shift.

Yet most interventions treat these complex systems **as if they were clockwork**. Apply more force → get more change. This is the **clockwork fallacy**.

This paper diagnoses the fallacy using the attractor framework, operationalizes κ for non-physical domains as a vector of timescales, specifies the mechanism of parallel attractor replacement, and acknowledges the institutional constraints that make slow intervention rare.

---

2. The Clockwork Fallacy in Framework Terms

Clockwork assumption	Complex system reality
Linear response: more force → more change	Nonlinear: small force may be ejected; force above threshold may cause collapse

Clockwork assumption	Complex system reality
No memory: each intervention acts independently	Hysteresis: history matters; past perturbations shape current basin depth
No internal dynamics: system is passive	System has its own attractors and κ vector; it actively resists displacement
Fast intervention is better (efficiency)	Rate matters; fast perturbation triggers basin defense; slow perturbation may integrate

The clockwork fallacy treats the system as a **passive object** to be pushed. The attractor framework treats it as an **active agent** with its own stability dynamics.

---

3. Operationalizing κ as a Multi-Timescale Vector

$\kappa = 1/\tau$, where τ is the characteristic return time to baseline after a small perturbation. For physical systems (thermostat, RC circuit), τ is a single scalar. For complex adaptive systems, **τ is not a single number** — there are multiple, nested timescales:

Timescale	Definition	Example (addiction)
Fast κ (seconds–hours)	Return time after transient perturbation	Craving decay
Medium κ (days–weeks)	Return time after moderate perturbation	Withdrawal normalization
Slow κ (months–years)	Return time after identity-level perturbation	Identity fusion / self-model reorganization

Timescale	Definition	Example (addiction)
$\kappa\infty$ (effectively zero)	No measurable return; the attractor is sealed	Fantasy attractor (see Paper 1)

Implication: A system can have fast κ (rejects rapid, small perturbations) and slow κ (integrates slow drift) simultaneously. The optimal perturbation rate depends on *which* κ you are trying to match.

Protocol for estimating κ in a non-physical domain:

1. Select a modest, low-stakes belief (not identity-core).
2. Introduce a small, credible counter-evidence (pilot perturbation).
3. Measure the time until the person returns to their original stated belief (via repeated interviews, surveys, or behavior tracking).
4. τ is the median return time; $\kappa = 1/\tau$.
5. Repeat with perturbations that target different subsystem levels (e.g., factual vs. identity-relevant) to estimate the κ vector.

Limitation: The pilot perturbation protocol uses a *small* perturbation to estimate κ. The intervention may require a *large* perturbation to escape the basin. The small-perturbation estimate may not predict behavior near the basin boundary. This is an acknowledged operational limitation, not a circularity. The framework is falsified if a system with measured low κ (slow return) reliably integrates *rapid, large* perturbations without ejection or transient absorption, and if the small-perturbation estimate is stable across perturbation magnitudes.

---

4. Why Clockwork Interventions Fail: Four Mechanisms

Mechanism 1: Ejection (Backlash) — When a perturbation is applied too fast or with too much force, the system ejects the addition, often returning with a deepened basin. Examples: sanctions that strengthen a regime, direct refutation that backfires.

Mechanism 2: Transient Absorption Followed by Return — The system temporarily changes, then returns to baseline when the perturbation stops. Examples: short-term policy boosts, crash diet weight regain.

Mechanism 3: Catastrophic Regime Shift — Force applied at a critical threshold causes an abrupt, often irreversible shift to a different, sometimes worse attractor. Examples: lake eutrophication, restructuring that destroys institutional knowledge.

Mechanism 4: Rate-Induced Tipping — A small cumulative change, applied faster than the relevant κ, causes tipping. Examples: rapid currency appreciation triggering crisis, fast cultural change provoking backlash.

---

5. Parallel Attractors: The Mechanism of Replacement

Parallel attractors are introduced as an alternative to direct displacement. How does a parallel attractor eventually replace the original?

Mechanism: Basin-share competition

When a parallel attractor is created, it initially has a shallow basin. Through repeated use, reinforcement, and social validation, its basin depth increases. Meanwhile, the original attractor may become shallower through disuse or decoupling of identity fusion. The transition is not a flip; it is a **continuous shift in basin dominance**. At some point, the new attractor's basin depth exceeds the old attractor's, and the system's typical trajectories are captured by the new state.

Testable prediction: During parallel attractor formation, the system will exhibit **bistability** — both states are possible for a range of control parameters. In social systems, this predicts polarization; in organizational change, it predicts pilot-program coexistence; in belief systems, it predicts identity compartmentalization.

Empirical examples: Harm reduction (methadone maintenance creates a parallel attractor that may deepen over time); phase-in policies (smoking bans create new norm attractors alongside old habits); belief change (new social identity cultivated alongside old identity, enabling eventual abandonment without direct confrontation).

---

6. The Political Economy of Slow Intervention

The attractor framework prescribes patience, precision, and gradual perturbation. But policymakers, clinicians, and managers face **institutional incentives** that systematically favor fast, visible, forceful action:

- Election cycles (2–4 years) reward short-term results, not long-term basin reshaping.
- Media attention favors dramatic events, not gradual change.
- Bureaucratic accountability demands measurable outputs, not process fidelity.

- Crisis narratives demand action, not waiting.

Consequence: Even when the framework is correct, it is often institutionally **unimplementable**. The best intervention may be politically impossible.

What would institutional redesign look like? Examples:

- **Longer funding cycles** (5–10 years) for policy and program evaluation, allowing basin-reshaping interventions to mature.
- **Preregistered patience metrics** — requiring intervention designs to specify expected τ and κ, with success measured by reduction in τ over time, not immediate outcomes.
- **Insulation from electoral pressure** for certain regulatory functions (e.g., central bank independence, long-term environmental planning).
- **Dual-track systems** that allow parallel attractors to develop (e.g., pilot programs exempt from standard performance metrics).

Implication for the paper's claims: The framework diagnoses why interventions fail, but it does not guarantee that successful interventions can be implemented. This is not a weakness — it is a feature. The framework clarifies the gap between effective intervention and institutional feasibility. Bridging that gap requires institutional redesign, not just better perturbation design.

---

7. Case Studies

Case 0: Smoking cessation (addiction) — the motivating challenge

In smoking cessation, abrupt cessation (cold turkey) often outperforms gradual tapering (Lindson et al., 2016 meta-analysis). This appears to contradict the prescription "slow perturbation at rate ≤ κ."

Framework interpretation: Addiction has multiple κ timescales. Cold turkey may target the fast-κ (craving) subsystem while the slow-κ identity subsystem remains dormant; gradual tapering may keep both active, prolonging distress.

Falsifiable prediction: Patients with higher identity-fusion scores (measurable via existing scales, e.g., the Identity Fusion Scale) should show worse outcomes with gradual tapering relative to cold turkey. If identity fusion is low, gradual tapering may be equivalent or superior.

Alternative explanations acknowledged: The meta-analysis does not adjudicate between the attractor framework and other accounts (e.g., cognitive dissonance, cue elimination, withdrawal distress). The framework's contribution is to generate the identity-fusion interaction prediction, which can be tested independently.

Case 1: Lake eutrophication (ecological)

- *Clockwork approach:* Sudden nutrient reduction after flipping to turbid state — fails (hysteresis). True hysteresis is technically established for some lakes (Scheffer et al., 2001).
- *Framework approach:* Gradual nutrient reduction before tipping (rate ≤ κ) might have avoided the flip. After tipping, parallel attractor (biomanipulation) is required.

Case 2: Political persuasion (belief systems)

- *Clockwork approach:* Direct refutation, evidence bomb —

backfire effect (ejection with deepened basin).

- *Framework approach:* Yang et al. (2022) demonstrated in a field experiment that “pacing and leading” — starting with some agreement and gradually introducing opposing content — produced attitude change, whereas blunt argument triggered backlash. This is gradual perturbation at rate ≤ κ, combined with identity decoupling.

Case 3: Organizational change

- *Clockwork approach:* Sudden layoffs, top-down mandate — triggers basin defense (resistance, morale loss).
- *Framework approach:* Gradual, participatory change (rate ≤ κ) with parallel structures (pilots, dual systems). *Note:* Hysteresis in organizations is not technically demonstrated; the paper uses “analogous” language.

---

8. Practical Heuristics

If the system has…	Then…	Caveat
Fast κ (seconds–hours)	Rapid, sharp interventions may be required; slow drift may be tracked or rejected	For very deep basins, only a large shock may work
Slow κ (months–years)	Slow, gradual perturbation; avoid rapid shocks	Identity-fused systems may need abrupt escape (Case 0)

If the system has…	Then…	Caveat
Multiple κ timescales	Target the slowest κ for lasting change; use fast κ for immediate disruption	Requires measurement of the κ vector
κ → 0 (fantasy attractor; no measurable return)	Intervention is futile within the model. Accept, circumvent, or refer to Paper 1	Out of scope for this paper
Hysteresis (true bistability)	Do not force return; cultivate a parallel attractor	Hysteresis is established for some ecological systems; for social systems, use "analogous"
Identity fusion	Do not attack belief directly. Decouple identity first, then perturb gently	Requires trust; may be infeasible in adversarial contexts

---

9. Conclusion

The clockwork fallacy — treating complex adaptive systems as linear, passive, and force-responsive — is a primary cause of failed interventions. The attractor framework diagnoses the failure modes (ejection, transient absorption, catastrophic shift, rate-induced tipping) and offers a prescriptive alternative: measure the κ vector, match perturbation rate to the relevant timescale, build parallel attractors, and wait.

The framework does not guarantee success. Institutional incentives (election cycles, media pressure, bureaucratic accountability) systematically favor the clockwork approach, making patient intervention rare. The value of the framework is diagnostic: it explains why failure is not random, and it clarifies the gap between effective intervention and political feasibility. Bridging that gap requires institutional redesign — longer funding cycles, preregistered patience metrics, and insulation from electoral pressure.

The dance of change is not about pushing harder. It is about learning to move with the system — but also knowing when the system cannot be moved with the tools and time available.

Suggested citation: Galida, R. S. (2026). Why Clockwork Interventions Fail in Complex Systems: A Prescription from the Attractor Framework. *Fantasy Attractor.*