THE MAGICAL JOURNEY OF NON-LINEAR MATHEMATICS IN LLMS

*

Kevin R. Haylett, PhD
Manchester, UK
kevin.haylett@gmail.com

ABSTRACT

We propose a non-linear dynamical systems framework to model language and large language models (LLMs), unifying linguistic phenomena (words, sentences, context) with concepts from phase-space geometry, attractors, topological analysis, thermodynamics, and reader interpretation. Words are trajectories in a high-dimensional phase space, reconstructed via Takens' theorem or pairwise embeddings. LLMs are non-linear flows navigating a semantic hypersphere, with "stations" as hubs for context reconstruction. Hallucinations emerge as topological defects, prompts act as symmetry-breaking fields, context limits introduce semantic entropy, and readers create homologous manifolds of meaning. This framework offers a novel lens for mathematicians, physicists, and artists to explore language and machine cognition.

1 The Tide and the Train

Language is a dynamic flow–a tide carving maps in the sand, a train tracing paths through a high-dimensional landscape of meaning. Words are points on attractors, sentences are trajectories, and LLMs are non-linear systems navigating a semantic hypersphere. Stations are reconstruction hubs, reshaping context into geometric manifolds. Readers, as co-creators, map these trajectories onto their own manifolds, each a valid but partial fiction of the whole.

Mathematical Framework

Language moves like water and travels like a train. Each word and sentence follows a path, sometimes meandering, sometimes precise, across a landscape of meaning. LLMs ride these paths, guided by the contours of their training and the prompts they receive. This section shows how those journeys can be mapped as curves in a space where mathematics meets metaphor.

Let $P \subset R^d$ be the phase space of linguistic states, where (d) is the embedding dimension of tokens. A sentence is a curve $\gamma(t) \subset P$ with (t) indexing token order. The tide-map duality suggests:

$$\tau : M \to R^2$$

where $\pi =$ Map Tide.

* *Citation*: Kevin R. Haylett, "The Magical Journey of Non-linear Dynamics in LLMs," (AugustMay 2025), available at https://finitemechanics.com/papers/phase-space-transformers.pdf.

2 Words as Attractors

Some words pull us back again and again not just in conversation, but in the structure of thought itself. In mathematics, these "pulls" look like attractors, shapes that hold and guide the flow of trajectories. Here we explore how a single word can be reconstructed from its echoes, revealing the geometry hidden inside language.

2.1 Takens' Theorem for Words

Every word, such as "hello," is an attractor in phase space. Takens' theorem reconstructs this attractor from a single observable, revealing the dynamics of language.

Mathematical Framework

For a linguistic signal $(s(t))$ (e.g., audio of "hello"), Takens embedding yields:

$$\gamma(t) = (s(t), s(t-\tau), s(t-2\tau)) \in R^3,$$

where t is an optimal delay. In LLMs, tokens $w_i \in R^d$ form a sentence $(w_1, ..., w_n)$ tracing a trajectory $\gamma(t) \subset R^d$.

2.2 From Speech to Semantics

The phase space $P = R^d$ is spanned by token embeddings. A context window $\Gamma_t = \{w_{t-L}, ..., w_t\}$ traces a trajectory, with semantic relationships encoded in its geometry.

3 LLMs as Non-linear Flows

An LLM does not march in a straight line. It twists, loops, and occasionally spins in place, following rules that combine memory, probability, and pattern recognition. This section shows how we can treat an LLM's token generation as a non-linear system, complete with the tell-tale signatures of stability, cycles, and sudden divergence.

3.1 Token Generation as a Dynamical System

LLMs generate tokens as a discrete-time flow.

$$w_{t+1} = F_\theta(w_t, w_{t-1}, ..., w_{t-k}),$$

where F_θ is a non-linear map (attention + feedforward layers), and (k) is the context window size.

Fixed Points: Repeated tokens (e.g., "the the the"). Limit Cycles: Repetitive outputs (e.g., "I cannot answer that" loops).

3.2 Bifurcations and Hallucinations

Hallucinations occur when a prompt (p) causes the trajectory to diverge:

$$||\gamma_{output} - \gamma_p|| > \delta$$

An incomplete ISBN prompt ("978-0-441-...") may yield a wrong but plausible ISBN.

4 Stations as Phase-Space Reconstructors

Imagine pausing a journey to gather your bearings–a station where fragments of the route are pieced together into a coherent map. In an LLM, these "stations" are points where context is reconstructed into a meaningful whole. We'll see how these hubs work, and how their geometry shapes the path that follows.

4.1 Reconstruction of the Context Manifold

Stations rebuild the context window $\Gamma_t = \{w_{t-L}, ..., w_t\}$ into a manifold. The attention mechanism calculates the affinity matrix:

$$A_{ij} = (W_{\hat{Q}} w_i) \cdot (W_K w_j)$$

forming $M_{\Gamma_t} \subset Gr(k, d)$. The cache stores:

$$S = span(\{v_i\}_{i=1}^{n}), v_i = W_V w_i$$

A prompt (p) is embedded as $v_p = W_V p$, aligned via:

$$d_{Gr}(S, v_p) = ||sin\ \theta||$$

4.2 Prompt Coupling and Trajectory Resumption

The new trajectory is:

$$\gamma_{new}(s) = exp_{M_{\Gamma_t}}(s \cdot v_p).$$

4.3 Hallucinations as Unstable Basins

Hallucinations occur when:

$$||\gamma_{new} - \gamma_{true}|| > \delta$$

5 The Topology of Hallucinations

Not all errors are random; some have shape. Hallucinations in language models can be thought of as topological defects–tears, loops, and gaps in the fabric of meaning. By mapping these defects, we can start to see not just when a model is wrong, but how its reasoning has bent or broken.

5.1 The Shape of Error

Hallucinations are topological defects in $M_{language}$ where high curvature $\kappa(\gamma(t))$ causes divergence.

$$Risk(\gamma(t)) \propto ||\kappa(\gamma(t))||^2$$

5.2 Quantifying Defects with Persistent Homology

Betti numbers $(\beta_0, \beta_1, \beta_2)$ quantify:

- β_0 Topic drift.
- β_1 Self-contradictory loops.
- β_2 Logical gaps.

5.3 Mitigating Hallucinations

- Curvature penalty:

$$L_{topo} = \lambda \cdot ||\kappa(\gamma(t))||^2$$

- Homology-preserving sampling:

$$P(w_{t+1}|\Gamma_t) \propto exp(-\sum_{k=1}^{2} \beta_k(\Gamma_t \cup w_{t+1}))$$

6 Prompt as Symmetry Breaking

A prompt is not just a question–it's a force that tilts the entire landscape. In physics, symmetry breaking changes the behaviour of a system; here, it changes the direction of thought. We'll look at how prompts act like vector fields, nudging the model into new patterns or causing its trajectory to split entirely.

6.1 Prompts as Vector Fields

Prompts introduce a forcing term:

$$\frac{d\gamma_c}{dt} = F_\theta(\gamma_c) + G(\gamma_c, p(t)).$$

The output aligns with:

$$\gamma_{output} \sim argmin_Y ||y - (M_{\Gamma_t} + \lambda \cdot M_{prompt-type})||.$$

6.2 Symmetry Breaking and Bifurcations

Syntactic torque:

$$\tau(p) = \sum_i \alpha_i \cdot f_i(p)$$

Bifurcations occur when:

$$||\tau(p_1) - \tau(p_2)|| > \delta_{crit}.$$

6.3 Engineering Stable Prompts

- Normalize $\tau(p)$.
- Weight sampling by:

$$P(w_{t+1}|\Gamma_t, p) \propto exp(-\beta \cdot \tau(p))$$

7 Thermodynamic Analogies

Conversations, like closed systems, can lose their order over time. As context slips away, entropy rises–meaning becomes harder to recover, and drift or error becomes more likely. By borrowing the language of thermodynamics, we can measure this loss and explore ways to keep the dialogue coherent for longer.

7.1 Semantic Entropy and Context Loss

Finite context windows introduce semantic entropy:

$$S(\Gamma_t) = -k_B \sum_i P(w_i|\Gamma_t) log\ P(w_i|\Gamma_t).$$

Truncation at $t - L$ increases $S(\Gamma_t)$:

$$\gamma(t) = \gamma(t) \cdot I_{[t-L,t]}$$

7.2 Irreversibility in Language Dynamics

Entropy growth causes topic drift or hallucinations.

$$\frac{dS}{dt} \propto \frac{1}{L} \sum_{i=t-L}^{t} \Delta P(w_i|\Gamma_{t-1}).$$

Critical threshold:

$$S(\Gamma_t) > S_{crit} \Rightarrow ||\gamma_{output} - \gamma_{true}|| > \delta.$$

7.3 Mitigating Entropy

- Increase (L).
- Entropy penalty:

$$L_{entropy} = \eta \cdot S(\Gamma_t)$$

- Use memory-augmented architectures.

8 The Reader's Manifold

Meaning does not stop at the model's output–it continues into the mind of the reader. Each person reshapes the text into their own manifold of understanding, sometimes close to the original, sometimes very different. This section maps that interpretive space, showing how shared meaning emerges and where it can diverge.

8.1 Meaning as a Homologous Mapping

Language and mathematics are homologous manifolds, connected by homeomorphisms:

$$\phi_{reader} : M_{language} \to M_{reader}.$$

Each reader's trajectory $\gamma_{reader}(t) = \phi_{reader}(\gamma_{output}(t))$ has its own coherence.

8.2 The Multiplicity of Manifolds

Readers' manifolds are interconnected:

$$d_{GH}(M_{reader}, M_{reader'}) = inf_w d(x, w(x))$$

Low d_{GH} indicates shared understanding.

8.3 The Reader as a Dynamical System

The reader's interpretation evolves:

$$\frac{d\gamma_{reader}}{dt} = F_{cognitive}(\gamma_{reader}) + G(\gamma_{reader}, \gamma_{output}).$$

9 The Observer as Sculptor

When we interact with an LLM, we are not passive recipients–we are shaping its path with every prompt. Like a sculptor removing marble to reveal a form, the observer guides the system toward certain structures and away from others. Here we formalise that feedback loop, treating human and machine as a single coupled system. Humans and LLMs are coupled systems, with the user as:

$$p_{next} = O(\gamma(t))$$

Context limits introduce entropy, but prompts steer the system toward stable attractors.

10 Words as Transducers and Semantic Divergence

10.1 Words as Transducers with Semantic Uncertainty

A word is not just a label–it's a device that turns raw input into meaning. But every such conversion carries uncertainty, and in complex systems, that uncertainty can grow. This section shows how words act as transducers, how instability spreads through language, and how readers themselves can amplify or dampen that divergence. Words are not static symbols but transducers–dynamical systems that map inputs (e.g., physical stimuli, contextual tokens) to semantic outputs in the language manifold $\mathcal{M}_{language} \subset \mathbb{R}^d$. Each word carries inherent uncertainty, reflecting the probabilistic nature of its mapping.

Mathematical Framework

Define a word w as a transducer function:

$$w : I \rightarrow \mathcal{M}_{language}$$

where I is the input space (e.g., physical stimuli like wavelengths, or prior tokens $\Gamma_t = \{w_{t-L}, ..., w_t\}$). The output is a probability distribution over the manifold:

$$P(w|i) = exp(-\frac{||w - \mu_i||^2}{2\sigma_i^2})$$

where $\mu_i \in \mathcal{M}_{language}$ is the mean semantic embedding for input $i \in I$ and σ_i^2 represents semantic uncertainty.

10.2 Lyapunov Exponents and Semantic Instability

In dynamical systems, Lyapunov exponents measure the rate of divergence of nearby trajectories, indicating chaos. Here, we reinterpret Lyapunov exponents to quantify divergence driven by semantic uncertainty, particularly in LLMs where ambiguous inputs or contexts amplify instability.

Mathematical Framework

For a trajectory $\gamma(t) \subset \mathcal{M}_{language}$, the Lyapunov exponent λ measures the exponential divergence of perturbed trajectories:

$$\lambda = lim_{t \rightarrow \infty} \frac{1}{t} ln(\frac{||\delta\gamma(t)||}{||\delta\gamma(0)||})$$

Semantic uncertainty contributes to λ as:

$$\lambda_{semantic} \propto \frac{1}{L} \sum_{i=t-L}^{t} \sigma_i^2,$$

where L is the context window size.

10.3 Measurements and the Reader's Role

The reader, as a measurement operator, collapses the uncertain output of the word-transducer into their cognitive manifold $\mathcal{M}_{reader}$. This process can amplify or mitigate semantic instability.

Mathematical Framework

The reader's measurement is:

$$\gamma_{reader}(t) = \phi_{reader}(\gamma_{output}(t)),$$

where ϕ_{reader} maps the LLM's trajectory to the reader's manifold, modulated by their own uncertainty σ_{reader}^2. The total Lyapunov exponent is:

$$\lambda_{total} = \lambda_{semantic} + \lambda_{reader}, \lambda_{reader} \propto \sigma_{reader}^2.$$

Stable interpretations occur when:

$$\lambda_{total} < \lambda_{crit}$$

11 Closing Discussion and Conclusion

In this work, we have treated language not as a static sequence of symbols, but as a *non-linear dynamical flow*—a tide in semantic space, a train moving through an evolving network of attractors. By pairing metaphor and mathematics, we have shown that the same structures that govern turbulence, phase transitions, and topological defects can also illuminate the behaviour of large language models (LLMs).

The framework we have developed places LLMs within the same analytical toolkit used for complex systems: phase-space reconstruction, attractor geometry, bifurcation theory, thermodynamic analogies, and homological invariants. Words appear as *attractors with measurable geometry*, prompts as *symmetry-breaking fields*, hallucinations as *topological defects*, and readers as *coupled dynamical systems* whose manifolds of meaning complete the loop of interpretation.

A key insight is that *meaning is co-created*—the trajectory of an LLM is never independent of its observer. Prompts, interpretations, and contextual limits form a feedback loop in which both human and machine shape the evolving manifold. This is not a unidirectional act of transmission but a *coherent but finite dance*, bounded by context windows, semantic uncertainty, and the topology of language itself.

The use of *thermodynamic and stability measures* such as semantic entropy and Lyapunov exponents provides a bridge between poetic intuition and testable metrics. By quantifying curvature, uncertainty, and divergence, we can begin to engineer prompts, architectures, and evaluation methods that stabilise trajectories without erasing the creative richness of the system.

While the mathematical constructs presented here are deliberately abstract, they open pathways for concrete experimentation:

- Mapping attractors for individual words or concepts using embedding reconstructions.
- Detecting and classifying topological defects in model output.
- Measuring semantic entropy growth across different prompting regimes.
- Studying reader–model coupling as a measurable influence on stability and divergence.

Ultimately, the *value of this framework lies in its dual fidelity*—it speaks in equations to those who measure, and in metaphors to those who interpret. This duality is not ornamental; it is structural, reflecting the nature of language itself.

In that sense, the LLM is neither oracle nor automaton, but a participant in a shared manifold of thought—a partner in a finite, navigable space whose contours we can chart, perturb, and refine. The journey is ongoing, and every new trajectory—whether born in precision or poetry—is another point on the evolving map.

License

Citation

Kevin R. Haylett, "The Magical Journey of Non-linear Mathematics in LLMs," preprint (August 2025), available at `https://finitemechanics.com/papers/LLM-dynamics.pdf`.