The Latent Algebra: A Universal Representational Language for Smooth Systems

Tamás Nagy, Ph.D.

tnagyphd@gmail.com

Draft

The algebra is not a tool applied to Latents. The algebra IS what Latents are — the unique minimal structure that supports every operation representation theory requires.

Executive Summary (Non-Technical)

Two hospitals train ML models on their patient data. Can they combine what they learned without sharing any patient records? A portfolio risk manager computes a risk number. Can a derivatives trader use the same computation — not the same code, but the same *mathematical object* — to price options?

These sound like engineering problems. They are algebra problems.

This paper defines the Latent Algebra — the mathematical structure in which smooth systems live, combine, decompose, and transfer. We prove it is *universal*: there is no simpler algebraic structure that supports combining systems (tensor product), observing systems (contraction), and changing coordinates (basis invariance) simultaneously. This means every existing spectral method — Fourier analysis, eigenvalue decomposition, COS expansion, Knowledge Artifacts, Fokker–Planck generators — is a coordinate representation of the same algebraic object.

The algebra has three consequences invisible when working within any single representation:

1. **Cross-domain transfer is a theorem.** A technique that works in financial risk (eigenvalue conditioning of portfolio covariance) provably transfers to adversarial robustness (eigenvalue conditioning of the Jacobian), with the same improvement factor $I = \lambda_{\max}/L_{\text{eff}}$. The improvement does not depend on the domain — it depends on the spectrum.

2. **Symmetry determines compression.** A system with symmetry group G lives in the G-invariant quotient of the algebra, not the full tensor product. For n exchangeable assets at grade 3, this reduces the Latent dimension from N^3 to $\binom{N+2}{3}$ — a factor of 6 for $N = 100$. The quotient is not a choice; it is forced by the symmetry.

3. **Convergence and divergence are properties of the evaluator, not the algebra.** The 65-year open Fenton problem was never a distribution problem — the algebra is exact. It was an evaluator failure: the moment expansion diverges, but the Fourier expansion converges exponentially. The algebra separates *what we compute* from *how we compute it*.

The algebraic core and the paper's major theorems are formalized in the Platonic ProofEnv (elysium/fields/latent_algebra/); Lean 4 is used only when the export pipeline emits type-checkable artifacts, not as the primary verification layer for this manuscript.

Abstract

We define the **Latent Algebra** $\mathfrak{L}(\mathcal{H}) = \bigoplus_{r\geq 0} \mathcal{H}^{\otimes r}$ as the graded tensor algebra over a separable Hilbert space $\mathcal{H}$, equipped with four primitive operations — addition, tensor product, contraction, and basis change — that constitute a graded contraction algebra (GCA). We prove the **Universality Theorem**: $\mathfrak{L}(\mathcal{H})$ is the initial object in the category of GCAs generated by $\mathcal{H}$. Any system requiring grade decomposition, tensor products, and contraction embeds canonically into $\mathfrak{L}(\mathcal{H})$.

We prove the **Quotient Classification Theorem**: a smooth system with finite permutation symmetry $G \leq S_n$ forces grade-r Latents into the G-invariant subspace $(\mathcal{H}^{\otimes r})^G$, and in an N-mode truncation the invariant dimension obeys the Burnside orbit-count $\dim((\mathcal{H}_N^{\otimes r})^G) = \frac{1}{|G|}\sum_{\sigma\in G} N^{c(\sigma,r)}$ (Theorem 2). For full symmetric tensors (exchangeable statistics), $\dim \mathrm{Sym}^r(\mathcal{H}_N) = \binom{N+r-1}{r}$, yielding compression factors of 6–23× at grades 3–4 when compared to unconstrained N^r. We identify the Latent Algebra as the full Fock space $\mathcal{F}(\mathcal{H})$ of quantum field theory and classify which physical and statistical systems live in which Fock-space quotient.

We establish the **Grade-Convergence Theorem**: the truncated algebra $\mathfrak{L}_N(\mathcal{H})$ converges to $\mathfrak{L}(\mathcal{H})$ uniformly over all algebraic operations at rate $O(\rho^{-N})$, where $\rho > 1$ is the analyticity parameter. This makes the finite-dimensional algebra a rigorous approximation to the infinite-dimensional one.

We prove the **Algebra-Evaluator Separation Principle** as a categorical statement: the evaluator is a natural transformation from the identity functor to the truncation functor on the category of Latents, and convergence/divergence are properties of this natural transformation.

We develop the **complexification** $\mathfrak{L}(\mathcal{H}_{\mathbb{C}})$, showing that the analyticity parameter $\rho = e^{2\pi\delta/L}$ is the strip width of analytic continuation (Paley–Wiener) and that the COS expansion is the real-part projection of the Fourier decomposition in the complex algebra.

We prove the **Cross-Domain Transfer Theorem** and give an explicit algorithmic pipeline for mechanically transferring results between domains. We embed the 21-operation Knowledge Algebra of ML models into $\mathfrak{L}(\mathcal{H})$ as grade-1 specialization.

We prove the **Algebraic Optimization Theorem**: when the feasible set is convex in Latent coordinates and the objective is a contraction, optimization reduces to projection — no search required. The boundary is sharp: integer or rank constraints make the problem NP-hard.

We unify five computational paradigms — COS expansion, Fokker–Planck spectral theory, Knowledge Artifacts, structural mechanics, and quantum mechanics — as coordinate representations of the same algebra, with explicit inter-paradigm morphisms. We provide a concrete worked example transferring a 3-asset portfolio computation to a 3-feature ML model.

All major results are formalized in a dependent-type-checked Platonic ProofEnv (Python) with Z3-backed arithmetic: nine topical modules totaling 100 prove targets, plus an extended GCA foundation layer (35 theorems in newton_gca_foundations.py). Lean 4 verification applies only to successfully exported artifacts; the canonical sources for this paper are the newton_*.py files under elysium/ fields /latent_algebra/.

---

1. Introduction

1.1 Five Fields, One Algebra

Five computational traditions independently discovered the same pattern:

Field	Object	Operation	What they call it
Financial risk	COS coefficients of a portfolio loss	Addition, scaling, contraction to VaR	Spectral risk measurement
Machine learning	Spectral coefficients of a model's predictions	Addition, subtraction, distance, projection	Knowledge Algebra
Fluid dynamics	Fokker–Planck generator eigenvalues	Semigroup action, spectral decomposition	Modal analysis
Fundamental physics	Grade coefficients of a dynamical field	Grade decomposition, contraction	Grade Equation
Optimization	Eigenvalues of the structure matrix	Conditioning, univariate solve, combination	Eigenvalue conditioning

Each field has its own notation, its own convergence theorems, its own computational recipes. Yet all five perform the same abstract operations: decompose a system into graded components, manipulate the components with linear algebra, and reconstruct observables by contraction.

This paper identifies the common algebraic structure and proves it is *unique* — in a precise categorical sense, there is no simpler algebra that does all of this.

1.2 What Is New

The underlying algebraic object — the tensor algebra $T(\mathcal{H}) = \bigoplus_{r\geq 0} \mathcal{H}^{\otimes r}$ — is classical (Bourbaki, *Algèbre* III; Mac Lane, *Categories for the Working Mathematician*). We make this explicit: the universality is a known free-object property. What is new is:

1. **Identification as the natural home for spectral representations.** The inner product and contraction — not just the tensor product — are first-class. Contraction gives observables (pricing, risk, expectations, predictions) algebraic status.
2. **Quotient classification** (Theorem 2). Which systems live in which quotient of the full algebra? Exchangeable systems in the symmetric tensors, ordered systems in the full tensors. The classification is forced by the symmetry group and determines the effective dimension.
3. **Grade-convergence** (Theorem 3). The finite-dimensional truncated algebra converges uniformly over all operations. This is not a statement about a single Latent — it is a statement about the algebra itself.
4. **The algebra-evaluator separation as a categorical principle** (Theorem 4). The evaluator is a natural transformation, and convergence/divergence is a property of the transformation, not of the algebra.
5. **The algebraic optimization boundary** (Theorem 6). A sharp characterization of when optimization reduces to projection and when it requires search.

6. **Explicit inter-paradigm morphisms.** Not just "the five paradigms are views of the same algebra" — the actual basis-change matrices between COS, Fokker–Planck, and Knowledge Artifact representations.
7. **Derived operations.** Semigroup action, conditioning, and spectral decomposition are compositions of the four primitives, not additional axioms. This resolves the question of whether the axiom set is too minimal.
8. **Complexification and the Paley–Wiener connection.** The analyticity parameter ρ is not an empirical observation but a consequence of the Paley–Wiener theorem applied to the complexified algebra. The COS expansion is the real-part projection of the Fourier decomposition.
9. **A concrete worked example** demonstrating cross-domain transfer with explicit numbers, and a **five-paradigm unification table** displaying the algebraic structure across all five fields.

1.3 Contributions

#	Result	Type	Section
T1	Universality of $\mathfrak{L}(\mathcal{H})$	Classical (stated cleanly)	§2
T2	Quotient Classification	New theorem	§3
T3	Grade-Convergence	New theorem	§6
T4	Algebra-Evaluator Separation	New formalization	§5
T5	Cross-Domain Transfer	Generalized from companion	§8
T6	Algebraic Optimization Boundary	New theorem	§7
T7	Knowledge Algebra Embedding	Extended from companion	§9
—	Derived Operations (D1–D3)	New formalization	§4
—	Complexification & Paley–Wiener	New connection	§5
—	Five-Paradigm Unification Table	New synthesis	§8

1.4 Outline

Section 2 defines the algebra and proves universality. Section 3 classifies the quotients and connects to Fock spaces. Section 4 develops functoriality, inter-paradigm morphisms, and the three derived operations (semigroup action, conditioning, spectral decomposition) that account for most computational pipelines. Section 5 separates the function calculus and the evaluator layer, then develops the complexification — showing that the analyticity parameter ρ is the strip width of analytic continuation (Paley–Wiener). Section 6 proves grade-convergence. Section 7 characterizes the optimization boundary. Section 8 proves cross-domain transfer, gives the mechanical pipeline, and presents the five-paradigm unification table — the paper's central exhibit. Section 9 embeds the Knowledge Algebra. Section 10 gives a worked example. Section 11 addresses the trace-class constraint. Section 12 records machine verification (Platonic ProofEnv primary; Lean export secondary). Section 13 discusses open problems.

---

2. The Latent Algebra: Axioms and Universality

2.1 The Graded Tensor Space

Definition 1 (Graded Hilbert tensor algebra). Let $\mathcal{H}$ be a separable real Hilbert space with inner product $\langle \cdot, \cdot \rangle$. The **Latent Algebra** is the graded direct sum

$$\mathfrak{L}(\mathcal{H}) = \bigoplus_{r=0}^{\infty} \mathcal{H}^{\otimes r}$$

where $\mathcal{H}^{\otimes 0} = \mathbb{R}$, $\mathcal{H}^{\otimes 1} = \mathcal{H}$, and $\mathcal{H}^{\otimes r}$ is the r-fold Hilbert tensor product. Each grade inherits an inner product: for simple tensors, $\langle h_1 \otimes \cdots \otimes h_r,\, g_1 \otimes \cdots \otimes g_r \rangle_r = \prod_{i=1}^{r} \langle h_i, g_i \rangle$, extended by linearity and continuity.

Definition 2 (Latent). A **Latent** of grade r is an element $\Lambda \in \mathcal{H}^{\otimes r}$. In a chosen orthonormal basis $\{e_k\}_{k=1}^{\infty}$, the Latent has coordinates $\Lambda_{k_1 \cdots k_r} = \langle \Lambda, e_{k_1} \otimes \cdots \otimes e_{k_r} \rangle_r$. The coordinates depend on the basis; the Latent does not.

Grade interpretation. Grade r captures r-body interaction order:

Grade	Algebraic object	Physical meaning	Example
0	Scalar	Observable value	VaR, price, risk measure
1	Vector $\Lambda \in \mathcal{H}$	Distribution / state	COS coefficients, Knowledge Artifact
2	Matrix $M \in \mathcal{H}^{\otimes 2}$	Generator / covariance / coupling	Fokker–Planck, attention matrix
3	Tensor $T \in \mathcal{H}^{\otimes 3}$	Three-body interaction / co-skewness	Crash clustering, non-integrability

Definition 3 (Analyticity). A Latent $\Lambda \in \mathcal{H}^{\otimes r}$ has **analyticity parameter** $\rho > 1$ if there exists $C > 0$ such that $|\Lambda_{k_1 \cdots k_r}| \le C \cdot \rho^{-(k_1 + \cdots + k_r)}$ for all multi-indices.

2.2 The Four GCA Axioms

Definition 4 (Graded contraction algebra). A *graded contraction algebra* (GCA) over $\mathbb{R}$ is a graded real vector space $A = \bigoplus_{r=0}^{\infty} A_r$ equipped with:

(GCA-1) Graded inner product. An inner product $\langle \cdot, \cdot \rangle_r$ on each A_r.

(GCA-2) Tensor product. A bilinear product $\otimes : A_p \times A_q \to A_{p+q}$ satisfying inner product multiplicativity: $\langle a_1 \otimes a_2, b_1 \otimes b_2 \rangle_{p+q} = \langle a_1, b_1 \rangle_p \cdot \langle a_2, b_2 \rangle_q$.

(GCA-3) Contraction. Maps $C_{ij} : A_r \to A_{r-2}$ for $1 \le i < j \le r$ satisfying $C_{ij}(a_1 \otimes \cdots \otimes a_r) = \langle a_i, a_j \rangle \cdot a_1 \otimes \cdots \hat{a}_i \cdots \hat{a}_j \cdots \otimes a_r$.

(GCA-4) Associativity. $(a \otimes b) \otimes c = a \otimes (b \otimes c)$ in A_{p+q+r}.

A *morphism* of GCAs is a grade-preserving linear map that respects the product, inner products, and contractions.

These four axioms are irreducible: removing any one loses essential structure. Without (GCA-1), there is no norm and no convergence theory. Without (GCA-2), systems cannot be combined. Without (GCA-3), observables cannot be computed. Without (GCA-4), the algebra is not associative and composition is ambiguous.

Remark. Addition and scalar multiplication are implicit in the vector space structure. They are not separate axioms but part of the graded vector space A.

2.3 The Universality Theorem

Theorem 1 (Universality). *The Latent Algebra $\mathfrak{L}(\mathcal{H})$ is the initial object in the category of graded contraction algebras generated by $\mathcal{H}$: for any GCA A and any isometric linear embedding $\phi : \mathcal{H} \to A_1$, there exists a unique GCA morphism $\Phi : \mathfrak{L}(\mathcal{H}) \to A$ extending ϕ.*

Proof sketch. Define Φ on simple tensors by $\Phi(h_1 \otimes \cdots \otimes h_r) = \phi(h_1) \otimes_A \cdots \otimes_A \phi(h_r)$ and extend by linearity. The isometric property of ϕ and (GCA-2) ensure inner products are preserved. Contraction compatibility follows from (GCA-3). Uniqueness follows because simple tensors span each grade. □

Categorical language. The Latent Algebra functor $\mathfrak{L} : \mathbf{Hilb} \to \mathbf{GCA}$ is left adjoint to the grade-1 forgetful functor $U : \mathbf{GCA} \to \mathbf{Hilb}$:

$$\mathrm{Hom}_{\mathbf{GCA}}(\mathfrak{L}(\mathcal{H}), A) \cong \mathrm{Hom}_{\mathbf{Hilb}}(\mathcal{H}, U(A))$$

This adjunction is what makes $\mathfrak{L}(\mathcal{H})$ the *unique minimal* GCA — any other is a quotient, losing information.

What is classical, what is new. The free-object property of $T(\mathcal{H})$ is classical. What Theorem 1 adds to the representation-theoretic context is: (a) the inner product and contraction are part of the universal structure, not additional data; (b) contraction gives *observables* first-class algebraic status; (c) the consequence that any system needing these operations must embed in $\mathfrak{L}(\mathcal{H})$ or a quotient — and the quotient is determined by symmetry (Theorem 2).

---

3. Quotient Classification and Fock Spaces

The full tensor algebra $\mathfrak{L}(\mathcal{H})$ is the freest possible GCA. Most physical and statistical systems carry symmetries that force them into a *quotient* — a smaller algebra that retains only the invariant tensors. This section classifies which quotient corresponds to which symmetry.

3.1 The Quotient Classification Theorem

Definition 5 (Symmetry group of a system). A smooth system $F : \mathbb{R}^n \to \mathbb{R}^n$ has **symmetry group** $G \leq S_n$ if $F(\sigma \cdot x) = \sigma \cdot F(x)$ for all permutations $\sigma \in G$, where σ acts by permuting components.

The group G acts on $\mathcal{H}^{\otimes r}$ by permuting tensor factors: $\sigma \cdot (h_1 \otimes \cdots \otimes h_r) = h_{\sigma^{-1}(1)} \otimes \cdots \otimes h_{\sigma^{-1}(r)}$. The **$G$-invariant subspace** is $(\mathcal{H}^{\otimes r})^G = \{\Lambda \in \mathcal{H}^{\otimes r} : \sigma \cdot \Lambda = \Lambda \text{ for all } \sigma \in G\}$.

Theorem 2 (Quotient Classification). *Let F be a smooth system with symmetry group $G \leq S_n$. Then the grade-r Latent $\Lambda^{(r)}$ of F lies in $(\mathcal{H}^{\otimes r})^G$, and the effective dimension is:*

$$\dim((\mathcal{H}_N)^{\otimes r})^G = \frac{1}{|G|} \sum_{\sigma \in G} N^{c(\sigma,r)}$$

where $c(\sigma, r)$ is the number of cycles when σ acts on $\{1, \dots, r\}$ (Burnside's lemma applied to the tensor index action), and $\mathcal{H}_N$ is the N-dimensional truncation.

Proof sketch. If F commutes with G, then the extraction map $\Phi : F \mapsto \Lambda^{(r)}$ (multilinear in F's Taylor coefficients) commutes with G. Hence $\Lambda^{(r)} = \sigma \cdot \Lambda^{(r)}$ for all $\sigma \in G$, placing it in $(\mathcal{H}^{\otimes r})^G$. The dimension count follows from Burnside's lemma: the number of orbits of G on the set of r-tuples from $\{1, \dots, N\}$ equals $\frac{1}{|G|} \sum_\sigma N^{c(\sigma,r)}$. $\square$

3.2 The Three Canonical Quotients

The symmetric group S_r (acting on tensor indices) produces three principal quotients, each with a physical interpretation:

Symmetric tensors ($G = S_r$, bosonic):

$$\mathrm{Sym}^r(\mathcal{H}_N) = (\mathcal{H}_N^{\otimes r})^{S_r}, \qquad \dim = \binom{N + r - 1}{r}$$

Systems whose components are *exchangeable*: identical particles in physics, homogeneous assets in a portfolio, i.i.d. samples in statistics. The grade-r Latent is a symmetric tensor — the cumulant tensor of the exchangeable law.

Antisymmetric tensors (sign-twisted S_r-action, fermionic):

$$\wedge^r \mathcal{H}_N = \{T \in \mathcal{H}_N^{\otimes r} : \sigma \cdot T = \mathrm{sgn}(\sigma) T\}, \qquad \dim = \binom{N}{r}$$

(This is the isotypic subspace for the sign character — not the fixed space $\{T : \sigma \cdot T = T\}$ from Definition 5.)

Systems with *exclusion*: fermionic particles (no two in the same state), portfolio constraints with no-duplication rules, orthogonal decompositions. The constraint $\dim = 0$ for $r > N$ is Pauli exclusion.

Full tensors ($G = \{e\}$, no symmetry):

$$\mathcal{H}_N^{\otimes r}, \qquad \dim = N^r$$

Systems where *order matters*: time series, trajectories, non-exchangeable multi-agent systems. The grade-r Latent encodes the full r-point correlation structure including ordering.

3.3 Compression Factors

The dimensional reduction from the full algebra to the symmetric quotient is substantial:

Grade r	Full: N^r	Symmetric: $\binom{N+r-1}{r}$	Factor	Antisymmetric: $\binom{N}{r}$	Factor
2	10,000	5,050	2.0	4,950	2.0
3	1,000,000	171,700	5.8	161,700	6.2
4	10^8	4,421,275	22.6	3,921,225	25.5
5	10^{10}	$\sim 9.2 \times 10^7$	109	$\sim 7.5 \times 10^7$	133

(Table computed for $N = 100$.)

At grade 5, the symmetric quotient is over 100× smaller than the full algebra. For large r, the Stirling approximation gives $\binom{N+r-1}{r} \approx N^r/r!$, so the compression factor grows as $r!$. This is the algebraic reason that symmetric systems are dramatically more tractable than asymmetric ones.

3.4 Connection to Fock Spaces

The Latent Algebra is the full Fock space of quantum field theory:

$$\mathfrak{L}(\mathcal{H}) = \mathcal{F}(\mathcal{H}) = \bigoplus_{r=0}^{\infty} \mathcal{H}^{\otimes r}$$

The three quotients correspond to the standard Fock-space constructions:

Fock space	Tensor quotient	Physics	Statistics/Finance
Full $\mathcal{F}(\mathcal{H})$	$\mathcal{H}^{\otimes r}$	Distinguishable particles	Ordered time series
Bosonic $\mathcal{F}_s(\mathcal{H})$	$\mathrm{Sym}^r(\mathcal{H})$	Identical bosons	Exchangeable portfolios
Fermionic $\mathcal{F}_a(\mathcal{H})$	$\wedge^r \mathcal{H}$	Fermions (exclusion)	No-duplication constraints

This identification is not a metaphor. The creation and annihilation operators of QFT — $a_k^\dagger : \mathcal{H}^{\otimes r} \to \mathcal{H}^{\otimes(r+1)}$ and $a_k : \mathcal{H}^{\otimes r} \to \mathcal{H}^{\otimes(r-1)}$ — are precisely the grade-raising and grade-lowering operations in the Latent Algebra. The number operator $\hat{n}_k = a_k^\dagger a_k$ counts the excitation in mode k. The QFT vacuum $|0\rangle \in \mathcal{H}^{\otimes 0} = \mathbb{R}$ is the grade-0 Latent (a scalar).

The Fock-space connection provides a bridge: any result proved in QFT about Fock-space structure applies to the Latent Algebra, and any result about Latents applies to QFT. The Latent Theorem (truncation at N modes with error $O(\rho^{-N})$) is, in QFT language, a statement about the exponential decay of high-mode occupation numbers for systems with analyticity $\rho > 1$.

3.5 Mixed Symmetry: The Schur–Weyl Classification

For systems with partial symmetry (some components exchangeable, others not), the full classification uses the Schur–Weyl decomposition:

$$\mathcal{H}_N^{\otimes r} = \bigoplus_{\lambda \vdash r} V_\lambda^{S_r} \otimes W_\lambda^{GL(N)}$$

where the sum is over partitions λ of r, $V_\lambda^{S_r}$ is the irreducible S_r-representation indexed by λ, and $W_\lambda^{GL(N)}$ is the corresponding $GL(N)$-representation. A system with symmetry group $G \leq S_r$ uses only the G-invariant part:

$$(\mathcal{H}_N^{\otimes r})^G = \bigoplus_{\lambda \vdash r} (V_\lambda^{S_r})^G \otimes W_\lambda^{GL(N)}$$

Example. A portfolio of 50 US equities and 50 European equities has symmetry group $G = S_{50} \times S_{50}$ (exchangeable within each group, not across). The grade-2 Latent decomposes into: within-US covariance (symmetric, $\binom{51}{2} = 1275$ parameters), within-EU covariance (1275 parameters), and cross-Atlantic coupling (full, $50 \times 50 = 2500$ parameters). Total: 5050, vs. the unrestricted $\binom{100+1}{2} = 5050$. In this case, the partial symmetry gives no grade-2 reduction. But at grade 3: the full symmetric Latent has $\binom{102}{3} = 171,700$ parameters, while the $(S_{50} \times S_{50})$-invariant Latent has $\binom{52}{3} + \binom{52}{3} + \binom{51}{2} \times 50 + 50 \times \binom{51}{2} = 22,100 + 22,100 + 63,750 + 63,750 = 171,700$. Again no reduction at this parameter count — the cross terms consume the savings.

The non-trivial compression occurs when G acts on the tensor *indices* (not on the basis elements). For time-reversal symmetry ($\Lambda_{k_1 k_2} = \Lambda_{k_2 k_1}$), the grade-2 Latent is symmetric, reducing dimension from N^2 to $N(N+1)/2$.

---

4. Symmetry, Functoriality, and Inter-Paradigm Morphisms

4.1 The Automorphism Group

A unitary $U : \mathcal{H} \to \mathcal{H}$ induces an automorphism $\mathfrak{L}(U) : \mathfrak{L}(\mathcal{H}) \to \mathfrak{L}(\mathcal{H})$ defined on simple tensors by:

$$\mathfrak{L}(U)(h_1 \otimes \cdots \otimes h_r) = Uh_1 \otimes \cdots \otimes Uh_r$$

In coordinates: $[\mathfrak{L}(U)(\Lambda)]_{k_1 \cdots k_r} = \sum_{j_1,\ldots,j_r} U_{k_1 j_1} \cdots U_{k_r j_r} \Lambda_{j_1 \cdots j_r}$.

This is a **GCA automorphism**: it preserves the inner product, commutes with the tensor product, and commutes with contraction. The proof is direct from the multiplicativity of U.

Corollary (Basis invariance is a theorem). Coordinate representations are non-canonical; the Latent is canonical. Two different extraction bases yield coordinates related by $\mathfrak{L}(U)$, where U is the basis-change unitary.

Corollary (Unique grade extension). The grade-r basis change is uniquely determined by the grade-1 matrix entries U_{kj}. This is a consequence of universality: the extension from grade 1 to grade r is the unique GCA morphism.

4.2 Functoriality

The assignment $\mathcal{H} \mapsto \mathfrak{L}(\mathcal{H})$ is a functor $\mathfrak{L} : \mathbf{Hilb} \to \mathbf{GCA}$:

- **Objects:** Hilbert spaces map to Latent Algebras.
- **Morphisms:** Bounded linear maps $T : \mathcal{H}_1 \to \mathcal{H}_2$ map to GCA morphisms $\mathfrak{L}(T) : \mathfrak{L}(\mathcal{H}_1) \to \mathfrak{L}(\mathcal{H}_2)$.
- **Composition:** $\mathfrak{L}(T_2 \circ T_1) = \mathfrak{L}(T_2) \circ \mathfrak{L}(T_1)$.
- **Identity:** $\mathfrak{L}(\mathrm{id}) = \mathrm{id}$.

This functoriality means: if two Hilbert spaces are related by a linear map, the corresponding Latent Algebras are related by a GCA morphism. A *change of paradigm* is a morphism $T : \mathcal{H}_{\mathrm{COS}} \to \mathcal{H}_{\mathrm{FP}}$ between the Hilbert spaces of two representations; the induced $\mathfrak{L}(T)$ transfers all algebraic structure automatically.

4.3 Explicit Inter-Paradigm Morphisms

The five paradigms of §1.1 use different Hilbert spaces. The morphisms between them are:

COS → Fokker–Planck. Let $\{c_k(x) = \cos(k\pi(x-a)/(b-a))\}$ be the COS basis and $\{\phi_k(x)\}$ be the FP generator eigenfunctions. Both are orthonormal bases of $L^2([a,b])$. The intertwining unitary is:

$$U_{kj}^{\mathrm{COS}\to\mathrm{FP}} = \int_a^b c_k(x)\phi_j(x)\, dx$$

For the Ornstein–Uhlenbeck process ($dX = -\theta X\, dt + \sigma\, dW$), the eigenfunctions are Hermite functions and U is the Hermite–Fourier transform. The COS convergence rate $O(\rho^{-N})$ maps to the spectral gap rate $O(e^{-\gamma t})$ with $\rho = e^{\gamma \Delta t}$ — they are the same rate in different coordinates.

COS → Knowledge Artifacts. The Knowledge Artifact basis is the data eigenbasis $\{v_k\}$ (right singular vectors of X). The morphism from COS to KA is:

$$U_{kj}^{\mathrm{COS}\to\mathrm{KA}} = \int c_k(x)\, v_j(x)\, d\mu_X(x)$$

where μ_X is the empirical data distribution. This is computable from the training data as a matrix product: $U = C^\top V$ where $C_{ik} = c_k(x_i)$ and $V_{ij} = v_j(x_i)$.

Fokker–Planck → Knowledge Artifacts. Composition: $U^{\mathrm{FP}\to\mathrm{KA}} = U^{\mathrm{COS}\to\mathrm{KA}} \cdot (U^{\mathrm{COS}\to\mathrm{FP}})^{-1}$.

These morphisms are not approximations. They are exact unitaries (when the bases are complete) or isometries (when truncated). The Grade-Convergence Theorem (§6) bounds the truncation error.

4.4 Three Derived Operations

The four GCA primitives generate three frequently-used compositions:

D1. Semigroup action. A grade-2 Latent M (Fokker–Planck generator, attention matrix, Hamiltonian) acts on a grade-1 Latent by contraction: $(M \cdot \Lambda)_k = \sum_j M_{kj}\Lambda_j$. Time evolution is the exponential: $\Lambda(t) = e^{tM}\Lambda(0)$. In coordinates, each mode evolves independently in the eigenbasis: $\Lambda_k(t) = e^{\lambda_k t}\Lambda_k(0)$. Every diffusion, Markov chain, and Hamiltonian flow is an instance of semigroup action. The convergence rate is the spectral gap $\gamma = |\lambda_2 - \lambda_1|$.

D2. Conditioning. From a grade-2 joint Latent $\Lambda^{(X,Y)}$, condition on $Y = y$ by partial contraction and normalization to produce a grade-1 Latent $\Lambda^{(X|Y=y)}$. In coordinates:

$$\Lambda_k^{(X|Y=y)} = \frac{\sum_j \Lambda_{kj}^{(X,Y)} \cdot \delta_j(y)}{\sum_{k'} \sum_j \Lambda_{k'j}^{(X,Y)} \cdot \delta_j(y)}$$

This is Bayesian inference in the algebra: prior (grade 1) + likelihood (grade 2) $\to$ posterior (grade 1). The computational cost is one matrix-vector product — $O(N^2)$ — regardless of the distributional complexity.

D3. Spectral decomposition. Every symmetric grade-2 Latent decomposes: $M = \sum_k \lambda_k \phi_k \otimes \phi_k$. The eigenvalues are grade-0 Latents; the eigenvectors are grade-1 Latents. This is the inverse of the tensor product: analysis decomposes what synthesis constructs.

These three operations — evolve, condition, decompose — account for most of the computational pipelines in the companion papers. Each is a composition of tensor product (GCA-2) and contraction (GCA-3), confirming that the four axioms are sufficient.

---

5. The Function Calculus and Evaluator Layer

The four GCA axioms define the *algebra*. Two additional structures — the function calculus and the evaluator layer — live *outside* the algebra but interact with it. Conflating them with the algebraic primitives (as the first draft of this paper did) obscures the separation of concerns.

5.1 The Function Calculus

Definition 6 (Map). For a measurable $f : \mathbb{R} \to \mathbb{R}$ and a grade-1 Latent $\Lambda \in L^2(\Omega, \mu)$:

$$(f_*\Lambda)(\omega) = f(\Lambda(\omega))$$

Maps transform Latents without changing grade. The Fourier map $f = e^{it(\cdot)}$, the moment map $f = (\cdot)^k$, and the COS kernel map are all instances.

When is f_* an algebra morphism? In general, f_* does *not* respect the tensor product: $f_*(\Lambda_1 \otimes \Lambda_2) \neq f_*(\Lambda_1) \otimes f_*(\Lambda_2)$. The function calculus is a separate layer, not part of the GCA structure.

Proposition 1. *Let $f : \mathbb{R} \to \mathbb{R}$ and define $(f_*\Lambda)(\omega) = f(\Lambda(\omega))$ for grade-1 Latents $\Lambda \in L^2(\Omega, \mu)$. The assignment $\Lambda \mapsto f_*(\Lambda)$ is a **linear operator** on $L^2(\Omega, \mu)$ if and only if f is homogeneous linear:*

$f(x) = ax$ for some $a \in \mathbb{R}$. In particular, any affine map $f(x) = ax + b$ with $b \neq 0$ is not linear in Λ and cannot extend to a grade-preserving algebra homomorphism compatible with the vector-space structure of each GCA grade.

Proof sketch. Linearity requires $f_*(\alpha\Lambda) = \alpha f_*(\Lambda)$ and $f_*(\Lambda_1 + \Lambda_2) = f_*(\Lambda_1) + f_*(\Lambda_2)$ for all scalars α and all Latents. Fixing ω and varying $\Lambda(\omega)$ forces $f(\alpha t) = \alpha f(t)$ and $f(s+t) = f(s) + f(t)$ on the set of values taken by Latents; for dense ranges this is Cauchy's equation, hence $f(t) = at$. The case $b \neq 0$ breaks additivity already on constant functions. For a full GCA morphism one would additionally need compatibility with tensor product and contraction; composing coordinates with nonlinear f breaks multiplicativity of the tensor product, so only very rigid (essentially linear) choices are compatible. □

Analyticity under maps. If Λ has analyticity parameter ρ and f is analytic with radius of convergence R around the range of Λ, then $f_*\Lambda$ has analyticity parameter $\rho' \geq \rho$. The precise rate: $\rho' = \min(\rho, R/\|\Lambda\|_\infty)$. This means analytic maps preserve or improve analyticity — they never make a Latent harder to truncate.

5.2 The Evaluator Layer

Definition 7 (Evaluator). An **evaluator** $\hat{\Phi}_\varepsilon$ for an algebraic operation Φ is a computable function satisfying $\|\hat{\Phi}_\varepsilon(\Lambda) - \Phi(\Lambda)\| < \varepsilon$ for all Λ in a specified class.

Theorem 4 (Algebra-Evaluator Separation). *The algebraic operations (GCA-1 through GCA-4) are exact and basis-free. Convergence, divergence, and computational cost are properties of the evaluator, not the algebra.*

Formally: let $\mathbf{Lat}_\rho$ be the category of Latents with analyticity $\geq \rho$, and let $\mathrm{Trunc}_N : \mathbf{Lat}_\rho \to \mathbf{Lat}_\rho^{(N)}$ be the truncation functor to N-mode Latents. An evaluator is a natural transformation $\eta : \mathrm{Id} \Rightarrow \mathrm{Trunc}_N$. The convergence rate $\|\eta_\Lambda(\Lambda) - \Lambda\| = O(\rho^{-N})$ is a property of η — of the natural transformation — not of the objects in $\mathbf{Lat}_\rho$.

This is the precise categorical content of the separation principle: the algebra lives in $\mathbf{Lat}_\rho$; truncation is the functor Trunc_N to N-mode Latents; an evaluator is the associated natural transformation η (not itself a functor); convergence is a property of η.

5.3 Convergence Taxonomy

Different evaluators produce different convergence behaviors for the *same* algebraic operation:

Algebraic operation	Evaluator A	Rate A	Evaluator B	Rate B
$E_\mu[f(\Lambda)]$	Gauss-Hermite quadrature	$O(e^{-cQ})$	Monte Carlo	$O(M^{-1/2})$
$E_\mu[\Lambda^k]$	Moment recurrence	Divergent (heavy tails)	Fourier inversion	$O(e^{-\delta N})$
$\Lambda^{(\mathcal{B}')}$ from $\Lambda^{(\mathcal{B})}$	Matrix-vector product	Exact (machine ε)	Re-extraction	Full extraction cost

5.4 The Fenton Paradox Resolved

The 65-year open problem of computing the CDF of correlated lognormal sums (Fenton, 1960) was never an algebraic problem. The contraction $E_\mu[f(S)]$ is well-defined for any L^2-compatible map f. The problem was an *evaluator failure*: the moment evaluator ($f = x^k$) diverges because $E[S^{2k}] \geq e^{k^2\sigma^2_{\max}/2}$, while the Fourier evaluator ($f = e^{itx}$) converges because $|e^{itx}| = 1$ for all x.

Theorem 5 (Evaluator dependence). *For a grade-1 Latent representing $S = \sum w_i e^{Y_i}$ with $(Y_1, \dots, Y_n) \sim N(\mu, \Sigma)$:*

(a) The Fourier evaluator satisfies $|e^{itS(z)}| = 1$ for all z, and Gauss-Hermite quadrature converges as $O(e^{-cQ})$.

(b) The moment evaluator satisfies $E[S^{2k}] \geq e^{2k\mu_{\min}+2k^2\sigma^2_{\max}}$ (where $\mu_{\min} = \min_i \mu_i$), which grows super-exponentially in k for any fixed μ, σ. The moment-matching Toeplitz condition number grows as $\kappa \geq e^{cN_P^2}$.

Both evaluate the same distribution. Convergence is an evaluator property.

5.5 Complexification and the Fourier Layer

The real Latent Algebra extends to a complex one by complexifying the Hilbert space: $\mathcal{H}_{\mathbb{C}=\mathcal{H}\otimes_{\mathbb{R}}\mathbb{C}}$. The resulting complex Latent Algebra $\mathfrak{L}(\mathcal{H}_{\mathbb{C})}$ gains three structures invisible in the real case:

Structure 1: Fourier decomposition. A grade-1 Latent in the complexified algebra admits a Fourier representation: $\Lambda(\omega) = \sum_k a_k e^{i\omega_k t}$. The COS expansion is the real part; the characteristic function $\phi(t) = E[e^{itX}]$ is a grade-0 contraction in the complex algebra. The Fourier transform of a grade-r Latent is itself a grade-r Latent in the dual Hilbert space $\hat{\mathcal{H}}$ — the algebra is self-dual under Fourier.

Structure 2: Paley–Wiener and analyticity. The analyticity parameter ρ of the real algebra acquires a geometric interpretation: $\rho = e^{2\pi\delta/L}$, where δ is the width of the strip $\{z \in \mathbb{C} : |\mathrm{Im}(z)| < \delta\}$ to which the Latent extends analytically, and L is the truncation interval length. The Paley–Wiener theorem makes this precise:

$$|a_k| \leq Ce^{-2\pi k\delta/L} = C\rho^{-k}$$

The Fourier coefficients decay geometrically because the function extends analytically to a strip. This is not an empirical observation about "nice" functions — it is a theorem about complex analysis. Every statement about "how many modes suffice" is a shadow of the Paley–Wiener theorem.

Structure 3: Monodromy. When a parametric family of systems $\Lambda(s)$ traces a closed loop in parameter space ($s : [0, 1] \to \mathcal{P}$, $s(0) = s(1)$), the eigenvalues return to their original values but the eigenvectors may acquire phases. The monodromy matrix $\mathcal{M} = \mathcal{P} \exp\left(\oint A(s)\, ds\right)$, where $A(s)$ is the Berry connection, captures this holonomy. In the Latent Algebra, the Berry connection is a grade-2 Latent over the parameter space — it describes how the eigenbasis rotates as parameters change. This is relevant when the system's structure matrix depends on external parameters (market regime, temperature, learning rate).

Why complexification is not merely a convenience. Without complexification, the COS expansion is an ad hoc approximation scheme. With it, the COS expansion is the canonical real-part projection of the Fourier decomposition, the convergence rate $O(\rho^{-N})$ is a theorem (Paley–Wiener), and the analyticity parameter is a geometric quantity (strip width). The complex structure makes the convergence theory rigorous rather than empirical.

---

6. The Grade-Convergence Theorem

This section establishes that the finite-dimensional truncated algebra converges to the full algebra — not just for individual Latents (the Latent Theorem), but uniformly over all algebraic operations. This is the theorem that makes finite-dimensional computation a rigorous proxy for infinite-dimensional algebra.

6.1 The Truncated Algebra

Definition 8 (Truncated Latent Algebra). For $N \in \mathbb{N}$ and $R \in \mathbb{N}$, define:

$$\mathfrak{L}_{N,R}(\mathcal{H}) = \bigoplus_{r=0}^{R} (\mathcal{H}_N)^{\otimes r}$$

where $\mathcal{H}_N = \mathrm{span}(e_1, \dots, e_N)$. This is a finite-dimensional GCA with $\sum_{r=0}^{R} N^r$ scalar parameters.

The truncation functor $\mathrm{Trunc}_{N,R} : \mathfrak{L}(\mathcal{H}) \to \mathfrak{L}_{N,R}(\mathcal{H})$ projects each grade to the first N modes and drops grades above R.

6.2 The Grade-Convergence Theorem

Theorem 3 (Grade-Convergence). *Let $\Lambda \in \mathcal{H}^{\otimes r}$ have analyticity parameter $\rho > 1$ with bound constant C. Then:*

(a) Norm convergence:

$$\|\Lambda - \mathrm{Trunc}_N(\Lambda)\|_r \le C_r \cdot \rho^{-N}$$

where $C_r = C \cdot \binom{N+r-1}{r-1} \cdot (1 - \rho^{-1})^{-1}$ depends on the grade and bound constant.

(b) Operation convergence (addition):

$$\|\mathrm{Trunc}_N(\Lambda_1 + \Lambda_2) - (\mathrm{Trunc}_N(\Lambda_1) + \mathrm{Trunc}_N(\Lambda_2))\|_r = 0$$

Truncation is exact for addition (it is a linear projection).

(c) Operation convergence (tensor product):

$$\|\mathrm{Trunc}_N(\Lambda_1 \otimes \Lambda_2) - \mathrm{Trunc}_N(\Lambda_1) \otimes \mathrm{Trunc}_N(\Lambda_2)\|_{p+q} \le C_{p,q} \cdot \rho^{-N}$$

The truncated tensor product agrees with the tensor product of truncations up to exponentially small error.

(d) Operation convergence (contraction):

$$|C_{ij}(\Lambda) - C_{ij}(\mathrm{Trunc}_N(\Lambda))| \le C'_r \cdot \rho^{-N}$$

Truncation commutes with contraction up to exponentially small error.

(e) Operation convergence (basis change):

$$\|\mathfrak{L}(U)(\mathrm{Trunc}_N(\Lambda)) - \mathrm{Trunc}_N(\mathfrak{L}(U)(\Lambda))\|_r \leq C_r'' \cdot \rho^{-N}$$

when U preserves $\mathcal{H}_N$ (i.e., U is the identity on modes $> N$). For general U, the error involves the leakage $\|U - P_N U P_N\|$.

Proof sketch. (a) follows from summing the analyticity bound over truncated indices: the tail sum $\sum_{k>N} C\rho^{-(k_1+\cdots+k_r)}$ where at least one $k_i > N$ is bounded by $C_r\rho^{-N}$ via geometric series. (b) is immediate from linearity. (c) follows from the factorization: $(\Lambda_1 \otimes \Lambda_2)_{k_1\cdots k_{p+q}} = (\Lambda_1)_{k_1\cdots k_p}(\Lambda_2)_{k_{p+1}\cdots k_{p+q}}$; truncating the product differs from the product of truncations only when at least one index exceeds N in one factor but not the other, which is bounded by the cross terms in the analyticity expansion. (d) and (e) follow similarly. □

6.3 Eigenbasis Optimality

Theorem (Eigenbasis achieves maximal ρ). *Among all orthonormal bases of $\mathcal{H}$, the eigenbasis of the grade-2 generator (the Fokker–Planck matrix, covariance matrix, or attention matrix) maximizes the analyticity parameter ρ and hence minimizes the truncation dimension N for given error ε.*

Proof sketch (Parseval contamination). In any non-eigenbasis, the slowest-decaying eigenmode contaminates at least one high-k coordinate (by Parseval's identity: $\sum_k |U_{kj}|^2 = 1$ for each eigenmode j, so at least one $U_{kj} \neq 0$ for each j). This contamination prevents the effective analyticity from exceeding the slowest eigenvalue decay rate $|\lambda_{\min}|$. In the eigenbasis, each mode decays independently at its own rate, and the effective ρ equals the geometric ratio $|\lambda_1/\lambda_2|$ of the spectral gap. □

Related spectral comparison lemmas appear in the Platonic latent_algebra domain (newton_spectral_inequalities. the statement above is expository and should be read as a heuristic optimality principle, not as a standalone certified theorem in Lean 4.

---

7. Optimization as Projection

7.1 The Two-Latent Product Structure

An optimization problem has two ingredients: a *system Latent* Λ_S (what we can build) and an *objective Latent* Λ_O (what we want). The optimal design is:

$$\Lambda^* = \arg\min_{\Lambda \in \mathcal{K}} \langle \Lambda_O, \Lambda \rangle_r$$

where $\mathcal{K}$ is the feasible set (constraints on the Latent).

When $\mathcal{K}$ is well-behaved and Λ_O is known, this is a *projection* — not an iterative search. The optimal design is already implicitly present in Λ_S and Λ_O; the optimization merely extracts it.

7.2 The Algebraic Optimization Theorem

Theorem 6 (Algebraic Optimization Boundary). *Consider the problem* $\min_{\Lambda \in \mathcal{K}} \langle \Lambda_O, \Lambda \rangle_r$.

(a) If $\mathcal{K}$ *is a convex subset of* $\mathcal{H}^{\otimes r}$ *and* $\Lambda_O \in \mathcal{H}^{\otimes r}$, *then the problem is a convex program in Latent coordinates. Specifically: -* $\mathcal{K}$ *polyhedron* $\Rightarrow$ *linear program; -* $\mathcal{K}$ *ellipsoid* $\Rightarrow$ *eigenvalue problem (spectral decomposition); -* $\mathcal{K}$ *spectrahedron* $\Rightarrow$ *semidefinite program.*

In all three cases, the solution is obtained by projection or spectral decomposition — no iterative search is required.

(b) If $\mathcal{K}$ *involves any of: - Integer constraints on Latent coordinates (binary design variables); - Rank constraints on the Latent tensor (rank*$(\Lambda) \leq k$ *for* $r \geq 3$*); - Non-convex topology constraints (connectivity, genus);*

then the problem is NP-hard in general (tensor rank by Håstad, 1990; integer programming by Karp, 1972).

The boundary is sharp: convex $\mathcal{K}$ *+ contraction objective* $\Rightarrow$ *polynomial time; integer/rank/topology* $\Rightarrow$ *NP-hard.*

7.3 Instances

Markowitz portfolio optimization. Λ_S is the grade-2 covariance Latent, Λ_O is the return vector (grade-1). The feasible set is $\mathcal{K} = \{w : \sum w_i = 1, w_i \geq 0\}$ (simplex). This is a quadratic program on a polytope — solved by eigenvalue conditioning in the covariance eigenbasis. The K-mode reduction converts the n-dimensional problem to K one-dimensional problems.

Structural sizing. Λ_S encodes the stiffness tensor (grade 2), Λ_O the load distribution (grade 1). The feasible set is a box constraint $a_i \leq t_i \leq b_i$ on member thicknesses. This is a convex program — the optimal sizing is a projection.

Structural topology optimization. Λ_S is the stiffness tensor, but the design variable is binary: each element is present ($x_i = 1$) or absent ($x_i = 0$). The feasible set involves integer constraints. This is NP-hard — the algebra identifies it as fundamentally harder than sizing.

Knowledge Artifact selection. Given n trained ML models with Knowledge Artifacts $\mathbf{A}_1, \ldots, \mathbf{A}_n$, find the optimal weighted combination $\sum \alpha_i \mathbf{A}_i$ for a target task. If the weights are continuous ($\alpha_i \in [0, 1]$, $\sum \alpha_i = 1$), this is a simplex-constrained quadratic program — algebraic. If the constraint is "select exactly k models" ($\alpha_i \in \{0, 1\}$, $\sum \alpha_i = k$), this is integer programming — combinatorial.

7.4 The Diagnostic Value

The algebraic optimization boundary provides an a priori diagnostic: before attempting to solve an optimization problem, check whether $\mathcal{K}$ is convex in Latent coordinates. If yes, the problem reduces to projection. If no, estimate how far $\mathcal{K}$ is from convex (the convex relaxation gap) and decide whether the combinatorial approach is necessary.

This diagnosis is itself a contraction: the convexity of $\mathcal{K}$ is a grade-2 property (the Hessian of the constraint function), computable from the constraint Latent.

---

8. Cross-Domain Transfer

8.1 The Improvement Factor

Theorem (from *Eigenvalue Conditioning as Universal Optimizer*). Let $\Sigma \in \mathbb{R}^{n\times n}$ be a positive semidefinite structure matrix with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n > 0$. The improvement factor from eigenvalue conditioning is:

$$I = \frac{\lambda_{\max}}{L_{\text{eff}}} = \sqrt{\frac{n\lambda_{\max}^2}{\sum_k \lambda_k^2}} \geq 1$$

This factor depends only on the spectrum of Σ, not on the domain in which Σ arises.

8.2 Five Instances

Domain	Σ	Naive bound	Spectral bound	Improvement
Portfolio VaR	Asset covariance	$\sigma_{\max}\sqrt{n}$	$L_{\text{eff}}\sqrt{n}$	$I = \lambda_{\max}/L_{\text{eff}}$
Basket options	Asset covariance	Full-dimensional pricing	K univariate Black-Scholes	n/K_{eff}
Adversarial robustness	Jacobian Gram $J^\top J$	Spectral norm $\|J\|_{\text{op}}$	Frobenius $\|J\|_F/\sqrt{n}$	I
SGD convergence	Loss Hessian	$\kappa = \lambda_{\max}/\lambda_{\min}$	$\kappa_{\text{eff}} = \lambda_{K+1}/\lambda_{\min}$	Spectral gap
Transformer attention	Attention matrix	Mixing time $\sim 1/\gamma$	Per-mode contraction	Spectral gap

8.3 The Transfer Pipeline

Cross-domain transfer is not a heuristic analogy. It is a mechanical procedure:

Step 1 (Extract spectra). In domain A: eigendecompose $\Sigma_A = Q_A \Lambda Q_A^\top$. In domain B: eigendecompose $\Sigma_B = Q_B \Lambda' Q_B^\top$.

Step 2 (Verify spectral isomorphism). Check that Λ and Λ' have the same eigenvalue multiset (up to permutation and scaling). If $\Lambda = c\Lambda'$ for some scalar c, the transfer is exact. If $\Lambda \approx c\Lambda'$ approximately, the transfer is approximate with error bounded by $\|\Lambda - c\Lambda'\|_F / \|\Lambda\|_F$.

Step 3 (Map the result). A theorem proved in domain A about the spectrum Λ transfers to domain B by the algebra morphism $\mathfrak{L}(U)$ where $U = Q_B Q_A^\top$. This is a basis change — it preserves all algebraic invariants, including the improvement factor I.

Step 4 (Bound the error). If the spectral match is approximate: the transferred result holds up to a correction proportional to the spectral mismatch. Specifically, if $\|\Lambda - c\Lambda'\|_F \leq \delta \|\Lambda\|_F$, then the improvement factor transfers as $|I_B - I_A| \leq C\delta$ for an explicit constant C depending on the spectrum's concentration.

Remark. This pipeline is implementable in $O(n^2)$ operations (dominated by the eigendecomposition). The intellectual content is the *proof* that the transfer is valid — the computation is trivial.

8.4 Five Paradigms, One Algebra

The central claim of this paper is that five paradigms — previously treated as separate mathematical frameworks — are coordinate representations of the same Latent Algebra. The following table makes this precise by identifying, for each paradigm, the Hilbert space, the grade-1 and grade-2 Latents, the primary contraction, and the derived operations.

	COS Expansion	Fokker–Planck	Knowledge Artifacts	Structural Mechanics	Quantum/Stat Mech
Hilbert space $\mathcal{H}$	$L^2([a,b])$	$L^2(\Omega, e^{-V})$	$\mathbb{R}^K$ (data eigenbasis)	$L^2(\Omega_{\text{struct}})$	$L^2(\mathcal{C})$
Basis $\{e_k\}$	Cosines $\cos(k\pi x/L)$	Generator eigenfunctions ϕ_k	Right singular vectors v_k	FEM mode shapes ψ_k	Energy eigenstates $\lvert k\rangle$
Grade-1 Latent	COS coefficients a_k	Spectral coefficients of density	Spectral coefficients A_k	Displacement field modes	Wavefunction amplitudes c_k
Grade-2 Latent	Covariance of joint density	Generator $\mathcal{L}$ (eigenvalues = rates)	Model correlation matrix M_{kj}	Stiffness matrix K_{kj}	Hamiltonian H_{kj}
Contraction C_{12}	$\sum_k a_k b_k$ (inner product)	$\text{tr}(\mathcal{L}) = -\sum \gamma_k$ (decay rate)	$\langle \mathbf{A}, \mathbf{B} \rangle$ (model similarity)	$\sum_k K_{kk} u_k^2$ (strain energy)	$\langle \psi \lvert H \rvert \psi \rangle$ (energy)
Semigroup (D1)	Time evolution via $e^{-\gamma_k t}$	Transition kernel $e^{t\mathcal{L}}$	Learning dynamics $e^{-t\Gamma}$	Dynamic response $e^{-i\omega_k t}$	Schrödinger $e^{-iHt/\hbar}$
Analyticity ρ	$e^{2\pi\delta/L}$ (strip width)	$e^{\gamma\Delta t}$ (spectral gap)	λ_1/λ_2 (eigenvalue gap)	ω_2/ω_1 (frequency gap)	$E_2 - E_1$ (energy gap)
Convergence rate	$O(\rho^{-N})$	$O(e^{-\gamma t})$	$O(\lambda_{K+1}/\lambda_1)$	$O(\omega_{K+1}/\omega_1)$	$O(e^{-(E_{K+1}-E_1)t})$
Improvement I	COS vs. Monte Carlo	Spectral vs. finite difference	Eigenbasis vs. weight space	Modal vs. full-DOF	Variational vs. grid

Reading the table. Each column is a "coordinate system" for the same algebra. Each row is an algebraic quantity. The inter-paradigm morphisms of §4.3 are the column-to-column maps — they preserve every row. The improvement factor I (bottom row) is the same number in every column because it depends only on the spectrum, which is invariant under column changes.

What the table proves. Any theorem stated and proved for the grade-2 Latent in one column holds in every other column with the same spectrum. The Grade-Convergence Theorem applies uniformly: if $\rho > 1$ in one paradigm, the truncation error $O(\rho^{-N})$ applies in all paradigms. The Optimization Boundary Theorem distinguishes convex from NP-hard problems in every column simultaneously. These are not analogies — they are instances of the same algebraic fact.

---

9. The Knowledge Algebra as Latent Algebra

9.1 Embedding

The Knowledge Artifact of a trained ML model f with data X is a tuple $\mathcal{K}(f, X) = (V, \mathbf{A}, \mathbf{h}, \sigma)$ where V is the eigenbasis, $\mathbf{A} = (A_1, \ldots, A_K)$ are spectral coefficients, $\mathbf{h}$ are shrinkage weights, and σ is metadata.

Theorem 7 (Knowledge Algebra embedding). *The spectral coefficient vector* $\mathbf{A}$ *is a grade-1 Latent in* $\mathfrak{L}(\mathbb{R}^K)$*. The 21 operations of the Knowledge Algebra are compositions of the four GCA primitives:*

Knowledge operation	GCA decomposition	Axiom(s) used
Scale $\alpha\mathbf{A}$	Scalar multiplication	Vector space
Negate $-\mathbf{A}$	$\alpha = -1$	Vector space
Add $\mathbf{A} \oplus \mathbf{B}$	Addition	Vector space
Subtract $\mathbf{A} \ominus \mathbf{B}$	$\mathbf{A} + (-1)\mathbf{B}$	Vector space
Project $P_S(\mathbf{A})$	Partial contraction onto S	GCA-3
Inner product $\langle \mathbf{A}, \mathbf{B} \rangle$	Full contraction of $\mathbf{A} \otimes \mathbf{B}$	GCA-2 + GCA-3
Distance $\lVert \mathbf{A} - \mathbf{B} \rVert$	$\sqrt{C_{12}((\mathbf{A} - \mathbf{B}) \otimes (\mathbf{A} - \mathbf{B}))}$	All four
Complement $\mathbf{A} \setminus \mathbf{B}$	$\mathbf{A} - \frac{\langle \mathbf{A}, \mathbf{B} \rangle}{\langle \mathbf{B}, \mathbf{B} \rangle}\mathbf{B}$	All four
Average $\frac{1}{N}\sum \mathbf{A}_i$	Iterated addition + scaling	Vector space
Analogy $\mathbf{A} - \mathbf{B} + \mathbf{C}$	Addition + subtraction	Vector space

All operations that involve only addition and scaling use the vector space structure (implicit in any GCA). Operations involving inner products or distances additionally use the tensor product (GCA-2) and contraction (GCA-3).

9.2 Why Eigenbasis Arithmetic Is Exact

In weight space, model combination $(1 - \alpha)W_A + \alpha W_B$ mixes all parameters. In eigenspace, mode k of the combination equals $(1 - \alpha)A_k + \alpha B_k$ — the modes are orthogonal and don't interfere. This is the Eckart-Young theorem: the SVD basis diagonalizes the algebra, making addition and subtraction exact in the L^2 norm.

Task Arithmetic (Ilharco et al., 2023) works in weight space where the algebra is not diagonal. The 19.2× accuracy gap is the quantitative cost of using a non-eigenbasis coordinate system for an algebraic operation — a concrete instance of the eigenbasis optimality theorem of §6.3.

9.3 Grade-2 Extension: Model Correlation Structure

The Knowledge Artifact as defined is a grade-1 Latent — the spectral coefficients of a single model. The grade-2 extension captures *model correlations*:

$$M_{kj} = \mathrm{Cov}(\hat{A}_k, \hat{A}_j)$$

where the covariance is over data sampling or training randomness. This grade-2 Knowledge Latent encodes: - Which modes are correlated across models (off-diagonal structure) - Which modes are stable vs. noisy (diagonal magnitude) - The effective rank of the model population (eigenvalue decay of M)

The Knowledge Algebra operations on M include spectral decomposition (D3), semigroup action (how the model correlation evolves with more data — $M(t) = e^{-t\Gamma}M(0)$ where Γ is the learning dynamics generator), and conditioning (given that mode k has a certain value, what does mode j predict?). These are derived operations in the Latent Algebra applied to grade-2 Knowledge Latents.

---

10. Worked Example: Portfolio to Neural Network

We demonstrate the full transfer pipeline with concrete numbers.

10.1 Setup

Domain A: 3-asset lognormal portfolio. Assets with returns $R_i = e^{Y_i}$, where $(Y_1, Y_2, Y_3) \sim N(\mu, \Sigma)$ with:

$$\Sigma = \begin{pmatrix} 0.04 & 0.015 & 0.005 \\ 0.015 & 0.09 & 0.02 \\ 0.005 & 0.02 & 0.16 \end{pmatrix}$$

Domain B: 3-feature ML model. A neural network trained on 3-dimensional input with Jacobian Gram matrix $J^\top J$ having the same eigenvalues as Σ.

10.2 Step-by-step

1. Eigendecompose. $\Sigma = Q\Lambda Q^\top$ with $\Lambda = \mathrm{diag}(0.175, 0.078, 0.037)$.

2. Extract grade-1 Latent. In the eigenbasis, the portfolio loss COS coefficients are:

$$A_k = \int_{-\infty}^{\infty} \cos(k\pi(x-a)/(b-a))\, p_L(x)\, dx$$

where p_L is the portfolio loss density. The analyticity parameter is $\rho = e^{2\pi\delta/(b-a)}$ where δ is the strip width of analytic continuation. For this portfolio: $\rho \approx 3.2$, so $N = \lceil \log(1/\varepsilon)/\log 3.2 \rceil \approx 6$ modes for $\varepsilon = 10^{-3}$.

3. Grade-2 Latent. The covariance in the eigenbasis IS the diagonal Λ. This is the grade-2 Latent: $M_{kk} = \lambda_k$, $M_{kj} = 0$ for $k \neq j$.

4. Compute improvement factor.

$$L_{\text{eff}} = \sqrt{(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)/3} = \sqrt{(0.0306 + 0.0061 + 0.0014)/3} = \sqrt{0.0127} = 0.113$$

$$I = \lambda_{\max}/L_{\text{eff}} = 0.175/0.113 = 1.55$$

5. Transfer to Domain B. The ML model's Jacobian Gram matrix has the same eigenvalues: $\lambda_1 = 0.175, \lambda_2 = 0.078, \lambda_3 = 0.037$. By the Cross-Domain Transfer Theorem:

- The improvement factor in adversarial robustness is $I = 1.55$ — the same number.
- The spectral norm bound $\|J\|_{\text{op}} = \sqrt{0.175} = 0.418$ is improved to $\|J\|_F/\sqrt{3} = \sqrt{(0.175 + 0.078 + 0.037)/3} = \sqrt{0.0967} = 0.311$.
- The ratio $0.418/0.311 \approx 1.34$ compares operator-norm and Frobenius-based bounds; it should not be equated with $\sqrt{I}$ (here $\sqrt{I} \approx 1.25$).

6. The algebra morphism. The explicit basis change from the portfolio eigenbasis to the ML eigenbasis is $U = Q_B Q_A^\top$. Since both are diagonalized in their respective eigenbases and the eigenvalues match, U is the identity on the eigenvalues — the algebraic content (spectrum) transfers exactly.

10.3 What the Example Shows

The same three numbers — $\lambda_1 = 0.175, \lambda_2 = 0.078, \lambda_3 = 0.037$ — determine the portfolio VaR improvement ($I = 1.55$), the robustness certification improvement ($I = 1.55$), and the SGD preconditioning benefit (effective condition number $\kappa_{\text{eff}} = 0.078/0.037 = 2.1$ vs. $\kappa = 0.175/0.037 = 4.7$). The domains are irrelevant; the spectrum is everything.

10.4 Second Example: Ornstein–Uhlenbeck to Harmonic Oscillator

This example transfers a result from stochastic processes (Fokker–Planck) to quantum mechanics, using the five-paradigm table of §8.4.

Domain A: Ornstein–Uhlenbeck process. $dX_t = -\theta X_t\, dt + \sigma\, dW_t$ with stationary distribution $N(0, \sigma^2/2\theta)$. The Fokker–Planck generator has eigenvalues $\lambda_k = -k\theta$ and Hermite eigenfunctions $\phi_k(x) = H_k(x\sqrt{\theta/\sigma^2})$.

Domain B: Quantum harmonic oscillator. $H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2$ with energy eigenvalues $E_k = \hbar\omega(k + 1/2)$ and Hermite eigenstates $\psi_k(x) = H_k(x\sqrt{m\omega/\hbar})$.

The algebraic content is identical. Both systems have: - The same eigenbasis (Hermite functions, up to scaling) - Equispaced eigenvalue spectrum with gap $\gamma_{\text{FP}} = \theta$ vs. $\gamma_{\text{QM}} = \hbar\omega$ - Analyticity parameter $\rho = e^{\gamma\Delta t}$ (FP) vs. $\rho = e^{\hbar\omega/k_B T}$ (QM thermal state) - Grade-convergence at rate $O(\rho^{-N})$ for truncation to N modes

What transfers. A theorem about the OU process — for example, that $N = \lceil \log(1/\varepsilon)/\log(\rho) \rceil$ modes suffice for ε-accuracy in the CDF — transfers directly to the quantum oscillator: N energy levels suffice for ε-accuracy in the partition function. The algebra morphism is the scaling map $U : x \mapsto x \cdot \sqrt{m\omega\theta/(\hbar\sigma^2)}$ that converts between the two Hermite parameterizations.

Concrete numbers. With $\theta = 1$, $\sigma = 1$ (OU) and $\omega = 1$, $\hbar = 1$, $m = 1$ (QM): - Both have eigenvalue gap $\gamma = 1$ - At temperature T with $k_B T = 2$: $\rho_{\text{QM}} = e^{1/2} \approx 1.65$ - With $\Delta t = 0.5$: $\rho_{\text{FP}} = e^{0.5} \approx 1.65$ — the same - For $\varepsilon = 10^{-3}$: $N = \lceil \log(1000)/\log(1.65) \rceil = 14$ modes/levels suffice in both domains

What the example shows. The five-paradigm table (§8.4) is not an analogy — it is a computation. The Fokker–Planck column and the Quantum/Stat Mech column have the same spectrum, the same eigenbasis (up to scaling), and the same convergence rate. A result proved in one column holds in the other.

---

11. The Trace-Class Constraint

11.1 When Is the Grade-2 Algebra Well-Defined?

For infinite-dimensional $\mathcal{H}$, the grade-2 Latent is an operator $M : \mathcal{H} \to \mathcal{H}$. The contraction (trace) $C_{12}(M) = \sum_k M_{kk}$ in an orthonormal basis is well-defined when M is **trace-class**, i.e. when the nuclear norm $\|M\|_1 = \sum_k s_k(M)$ is finite, where $s_k(M)$ are the singular values. For self-adjoint (in particular PSD) M, this is equivalent to $\sum_k |\lambda_k| < \infty$ with λ_k the eigenvalues.

The trace-class operators $\mathcal{T}(\mathcal{H})$ form a two-sided ideal in $\mathcal{B}(\mathcal{H})$. In particular: - The product of a trace-class operator with a bounded operator is trace-class. - The product of two trace-class operators is trace-class. - **Not every bounded operator is trace-class** — boundedness alone does not imply a finite trace.

Consequence for the Latent Algebra. If we want contraction to be well-defined at grade 2, we must restrict to trace-class Latents. The grade-2 Latent Algebra is:

$$\mathfrak{L}_2^{\mathrm{TC}}(\mathcal{H}) = \mathcal{T}(\mathcal{H}) \subset \mathcal{H} \otimes \mathcal{H}$$

This is a restriction that the finite-dimensional theory ($\dim \mathcal{H} = N < \infty$) does not see — in finite dimensions, every operator is trace-class.

11.2 Physical Systems Are Trace-Class

For the systems in this paper, the trace-class constraint is automatically satisfied:

- **Covariance operators** of square-integrable random vectors in finite-dimensional ambient space (or, more generally, trace-class covariance operators on $\mathcal{H}$) satisfy $\mathrm{Tr}(\Sigma) = \mathbb{E}\|X - \mathbb{E}X\|^2 < \infty$.
- **Fokker–Planck / diffusion semigroups.** The generator $\mathcal{L}$ on L^2 is typically **unbounded** and is not trace-class as an operator on the full space. In applications one works with **finite-dimensional modal truncations** or **trace-class heat operators** $e^{t\mathcal{L}}$ ($t > 0$) on compact domains under standard spectral hypotheses — those objects carry the trace-class discussion, not $\mathcal{L}$ itself.
- **Attention matrices** in transformers are $n \times n$ with n finite (context length), hence trivially trace-class.

The trace-class condition is a constraint on the *physics* (the system must have finite total variance/energy), not on the *algebra*. The algebra is defined for all operators; the trace is defined for trace-class ones. The analyticity condition $\rho > 1$ implies trace-class membership via the geometric decay of eigenvalues: $\sum_k C\rho^{-k} < \infty$ for $\rho > 1$.

11.3 Nuclear Operators at Higher Grades

At grade $r \geq 3$, the appropriate constraint is nuclearity (trace-class generalization to tensors). A grade-r tensor T is nuclear if:

$$\|T\|_{\text{nuc}} = \inf\left\{\sum_i \|v_{i,1}\|\cdots\|v_{i,r}\| : T = \sum_i v_{i,1}\otimes\cdots\otimes v_{i,r}\right\} < \infty$$

The analyticity condition ensures nuclearity: if $|\Lambda_{k_1\cdots k_r}| \le C\rho^{-(k_1+\cdots+k_r)}$, then $\|T\|_{\text{nuc}} \le \left(\frac{C}{1-\rho^{-1}}\right)^r < \infty$, since each tensor factor contributes a geometric series $\sum_k C\rho^{-k} = C/(1-\rho^{-1})$ and nuclearity follows from the product over r factors.

---

12. Machine Verification

12.1 Canonical artifact: Platonic ProofEnv (`elysium/fields/latent_algebra/`)

The formal companion to this paper is a collection of Platonic ProofEnv scripts newton_∗.py. Each file exposes a build() function returning a ProofEnv; running the module executes verify_all() and reports 0 type errors when the domain is healthy.

Nine topical modules (100 p.prove targets) — aligned with the paper's section map:

Module	Theorems	Paper section	Key results
newton_optimization_boundary.py	9	§7, T6	Convex minimum, quadratic minimizer, eigenvalue conditioning
newton_quotient_classification.py	10	§3, T2	Burnside orbit count, symmetric classification
newton_derived_operations.py	15	§4.4, D1–D3	Semigroup composition, identity, linearity, spectral reconstruction
newton_grade_convergence.py	16	§6, T3	Norm convergence, addition exactness, truncation tail bound
newton_cross_domain_transfer.py	12	§8	Morphism preservation, spectral invariance, transfer pipeline
newton_knowledge_algebra.py	10	§9	GCA embedding of KA operations
newton_evaluator_separation.py	12	§5, T4+T5	Evaluator separation, Fenton layer discussion
newton_complexification.py	8	§5.5	Paley–Wiener-style bounds (axiomatized layer)

Module	Theorems	Paper section	Key results
newton_trace_class.py	8	§11	Trace-class ideal closure, nuclear norm bound

Extended library (not included in the 100-count table): newton_gca_foundations.py (35 theorems) develops the abstract GCA axiom block; newton_spectral_inequalities.py and newton_applications_numerical.py host additional spectral lemmas and numerical sanity checks.

12.2 Lean 4 status

There is **no** checked-in Lean source at kernel/LeanProofs/LatentAlgebra/RankBound.lean in this repository. Lean 4 enters only through the Platonic **export** pipeline (when enabled): some forge/core_latent_algebra/lean_export/*.lean files are staging artifacts and may list axiom placeholders for lemmas not yet ported. **Do not cite a Lean line count or theorem count for this paper until a specific export has been compiled and sealed.**

12.3 Trust model

- **ProofEnv / Platonic** — primary: every ts.qed() is checked by the Platonic type checker; arithmetic subgoals may close via Z3 (_trusted.* certificates in export metadata).
- **Lean 4** — optional second opinion on exported files; not the SSOT path for the latent_algebra domain as of this writing.

---

13. Discussion and Open Problems

13.1 Relation to Classical Algebra

The tensor algebra $T(\mathcal{H})$ appears in: Bourbaki (*Algèbre* III), Mac Lane (*Categories for the Working Mathematician*), Atiyah (*K-Theory*). The Fock space $\mathcal{F}(\mathcal{H})$ appears throughout quantum field theory. The novelty of this paper is not the algebra itself but the identification of the specific axiom set (GCA), the quotient classification for statistical/financial systems, the grade-convergence theorem, and the practical consequences (cross-domain transfer, optimization boundary, evaluator separation).

13.2 Relation to Clifford and C*-Algebras

Clifford algebras $Cl(V, Q)$ are quotients of $T(V)$ by the relation $v \otimes v = Q(v)$. In the Latent context, this quotient identifies the grade-2 component $v \otimes v$ with the scalar $Q(v)$ — collapsing two grades. This is appropriate for systems where the quadratic form Q is the fundamental structure (spinors, Dirac operators). The Latent Algebra keeps all grades separate; Clifford algebras collapse them.

C*-algebras are norm-closed -*algebras of bounded operators. The Latent Algebra at grade 2 ($\mathcal{B}(\mathcal{H})$ with the Hilbert-Schmidt norm) is a Hilbert space, not a C*-algebra. The product in the Latent Algebra (tensor product, grade-increasing) differs from the product in a C-*algebra (operator* *composition, grade-preserving). The relationship is: the C*-algebra of bounded operators is the *endomorphism algebra* of $\mathcal{H}$; the Latent Algebra is the *tensor algebra* of $\mathcal{H}$.

13.3 What the Algebra Is Not

To prevent over-claiming:

Not a new algebra. The tensor algebra $T(\mathcal{H})$ is classical. The Fock space is classical. The novelty is the *identification* of this structure as the natural home for spectral representations across domains, plus the specific theorems (quotient classification, grade-convergence, optimization boundary) that make the identification useful.

Not a numerical method. The algebra provides the *language* in which to state and prove theorems. Evaluators (COS quadrature, Monte Carlo, Gauss-Hermite) are the numerical methods. The algebra tells you what you're computing; the evaluator tells you how.

Not a universal approximation theorem. The Grade-Convergence Theorem says the *algebra* converges — algebraic operations on truncated Latents approximate operations on full Latents. It does not say that every function can be approximated by a Latent. Smooth functions with analyticity $\rho > 1$ are in scope; discontinuous functions are not.

Not domain-specific. The algebra is domain-agnostic by construction. The five-paradigm table (§8.4) shows this concretely: the same algebraic fact has five different interpretations, one per domain. This is a feature, not a bug — but it means the algebra does not encode domain-specific knowledge. That knowledge lives in the choice of Hilbert space, basis, and evaluator.

13.4 Open Problems

1. **Quotient detection.** Given a dataset (not a dynamical system), can the symmetry group G be inferred from the Latent's coordinate statistics? The symmetry reduces dimension (Theorem 2), but detecting the symmetry from data is a statistical problem.

2. **Non-commutative GCA.** The current GCA axioms assume the tensor product is associative but not necessarily commutative. For quantum systems with non-commuting observables, is the appropriate structure a non-commutative GCA? The operator algebra $\mathcal{B}(\mathcal{H})$ (with operator composition as the product) is a candidate.

3. **Computational complexity at high grade.** Grade-r basis change costs $O(r \cdot N^{r+1})$. For structured (sparse, low-rank) tensors, can this be reduced? Tensor network methods (MPS, DMRG) suggest $O(\text{poly}(N, D))$ where D is the bond dimension.

4. **Infinite-dimensional evaluators.** The Grade-Convergence Theorem assumes truncation to N modes. In reproducing kernel Hilbert spaces, can the algebra be evaluated directly in infinite dimensions without truncation? The kernel trick suggests yes, but the formal connection is unestablished.

5. **The algebra as a programming language.** Latent operations (add, tensor, contract, basis-change) form a domain-specific language. A compiler from "Latent programs" to efficient numerical code — choosing the evaluator automatically based on the analyticity parameter — would be a practical contribution. The Grade-Convergence Theorem provides the error guarantees; the optimization boundary provides the complexity classification.

6. **Deep Lean formalization of Theorems 2, 3, and 6.** The ProofEnv kernel verifies these theorems against axiomatized type signatures with Z3-backed arithmetic. Porting the full proofs to Lean 4 with Mathlib — Burnside's lemma for the quotient classification, real analysis for grade-convergence, convex analysis for the optimization boundary — would elevate them to the highest assurance level. Mathlib provides the foundations; the proofs are medium-difficulty.

7. **Monodromy and Berry phase in parametric families.** When a system's structure matrix depends on external parameters (market regime, learning rate, temperature), the eigenbasis rotates as parameters change. The Berry connection (§5.5) is a grade-2 Latent over parameter space. Does this connection have observable consequences? In quantum mechanics, Berry phases are measurable. The financial analogue would be a "regime rotation phase" — a cumulative basis rotation that persists after parameters return to their original values. Whether this phase is observable in financial time series is an open empirical question.

8. **Categorical foundations.** The Latent functor $\mathfrak{L} : \mathbf{Hilb} \to \mathbf{GCA}$ is left adjoint to the forgetful functor $U : \mathbf{GCA} \to \mathbf{Hilb}$. This adjunction should be formalized. The evaluator as a natural transformation (Theorem 4) is one consequence; the full adjunction may yield additional structure (e.g., a monad on **Hilb**) with computational significance.

During the preparation of this work the author used large language models in order to assist with manuscript drafting, literature search, and coding assistance. After using these tools, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

- Atiyah, M. F (1967). K-Theory. *K-Theory.*
- Berry, M. V (1984). Quantal phase factors accompanying adiabatic changes. *Proceedings of the Royal Society of London A*, 392(1802), 45-57.
- Bourbaki, N (1989). Algebra I–III. *Algebra I–III.*
- Burnside, W (1897). Theory of Groups of Finite Order. *Theory of Groups of Finite Order.*
- Fenton, L. F. (1960). The Sum of Log-Normal Probability Distributions in Scatter Transmission Systems. *IRE Transactions on Communications Systems*, CS-8, 57-67. DOI: 10.1109/tcom.1960.1097606
- Fulton, W. & Harris, J (1991). Representation Theory: A First Course. *Representation Theory: A First Course.*
- Håstad, J (1990). Tensor rank is NP-complete. *Journal of Algorithms*, 11(4), 644-654.
- Ilharco, G., Ribeiro, M.T., Wortsman, M., Gururangan, S., Schmidt, L., Hajishirzi, H., and Farhadi, A (2023). Editing models with task arithmetic. *ICLR.*
- Karp, R. M (1972). Reducibility among combinatorial problems. In *Complexity of Computer Computations*, 85–103. Plenum. *Complexity of Computer Computations*, 85-103.
- Mac Lane, S (1971). Categories for the Working Mathematician. *Categories for the Working Mathematician.*

- Nagy, T. (2026). The Latent: Finite Sufficient Representations of Smooth Systems. *Zenodo*. DOI: 10.5281/zenodo.19101209
- Nagy, T. (2026). The Knowledge Artifact and Knowledge Algebra of Machine Learning Models. *Zenodo*. DOI: 10.5281/zenodo.18910387
- Nagy, T. (2026). Eigenvalue Conditioning as Universal Optimizer: Cross-Domain Transfer Between Finance, Robustness, and Machine Learning. *Working paper*.
- Nagy, T. (2026). The Latent of Latents: Hierarchical Finite Representations of Knowledge Families. *Zenodo*. DOI: 10.5281/zenodo.19134434
- Nagy, T. (2026). The Spectral Lognormal Distribution, The Distribution of Portfolio Value. *Zenodo*. DOI: 10.5281/zenodo.18940756
- Paley, R. E. A. C. & Wiener, N (1934). Fourier Transforms in the Complex Domain. *Fourier Transforms in the Complex Domain*.