The RSVP Trilogy: A Modal-Thermodynamic Paradigm for Neural Representation Dynamics

Abstract

This document presents a unified framework for the RSVP (Representation, Semantic, Vector, Potential) field theory, developed across three interconnected papers: (1) *From Fractured Representations to Modal Coherence*, (2) *Diagnosing Representation Fracture via Scalar-Vector-Entropy Field Dynamics*, and (3) *Beyond Gradient Descent: A Modal-Thermodynamic Paradigm for AI*. The RSVP framework redefines neural learning as convergence in a semantic field space, defined by the triplet $\mathcal{F}(x,t) = \{\Phi(x,t), \vec{v}(x,t), \mathcal{S}(x,t)\}$. We provide theoretical foundations, diagnostic tools, and a visionary paradigm shift, challenging the limitations of gradient descent and offering a path toward modular, interpretable, and generalizable AI systems.

1 Introduction

Modern neural networks excel in performance but suffer from fractured, entangled representations (FER) that hinder interpretability and generalization. The RSVP field theory addresses this crisis by modeling representations as dynamic fields governed by scalar potential (Φ), vector flow ($\vec{v}$), and entropy ($\mathcal{S}$). This trilogy establishes: (1) a theoretical foundation for semantic coherence, (2) diagnostic metrics for representation quality, and (3) a new learning paradigm rooted in modal logic and thermodynamics. Together, these papers propose a unified path toward AI systems that learn with intrinsic meaning and stability.

2 Paper 1: From Fractured Representations to Modal Coherence

2.1 Objective

Develop RSVP field theory to explain representational quality, contrasting fractured entangled representations (FER) with unified factored representations (UFR) via modal and thermodynamic field structures.

2.2 Framework

The RSVP field is defined as:

$$\mathcal{F}(x,t) = \{\Phi(x,t), \vec{v}(x,t), \mathcal{S}(x,t)\}$$

where:

- $\Phi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$: Scalar semantic potential.
- $\vec{v} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$: Vector semantic flow.
- $\mathcal{S} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}_{\geq 0}$: Entropy field.

Field evolution follows:

$$\frac{\partial \Phi}{\partial t} + \nabla \cdot (\Phi \cdot \vec{v}) = -\delta \mathcal{S}$$

Modal coherence is achieved via the fixpoint operator $\Box A$, satisfying Löb's theorem:

$$\Box(\Box A \to A) \to \Box A$$

Fracture is quantified by torsion:

$$T_{\text{ent}} = \int_\Omega \|\nabla \times \vec{v}\|^2 \, dx$$

2.3 Key Insights

Fractured representations arise from Gödelian loops and non-converging field dynamics, while factored representations align with thermodynamic attractors, achieving modal stability.

3 Paper 2: Diagnosing Representation Fracture via Scalar-Vector-Entropy Field Dynamics

3.1 Objective

Provide a toolkit for measuring representational quality using RSVP-based metrics, enabling detection of fracture, entanglement, and modularity in deep networks.

3.2 Diagnostic Metrics

From activations $h_i \in \mathbb{R}^d$:

- Scalar field: $\Phi_i = \|h_i\|$
- Vector field: $\vec{v}_i = h_{i+1} - h_i$
- Entropy: $\mathcal{S}_i = \mathbb{H}(p(y|x, h_i))$ or $\text{Var}(h_i)$

Key metrics:

- Torsion: $T(x) = \|\nabla \times \vec{v}(x)\|, \quad T_{\text{avg}} = \frac{1}{|\Omega|} \int_\Omega T(x)^2 dx$
- Modal closure depth: $D_\Box = \min \left\{ t \mid \|\Phi^{(t+1)} - \Phi^{(t)}\| < \varepsilon \right\}$
- Redundancy: $\bar{R} = \frac{1}{n(n-1)} \sum_{i \neq j} I(h_i; h_j)$

3.3 Experimental Validation

Metrics are applied to MLPs, CNNs, and transformers, revealing patterns of fracture and coherence across tasks like modular classification and symbolic reasoning.

4 Paper 3: Beyond Gradient Descent: A Modal-Thermodynamic Paradigm for AI

4.1 Objective

Propose RSVP as a new learning paradigm, replacing gradient descent with field-based convergence toward modal closure and semantic stability.

4.2 Learning Rule

Learning is defined as achieving $\Box A$, with updates:

$$\Phi_{t+1} = \Phi_t - \eta \left(\nabla \mathcal{S}(\Phi_t) - \nabla \cdot (\Phi_t \cdot \vec{v}_t) \right)$$

The RSVP energy functional is:

$$\mathcal{L}_{\text{RSVP}} = \int_\Omega \left[\frac{1}{2} \|\nabla \Phi\|^2 + \alpha \cdot \|\nabla \times \vec{v}\|^2 + \beta \cdot \mathcal{S} \right] dx$$

Property	SGD	RSVP Descent
Objective	Minimize loss	Minimize semantic tension
Update Rule	$\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}$	$\Phi \leftarrow \Phi - \eta \nabla \mathcal{S}_{\text{eff}}$
Field Structure	Flat weight space	Recursive modal geometry
Interpretability	Post hoc	Intrinsic
Generalization	Empirical	Emerges from field stability

Table 1: SGD vs. RSVP Descent.

4.3 Comparison with SGD

4.4 Future Directions

RSVP enables modular, interpretable architectures, with applications in AGI alignment and continual learning.

5 Conclusion

The RSVP trilogy redefines neural learning as a process of semantic field convergence, offering theoretical rigor, practical diagnostics, and a visionary alternative to gradient descent. Future work will implement RSVP-inspired architectures and validate their generalization in real-world tasks.