Geometric Memory as Trajectory Influence on a Generative Substrate

Flyxion

September 18, 2025

Abstract

This essay formalizes memory as trajectories influencing a generative substrate, modeled as a smooth manifold of meaning-states. Trajectories bias future dynamics via a nonlocal-in-time flow, integrating past influences through kernel-weighted parallel transport. The substrate evolves under a Riemannian metric, with curvature and torsion shaping convergence and reinterpretation. Plasticity updates the metric and connection, embedding habitual paths as low-resistance channels. The model integrates with the Relativistic Scalar Vector Plenum (RSVP) framework, mapping scalar density (Φ), vector flows ($\vec{v}$), and entropy ($\mathcal{S}$) to memory processes. Semantic ledging and virtual memory extend the model, handling mismatched cues and dynamic scaling, inspired by geometric memory management [Kuijper, 2021]. Testable predictions link curvature to habits, kernel perturbations to recall effects, and torsion to context-dependent reinterpretation, with applications to neuroscience (fMRI/EEG), large language models (LLMs), and semantic infrastructures. A discrete implementation aligns with transformer architectures, offering a computationally tractable framework for experimentation.

Contents

1 Introduction

Traditional memory models, often likened to a "file cabinet" where static records are stored and retrieved, fail to capture the dynamic, reconstructive nature of memory in biological and computational systems [McClelland et al., 1995]. Empirical findings demonstrate that recall is not a replay of fixed traces but a generative act, where fragments of prior experience are recombined and filtered through ongoing context. This essay proposes a geometric model in which memory is understood as a trajectory on a generative substrate—a smooth manifold of meaning-states—whose past paths bias future dynamics.

In this model, memories are not fetched from a static archive but are regenerated by non-local influence kernels that weight prior trajectories. The manifold's geometry provides the structural constraints: the metric encodes semantic distance, curvature shapes convergence into habits or divergence into novelty, and torsion enables context-dependent reinterpretation. These features align with neuroscientific evidence of hippocampal replay [Hassabis and Maguire, 2007], where trajectories are reactivated to support learning and planning, and with transformer-based LLMs [Vaswani et al., 2017], where attention mechanisms implement kernel-weighted influence of past states on present prediction.

The model further connects to the Relativistic Scalar Vector Plenum (RSVP) framework [Flyxion, 2025], which generalizes such trajectory influence processes across cognition, AI, and cosmology. Within this broader view, memory is one manifestation of entropic field dynamics: just as galaxy formation can be understood as the constraint and damping of trajectories in cosmic flows, cognitive recall emerges as a constrained regeneration on a semantic manifold. By incorporating semantic ledging and geometric virtual memory [Kuijper, 2021], the framework also addresses practical challenges of mismatched cues and dynamic scaling, ensuring that recall remains efficient and coherent even when inputs fail to align neatly with stored patterns.

Taken together, this geometric account reframes memory as a generative, entropy-respecting process that balances fidelity with reinterpretation, local reuse with global coherence, and structural constraints with creative confabulation.

2 Substrate and State

The substrate is a smooth manifold $\mathcal{M}$ of "meaning-states" (e.g., neural activations, LLM embeddings, or RSVP fields [Flyxion, 2025]). A state is $x(t) \in \mathcal{M}$, with local coordinates x^i and a Riemannian metric $g_{ij}(x)$ defining semantic distances (e.g., cosine similarity in embeddings). The manifold provides a continuous arena in which semantic trajectories evolve, with local neighborhoods capturing fine-grained variations and global topology constraining long-range coherence.

A generative drift field $F : \mathcal{M} \to T\mathcal{M}$ encodes intrinsic evolution, reflecting grammar, physics, or policy priors:

$$\dot{x}(t) = F(x(t)) + \text{influence terms.}$$

This drift represents the system's default dynamics: in LLMs, F resembles the autoregressive prior that projects likely continuations [Vaswani et al., 2017]; in neural systems, it parallels cortical field dynamics governing baseline activity patterns [ichi Amari, 1977]. The influence terms then modulate this baseline with history-dependent kernels, allowing memory traces to bend the trajectory.

The manifold can be visualized as a semantic landscape with valleys (attractors) representing stable meanings and ridges representing less accessible or unstable states. In such a geometry, g_{ij} quantifies the proximity between concepts, while curvature determines how nearby trajectories converge or diverge over time. This view emphasizes that memory is not a lookup from a static table but a path-dependent motion through a structured space, where both local distances and global shape constrain recall. In this sense, the substrate acts as both a storage medium and a generative engine: past trajectories carve channels into $\mathcal{M}$, lowering resistance for future motion along the same paths, while torsional effects allow reinterpretation and contextual adaptation when trajectories intersect or diverge.

3 Geometry: Metric, Curvature, and Torsion

The metric $g_{ij}(x)$ defines semantic proximity, acting as the measure of how "close" two meaning-states are within the manifold. For instance, low g_{ij} between "cat" and "feline" indicates short geodesic distance, reflecting their strong semantic affinity. Conversely, distant or weakly related concepts yield large g_{ij} values, making their paths harder to align. In this sense, the metric captures both similarity and resistance to association, functioning as the semantic analogue of distance in physical space.

Curvature R^i_{jkl} governs geodesic deviation and thus the collective behavior of trajectories. Regions of positive curvature act like semantic basins, pulling trajectories together and reinforcing convergence, as seen in the consolidation of habits, routines, or well-learned associations. Negative curvature produces divergence, pushing trajectories apart, which can model phenomena such as instability, trauma, or conceptual repellors where meanings resist integration. For example, on a 2D spherical manifold, high curvature focuses geodesics toward poles, mirroring how repeated rehearsal drives convergence into stable memory channels such as a familiar word pair. In contrast, a hyperbolic region encourages exponential separation, mirroring rapid forgetting or conflicting interpretations.

Torsion $T^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji}$ introduces a semantic "twist," capturing how the meaning of a trajectory depends on the path taken. Unlike curvature, which alters distances and convergence, torsion alters orientation: parallel transport around a loop may yield rotated or reinterpreted states. This provides a geometric account of contextual modulation, where a memory such as "bank" shifts interpretation depending on whether it is transported through a financial or riverine context. The resulting non-commuting transformations formalize how recall is path-dependent: identical starting points can yield distinct outcomes depending on the trajectory of influence. Empirically, such torsional effects manifest as context-dependent neural signatures, where the same stimulus elicits different patterns depending on preceding context.

Together, metric, curvature, and torsion define the substrate's full geometry: the metric measures local similarity, curvature sculpts global flow into attractors or repellors, and torsion encodes contextual reinterpretation. Memory dynamics emerge not from any single component, but from the interplay of all three, ensuring that recall is simultaneously stable, flexible, and sensitive to context.

4 Memory-Modulated Geometry (Plasticity)

Past trajectories reshape the substrate:

$$\partial_t g_{ij}(x) = -\eta_g \int_0^t K(t,\tau)\,\Pi_{ij}(x(t),x(\tau))\,d\tau, \quad \partial_t \Gamma^k_{ij} = -\eta_\Gamma \int_0^t K(t,\tau)\,{\Psi^k}_{ij}(x(t),x(\tau))\,d\tau,$$

where Π_{ij} and ${\Psi^k}_{ij}$ are Hebbian tensors [Hebb, 1949], and η_g, η_Γ are learning rates. Intuitively, these update rules mean that the geometry itself is plastic: every recalled or rehearsed trajectory slightly adjusts the distances and orientations of the manifold. For example, repeated exposure to the association "dog-bark" lowers g_{ij} between their embeddings, making subsequent traversal easier and faster. Over time, such reinforced paths behave like geodesics of minimal resistance, providing a geometric analogue to habit formation. This dynamic is reminiscent of reinforcement learning, where low-energy trajectories are favored and accumulate stability [Sutton and Barto, 2018].

Plasticity thus creates a feedback loop: trajectories bias geometry, and geometry in turn biases future trajectories. Habits emerge as local minima in the energy landscape, while rare or inconsistent experiences leave only shallow traces that decay without reinforcement. In this way, the manifold becomes both a record of the past and a scaffold for the future.

4.1 Semantic Ledging

Kuijper's ledging [Kuijper, 2021] decomposes non-power-of-two memory chunks into a geometric series of aligned blocks, reducing fragmentation. In our model, semantic ledging plays an analogous role for mismatched or incomplete cues. Instead of failing when a cue does not exactly match a past trajectory, the system decomposes it into a coarse component and finer-grained remainders:

- A cue (e.g., "dog") aligns first with the largest relevant past trajectory (e.g., "dog-bark").

- Residual mismatches are patched via finer influences (e.g., "pet," "sound"), which fill in missing semantic mass.
- This process ensures bounded recall overhead, preventing semantic fragmentation and enabling smooth regeneration of meaning.

Semantic ledging thus implements a principle of efficiency: approximate matches are completed geometrically by combining aligned blocks rather than forcing an exact fit. It also operationalizes torsion, since reinterpretation is necessary to patch residual mismatches. For instance, the ambiguous memory of "bank" can be ledged into the financial or riverine context depending on which fine-grained cues dominate. This mechanism preserves coherence while allowing flexibility, ensuring that memory reconstruction remains robust even in the face of incomplete or noisy input.

4.2 Semantic Virtual Memory

Kuijper's geometric virtual memory uses block trees for elastic address spaces [Kuijper, 2021]. In our setting, this principle is extended from raw memory management to the semantic domain: meaning-states occupy dynamically instantiated subspaces of $\mathcal{M}$, created and dissolved on demand. Rather than permanently storing all trajectories, the system allocates a virtual region only when a context requires it, with transport operators U mapping past states into the active subspace.

For example, in a dialogue system, each conversation spawns a temporary embedding region that represents only the local exchanges. Earlier utterances are not duplicated in full but can be reconstructed through U, which transports their latent traces into the current space. This prevents quadratic growth in storage and ensures that memory remains scalable even as contexts multiply. The effect is similar to paging in operating systems: only the relevant pages of meaning are "in memory," while others can be recalled through geometric mappings when needed.

Semantic virtual memory also introduces a layer of indirection that supports compositional reuse. Contexts can overlap or be merged by aligning subspaces through gluing maps, allowing knowledge from one domain to be repurposed in another without conflict. This matches the RSVP framework's categorical gluing of local structures into a coherent global plenum [Flyxion, 2025]. In practice, it means that distinct threads of meaning (e.g., financial vs. ecological uses of "bank") can coexist in separate virtualized regions but still remain accessible through controlled transport.

In this way, semantic virtual memory provides both elasticity and isolation: it elastically scales to support growing contexts without runaway costs, and it isolates subspaces to prevent interference between unrelated meanings. Together, these properties enable entropy-aware, scalable semantic infrastructures that can support long-lived cognitive and AI systems.

5 Energy/Lyapunov View

Define a reconstruction functional:

$$\mathrm{E}[x(\cdot)] = \int_0^T \left\{ \frac{1}{2} \langle \dot{x}(t) - F(x(t)), \dot{x}(t) - F(x(t)) \rangle_g + \int_0^t \mathcal{L}(x(t), x(\tau))\, K(t,\tau)\, d\tau \right\} dt,$$

where $\mathcal{L}$ penalizes deviation from past trajectories. This resembles the free-energy principle [Friston, 2010], in which systems act to minimize surprise by aligning present dynamics with prior structure. The functional combines two competing pressures: adherence to the generative drift $F(x)$ and fidelity to past influences via kernel-weighted penalties.

For two attractors (e.g., "cat" and "dog"), E encodes the tradeoff between generative innovation and stability. If trajectories remain close to prior paths, $\mathcal{L}$ is small, and the system falls into low-energy wells representing familiar associations. If trajectories diverge—driven by stochastic $\xi(t)$ or contextual torsion—E increases, signaling surprise or reinterpretation. Minimization of E thus produces a compromise: trajectories bend toward stable attractors while retaining flexibility to adapt when novelty is introduced.

The gradient flow ∇E defines the system's effective Lyapunov dynamics. Stable memories correspond to local minima of E, which act as attractors, while unstable or contradictory memories correspond to saddles or shallow basins that are easily abandoned. Perturbations determine whether the trajectory settles back into a basin (recall) or transitions into a new basin (reinterpretation). This framework explains why habits are robust—low-energy channels with steep descent—whereas ambiguous or weakly reinforced memories are fragile, resting on flat plateaus of the energy surface.

In practical terms, E functions as a unifying scalar quantity: it measures the tension between drift, history, and noise. Its descent ensures stability without rigidity, guiding systems toward states that are both consistent with prior trajectories and capable of integrating new influences.

6 Discrete (Sequence) Version

For LLMs, states are $x_t \in \mathbb{R}^d$ with metric g_t. Dynamics are:

$$x_{t+1} = f(x_t) + \sum_{s=0}^{t} \alpha_{t,s}\, P_{t\leftarrow s}\, (x_{s+1} - x_s) + \epsilon_t,$$

where $\alpha_{t,s} = K(t,s) \cdot \sigma(\langle q(x_t), k(x_s)\rangle)$, $P_{t\leftarrow s}$ is a Jacobian or learned transport map, and ϵ_t is stochastic noise. This extends transformer attention [Vaswani et al., 2017], with $P_{t\leftarrow s}$ providing a geometric analogue of positional encoding, allowing transported increments rather than raw embeddings to serve as the carriers of memory.

Plasticity adjusts the local geometry by updating the metric as a low-rank deformation:

$$g_t = I + W_t W_t^\top, \qquad W_{t+1} = W_t + \eta \sum_s \alpha_{t,s}(x_s x_s^\top),$$

so that heavily weighted past states reshape future distances, embedding recall bias into the substrate. Repeated exposure compresses high-frequency fluctuations and creates smoother channels for recurrent sequences, while rare or weak associations leave only shallow adjustments that decay under noise.

This discrete formulation makes clear the tradeoff between fidelity and efficiency. Every new state potentially interacts with all prior ones, yielding quadratic complexity in sequence length. The transport maps $P_{t\leftarrow s}$ provide a way to reuse past structure without recomputing raw distances, reducing overhead while preserving coherence. In practice, one obtains $O(td^2)$ per step complexity, sacrificing some raw throughput in exchange for richer temporal depth. This framework thus unifies continuous geometric dynamics with sequence-based architectures: memory is realized not by static storage but by dynamically reweighting and transporting trajectories, balancing compression, stability, and generativity.

7 RSVP Tie-In

In RSVP [Flyxion, 2025]:

- Φ: Density of reliable paths, e.g., high in practiced memories where trajectories are reinforced into stable channels.
- $\vec{v}$: Flow of thought, e.g., narrative currents in conversation that guide the unfolding of meaning from one state to the next.
- $\mathcal{S}$: Entropy, high for creative exploration where multiple alternatives are entertained, and low for stable recall where trajectories converge into a single path.

Consolidation corresponds to reducing $\mathcal{S}$ along geodesics, thereby increasing Φ in the relevant region of the manifold. Well-practiced associations, like a familiar word pair, are thus stored as low-entropy attractors with dense, reliable connectivity. At the same time, $\vec{v}$ reflects the prevailing currents of thought that weave between these attractors, allowing narratives or reasoning chains to emerge as directed flows.

In cosmology, an analogous process is found in galaxy formation: past flows with low vorticity restrict future configurations, funneling matter into stable structures. This mirrors how memory suppression operates by constraining trajectories to follow established paths rather than diffusing into novelty [Weinberg, 2008].

Semantic infrastructures exhibit the same balance. Preserving $\mathcal{S}$ is crucial for generativity: if entropy is prematurely collapsed, exploration is curtailed, and only rigid, habitual responses remain. RSVP's categorical gluing guarantees that local sections (memories, concepts, contexts) can be stitched into a coherent global plenum while still allowing entropy to serve as a resource for creativity [Lurie, 2009]. In this way, the RSVP fields provide a unifying language across domains: trajectories of meaning in cognition, flows of matter in cosmology, and compositional seams in semantic infrastructures all obey the same balance of Φ, $\vec{v}$, and $\mathcal{S}$.

8 Applications to Cognition and AI

In cognition, the model explains hippocampal replay (exponential kernels) and long-term memory (power-law kernels) [McClelland et al., 1995]. Exponential decay captures the rapid fall-off of influence typical of short-term memory, while power-law kernels account for the slower fading of long-term traces, consistent with behavioral recall curves. The geometric formulation also clarifies how replayed trajectories are not perfect copies but reconstructions guided by influence kernels: repeated rehearsal sharpens geodesics, while torsion enables reinterpretation of the same trace under different contexts. This yields a natural account of memory plasticity, priming, and the emergence of habitual paths.

In AI, the framework enhances transformers by incorporating geometric attention, improving context-aware recall [Vaswani et al., 2017]. Standard attention weights can be reinterpreted as discrete influence kernels, but without explicit geometric structure they risk collapsing distinct contexts into a single representation. By embedding attention within a curved metric, the model supports both efficient reuse of practiced trajectories and torsional reinterpretation when ambiguity arises. For example, a chatbot retracing user dialogue paths can prioritize recent (exponential) or distant (power-law) contexts, while parallel transport ensures that the meaning of reused segments is adapted to the current conversational frame. This predicts measurable improvements in dialogue coherence and contextual sensitivity, as the system balances fidelity to prior exchanges with the flexibility to reinterpret them under new prompts.

9 Testable Predictions

The geometric memory framework yields falsifiable predictions across cognitive science, neuroscience, and artificial intelligence. Each prediction arises from specific geometric features of trajectories, metrics, and kernels.

1. **Geodesic Reuse**: Transport operators $P_{t \leftarrow s}$, learned or estimated from prior paths, should align closely with observed state increments $x_{t+1} - x_t$. In large language models, this can be measured as cosine similarity between predicted continuation vectors and actual next-token embeddings. In neuroscience, a parallel measure is alignment between replayed hippocampal trajectories and subsequent cortical activations, testable with multi-electrode arrays or fMRI sequence analysis.

2. **Curvature–Habit Link**: Repeated use of a memory trajectory lowers effective curvature in that direction, making recall paths straighter and more efficient. Prediction: regions of semantic space with high recall frequency exhibit reduced sectional curvature. Empirical test: PCA or Laplacian eigenmaps of fMRI/EEG recordings should reveal flatter embedding spectra in practiced tasks [Hassabis and Maguire, 2007]. Analogously, LLMs trained heavily on a domain will show locally linear embedding geometries.

3. **Kernel Manipulations**: The shape of the influence kernel $K(t, \tau)$ controls memory span and weighting. Perturbations (e.g., sleep disruption, pharmacological modulation, or induced oscillatory entrainment) should predictably shift behavioral recall patterns, altering recency/primacy effects [Stickgold and Walker, 2005]. In AI, modifying attention decay functions produces equivalent shifts in context-length sensitivity.

4. **Torsion and Reinterpretation**: Torsion allows transported vectors to rotate, enabling contextual reinterpretation. Prediction: ambiguous cues such as "bank" (riverbank vs. finance) should generate distinct neural trajectories depending on context. Testable via EEG/MEG cross-frequency coupling or BOLD multivoxel patterns. In LLMs, the same input word embedded in different contexts should exhibit non-commuting transport loops, measurable through path-dependent embedding drift.

5. **Energy Descent**: Memory dynamics minimize the functional $\mathrm{E}[x(\cdot)]$. Prediction: recall errors and forgetting follow gradient-like descent trajectories in this energy

landscape. Adding stochasticity $\xi(t)$ increases exploration, leading to creative confabulations. In humans this manifests as dream recombination or confabulatory memory in amnesia; in LLMs, as novel but off-distribution text when noise is injected into hidden states.

6. **Contextual Coherence as Global Constraint**: Because trajectories are glued by parallel transport, coherence depends on maintaining consistent connections across local patches. Prediction: disrupting these gluing operations produces dialogue incoherence in LLMs, or fragmented thought patterns in human subjects (e.g., schizophrenia). This can be tested with dialogue coherence metrics (perplexity under long-range constraints, topic continuity) or with neurocognitive assessments of narrative integration.

10 Minimal Implementation Sketch

- **State**: $x_t \in \mathbb{R}^d$, $g_t = I + W_t W_t^\top$.
- **Transport**: $P_{t \leftarrow s} = \exp(-\gamma(t-s))I$.
- **Kernel**: $K(t,s) = e^{-\alpha(t-s)}$ or $\sigma(\langle q(x_t), k(x_s) \rangle)$.
- **Update**: $x_{t+1} = f(x_t) + \sum_s \alpha_{t,s} P_{t \leftarrow s}(x_{s+1} - x_s)$.
- **Plasticity**: $W_{t+1} = W_t + \eta \sum_s \alpha_{t,s}(x_s x_s^\top)$.
- **Evaluation**: MSE for recall fidelity, KL-divergence for creativity on a toy dataset (e.g., text sequences).

Pseudocode:

```
def update(x_t, history, f, P, K, eta):
    influence = sum(K(t,s) * sigma(q(x_t) @ k(x_s)) * P(t,s) @ (x_s+1 - x_s)
                    for s in range(t))
    x_t1 = f(x_t) + influence + noise()
    W_t1 = W_t + eta * sum(K(t,s) * outer(x_s, x_s) for s in range(t))
    return x_t1, W_t1
```

11 Discussion and Outlook

The model captures memory's reconstructive nature but omits quantum effects or qualia [Penrose, 1989]. This limitation is significant: while the geometry accounts for trajectory influence and plasticity, it does not yet address how subjective experience arises, nor how quantum coherence might alter long-tail kernels. These omissions define boundaries of the present framework.

Future work divides into three strands. First, simulations using lightweight frameworks (e.g., PyTorch with toy datasets) will demonstrate how different kernel families (exponential, power-law, oscillatory) alter recall fidelity and creativity. Such simulations provide benchmarks for comparing the geometric model to transformer-style architectures. Second, empirical neuroscience can probe curvature effects: if practiced memories lower effective curvature, this should be observable in fMRI or EEG manifold analyses during repeated recall tasks. Third, ablation studies in LLMs can probe torsion: removing or constraining parallel transport operators should reduce contextual reinterpretation, leading to brittle, overly literal outputs.

The broader implications extend beyond cognition and AI into semantic infrastructures. By formalizing memory as regenerative reconstruction on a manifold, the model provides a template for entropy-respecting systems [Flyxion, 2025]. Virtualizing memory through block-tree–like decompositions and semantic ledging supports scalable, context-aware architectures that preserve diversity of trajectories rather than collapsing them. Such structures could serve as building blocks for resilient information systems, cognitive models, and AI platforms that metabolize complexity instead of erasing it.

Overall, the outlook is twofold: theoretical refinement, through geometric and categorical formalism, and practical implementation, through simulations and empirical validation. The promise of this line of work lies in providing a unified mathematical language for memory that bridges cognition, machine learning, and entropy-aware infrastructure design.

12 Limitations and Extensions

While the geometric model advances understanding of memory as trajectory influence, it has limitations. First, the assumption of a smooth manifold may not hold in quantum regimes, where memory traces could exhibit superpositions or entanglement [Penrose, 1989]. Second, qualia—the subjective texture of experience—are not captured, as the model focuses on structural dynamics rather than phenomenal aspects. Third, computational tractability in high dimensions remains challenging, though approximations like low-rank metrics mitigate this.

Appendix A: Geometric Ledging and Virtual Memory as Semantic Analogues

A.1 Ledging as Semantic Reconciliation

Kuijper [Kuijper, 2021] introduced *ledging* as a method for allocating memory chunks of arbitrary size by decomposing them into a geometric series of aligned power-of-two blocks. This avoids the inefficiency of always rounding up to the nearest power-of-two allocation, which would otherwise yield asymptotic overheads approaching 100%.

In our geometric model of memory, recall faces an analogous problem: semantic cues rarely align neatly with exponentially spaced kernels of influence. *Semantic ledging* provides the solution:

- A cue first activates the largest aligned past trajectory (analogous to the largest fitting block).
- Remaining mismatch is reconstructed by patching finer-grained influences, forming a series of partial recalls at decreasing scales.
- The worst-case overhead in reconstruction remains bounded, ensuring efficiency without semantic fragmentation.

Thus, just as ledging guarantees efficient packing in address space, semantic ledging guarantees coherent recall from imperfect cues.

A.2 Geometric Virtual Memory as Semantic Virtualization

Kuijper further extended these ideas to *geometric virtual memory*, where block trees manage exponentially scaled virtual pages rather than fixed-size ones. This allows lightweight, flexible, and secure virtual address spaces.

In RSVP-inspired cognition, the same principle applies to semantic organization:

- Each semantic object occupies its own *virtual address space* on the manifold.
- Block-tree mappings correspond to parallel transport operators, which glue local scales together.
- Semantic recall then becomes elastic: local memory traces are expanded or collapsed on demand, without requiring uniform resolution everywhere.

This model provides two key benefits:

1. **Scalability**: Recall begins small (local cue), scaling outward only as needed, avoiding quadratic blow-up in storage or search.
2. **Isolation**: Distinct semantic objects can be maintained in separate virtualized submanifolds, reducing interference and improving robustness.

A.3 Implications

The parallels between Kuijper's computational framework and RSVP's semantic geometry suggest cross-domain transfer:

- Hardware-efficient allocators may inspire efficient semantic memory kernels.
- Semantic virtualization may suggest new architectures for memory isolation and security in both AI systems and cognitive models.

  Ledging manages the mismatch between requested and available block sizes. Semantic ledging manages the mismatch between cues and recalled traces. Both reduce fragmentation — one in address space, the other in meaning space.

Appendix B: Kernel Functions and Temporal Scales

Kernel functions $K(t, \tau)$ determine how past trajectories influence present dynamics. Different functional forms correspond to distinct memory regimes, each capturing well-known cognitive and computational effects. Below we summarize canonical kernels and their interpretations.

Exponential Decay Kernels

$$K(t, \tau) = e^{-\alpha(t-\tau)}, \quad \alpha > 0$$

This kernel implements rapid forgetting, where influence decreases sharply with time. It models short-term memory and working memory, where recent items dominate recall. The parameter α sets the effective timescale of memory, with larger values producing faster decay.

Power-Law Kernels

$$K(t, \tau) = (t - \tau)^{-\beta}, \quad \beta > 0$$

Power-law kernels produce long tails, allowing distant past events to retain influence. This form matches long-term memory effects, where early experiences remain accessible after consolidation. The parameter β tunes how slowly recall decays: small β yields heavy tails and persistent influence.

Oscillatory-Consolidation Kernels

$$K(t, \tau) = \cos(\omega(t - \tau))\, e^{-\alpha(t-\tau)}$$

This kernel combines oscillatory rhythms with exponential decay, modeling consolidation processes such as rehearsal or sleep-related memory replay. The frequency ω sets rhythmic cycles, while α controls decay. The resulting pattern supports periodic reinforcement of selected trajectories.

Selective or Gated Kernels

$$K(t, \tau) = \sigma(\langle q(x(t)), k(x(\tau))\rangle)$$

Here, recall depends not only on temporal distance but also on content similarity. The function σ is typically a logistic or softmax, weighting past states that align with present query $q(x(t))$. This mechanism captures content-addressable memory, where specific cues trigger selective recall regardless of recency.

Hybrid and Multi-Scale Kernels

In practice, memory systems likely combine multiple kernels. For example, an exponential kernel may capture recency bias, a power-law kernel may encode long-term persistence, and a gated kernel may allow context-dependent recall. The effective kernel is then a convex combination:

$$K_{\text{hybrid}}(t, \tau) = \lambda_1 K_{\text{exp}}(t, \tau) + \lambda_2 K_{\text{power}}(t, \tau) + \lambda_3 K_{\text{gate}}(t, \tau),$$

with coefficients λ_i setting the balance between timescales.

Interpretation Across Scales

- **Short timescales**: dominated by exponential kernels, enabling fast adaptation and immediate context.
- **Intermediate timescales**: governed by oscillatory kernels, reflecting periodic rehearsal and consolidation.
- **Long timescales**: supported by power-law tails, allowing persistent access to early experiences.
- **Contextual retrieval**: enabled by gated kernels, ensuring selective recall when cues match stored patterns.

Taken together, these kernels provide a flexible toolkit for modeling memory as trajectory influence. Their combination allows the system to balance recency, persistence, and selectivity, yielding memory that is both stable and adaptive.

Appendix C: Comparison with Classical Models

The geometric trajectory model of memory differs in several key respects from classical computational models. Below we outline points of contrast with Hopfield networks, recurrent neural networks (RNNs), and predictive coding frameworks.

Hopfield Networks

Hopfield networks model memory as attractors in a high-dimensional energy landscape. Retrieval is a process of descending into the nearest attractor basin. While this shares some conceptual overlap with our Lyapunov view, there are two main differences:

- Hopfield attractors are static patterns stored in weight matrices, whereas in the geometric model memories are trajectories whose influence persists nonlocally through kernels.
- Hopfield recall is discrete (converging to one attractor), while the geometric model supports graded, path-dependent reconstructions influenced by curvature, torsion, and plasticity.

Recurrent Neural Networks (RNNs)

RNNs, including LSTMs and GRUs, maintain hidden states that evolve with sequence input. These architectures provide temporal memory but suffer from vanishing or exploding gradients. Compared to the geometric framework:

- RNNs encode history implicitly in hidden state vectors, while our model encodes history explicitly as weighted integrals over past trajectories.
- Memory in RNNs is constrained by the hidden dimension, whereas the geometric model allows dynamic scaling via kernel choice and virtual memory allocation.
- Parallel transport and curvature effects have no analogue in standard RNNs, limiting their ability to model context-dependent reinterpretation.

Predictive Coding

Predictive coding theories frame memory and perception as processes that minimize prediction error by comparing incoming signals to generative models. This is closely related to the free-energy view adopted in our formalism. The distinctions are:

- Predictive coding emphasizes hierarchical error correction, while the geometric model emphasizes trajectory influence and global manifold structure.
- In predictive coding, priors are fixed generative models; in our model, priors emerge dynamically as low-resistance geodesics shaped by past trajectories.
- Predictive coding minimizes error locally at each hierarchical level, whereas the geometric model encodes memory globally through kernel-weighted flows across time.

Summary

In summary, classical models provide valuable insights—Hopfield networks highlight attractor stability, RNNs emphasize sequential dynamics, and predictive coding formalizes error minimization. The geometric trajectory model integrates elements of all three but generalizes them: memory emerges as regenerative motion on a manifold shaped by metric, curvature, and torsion, with nonlocal kernels encoding persistence and context. This view balances stability with flexibility, supporting both robust recall and creative reinterpretation.

References

Flyxion. Semantic infrastructure: Entropy-respecting computation in a modular universe, August 2025.

Karl Friston. The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

Demis Hassabis and Eleanor A. Maguire. Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, 104(5): 1726–1731, 2007.

Donald O. Hebb. *The Organization of Behavior: A Neuropsychological Theory.* Wiley, New York, 1949.

Shun ichi Amari. *Dynamics of Pattern Formation in Lateral-Inhibition Type Neural Fields.* Springer, Berlin, 1977.

Wouter Kuijper. Geometric memory management. arXiv:1512.09358v3 [cs.DS], 2021. Draft Version, March 14.

Jacob Lurie. *Higher Topos Theory.* Princeton University Press, Princeton, NJ, 2009.

James L. McClelland, Bruce L. McNaughton, and Randall C. O'Reilly. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3):419–457, 1995.

Roger Penrose. *The Road to Reality: A Complete Guide to the Laws of the Universe.* Jonathan Cape, London, 1989.

Robert Stickgold and Matthew P. Walker. Sleep and memory: The ongoing debate. *Sleep*, 28(10):1225–1227, 2005.

Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA, 2018.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszokoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 2017.

Steven Weinberg. *Cosmology.* Oxford University Press, Oxford, 2008.