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—$D_T(x)=\int_0^T d(\phi_\tau(x),\mathcal{A})\,d\tau$—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: $D_\infty(x)\leq\frac{C}{\kappa}d(x,\mathcal{A})$. 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 $\phi_0 = \mathrm{id}$. 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 $t \geq 0$: $P_{topo}(t) = \int_0^t \sum_{k \geq 0} \sum_{(b,d) \in PH_k(X_\tau)} (d-b)\, d\tau$

The map $\tau \mapsto PH_k(X_\tau)$ 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: $P_{topo}(t)$ 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 D_T; 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 D_T, 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: $D_{T_2}(x) \geq D_{T_1}(x)$ 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:
$$D_T(x) = D_{\tau_0}(x) \quad \text{for all } T \geq \tau_0$$

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

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 $V(x) = \inf_u D_\infty(x)$

$D_\infty(x)$ 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., $\|\phi_\tau(x) - \phi_\tau(y)\| \leq e^{L\tau}\|x - y\|$. Then for any x, y in the basin of A: $$|D_T(x) - D_T(y)| \leq \int_0^T e^{L\tau}\, d\tau \, \|x - y\| = \frac{e^{LT} - 1}{L}\|x - y\|$$

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

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

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 $\kappa > 0$: $d(\phi_\tau(x), A) \leq C e^{-\kappa \tau} d(x, A)$

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

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

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

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

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

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 $T > T_\epsilon$: $\|D_T(x) - D_\infty(x)\| \leq \epsilon$

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

Property	Statement		
Non-negativity	$D_T(x) \geq 0$		
Monotonicity	$D_{T_2}(x) \geq D_{T_1}(x)$ for $T_2 \geq T_1$		
Additivity	$D_{T+S}(x) = D_T(x) + D_S(\phi_T(x))$		
Lipschitz continuity	(	D_T(x) − D_T(y)	\leq \frac{e^{LT} − 1}{L} |x − y|)
Instantaneous growth	$\frac{d}{dT} D_T(x) = d(\phi_T(x), A)$		
Ergodic limit	$\lim_{T \to \infty} \frac{1}{T} D_T(x) = \int d(y, A)\, d\mu(y)$		

Property	Statement	
Exponential stability implies finite D_∞	$D_\infty(x) \le \frac{C}{\kappa} d(x,A)$	
Recovery bound (general)	$\kappa \le \frac{C\, d(x,A)}{D_\infty(x)}$	
Recovery bound (C=1)	$\kappa \le \frac{d(x,A)}{D_\infty(x)}$	
Finite horizon approximation	$D_T(x) \to D_\infty(x)$ as $T \to \infty$	

---

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 $d(x,A)$, $\kappa \le \frac{C\, d(x,A)}{D_\infty(x)}$.

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 λ_{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:$$\tilde{B} = \min_{x \in \partial B} D_T(x)$$

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 $P_{\text{escape}} \propto e^{-B/T}$, 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 D_T, 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 D_T, P_{topo}, and $E(t)$

Functional	What It Measures	Regime
$D_T(x)$	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, D_T 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: $\mathrm{Corr}(E(t), \kappa) > 0 \implies \lambda_{\max} < 0$ $\mathrm{Corr}(E(t), \gamma) > 0 \implies \lambda_{\max} \approx 0$

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, D_T and P_{topo} are negatively correlated late in learning; in systems with low R, they are uncorrelated or positively correlated.

Falsification: If D_T 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 D_T, 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*.

The Persistence Functional: A Mathematical Measure of Attractor Resilience

Author: Robert Galida
Date: May 2026

Abstract

The attractor framework says that **persistence under disturbance** is the basic mark of reality.
To turn this idea into a formal science, we introduce the **persistence functional** P(x)$P(x)$.

P(x)$P(x)$ is a single number that measures:

- How deep a state is inside an attractor basin.
- How quickly it returns after a knock.

We define P(x)$P(x)$ for three different kinds of systems:

1. **Deterministic dissipative systems** — here P is linked to Lyapunov exponents and basin stability.
2. **Stochastic systems** — here P is linked to escape time and quasipotential.
3. **Information-theoretic systems** — here P is linked to negative free energy or mutual information.

The **recovery rate** $-\dot{P}/P$ is a universal sign of **critical slowing down** — a warning that a system is about to tip.

We also discuss limitations: resilience may depend on direction ("anisotropic"), and multiple timescales may need **vector** or **tensor** persistence. We list open mathematical problems.
This paper is a **roadmap**, not a finished theory.

---

1. Introduction

In the attractor framework, **persistence under disturbance** is central. But we have not had a single number to say *how persistent* a state is.
The **persistence functional** $P(x)$ aims to fill that gap.

What $P(x)$ should do:

- $P(x)>0$ for states inside an attractor basin.
- For a **conservative attractor** (like a free electron), P is maximal (normalised to 1).
- For a **dissipative attractor**, P drops after a disturbance and then recovers.
  The recovery rate $-\dot{P}/P$ equals:
  - the negative of the largest Lyapunov exponent (for deterministic systems)
  - the inverse return time (for stochastic systems)

- the rate of information loss (for informational systems)

- P falls as the system approaches a **bifurcation**, giving early warning.

We do **not** give one universal formula. Instead, we give a **family** of definitions, each suited to a different type of system, all united by the same purpose — measuring resilience.

---

2. Deterministic Dissipative Systems

Consider a smooth system $\dot{x}=f(x)$ with a stable attractor A and its basin $B(A)$.

A natural candidate for $P(x)$ uses a **Lyapunov function** $V(x)$ — a kind of energy that always decreases inside the basin ($\dot{V}<0$).

We define:$$P(x)=1-\frac{V(x)-V_A}{V_{\max}-V_A}$$

This gives $P=1$ on the attractor and $P\to 0$ at the basin boundary.
Near the attractor, the recovery rate is related to the **largest Lyapunov exponent** λ_1:$-\dot{P}/P\approx-\lambda_1$

When the system approaches a tipping point, $\lambda_1\to 0^-$, so the recovery rate slows down — this is **critical slowing down**.

Conclusion: For deterministic systems, P can be built from a Lyapunov function. The recovery rate equals the negative of the largest Lyapunov exponent.

---

3. Stochastic Systems

When noise is present, persistence is about how long it takes to escape from the basin.
The **mean first passage time** $\tau(x)$ — the average time to leave — is a natural measure.

We define: $$P(x) = \frac{\tau(x)}{\tau_{\max}}$$

where $\tau_{\max}$ is the value at the attractor.

For weak noise, $\tau(x)$ grows exponentially with the **quasipotential** $U(x)$ (Freidlin–Wentzell theory): $$\tau(x) \sim e^{U(x)/\epsilon}$$

So: $$P(x) \propto e^{-(U_{\max} - U(x))/\epsilon}$$

The recovery rate is the inverse of the return time. As a tipping point is approached, the return time diverges, and the recovery rate goes to zero. This again gives **critical slowing down** — rising variance and autocorrelation.

Conclusion: For stochastic systems, P is proportional to the mean exit time (or the exponential of the quasipotential). This connects persistence to large deviation theory.

---

4. Information-Theoretic Systems

For systems where information matters (neural, cognitive, social), we can define persistence using **mutual information** between past and future.

Let $I_{\text{past,future}}$ be the **predictive information**. Then: $$P(t) = I(\text{past};\text{future at time } t) \quad \text{or} \quad P = e^{-\text{surprisal}}$$

uture at time t) or $P = e^{-\text{surprisal}}$

The decay of $P(t)$ over time measures **memory loss**.
Landauer's principle connects information loss to entropy production: $\dot{P}/P \leq -\frac{\dot{S}}{k_B \ln 2}$

Alternatively, in the **free energy principle** (Friston), the negative free energy $-F$ acts like a Lyapunov function. We can set: $P = e^{-F/kT}$ or $P = -F$

Then $-\dot{P}/P$ is the rate of free energy minimisation, which slows near bifurcations.

Conclusion: For information-theoretic systems, P can be defined via mutual information decay or negative free energy, linking persistence to entropy production and predictive coding.

---

5. Unifying Recovery Rate and Critical Slowing Down

Across all types of systems, the **recovery rate** $\lambda_{rec} = -\dot{P}/P$ (just after a small disturbance) is a universal indicator:

- **Deterministic dissipative:** $\lambda_{rec} = |\lambda_1|$ (absolute value of the largest Lyapunov exponent)
- **Stochastic:** λ_{rec} = inverse of the return time, related to the quasipotential's curvature
- **Information-theoretic:** λ_{rec} = rate of free energy minimisation or information loss

As the system approaches a bifurcation, $\lambda_{rec} \to 0$. This is **critical slowing down**.
It shows up as rising lag-1 autocorrelation and variance

(Scheffer et al., 2009).
So P and its recovery rate give early warnings.

---

6. Normalisation for Conservative Attractors

For a perfect **conservative attractor** (e.g., an electron in its ground state, no decay), the persistence functional should be constant and maximal:

$$P_{\text{cons}} = 1 \quad \text{for all times}$$

No recovery rate is defined (or it is zero). This anchors the scale.

For **emergent approximate conservative systems** (like atomic clocks), P is very close to 1 and decays extremely slowly.

---

7. Limitations – Scalar Collapse and Anisotropic Resilience

A single scalar $P(x)$ may not be enough for systems where resilience is **anisotropic** – that is, recovery speed depends on the direction of the perturbation.
High-dimensional systems can have **multiple timescales** (fast and slow modes). A scalar average can miss important structure.

Future work may need:

- **Vector persistence** – a list of recovery rates along different directions.
- **Tensor persistence** – a metric that captures the full

shape of the basin.
- **Persistence manifold** – the geometry of the basin in state space.

We accept this limitation. The scalar P is a useful first approximation for systems with isotropic resilience or for early-warning applications where a single number is enough. For complex systems, a multidimensional generalisation is an open research problem.

---

8. Open Mathematical Problems

1. **Derive $P(x)$ from first principles** for a given class of systems (e.g., from a variational principle).
2. **Prove that $-\dot{P}/P = \langle \lambda_1 \rangle$** for a wide class of dissipative systems.
3. **Extend the definition to systems with multiple attractors and chaotic basins** (where basin stability is fractal).
4. **Establish a rigorous relationship between P and the mutual information decay rate** for non-equilibrium processes.
5. **Formulate a universal persistence functional** that works across all regimes – or prove it's impossible.
6. **Test the predictive power of P** in controlled experiments (e.g., ecological microcosms, neural cultures, social media sentiment).
7. **Develop vector/tensor persistence** for anisotropic resilience.

---

9. Conclusion

The persistence functional $P(x)$ gives a mathematical language for attractor resilience.

We have given **operational definitions** for three regimes:

- **Deterministic dissipative** → Lyapunov / basin stability
- **Stochastic** → escape time / quasipotential
- **Information-theoretic** → mutual information / free energy

The **recovery rate** $-\dot{P}/P$ unifies critical slowing down across all these domains.

We have explicitly noted **limitations** (scalar collapse, anisotropy) as open problems.

This paper is a **roadmap**, not a final theory. The framework now has a quantitative step.

Suggested citation: Galida, R. S. (2026). *The Persistence Functional: Towards a Mathematical Measure of Attractor Resilience (Reader-Friendly Version).* Fantasy Attractor.