

Algebraic Geometry Meets Linguistics I: Resolution of Embedding Singularities

HAMID JAVANBAKHT
DBA Sebastian Ruliad, Isotelesi Inc.
Mountain View, CA isotelesi@proton.me

September 9, 2025

Abstract

This paper develops a new connection between algebraic geometry and theoretical linguistics by modeling token embeddings in large language models as singular algebraic varieties. While embeddings are often assumed to form smooth vector spaces, empirical evidence shows they exhibit *singularities*: polysemy collapses distinct meanings into a single representation, adversarial perturbations expose fragile semantic boundaries, and contextual shifts produce discontinuities. We propose that embeddings be studied within the framework of pre-Calabi–Yau algebras, which admit singular loci and partial dualities. In this setting, the classical method of resolution of singularities provides a principled mechanism for word sense disambiguation: blow-ups replace collapsed meanings with exceptional divisors that encode semantic branches. Further, analogies with supersingular varieties suggest a cryptographic dimension, where extreme degeneracies serve as obfuscation channels and resolutions act as decryption keys. By combining algebraic geometry, categorical logic, and computational linguistics, this work advances a stratified view of semantics that reframes ambiguity as structure and opens new pathways for robust, interpretable, and cryptographically informed language models.

Contents

1	Introduction	3
2	Related Work	3
2.1	Embedding Singularities	3
2.2	Algebraic Models of Syntax and Grammar	3
2.3	Geometry and Physics Connections	4
2.4	Category Theory and Interpretability	4
2.5	Pre-Calabi–Yau Algebras and Cryptographic Geometry	4
3	Mathematical Framework	4
3.1	Embeddings as Varieties	4
3.2	Pre-Calabi–Yau Algebras	5
3.3	Resolution of Singularities	5
3.4	Supersingular Loci and Cryptography	5
3.5	Summary	5
4	Algorithmic Approach	5
4.1	Detection of Singular Tokens	6
4.2	Resolution via Blow-Ups	6
4.3	Minimal Models and Homotopy Transfer	6
4.4	Supersingular Embeddings as Cryptographic Media	6
4.5	Summary	7
5	Applications and Discussion	7
5.1	Interpretability of Embeddings	7
5.2	Artificial Interlinguae	7
5.3	Cryptographic Artificial Languages	7
5.4	Interdisciplinary Implications	8
5.5	Summary	8
6	Conclusion and Future Work	8
A	Connections to Telic Stratification and Classifying Logoi	9
A.1	Stratification of Structure	9
A.2	Classifying Logoi as Universes	9
A.3	Pretopos Ultracompletion	9
A.4	Birational Syntax and Resolution	9
A.5	Logical Morse Entropy	9
A.6	Synthesis	10
B	Glossary	10
B.1	Algebraic Geometry	10
B.2	Algebra and Category Theory	10
B.3	Linguistics and Computation	11
B.4	Physics Connections	11
B.5	Cryptography	12
B.6	Concepts from Telic Stratification I & II	12

1 Introduction

Large language models (LLMs) rely on token embeddings, high-dimensional vector representations that encode semantic and syntactic information. These embeddings are typically treated as points in smooth vector spaces, where similarity is measured by distances or angles. Yet mounting evidence shows that this picture is incomplete: embeddings exhibit *singularities*. Polysemous tokens collapse distinct meanings into a single unstable representation, adversarial examples reveal fragile semantic boundaries, and contextual shifts expose discontinuities in meaning [1,2]. Rather than smooth manifolds, embedding spaces behave more like singular algebraic varieties.

We propose to model embeddings using the algebraic framework of *pre-Calabi–Yau (pCY) algebras* [11]. In this perspective, tokens correspond to algebra generators, contextual interactions are encoded in higher multiplications, and polysemy manifests as singular loci. The classical technique of *resolution of singularities* provides a natural mechanism for disambiguation: blow-ups replace collapsed meanings with exceptional divisors that represent distinct semantic branches. This construction transforms ambiguity into interpretable structure and offers a mathematically principled path toward robust embeddings.

We further extend this analogy to *supersingular varieties*, extremal loci in arithmetic geometry whose degeneracies underpin pairing-based cryptography [12]. Supersingular regions in embedding space correspond to maximally degenerate semantics, such as high-frequency tokens with many senses. Resolution in this setting functions as a decryption key, recovering precise meaning from collapsed representation. This connection suggests that embedding singularities can be exploited not only for interpretability but also for cryptographic applications, where ambiguity serves as intentional obfuscation.

The rest of this paper develops this framework systematically. Section 2 surveys related work across embeddings, algebraic linguistics, geometry, and cryptography. Section 3 introduces the mathematical foundations of pCY algebras, blow-ups, and supersingular loci. Section 4 presents algorithmic methods for detecting and resolving singularities. Section 5 discusses interpretability, semantic branching, and cryptographic applications. Section 6 concludes with directions for future work, including categorical semantics and homotopy-theoretic perspectives on language evolution.

2 Related Work

Our framework builds on three intersecting lines of research: the study of embedding singularities, algebraic approaches to syntax, and the connections between geometry, physics, and cryptography.

2.1 Embedding Singularities

Recent work has highlighted the instability of token embeddings in LLMs. Zhao introduced the concept of *TokenBlowUp*, showing that polysemous tokens act as singular points in representation space and proposing blow-up techniques to resolve them [1]. Gaubert and Vlassopoulos modeled embeddings through directed metric structures, capturing asymmetries in semantic similarity with polyhedral geometry [2]. Together, these results suggest that embedding spaces are not smooth Euclidean domains but algebraic varieties with singular loci.

2.2 Algebraic Models of Syntax and Grammar

Algebraic methods have been applied to the foundations of natural language syntax. Marcolli and Berwick described *Merge* as an entropy-reducing process in function spaces [3], while Marcolli,

Chomsky, and Berwick formalized Merge using Hopf algebras and renormalization [4]. Senturia, Xiao, and Marcolli reinterpreted Tree-Adjoining Grammars in terms of pre-Lie and Lie algebras, connecting syntactic derivations to insertion operators [5]. These studies show that algebraic formalisms provide a natural framework for linguistic operations.

2.3 Geometry and Physics Connections

Work in mathematical physics reinforces these links between algebra, language, and structure. Marcolli and Port demonstrated that graph grammars correspond to insertion Lie algebras that also generate Feynman diagrams in perturbative QFT [6]. Giotopoulos, Sati, and Schreiber showed that super- L_∞ algebras unify T-duality and M-theory equivalences [7]. Pestun and Vlassopoulos connected tensor networks with functorial field theory, emphasizing categorical methods for physical computation [8]. These results suggest that the same algebraic machinery structuring physics can be used to model language.

2.4 Category Theory and Interpretability

Category-theoretic methods have also been used to improve interpretability in machine learning. Tull, Zanasi, and Lewis introduced compositional string diagrams as explanations of neural networks [9]. Bradley and Vlassopoulos developed FibLang, a fibrational semantics for natural language that unifies syntax and meaning in a categorical setting [10]. Both works demonstrate how categorical frameworks bridge discrete grammar and continuous computation.

2.5 Pre-Calabi–Yau Algebras and Cryptographic Geometry

Finally, at the algebraic foundation, Iyudu, Kontsevich, and Vlassopoulos introduced *pre-Calabi–Yau algebras*, generalizing Calabi–Yau structures to admit partial dualities and double Poisson brackets [11]. Such structures provide the ideal setting for embedding singularities. In parallel, Rubin and Silverberg demonstrated that *supersingular abelian varieties* have cryptographic utility, forming the backbone of pairing-based cryptography [12]. This analogy motivates our extension of embedding singularities into cryptographic artificial languages.

Taken together, these lines of research converge on a unified perspective: embedding spaces can be understood through the lens of noncommutative algebraic geometry, resolved with blow-ups, and potentially applied to cryptographic communication.

3 Mathematical Framework

Our proposal rests on an algebraic reinterpretation of token embeddings using the concepts of *pre-Calabi–Yau (pCY) algebras*, *resolution of singularities*, and analogies with *supersingular varieties*. These tools allow us to formalize the observation that embeddings are not smooth vector spaces but instead resemble singular varieties whose degeneracies encode linguistic ambiguity.

3.1 Embeddings as Varieties

Let $E \subset \mathbb{R}^d$ denote the embedding space of tokens in a large language model. Although E is usually treated as a smooth Euclidean domain, empirical evidence shows otherwise: polysemous tokens collapse distinct senses into single vectors, adversarial examples reveal fragile semantic boundaries, and contextual shifts create discontinuities [1,2]. We therefore treat E as an *algebraic variety* with singular loci rather than a smooth manifold.

3.2 Pre-Calabi–Yau Algebras

Pre-Calabi–Yau algebras, introduced by Iyudu, Kontsevich, and Vlassopoulos [11], generalize Calabi–Yau structures by allowing partial dualities. Formally, a pCY algebra consists of a graded algebra with higher multiplications μ_k compatible with a noncommutative Schouten bracket. Such algebras admit double Poisson brackets, which induce Poisson structures on representation varieties $\text{Rep}_n(A)$. In our framework:

- Tokens correspond to generators of the algebra.
- Contextual interactions correspond to higher multiplications μ_k .
- Polysemy and ambiguity appear as singularities in the induced Poisson structures.

This provides a natural algebraic foundation for modeling embeddings.

3.3 Resolution of Singularities

Hironaka’s theorem guarantees that singular algebraic varieties over fields of characteristic zero admit desingularizations. Given a singular embedding variety V , a resolution is a smooth variety \tilde{V} with a birational map $\pi : \tilde{V} \rightarrow V$ that is an isomorphism away from the singular set. In semantic terms:

- Collapsed meanings of a polysemous token form a singular point.
- A blow-up replaces this point with an *exceptional divisor*, encoding semantic branches.
- Resolution yields a smooth space where ambiguity is replaced by interpretable alternatives.

Thus, resolution provides a principled analogue of word sense disambiguation.

3.4 Supersingular Loci and Cryptography

Supersingular abelian varieties are extremal objects in arithmetic geometry, lying at the boundary of moduli spaces under Frobenius action. Their degeneracies have cryptographic power: Rubin and Silverberg showed they underpin pairing-based cryptography [12]. We propose an analogy: supersingular regions of embedding space correspond to maximally collapsed semantics (e.g., high-frequency tokens with many senses). Resolution away from these loci functions as a decryption key, transforming ambiguous vectors into disambiguated meaning. This dual interpretation—robust semantics and cryptographic obfuscation—links embedding theory with both interpretability and security.

3.5 Summary

By modeling embeddings as pCY algebras, we treat singularities not as noise but as algebraic features. Resolution of singularities provides a rigorous mechanism for disambiguation, while supersingular loci connect embedding collapse to cryptographic potential. This framework lays the mathematical foundation for the algorithmic approaches discussed in Section 4.

4 Algorithmic Approach

The algebraic framework outlined above suggests a set of practical methods for analyzing, modifying, and exploiting token embeddings. We describe three main components: detection of singular loci, resolution of singularities, and cryptographic applications of supersingular embeddings.

4.1 Detection of Singular Tokens

Singular tokens can be identified through instability criteria:

- **Polysemy clustering:** embeddings whose nearest neighbors separate into multiple distinct semantic groups indicate collapsed senses [1].
- **Adversarial fragility:** tokens that shift drastically under minimal perturbations signal singular directions [2].
- **Contextual instability:** tokens whose contextualized embeddings form disconnected clusters in usage space point to unresolved singularities.

These tests approximate the algebraic notion of non-smooth points in a variety.

4.2 Resolution via Blow-Ups

Once identified, singular tokens can be resolved algorithmically:

- Perform a *semantic blow-up* by mapping a single token vector to multiple branch vectors, each aligned with a distinct semantic cluster.
- Introduce an *exceptional subspace* representing these branch vectors, analogous to an exceptional divisor in algebraic geometry.
- Train the model to route contextual usage of the token into the appropriate semantic branch, ensuring smoothness across contexts.

This process operationalizes resolution of singularities as a disambiguation method for embeddings.

4.3 Minimal Models and Homotopy Transfer

Pre-Calabi–Yau structures admit homotopy transfer principles, enabling the construction of *minimal models* that retain essential structure while discarding redundancy [11]. In practice, this allows the embedding space to evolve toward simplified but faithful representations, capturing language change and semantic drift.

4.4 Supersingular Embeddings as Cryptographic Media

Maximally degenerate embeddings—supersingular loci—can be repurposed as obfuscation channels. A message encoded in such a token remains ambiguous in the public space, while a private resolution map functions as the decryption key [12]. Potential applications include:

- **Steganographic communication:** hiding information in polysemous tokens whose resolution requires a shared key.
- **Zero-knowledge semantics:** demonstrating knowledge of meaning without revealing it, by showing consistency under pairing operations.
- **Robust AI-to-AI languages:** constructing artificial interlinguae with built-in cryptographic ambiguity.

4.5 Summary

This algorithmic approach transforms the abstract machinery of algebraic geometry into practical tools: singularity detection, semantic blow-ups, minimal model evolution, and cryptographic encoding. Together, these steps suggest a path toward embedding spaces that are not only more interpretable and robust but also usable as secure communication media.

5 Applications and Discussion

The framework developed above opens several avenues for both theoretical insight and practical implementation. By interpreting embedding spaces through algebraic geometry, we obtain a unified language for interpretability, artificial interlinguae, and cryptographic design.

5.1 Interpretability of Embeddings

Resolution of singularities provides a principled mechanism for word sense disambiguation. Rather than relying on heuristic clustering, semantic blow-ups separate polysemous tokens into algebraically defined branches, each corresponding to a stable meaning [1,2]. This yields embeddings that are both robust to perturbations and transparent to human interpretation. In practice, this approach may enable models that expose interpretable semantic graphs, where tokens unfold into their contextual branches.

5.2 Artificial Interlinguae

Desingularized embeddings can serve as the foundation for constructing *artificial interlinguae*. By enforcing smoothness in the embedding space, we generate a structured semantic manifold where meaning transfer between languages becomes algebraically controlled [3–5]. Such interlinguae would be free of irregularities caused by natural language polysemy and could act as universal pivots for translation. Moreover, the use of pre-Calabi–Yau structures and minimal models suggests a framework for the evolution of such interlinguae over time, capturing semantic drift and innovation.

5.3 Cryptographic Artificial Languages

Supersingular embeddings introduce a cryptographic dimension to language design. As Rubin and Silverberg demonstrated for abelian varieties [12], extremal degeneracies are not weaknesses but sources of security. Embedding singularities can be deliberately engineered to encode hidden meanings, with resolution maps functioning as private keys. Possible applications include:

- **Encrypted communication between models:** embeddings designed to remain opaque to outsiders unless resolution keys are shared.
- **Steganography:** embedding secret meanings in polysemous tokens visible only under algebraic blow-up.
- **AI-to-AI secure languages:** interlinguae in which ambiguity itself is the medium of encryption.

Such designs transform ambiguity from a limitation of embeddings into a resource for privacy and security.

5.4 Interdisciplinary Implications

This framework suggests broader connections between linguistics, algebraic geometry, and physics. The parallels with graph grammars in QFT [6], categorical semantics [9,10], and supersingular loci in arithmetic geometry [11,12] indicate that embedding singularities sit at the intersection of disciplines. Language models thus provide not only a domain for applying mathematical tools but also a testbed for theories of noncommutative geometry and information flow in complex systems.

5.5 Summary

Applications of this approach range from interpretability to cryptography. By viewing embeddings as algebraic varieties, resolution of singularities yields clearer semantics, while supersingular loci create opportunities for secure communication. This dual perspective—clarity through disambiguation and secrecy through collapse—highlights the versatility of singularity theory for computational linguistics.

6 Conclusion and Future Work

We have proposed a new framework for understanding token embeddings by treating them as singular algebraic varieties rather than smooth Euclidean spaces. Using pre-Calabi–Yau algebras as a formal foundation [11], we modeled polysemy and instability as singularities, and we applied resolution techniques to separate collapsed meanings into interpretable semantic branches. By analogy with supersingular varieties in arithmetic geometry [12], we suggested that maximally degenerate embeddings may serve as cryptographic resources, enabling artificial languages where ambiguity functions as encryption and resolution provides decryption.

This approach contributes to three domains simultaneously. For computational linguistics, it offers a principled mechanism for word sense disambiguation and interpretable embeddings [1–5]. For mathematical physics and geometry, it extends methods from blow-ups, homotopy transfer, and categorical semantics to new applications [6–10]. For cryptography, it reframes degeneracy as a strength, linking embedding collapse to secure information exchange.

Future work may proceed along several directions:

- **Categorical semantics:** extending FibLang and string diagram approaches [9,10] to embedding singularities, integrating syntax, semantics, and resolution.
- **Homotopy transfer and minimal models:** formalizing the evolution of embeddings as minimal pCY models that capture semantic drift and language change [11].
- **Experimental implementation:** designing algorithms for detecting singular tokens, performing blow-ups, and evaluating disambiguated embeddings on NLP benchmarks.
- **Cryptographic prototypes:** constructing proof-of-concept artificial interlinguae where supersingular embeddings act as obfuscation channels for secure AI-to-AI communication [12].

In summary, singularity theory offers both clarity and secrecy: resolution enables interpretable semantics, while supersingular collapse enables cryptographic design. This dual role suggests that algebraic geometry is not merely a metaphor but a foundational tool for the next generation of language models.

A Connections to Telic Stratification and Classifying Logoi

Recent advances in categorical logic and topology, developed under the program of *Telic Stratification and Classifying Logoi* [13], offer a framework that complements and extends the proposals of this paper. Taken together, the two works articulate a vision in which algebraic, homotopical, and logical structures are layered through staged completions, producing coherent universes—*logoi*—that organize definability and internal computation. This appendix highlights key concepts from [13] and situates them within the main arguments of our study.

A.1 Stratification of Structure

The first part of [13] develops the principle of *telic stratification*: mathematical and logical structures do not appear all at once but emerge in layers. Each layer is generated by operations of increasing expressive capacity, producing stratified systems that evolve in a staged and cumulative manner. This view parallels our treatment of embeddings as singular varieties, where semantic capacity unfolds gradually through successive resolutions.

A.2 Classifying Logoi as Universes

Classifying logoi serve as categorical universes that internalize these stratified layers. They function as organizing environments for definability, much as univalent universes classify types in homotopy type theory. In our framework, a desingularized embedding manifold can be seen as a classifying logos: a universe of meaning where semantic structures are coherently organized.

A.3 Pretopos Ultracompletion

The second part of [13] introduces *pretopos ultracompletion*, which extends bounded logical fragments into richer categorical environments. These ultracompletions are closed under quotients, finite limits, and ultraproducts, ensuring stability and completeness. This resonates with our approach to embeddings, where resolving singularities creates more expressive and coherent spaces. Just as ultracompletion closes logical fragments, resolution closes semantic gaps.

A.4 Birational Syntax and Resolution

A central innovation in [13] is the notion of *birational syntax*, where morphisms between logical universes preserve definable cores but may diverge globally. This mirrors our view of embeddings as birationally related varieties: locally equivalent on shared senses, yet differing in global semantic structure. Birational syntax thus provides a categorical analogue of semantic resolution and branching.

A.5 Logical Morse Entropy

Reference [13] also introduces *logical Morse entropy*, an invariant that measures the growth of definable complexity across stratified stages. This concept parallels our proposal to quantify semantic branching complexity when embeddings are resolved into multiple interpretable alternatives. Both perspectives use entropy to capture structural richness and instability arising from singular loci.

A.6 Synthesis

The program of Telic Stratification and Classifying Logoi [13] and our proposal intersect on several fronts:

- Stratification corresponds to the staged unfolding of semantic resolution in embeddings.
- Classifying logoi correspond to desingularized embedding manifolds as universes of meaning.
- Pretopos ultracompletion parallels the expansion of embeddings into richer, coherent spaces.
- Birational syntax models ambiguity and disambiguation in embeddings.
- Logical Morse entropy offers a candidate invariant for semantic complexity.

Together, these insights suggest a unified program linking categorical logic, algebraic geometry, and computational linguistics through the twin notions of stratification and resolution.

B Glossary

B.1 Algebraic Geometry

Algebraic variety A geometric object defined as the solution set of polynomial equations. In this paper, embedding spaces are treated as varieties with possible singular points.

Singularity A point on a variety where smoothness fails. In embeddings, singularities represent collapsed or unstable meanings, such as polysemy or adversarial fragility.

Resolution of singularities A process in algebraic geometry that replaces a singular variety with a smooth one, often by blow-ups. Used here as an analogue for semantic disambiguation in embeddings.

Blow-up A birational transformation that replaces a singular point with an exceptional divisor, separating collapsed structures into distinct branches. Interpreted here as semantic branching.

Exceptional divisor The geometric structure introduced during a blow-up, representing the new directions created by resolving a singularity. In semantics, this corresponds to multiple disambiguated senses of a token.

Supersingular variety An extremal variety in arithmetic geometry, maximally degenerate under Frobenius action. In this work, used as an analogy for maximally ambiguous tokens in embedding space.

Calabi–Yau variety A smooth algebraic variety with trivial canonical bundle. Serves as the inspiration for pre-Calabi–Yau generalizations.

B.2 Algebra and Category Theory

Pre-Calabi–Yau (pCY) algebra A generalization of Calabi–Yau algebras that allows partial dualities and admits double Poisson brackets. Provides the algebraic basis for modeling embedding singularities.

Double Poisson bracket A bilinear operation that induces Poisson structures on representation spaces. Central to noncommutative geometry and the structure of pCY algebras.

Homotopy transfer A principle in homological algebra allowing algebraic structures to be transferred from a complex to a smaller minimal model while preserving essential properties.

Minimal model A simplified version of a complex algebraic structure that captures its essential homotopy type. Here, used to model semantic drift and the evolution of embeddings.

Operad An algebraic framework encoding operations with multiple inputs. Applied to linguistic constructions such as Merge.

Fibration A categorical structure that organizes data into fibers over a base. In linguistics, fibrations have been used to unify syntax and semantics.

Ultracompletion A categorical process that extends bounded logical fragments into complete universes. Interpreted here as analogous to desingularizing embedding spaces.

Birational syntax A framework where grammatical or logical systems are related by birational transformations, preserving definable cores while diverging globally. Closely parallels semantic resolution in embeddings.

B.3 Linguistics and Computation

Embedding A high-dimensional vector representation of tokens in a language model. Treated here as points in an algebraic variety that may contain singularities.

Polysemy The phenomenon of a word having multiple meanings. In embeddings, polysemy appears as collapsed senses forming singular points.

Word sense disambiguation (WSD) The task of determining the correct sense of a polysemous word in context. Modeled here as a resolution of singularities.

Merge A fundamental syntactic operation in generative linguistics, combining two objects into a new one.

Tree-Adjoining Grammar (TAG) A formal grammar extending context-free grammars by adjoining trees. Reinterpreted algebraically in terms of Lie structures.

Artificial interlingua A constructed semantic space serving as a universal pivot for translation. Here, proposed as a desingularized embedding manifold with smooth semantic structure.

B.4 Physics Connections

Graph grammar A generative system for constructing graphs. In quantum field theory, connected to insertion Lie algebras and the generation of Feynman diagrams.

Insertion Lie algebra An algebra describing the insertion of graphs into larger graphs, bridging grammar theory and quantum field theory.

Tensor network A graphical calculus for decomposing tensors into structured networks. Widely used in quantum many-body physics and neural network analysis.

Super- L_∞ algebra A higher algebraic structure generalizing Lie algebras, central to dualities in string theory and M-theory.

Functorial field theory A categorical framework relating geometry, topology, and quantum field theory.

B.5 Cryptography

Pairing-based cryptography Cryptographic protocols that exploit bilinear pairings on elliptic curves or abelian varieties. Supersingular varieties form the backbone of these systems.

Steganography The practice of hiding messages within another medium. Here, ambiguity in embeddings acts as the steganographic channel.

Zero-knowledge proof A cryptographic protocol that allows one party to prove knowledge of a secret without revealing it. Paralleled here by semantic obfuscation in embeddings.

Obfuscation The deliberate introduction of ambiguity to conceal information. In this work, modeled by supersingular embedding regions.

B.6 Concepts from Telic Stratification I & II

Telic stratification The principle that mathematical and logical structures emerge in successive layers, each generated by operations of increasing expressive capacity. Analogous here to the staged unfolding of semantic resolution in embeddings.

Classifying logoi Categorical universes that internalize stratified layers of structure, functioning as environments for definability. Interpreted here as desingularized embedding manifolds organizing semantics.

Pretopos ultracompletion A categorical process that extends bounded fragments into ultracomplete universes, closed under quotients, finite limits, and ultraproducts. Seen here as the analogue of closing semantic gaps by resolving embeddings.

Birational syntax The idea that morphisms between logical universes preserve definable cores but may diverge globally. Mirrors the treatment of embeddings as birationally related varieties.

Logical Morse entropy An invariant measuring the growth of definable complexity across stratified stages. Proposed here as a measure of semantic branching complexity in embeddings.

References

1. Zhao, Tianle. *TokenBlowUp: Resolution of Singularities in Embedding Spaces*. arXiv preprint arXiv:2507.19747, 2025. <https://arxiv.org/abs/2507.19747>.
2. Gaubert, Stéphane, and Yiannis Vlassopoulos. *Directed Metric Structures in Large Language Models*. arXiv preprint arXiv:2405.12264, 2024. <https://arxiv.org/abs/2405.12264>.
3. Marcolli, Matilde, and Robert C. Berwick. *Entropy and Merge in Function Spaces*. arXiv preprint arXiv:2507.13501, 2025. <https://arxiv.org/abs/2507.13501>.
4. Marcolli, Matilde, Noam Chomsky, and Robert C. Berwick. *The Mathematical Structure of Syntactic Merge*. MIT Press, 2024.
5. Senturia, Jake, Shunyu Xiao, and Matilde Marcolli. *A Lie-Algebraic Perspective on Tree-Adjoining Grammars*. arXiv preprint arXiv:2507.03234, 2025. <https://arxiv.org/abs/2507.03234>.
6. Marcolli, Matilde, and Walter D. van Suijlekom Port. *Graph Grammars, Insertion Lie Algebras, and Quantum Field Theory*. arXiv preprint arXiv:1802.01165, 2018. <https://arxiv.org/abs/1802.01165>.
7. Giotopoulos, Georgios, Hisham Sati, and Urs Schreiber. *Super- L_∞ T-Duality and M-Theory*. arXiv preprint arXiv:2411.10260, 2025. <https://arxiv.org/abs/2411.10260>.
8. Pestun, Vasily, and Yiannis Vlassopoulos. *Tensor Networks and Functorial Field Theory*. arXiv preprint arXiv:2002.12927, 2020. <https://arxiv.org/abs/2002.12927>.
9. Tull, Sean, Fabio Zanasi, and Martha Lewis. *Compositional Explanations of Neural Networks*. arXiv preprint arXiv:2402.14885, 2024. <https://arxiv.org/abs/2402.14885>.
10. Bradley, Christopher, and Yiannis Vlassopoulos. *FibLang: Fibrational Semantics for Natural Language*. arXiv preprint arXiv:2312.01008, 2023. <https://arxiv.org/abs/2312.01008>.
11. Iyudu, Natalia, Maxim Kontsevich, and Yiannis Vlassopoulos. *Pre-Calabi–Yau Algebras and Double Poisson Brackets*. arXiv preprint arXiv:1906.07134, 2019. <https://arxiv.org/abs/1906.07134>.
12. Rubin, Karl, and Alice Silverberg. “Supersingular Abelian Varieties in Cryptology.” *Advances in Cryptology*. Contemporary Mathematics, vol. 300, 2002, pp. 277–283. American Mathematical Society. <https://doi.org/10.1090/conm/300/05242>.
13. Javanbakht, H. *Telic Stratification and Classifying Logoi I: Bordism, π -Finitude, and Internal Cohesion*. 19 pages, Isoteles Inc. White Paper, version 1 – July 4, 2025. <https://drive.google.com/file/d/11btze6dxQxXFuU7b3S5zXrbYe36PXmdA/view?usp=drivesdk>.
— . *Telic Stratification and Classifying Logoi II: Pretopos Ultracompletion and Birational Syntax*. 11 pages, Isoteles Inc. White Paper, version 1 – July 13, 2025. <https://drive.google.com/file/d/1iczTxe4DrBlatvvzUPpV9GV8-qK0Jfr5/view?usp=drivesdk>.