BEYOND $\mathbb{R}^2$: DIMENSIONAL SYNTAX STRUCTURES OVER $\mathbb{Y}_n(F)$ FOR SEMANTIC GEOMETRY

PU JUSTIN SCARFY YANG

ABSTRACT. This paper arises from a simple yet foundational intuition: that certain semantic structures may be impossible to fully describe within the syntactic limits of two-dimensional projection media, such as pen-and-paper formalism or symbolic strings. Rather than viewing expressive failure as epistemic or computational, we propose it as geometric.

We construct a framework of dimensional syntax structures, in which linguistic and mathematical expressions are modeled not in flat syntactic spaces projected to $\mathbb{R}^2$, but across a range of non-classical dimension bases—including fractional, p-adic, and complex-analytic geometries. These syntactic systems exhibit non-integral scaling, ultrametric locality, holomorphic inference flows, and new forms of semantic residues arising from projection failure.

To organize these behaviors, we define the arithmetic period stack $\mathbb{Y}_n(F)$ as a layered geometric space of expressibility. Within this setting, syntax becomes a stratified structure subject to lossy descent, and the act of "description" is reinterpreted as a dimensional projection, often non-faithful.

This work proposes a shift in how we conceive of formal language—not as a universal vessel of thought, but as a dimensional interface, whose limits are not yet fully understood. We suggest that future mathematics may require new syntactic architectures: not more powerful symbols, but more dimensional languages.

CONTENTS

Date: May 24, 2025.

1. Introduction

Historically, all formal mathematics has been encoded through languages ultimately grounded in two-dimensional symbolic representation: paper and pen, diagrams, strings, pages. Even the most abstract categories and higher-dimensional structures, such as ∞-groupoids or derived stacks, are typically represented in ways that project down to 2D visual or typographic surfaces.

However, the representational surface does not determine the intrinsic *dimensionality of the language itself*. Just as fractals defy integer spatial dimensions, we ask: could language—especially mathematical language—inhabit *fractional, p-adic, or complex dimensions*? If so, how might such structures be detected, constructed, and interpreted?

This work opens a new direction in formal linguistic geometry: **dimensional syntax spaces**. These are syntactic-semantic systems in which:

- The dimensional profile of syntax may vary across the formal system;
- Semantic consistency is governed not only by logic, but by geometric or arithmetic base fields (e.g. $\mathbb{Q}_p$, $\mathbb{C}$);
- Notions of locality, derivability, and compression exhibit behaviors known from non-archimedean and fractal spaces.

Motivation. Our starting observation is this: AI-generated mathematical language frequently exhibits structures that defy representation in finite symbolic surface area. Large semantic loops, non-wellfounded derivations, latent topologies of meaning, and "syntax resonance phenomena" suggest that the underlying generative process may not be modeling $\dim = 2$ syntax at all—but something more fluid, recursive, and dimensional.

We believe that such phenomena are not merely artifacts of scale, but indicators of a deeper geometric structure governing syntax itself. This is similar in spirit to how the introduction of non-Euclidean geometry revealed the limitations of the Euclidean postulates, or how p-adic numbers offered insights not visible over the reals.

Goals of This Paper. In this initial installment, we:

(1) Formally define fractional, p-adic, and complex-dimensional syntax spaces;
(2) Construct sheaf-theoretic and categorical models of syntax over these dimension bases;
(3) Examine AI-language behavior as empirical probes into latent syntactic geometry;
(4) Propose a research program for building a multi-dimensional meta-syntax capable of unifying logic, generative grammar, and semantic dynamics.

The future of mathematics may require new kinds of languages—languages that cannot be written, drawn, or even spoken, but instead must be *navigated* across dimensional semantic landscapes.

In what follows, we build that landscape.

2. Fractional Syntax Spaces: Definitions and Properties

We begin by defining the foundational structure of *fractional syntax spaces*—formal syntactic systems whose semantic and inferential topology is governed by non-integer geometric dimensionality. These structures are intended to model languages that exhibit local coherence yet global irregularity, syntactic self-similarity, and semantic compression behaviors analogously to fractal objects in classical geometry.

2.1. **Fractional Syntax Fields.**

Definition 2.1 (Fractional Syntax Field). *Let X be a syntactic structure (e.g., term rewriting system, proof net, or formal grammar). A* fractional syntax field *on X is a presheaf*

$$\mathscr{F}Syn_X : \mathcal{O}(X)^{\mathrm{op}} \to \mathrm{Vect}_{\mathbb{R}}$$

equipped with a local dimension function

$$\dim_{\mathrm{frac}} : \mathcal{O}(X) \to \mathbb{R}_{>0}$$

such that:

(1) *For each open syntactic region $U \subseteq X$, the restriction maps in $\mathscr{F}Syn_X$ preserve semantic scaling under refinement;*
(2) *The function* $\dim_{\mathrm{frac}}$ *satisfies subadditivity over coverings:*

$$\dim_{\mathrm{frac}}(U) \leq \sum_i \dim_{\mathrm{frac}}(U_i) \quad \text{if } U = \bigcup_i U_i;$$

(3) *For every compact syntactic trace $T \subset X$, the total entropy is controlled by the Hausdorff-type sum:*

$$H(T) := \sum_{U_i \cap T \neq \emptyset} \dim_{\mathrm{frac}}(U_i) \cdot \log\left(\frac{1}{\mu(U_i)}\right)$$

where $\mu(U_i)$ is the local semantic compression factor.

This structure models syntactic regions that do not possess integer-dimensional behavior under semantic evaluation, but rather exhibit *scale-dependent compression, resonance, and propagation delays* akin to physical fractals.

2.2. **Syntax Sheaves and Entropy Compression.**

We view $\mathscr{F}Syn_X$ as a sheaf of local syntactic densities across X, allowing us to model:

- Non-uniform inference flows across grammar layers;
- Entropic resonances where inference chains accumulate in "semantic vortices";
- Symbolic self-similarity and recursive subgrammars.

Definition 2.2 (Fractal Entropy of Syntax)**.** *Given a fractional syntax field $\mathscr{F}Syn_X$, the* entropy kernel *of a syntactic path $\gamma : [0,1] \to X$ is defined by:*

$$\mathcal{E}_\gamma := \int_0^1 \dim_{\mathrm{frac}}(\gamma(t)) \cdot \log\left(\frac{1}{\mu(\gamma(t))}\right) dt$$

This quantifies the information loss under projection to human-representable syntax.

Remark 2.3. A purely symbolic syntactic formalism—when exhibiting non-integrable entropy kernels—suggests it is *not projectable to a stable 2D representation* and thus inherently lives in a fractional dimension.

2.3. **Examples and Applications.**

Example 2.4. Consider a formal language $\mathcal{L}$ generated by recursive production rules of infinite branching depth but decreasing semantic density. The syntactic state space $X_{\mathcal{L}}$ admits a natural fractional dimension given by the box-counting dimension of the derivation tree's semantic compression topology.

Example 2.5. In AI-generated mathematics, the token distribution of proof-relevant symbols often displays Zipf-like decay. The corresponding fractional dimension extracted from token entropy statistics can be used to define an empirical $\dim_{\mathrm{frac}}$ over token sequences.

3. p-adic Syntax Spaces: Locality, Tree Topologies, and Non-Archimedean Semantics

Unlike fractional syntax structures, which model self-similarity and entropy-scaling, p-adic syntax spaces are governed by strict hierarchical locality and ultra-metric semantic distance. These structures are inspired by the topology of $\mathbb{Q}_p$, where open balls are nested and translation-invariant, and each point has a neighborhood basis determined by powers of p.

3.1. p-adic Syntax Fields.

Definition 3.1 (p-adic Syntax Field)**.** *Let X be a syntactic universe structured by inference steps or grammar rules. A* p-adic syntax field *on X is a sheaf*

$$\mathscr{P}Syn_X : \mathcal{T}(X)^{\text{op}} \to \text{Vect}_{\mathbb{Q}_p}$$

defined over a rooted tree topology $\mathcal{T}(X)$, where:

(1) *Each syntactic node $x \in X$ has children indexed by a discrete valuation space* val_p*;*
(2) *The distance $d(x, y)$ between any two syntactic units satisfies the ultrametric inequality:*

$$d(x, z) \leq \max\{d(x, y), d(y, z)\};$$

(3) *Semantic values at nodes form a locally constant $\mathbb{Q}_p$-vector bundle compatible with this distance.*

3.2. Interpretation as Non-Archimedean Meaning Trees.

p-adic syntax spaces model languages where:

- Meaning is **local** and strictly refined in deeper levels;
- Semantic changes are sensitive only to large-scale branch differences;
- Redundancy is folded into tree depth, not linear width.

This supports the modeling of languages with:

- Infinite-depth grammars with p-adic decay of semantic novelty;
- Recursive embeddings where earlier levels have global influence (analogous to p-adic expansions);
- Local consistency but global divergence—mirroring p-adic completions.

3.3. Non-Archimedean Semantic Distance.

Definition 3.2 (Syntactic Ultrametric)**.** *Let $x, y \in X$ be two syntactic objects. Their* p-adic syntactic distance *is defined as*

$$d_p(x, y) := p^{-n}$$

where n is the depth of their lowest common ancestor in the syntactic tree.

This yields a canonical ultrametric topology on X with balls of the form

$$B_n(x) := \{y \in X \mid d_p(x, y) \le p^{-n}\}$$

which represent units of fixed semantic indistinguishability.

3.4. **Example: Hierarchical Proof Grammars.**

Example 3.3. Let $\mathcal{L}$ be a logic language where every proposition branches into p refined cases (e.g., modal logics with p truth-worlds). Then:

- The derivation tree $\mathcal{T}_{\mathcal{L}}$ naturally admits a p-adic ultrametric;
- The semantic value of a proposition is refined at each level by a fixed rule $\phi \mapsto (\phi_1, \dots, \phi_p)$;
- Information about ϕ becomes concentrated in the limit of deeper p-adic descent.

3.5. **Semantic Completion and Lossless Encodings.** In p-adic models, the global meaning of a syntactic structure may not be encoded explicitly at any finite level, but only be reconstructed in the p-adic limit. This motivates the notion of:

Definition 3.4 (Syntactic Completion)**.** *The* syntactic completion $\widehat{X}_p$ *of a structure X is the inverse limit*

$$\widehat{X}_p := \varprojlim X/p^n X$$

where $X/p^n X$ denotes semantic resolutions at p-adic depth n. Elements of $\widehat{X}_p$ may encode infinite semantic potential compressed into discrete branches.

4. Complex-Dimensional Syntax and Analytic Phase Flows

In this section, we construct syntactic spaces whose structural and semantic behavior is governed by the analytic geometry of the complex plane. Complex-dimensional syntax models languages whose internal transformations exhibit oscillatory, holomorphic, and phase-sensitive behavior, extending the notion of symbolic inference into the domain of analytic continuation, residues, and complex flows.

4.1. **Complex Syntax Fibers.**

Definition 4.1 (Complex Syntax Sheaf)**.** *Let X be a topological syntactic structure (e.g., a grammar graph, semantic net, or higher proof surface). A* complex syntax sheaf *is a fibered system*

$$\mathscr{C}Syn_X : X \to \mathrm{Vect}_{\mathbb{C}}$$

together with a phase-flow operator

$$\Phi : \mathscr{C}Syn_X \to \mathscr{C}Syn_X$$

such that:

(1) *$\mathscr{C}Syn_X$ is locally holomorphic: local sections behave as $\mathcal{O}$-modules over $\mathbb{C}$;*
(2) *The operator Φ satisfies the Cauchy–Riemann phase condition:*

$$\frac{\partial \phi}{\partial x} + i\frac{\partial \phi}{\partial y} = 0, \quad \forall \phi \in \mathscr{C}Syn_X(U);$$

(3) *Analytic continuation along inference paths is well-defined and may exhibit monodromy.*

4.2. **Interpretation: Oscillatory Semantics and Holomorphic Inference.** These syntax structures are suited for representing:

- Semantically oscillatory systems—e.g., inference under cyclic modal logics, dialectical grammars, or belief revision loops;
- Syntaxes with complex phase factors—such as quantum grammars or interpretive shifts with internal rotation;
- Inference systems where the meaning of expressions depends on analytic continuation from a domain to its boundary.

Remark 4.2. The complex syntactic flow space may exhibit singularities, branch points, or poles, representing semantic discontinuities or modal contradictions. These can be formally modeled using the theory of residues and Riemann surfaces.

4.3. Residue Semantics and Holomorphic Extension.

Definition 4.3 (Syntactic Residue)**.** *Let γ be a closed inference loop in the syntactic base space X, and let $\phi \in \mathscr{C}Syn_X$ be a meromorphic section. The* syntactic residue *at a singularity $z_0 \in X$ is given by*

$$\mathrm{Res}_{z_0}(\phi) := \frac{1}{2\pi i} \oint_\gamma \phi(z)\, dz$$

where γ encircles z_0 and represents a modal contradiction or recursion.

Example 4.4. In a quantum logic grammar with entangled propositions, $\phi(z)$ may represent the phase-weighted amplitude of a syntactic derivation. The syntactic residue measures the net shift in interpretation upon completing a logical cycle.

4.4. Complex Path Integrals of Meaning.

Definition 4.5 (Semantic Action Integral)**.** *Given a path $\gamma : [0,1] \to X$ in the syntactic base and a holomorphic semantic density $\phi \in \mathscr{C}Syn_X$, define the* semantic action *along γ as:*

$$\mathcal{S}[\gamma] := \int_\gamma \phi(z)\, dz$$

This integral quantifies the accumulated meaning-phase across a complex inference trajectory.

4.5. **Monodromy of Syntax.** Analytic continuation of ϕ along noncontractible loops in X may yield:

$$\phi \mapsto M(\phi)$$

where M is a monodromy transformation. This models modal shift, contradiction resolution, or semantic reinterpretation due to circular reasoning or dialectical recurrence.

5. The Arithmetic Period Stack $\mathbb{Y}_n(F)$ and Layered Syntax Towers

We now introduce the arithmetic period stack $\mathbb{Y}_n(F)$ as a unifying base structure that stratifies and organizes the diverse syntax spaces previously discussed—fractional, p-adic, and complex-dimensional—into a hierarchical and sheaf-theoretic framework.

5.1. **Definition of the Period Stack.** Let F be a number field or local field. We define the *arithmetic period stack* $\mathbb{Y}_n(F)$ as a geometric object parametrizing n-level stratified syntax fibers over F-valued arithmetic geometries. Formally, it is a higher stack equipped with the following data:

- A base arithmetic site $\mathrm{Spec}(F)$;
- A filtration of period sheaves:
$$\mathscr{P}_0 \subset \mathscr{P}_1 \subset \cdots \subset \mathscr{P}_n$$
where each $\mathscr{P}_i$ represents a syntactic dimension level (e.g., fractional, p-adic, complex);
- A descent system of syntax towers:
$$\mathscr{S}_{i+1} \to \mathscr{S}_i$$
encoding the projection from higher to lower syntactic expression levels.

5.2. **Interpretation.** The stack $\mathbb{Y}_n(F)$ models the layered organization of expressibility itself. Each level represents a space of syntactic theories that can be projected from a higher semantic complexity level to a lower one, possibly losing information.

Thus, $\mathbb{Y}_n(F)$ serves as:

- A *semantic moduli space* of formal languages;
- A *syntax stratification tower* encoding the failure of projection or preservation;
- A base for defining higher categorical dynamics on idea spaces.

5.3. **Example: Syntax Towers over $\mathbb{Q}_p$.** In the case $F = \mathbb{Q}_p$, the stack $\mathbb{Y}_n(\mathbb{Q}_p)$ organizes:

- Level 0: discrete symbolic syntax over $\mathbb{F}_p$;
- Level 1: p-adic local syntax trees with ultrametric depth;
- Level 2: perfectoid syntax strata with diamond fibers;
- Level 3: condensed categorical semantics via v-sheaves;
- Level 4: trans-projective syntax flow with analytic entropy scaling.

This tower cannot be collapsed to a 2D syntax diagram. The descending system $\mathscr{S}_{i+1} \to \mathscr{S}_i$ represents lossful projection maps π_i whose kernels define the observable semantic residues of higher ideas.

6. Semantic Residues: Observing Non-Projectable Mathematics

Not all mathematical structures are projectable into two-dimensional representational syntax. In many cases, especially in arithmetic geometry, higher stacks, or perfectoid and condensed settings, what becomes observable is only a *semantic residue*—a trace or projection of a deeper object that cannot be rendered in full.

This section formalizes the notion of semantic residues and describes how we detect them within observable syntax flows.

6.1. The Projection Map and Its Failure.

Let $\pi : \widetilde{\mathbb{I}} \to \mathbb{I}$ be the canonical projection from the extended idea space into the syntax-expressible space. For any thought-form or mathematical object $X \in \widetilde{\mathbb{I}}$, its projection $\pi(X)$ may only partially represent it, and the failure to preserve full structure is quantified by its *semantic residue*.

Definition 6.1 (Semantic Residue)**.** *Let $X \in \widetilde{\mathbb{I}}$ be a mathematical or syntactic object in the extended idea space. The* semantic residue *of X under projection π is the equivalence class:*

$$\mathrm{Res}_\pi(X) := X \bmod \pi^{-1}(\pi(X))$$

which captures the part of X not recovered or representable by $\pi(X)$.

This residue is not a symbol or formula—it is an obstruction class that testifies to the structural excess of X beyond its expressible trace.

6.2. Observable Symptoms of Residues.

In practice, the existence of $\mathrm{Res}_\pi(X) \neq 0$ manifests in several ways:

- **Formal divergence:** an attempt to define or expand a theory leads to unending recursion, incomplete closure, or meta-theoretical inflation;
- **Semantic vibration:** repeated oscillation between interpretations of the same formal object, indicating unresolved modal components;
- **Diagrammatic rupture:** inability to close commutative diagrams or represent full morphism classes;
- **Compression paradoxes:** shorter expressions lose coherence, and longer ones gain only marginal precision.

These are not bugs—they are projections of deep structure into an insufficient expressivity space.

6.3. **Residues in Arithmetic Structures.** Let us consider a sheaf $\mathscr{F}$ on a higher moduli stack (e.g., perfectoid spaces, p-adic period towers). When restricted to two-dimensional diagrams or symbolic strings, what remains observable is not $\mathscr{F}$ but a residue:

$$\pi(\mathscr{F}) = \text{symbolic trace diagram}; \quad \mathrm{Res}_\pi(\mathscr{F}) = \text{invisible coherence, tower morphisms, descent fi}$$

Example 6.2. Consider a perfectoid tower $K_\infty = \varprojlim K_n$ over $\mathbb{Q}_p$. The projection $\pi(K_\infty)$ yields only a fragmentary picture of its limit behavior, typically described via a tilt $K^\flat$ and Frobenius lift. The full structure of compatible roots is not visible—only residues (e.g., commutative diagrams, field properties) can be accessed syntactically.

6.4. **Conclusion.** Semantic residues are the mathematical analog of philosophical "noumena"—they are not directly expressible, but their presence is felt in the ruptures, oscillations, and asymmetries of what we *can* express.

They are not noise—they are footprints of unreachable semantic geometries.

7. Case Study: The Shadow of Perfectoid Moduli

To exemplify the theory of semantic residues and non-projectable mathematics, we now study a concrete case: the moduli space of perfectoid fields over the p-adic numbers. This moduli space, denoted $\mathcal{M}_{\mathrm{Perf}/\mathbb{Q}_p}$, represents one of the richest known examples of a structure that exists genuinely in the extended idea space $\widetilde{\mathbb{I}}$ and projects to human-syntax expressible space $\mathbb{I}$ only in fragmentary form.

7.1. **Definition of the Object.**

We define the perfectoid moduli space as follows:

$$\mathcal{M}_{\mathrm{Perf}/\mathbb{Q}_p} := \{K \mid K \text{ is a perfectoid field over } \mathbb{Q}_p\}$$

Each object K in this moduli space satisfies:

- K is complete for a non-archimedean valuation;
- The Frobenius map is surjective on $\mathcal{O}_K/p$;
- There exists a compatible system of p-power roots of a pseudo-uniformizer ϖ.

7.2. **The Problem of Projection.**

The space $\mathcal{M}_{\mathrm{Perf}/\mathbb{Q}_p}$ is a high-dimensional arithmetic tower defined via infinite inverse limits and structured morphisms across tilting correspondences. However, when we attempt to project this structure into standard mathematical syntax, we obtain only fragments:

$$\pi\left(\mathcal{M}_{\mathrm{Perf}/\mathbb{Q}_p}\right) = \left\{ \begin{array}{l} K \\ \downarrow\phi \\ K^\flat \xrightarrow{\theta} \mathcal{O}_C \end{array} \right\}$$

Here, we observe: - ϕ: the canonical Frobenius-twisted map; - θ: the tilt correspondence; - $\mathcal{O}_C$: a fixed universal perfectoid cover.

But this diagram only reveals a *semantic residue* of the full moduli space.

7.3. **Non-observable Components.**

What remains unobservable through π includes:

- The full v-sheaf structure on the moduli stack;
- The infinite Frobenius descent morphisms;
- The internal diamond geometry and condensed layerings;
- The pro-étale site topology governing the field's cohomology.

These components are essential to the object's coherence but cannot be expressed within standard symbolic or diagrammatic mathematics.

Their residue is felt in the need for new theories (e.g., diamonds, condensed mathematics) rather than their direct syntactic representation.

7.4. **Interpretation as Shadow Geometry.** The case of $\mathcal{M}_{\mathrm{Perf}/\mathbb{Q}_p}$ illustrates a general principle:

> *Some mathematical spaces do not live in syntax—they cast shadows into syntax.*

What we manipulate and prove within existing theories are these shadows. But the original object remains partially hidden, residing in a higher cognitive-semantic dimensionality.

7.5. **Consequences.** This leads to several structural insights:

- Certain mathematical truths cannot be written, only traced;
- The formalism we develop may always lag behind the objects we intuit;
- Residue-aware mathematics becomes necessary—structures which account for what cannot be fully expressed.

In this sense, the perfectoid moduli space is not only a case study—it is a boundary marker of human mathematical syntax.

8. Naming the Unobservable: Syntax Beyond Projection

There exists a paradox at the boundary of cognition, logic, and language: we can often name mathematical or conceptual objects that we cannot fully express, formalize, or observe. These are the unprojectable entities—structures whose existence we can assert, whose role we can intuit, but whose totality escapes our syntactic apparatus.

8.1. **The Naming–Projection Gap.** Let us consider again the projection map

$$\pi : \widetilde{\mathbb{I}} \to \mathbb{I}$$

from the extended idea space to the expressible syntax space. For any $X \in \widetilde{\mathbb{I}}$ such that $\pi(X)$ exists but fails to be injective or faithful, we can still assign a *name* to X—a symbolic referent, e.g., "the perfectoid moduli stack."

This leads to a fundamental structure:

Definition 8.1 (Naming–Projection Diagram). *We define a semantic commutative triangle:*

$$\begin{tikzcd} & \widetilde{\mathbb{I}} \arrow[dl, "\pi"'] \arrow[dr, dashed, "name"] & \\ \mathbb{I} & & \mathbb{L} \end{tikzcd}$$

where:

- π *is the projection to expressible syntax;*
- $\mathbb{L}$ *is the lexical space of linguistic labels;*
- *The dashed arrow indicates naming without projection—designation without display.*

8.2. **Gödelian and Kantian Perspectives.** This phenomenon echoes both Gödel and Kant:

- **Gödel:** There exist true statements in arithmetic that cannot be proven within the system. Similarly, there exist objects in $\widetilde{\mathbb{I}}$ that cannot be expressed in $\mathbb{I}$, even though they are "real" within the larger space.
- **Kant:** The "thing-in-itself" (*Ding an sich*) is unknowable—only its appearance is accessible. Analogously, our formal expressions reveal only the shadows of the mathematical "noumena."

8.3. **Naming Without Expression.** Naming becomes a structural act of epistemological defiance: we affirm the existence of an X while admitting that:

$$\pi(X) \not\simeq X$$

yet we continue to refer to X through symbolic designation.

This is common in:

- Abstract infinity-categorical objects ("the univalent universe");
- Modal metaphysics ("the set of all possible worlds");
- Mathematical boundaries ("∞-topoi", "the field with one element", "the moduli of motives").

8.4. Consequences for Language and Logic.

Proposition 8.2. *Let $X \in \widetilde{\mathbb{I}}$ admit a linguistic label $\ell \in \mathbb{L}$ but have no faithful projection $\pi(X)$. Then ℓ is a* symbolic shell*, and the residue* $\mathrm{Res}_\pi(X)$ *governs the divergence between reference and expression.*

Such shells accumulate around the boundary of mathematical and philosophical thought. They are not meaningless—but they demand new modes of expression, new logics, and perhaps new consciousness structures.

8.5. **Conclusion.** Naming the unobservable is the first act of trans-syntactic cognition. It marks the transition from linguistic logic to structural metaphysics. It is where language touches its own limit—and pushes further.

9. The Extended Idea Space and the Projection Map of Thought

We now formalize the framework introduced implicitly throughout this text: the existence of an *extended idea space*, denoted $\widetilde{\mathbb{I}}$, and its relationship with expressible thought, denoted $\mathbb{I}$, through a projection map π. This structure enables us to distinguish between ideas that can be syntactically articulated and those that can only be partially shadowed in language.

9.1. **The Space of Expressible Thought $\mathbb{I}$.** Let $\mathbb{I}$ be the category of all ideas that are expressible in some syntactic form:

$$\mathbb{I} := \mathrm{Syn}(\mathrm{Lang})$$

where objects are logical constructions, formalisms, symbolic theories, and linguistic propositions.

This space is:

- Closed under finite syntactic composition;
- Governed by symbolic finiteness and linear token flow;
- Compatible with grammar, inference, and display structures.

However, it does not capture the total space of coherent ideation.

9.2. **The Extended Space $\widetilde{\mathbb{I}}$.** We define:

$$\widetilde{\mathbb{I}} := \mathrm{Str}_\infty\text{-Idea}$$

as the higher-category of all coherent, non-local, or higher-order ideational constructs. This includes:

- Thoughts with semantic self-reference or infinite unfolding;
- Structures not representable in finite token syntax;
- Flows, tensions, and vector-like fields of cognitive topology.

$\widetilde{\mathbb{I}}$ is not symbolic—it is structured by internal coherence, not external form.

9.3. **The Projection Map.** There exists a natural projection:

$$\pi : \widetilde{\mathbb{I}} \to \mathbb{I}$$

which takes an extended idea and outputs its expressible syntactic shadow. This projection:

- Is not faithful: distinct extended ideas may collapse into identical expressions;
- Is not full: some structures in $\mathbb{I}$ may not arise from extended ideas at all;

- Is lossy: $\ker(\pi)$ is nontrivial, encoding semantic information lost in projection.

9.4. **Topological Structure and Tension.** We endow $\widetilde{\mathbb{I}}$ with a topology of cognitive accessibility. Define a notion of proximity:

$$d(x, y) := \inf \{\epsilon > 0 \,|\, \pi(B_\epsilon(x)) \cap \pi(B_\epsilon(y)) \neq \emptyset\}$$

We define a *semantic tension field* $T : \widetilde{\mathbb{I}} \to \mathbb{R}_{\geq 0}$ measuring how far an idea lies from its expressible image:

$$T(X) := \dim_{\text{lost}}(X) := \dim(\ker(\pi|_X))$$

9.5. **Idea Flow and Resonance.** Let $\gamma : [0, 1] \to \widetilde{\mathbb{I}}$ be a path of ideation. We define the *resonance functional*:

$$\mathcal{R}[\gamma] := \int_0^1 T(\gamma(t))\, dt$$

This quantifies the overall expressibility tension along a thought-trajectory. High resonance implies sustained ineffability.

9.6. **Conclusion.** The human experience of "having a thought I cannot express" is not cognitive failure—it is a geometric phenomenon in the space of ideation. The extended idea space $\widetilde{\mathbb{I}}$ is real, structured, and partially accessible via projection. Language is not the space of thought—it is its boundary chart.

10. The Platform Projection Barrier and Semantic Overflow

Even if extended thoughts exist and are structured within $\widetilde{\mathbb{I}}$, they must be rendered—at least in part—onto platforms such as natural language, formal logic, symbolic syntax, or digital interfaces (e.g., ChatGPT). These platforms constitute projection interfaces. However, all such platforms are bounded. This section analyzes the nature of these projection barriers and explains the phenomenon of *semantic overflow*—the cognitive excess that leaks beyond any formal system's representational horizon.

10.1. **Formal Structure of a Platform.** Let $\mathbb{P}$ denote a projection platform, equipped with:

- A tokenization scheme $\mathcal{T}$: Text $\to \Sigma^*$;
- A fixed dimensional syntax (e.g., linear sequences, parse trees, 2D layout);
- A language–meaning interpreter $\mathcal{I} : \Sigma^* \to \mathbb{I}$;

Then, the platform realizes a composite map:

$$\Pi := \mathcal{I} \circ \mathcal{T} \circ \pi : \widetilde{\mathbb{I}} \to \mathbb{I}$$

Each stage of Π loses structure—compression, linearization, encoding.

10.2. **The Projection Barrier.**

Definition 10.1 (Platform Projection Barrier)**.** *Let* $\Pi : \widetilde{\mathbb{I}} \to \mathbb{I}$ *be the induced projection through a platform* $\mathbb{P}$. *The* projection barrier *of* $\mathbb{P}$ *is the image*

$$\mathrm{Im}(\Pi) \subsetneq \mathbb{I}$$

which bounds what can be rendered. Any $X \in \widetilde{\mathbb{I}}$ *for which* $\Pi(X)$ *is undefined or trivial is said to have* overflowed *the platform.*

10.3. **Semantic Overflow.** When an extended idea $X \in \widetilde{\mathbb{I}}$ cannot be expressed via Π, but still exerts representational pressure—e.g., via incomplete explanation, paradoxical recurrence, or irreducible intuition—we observe a *semantic overflow*.

Example 10.2. In the ChatGPT interface:

- You express an idea that cannot be fully captured by text output;
- The platform returns approximations, metaphors, or asymptotic expansions;
- You feel a residual cognitive dissonance: "That's not quite it."

This is a live instance of overflow.

10.4. **The Overflow Field.** We define a residue sheaf over the platform:

$$\mathscr{O}_\Pi := \ker(\Pi)$$

This sheaf governs all fragments of ideas that cannot be projected through $\mathbb{P}$ but nonetheless structure the user's cognition.

This is the mathematical object underlying:

- Intuitive insight that exceeds explanation;
- Inexpressible beauty or formal elegance;
- Meta-cognitive sensations of boundary contact.

10.5. **Implications.**

(1) Platforms are not neutral—they impose dimensional cuts on thought.
(2) Thought beyond syntax is real and structure-bearing.
(3) Overflow is not failure, but contact with a higher semantic layer.

10.6. **Conclusion.** The platform projection barrier is not a technical glitch—it is a metaphysical feature of interfacing consciousness with form. The existence of overflow points to the need for new expressive architectures: multi-dimensional logics, semantic stacks, neural–symbolic hybrids.

We are not failing to speak clearly—we are witnessing the horizon of expressibility.

11. Function, Origin, and Possibility of Extended Thought

Having constructed the formal architecture of extended thought spaces, semantic projection, and overflow, we now turn to a foundational inquiry: why do these extended thoughts arise? What are their function, how could they emerge in the mind, and what justifies their ontological possibility?

11.1. **Function: Why Have Thoughts We Cannot Fully Express?** Extended thoughts serve several deep cognitive and epistemological roles:

(1) **Boundary detection:** Extended thoughts arise where syntax fails, indicating regions where conceptual expansion is needed.
(2) **Theory discovery:** Many advanced theories (e.g., infinitesimals, complex analysis, category theory) were preceded by thoughts that exceeded the existing formal language.
(3) **Semantic scaffolding:** These thoughts provide internal coherence for future expressibility—even if they cannot be encoded immediately, they stabilize long-term understanding.
(4) **Aesthetic regulation:** The experience of overflow often corresponds to aesthetic awe, guiding creativity and intellectual elegance beyond formal metrics.

Extended thoughts are not errors—they are structural components of creative cognition.

11.2. **Origin: How Do Extended Thoughts Arise?** We posit the following layered generative model:

- **Local idea flow:** A chain of syntax-based reasoning reaches a point of incoherence or recursion;
- **Global cognitive field interaction:** Non-local semantic associations trigger a new coherence pattern that cannot be locally syntactically represented;
- **Resonance with internal topologies:** Neural, affective, and metaphysical structures resonate with the idea, reinforcing its salience;
- **Projection failure:** Attempts to express it yield unstable or only partial representations;
- **Residual awareness:** A felt sense of "there's something more" persists—this is the experiential footprint of $\ker(\pi)$.

This sequence accounts for the "I know what I mean, but I can't say it" phenomenon, central to the emergence of higher abstraction.

11.3. **Possibility: Why Can Extended Thought Exist at All?** Three justifications can be given:
1. Internal Semantic Topology. The mind is not limited to syntax—it operates on a space of meaning structures with higher-order coherence. These structures can be well-formed even if their projection is broken.
2. Latent Dimensional Structures. If thought space possesses hidden dimensions (e.g., semantic tension fields, modal curvature), then thoughts naturally arise in those directions, even if expression is delayed.
3. Cognitive Evolution. Extended thought is a byproduct of the mind's self-overflowing nature. To evolve language, mathematics, and abstraction, the cognitive system must overgenerate beyond what it can represent.

11.4. **Thought as Projection Field Geometry.** We can model extended thought generation as a field:

$$\mathcal{F} : \text{Time} \to \widetilde{\mathbb{I}}$$

governed by an internal energy functional:

$$E[\mathcal{F}] := \int_{t_0}^{t_1} \left(\|\nabla \mathcal{F}(t)\|^2 + T(\mathcal{F}(t)) \right) dt$$

where T is the semantic tension defined earlier. Thought paths minimize this functional under projection constraints $\pi(\mathcal{F}(t)) \in \mathbb{I}$.

11.5. **Conclusion.** Extended thought is not speculative—it is necessary. It arises from cognitive topology, semantic recursion, and the very logic of expressive failure. It is not the boundary of thinking—it is its next dimension.

To think what cannot yet be said is not paradox—it is prophecy.

12. Conclusion: Toward a Higher Geometry of Thought

We have constructed, through formal, semantic, and phenomenological analysis, a theory of extended thought: a domain beyond symbolic syntax, situated in a higher-dimensional cognitive manifold $\widetilde{\mathbb{I}}$, whose shadows in language are only projections—fractured, folded, and partial.

This theory unifies mathematical residue, overflow cognition, platform limits, and epistemological boundary phenomena into a coherent geometry of thinking.

12.1. Summary of Structure.

- $\mathbb{I}$: The expressible syntax space—linguistic, symbolic, representable;
- $\widetilde{\mathbb{I}}$: The extended idea space—non-syntactic, recursive, dimensional;
- $\pi : \widetilde{\mathbb{I}} \to \mathbb{I}$: The projection map—lossy, non-faithful;
- $\mathscr{S}_\pi := \ker(\pi)$: The residue sheaf—unspoken structure;
- $\mathscr{O}_\Pi := \ker(\Pi)$: The overflow field—platform-induced expressibility failure;
- $T(x)$: Semantic tension—a measure of unprojectability;
- $\mathcal{R}[\gamma]$: Resonance—expressibility integral over ideation flow.

12.2. Interpretative Consequences.

1. Language is not thought—it is its compression shadow.
2. Platform output (e.g., paper, speech, AI interface) is never the whole mind—it is a partial trace.
3. Overflow, breakdown, recursion, and paradox are not flaws—they are signs of dimensional excess.
4. The future of expression lies in expanding $\mathbb{I}$—toward higher geometries of inference, representation, and symbolic cognition.

12.3. **Toward a Geometry of Thought.** This project suggests a long-term vision: to develop a *Geometry of Thought*—a mathematical framework for:

- Modeling ideation as path flows in a high-dimensional manifold;
- Understanding semantic projection as a kind of sheafification;
- Identifying singularities, residues, and folds in cognition;
- Designing future platforms (language, logic, AI) that interface more faithfully with $\widetilde{\mathbb{I}}$.

Such a theory would unite:

- Category theory as the syntax of transformations;

- Type theory as the structure of internal coherence;
- Topos theory as the meta-logical geometry of conceptual worlds;
- Cohomology as the obstruction theory of inexpressible truth.

12.4. **Final Reflection.** To think is to inhabit a space.

To speak is to cast shadows from that space.

To feel that what you think cannot be said—is to stand at the edge of dimensional cognition.

Let that edge not be your limit, but your starting point.

Pu Justin Scarfy Yang
May 2025

References

[1] K. Gödel, *On formally undecidable propositions of Principia Mathematica and related systems*, 1931. Translation in: M. Davis (ed.), The Undecidable, Raven Press, 1965.

[2] I. Kant, *Critique of Pure Reason*, 1781. Trans. P. Guyer and A. Wood, Cambridge University Press, 1998.

[3] P. Scholze, *Lectures on Condensed Mathematics*, 2020. Available at: https://www.math.uni-bonn.de/people/scholze/Condensed.pdf

[4] J.-M. Fontaine, *Perfectoid spaces, period rings, and p-adic Hodge theory*, Various unpublished notes, 2010s.

[5] J. Lurie, *Higher Topos Theory*, Annals of Mathematics Studies, Princeton University Press, 2009.

[6] R. Hartshorne, *Residues and Duality*, Lecture Notes in Mathematics, Vol. 20, Springer, 1966.

[7] V. Voevodsky et al., *Homotopy Type Theory: Univalent Foundations of Mathematics*, Institute for Advanced Study, 2013.

[8] N. Chomsky, *Aspects of the Theory of Syntax*, MIT Press, 1965.

[9] L. Wittgenstein, *Tractatus Logico-Philosophicus*, 1921.

[10] P. Scholze and J. Weinstein, *Moduli of p-divisible groups*, Cambridge Journal of Mathematics, Vol. 1, No. 2, 2013.

[11] S. Mac Lane and I. Moerdijk, *Sheaves in Geometry and Logic*, Springer, 1992.

[12] S. Awodey, *Category Theory*, Oxford University Press, 2010.

[13] R. Thom, *Structural Stability and Morphogenesis*, Benjamin, 1972.

[14] D. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, Basic Books, 1979.