The Latent Solution: A Finite Sufficient Representation Framework for Partial Differential Equations

Tamás Nagy, Ph.D.

tnagyphd@gmail.com

Submission Ready • April 2026

Classical PDE theory asks: does a solution exist, and is it smooth? The Latent asks: how many numbers does the solution contain? For analytic PDEs, the answer is finite.

Executive Summary (Non-Technical)

A partial differential equation describes how a function — a temperature field, a velocity field, a wave — evolves in space and time. The function lives in an infinite-dimensional space: at every point in space, there is a number, and there are uncountably many points. The central question of PDE theory is: **what is the solution?**

For 300 years, mathematicians have given increasingly general answers:

- **Classical solution** (18th–19th century): a smooth function satisfying the equation at every point. Ideal, but often doesn't exist.
- **Weak solution** (Leray, 1934; Sobolev, 1938): a function satisfying the equation "on average." Always exists, but may not be unique or smooth.
- **Distributional solution** (Schwartz, 1950): a generalized function. Even more general, but even less physical content.

Each expansion weakens the meaning of "solution" to guarantee existence. The gap between weak and strong — between "a solution exists" and "the solution is the one nature actually computes" — is where every millennium-class PDE problem lives.

The Latent framework asks a different question entirely. Not "in what function space does the solution live?" but:

How many numbers does the solution contain?

For an analytic PDE — one whose data is smooth and extends to the complex plane — the answer is finite:

$$N^* = \left\lceil \frac{\log(C_0/\varepsilon)}{\log \rho} \right\rceil$$

where ε is the desired accuracy and ρ is the analyticity radius, a single number measuring the solution's smoothness. The solution $u(x,t)$ — nominally an infinite-dimensional object — is completely encoded by N^* spectral coefficients $(c_0(t), \dots, c_{N^*}(t))$. These coefficients evolve according to a finite ordinary differential equation derived from the PDE.

This is not an approximation that happens to work. It is a representation theorem: the solution IS these N^* numbers, to certified accuracy ε. The infinite-dimensionality of the PDE is an artifact of the continuum representation, not a feature of the dynamics.

The Latent solution builds on the classical Galerkin method (projecting the PDE onto finitely many modes) but adds four structural properties that standard Galerkin lacks:

1. **Self-determined truncation.** The number of modes N^* is determined by the PDE itself, not chosen by the analyst.
2. **Exponentially certified error.** The truncation error is guaranteed to be at most ε, decaying exponentially in N^*.
3. **Grade-aware coupling.** The polynomial structure of the PDE determines which coefficient interactions matter, yielding provable sparsity.
4. **Faithful truncation.** The truncated system cannot introduce spurious dynamics — unresolved modes affect resolved modes at amplitude $O(\varepsilon)$.
5. **Perturbation stability.** Modifying the PDE changes ρ by a computable amount. The framework absorbs perturbations without rebuilding.

Abstract

We introduce the **Latent solution** as a quantitative framework for finite-dimensional representation of PDE solutions. The framework builds on the Galerkin method but exploits the grade structure of analytic PDEs (Nagy, 2026) to provide four structural guarantees absent from standard Galerkin theory.

For an evolution PDE $\partial_t u = F(u)$ with analytic data and analyticity radius $\rho > 1$, the Latent solution is the trajectory $\Lambda(t) = (c_0(t), \dots, c_{N^*}(t))$ in the truncated spectral coefficient space, where $N^* = \lceil \log(C_0/\varepsilon)/\log\rho \rceil$ modes are determined by the PDE's own analyticity radius. We prove:

(I) Self-determined truncation (Theorem 1). The analyticity radius ρ of the PDE data uniquely determines the truncation level $N^*(\varepsilon, \rho)$ with certified error bound $\|u - u_{N^*}\|_{L^2} \leq \varepsilon$.

(II) Grade-structured coupling (Theorem 2). For a PDE whose right-hand side is a polynomial of degree K in u, the Latent ODE on $\mathbb{R}^{N^*+1}$ has at most K-body couplings. High-mode self-interactions are exponentially suppressed by the grade bound, giving effective sparsity $O((N^*)^K)$ instead of $(N^*)^{N^*}$.

(III) Exponential closure (Theorem 3). The closure error of the truncated system — the effect of unresolved modes ($k > N^*$) on the resolved modes ($k \leq N^*$) — is bounded by $O(\varepsilon)$. The truncated system is a certified approximation of the full dynamics, not an uncontrolled projection.

(IV) Spectral interpolation (Theorem 4). The framework interpolates continuously between three regimes:

Data regularity	Coefficient decay	N^*	Solution quality
Analytic (strip width $\sigma > 0$)	$\lvert c_k\rvert \leq Ce^{-\sigma\lvert k\rvert}$	$O(\log 1/\varepsilon)$	Classical, analytic
C^s (finite smoothness)	$\lvert c_k\rvert \leq C\lvert k\rvert^{-s}$	$O(\varepsilon^{-1/s})$	Classical, C^s

Data regularity	Coefficient decay	N^*	Solution quality
L^2 (no smoothness)	No guaranteed decay	$O(\varepsilon^{-d})$	Weak (standard Galerkin)

The transition from exponential to algebraic compression at the analyticity boundary is a phase transition in the solution's representability.

---

1. Introduction

1.1 The fundamental tension

Partial differential equations govern virtually every continuous physical process. The central questions are always the same: does a solution exist? Is it unique? Is it smooth?

For linear PDEs, these questions are settled by the spectral theory of self-adjoint operators. For nonlinear PDEs, they are largely open. The fundamental tension:

Existence is easy to prove for WEAK solutions — functions in L^2 or Sobolev spaces that satisfy the equation in an averaged (distributional) sense. The Galerkin method, energy estimates, and compactness arguments reliably produce weak solutions.

Regularity and uniqueness are hard to prove. Showing that weak solutions are smooth (classical, strong) requires equation-specific bootstrap arguments that break under perturbation. Each new equation, each new coupling term, demands a fresh proof.

Problem	Weak existence	Strong regularity	Status
3D Navier–Stokes	Leray, 1934	Open	Millennium Prize
3D Euler	De Lellis–Székelyhidi	Open (incompressible)	Onsager conjecture partially resolved
Boltzmann	DiPerna–Lions, 1989	Near-equilibrium only	Long-range open
SQG equation	Known	Open	Critical regularity

1.2 What IS a PDE solution?

Before introducing the Latent framework, we must be precise about what "solution" means for a PDE. This is not a pedantic question — the entire history of PDE theory is a sequence of expansions of this concept, each expansion motivated by the failure of the previous one.

Consider a PDE $\partial_t u = F(u)$ with initial data u_0.

Classical (strong) solution. A function $u \in C^k(\Omega \times [0,T])$ (with k matching the order of F) that satisfies the equation pointwise:

$$\partial_t u(x,t) = F(u)(x,t) \quad \text{for every } (x,t) \in \Omega \times (0,T)$$

This is the "obvious" solution concept. The problem: for most nonlinear PDEs, we cannot prove these exist globally.

Mild solution. A function $u \in C([0,T]; X)$ satisfying the integral equation (Duhamel):

$$u(t) = S(t)u_0 + \int_0^t S(t-s)F(u(s))\,ds$$

where $S(t)$ is the semigroup of the linear part. This avoids requiring spatial derivatives of u — the semigroup handles the differentiation. Mild solutions exist for longer times than classical solutions, but may not be smooth.

Weak solution. A function $u \in L^2([0,T]; H^1(\Omega))$ satisfying the equation tested against smooth functions φ:

$$\int_0^T \int_\Omega [-u \cdot \partial_t \varphi + \langle F(u), \varphi \rangle]\,dx\,dt = \int_\Omega u_0 \cdot \varphi(0)\,dx$$

All derivatives are transferred from u to the test function φ via integration by parts. This is the concept that guarantees existence (via Galerkin + compactness) for virtually all physically relevant PDEs. The price: weak solutions may be non-unique and non-smooth.

Distributional solution. A distribution $u \in \mathcal{D}'(\Omega \times [0,T])$ satisfying the equation in the distributional sense. This includes measure-valued solutions, Young measure solutions, and varifold solutions — progressively weaker notions used when even L^2 membership fails.

The hierarchy, from strongest to weakest:

$$\text{Classical} \subset \text{Mild} \subset \text{Weak} \subset \text{Distributional}$$

Each level answers the question "does a solution EXIST?" with increasing generality. Each level pays an increasing price in information loss: uniqueness may fail, regularity is weaker, and the physical meaning becomes diluted.

Where does the Latent fit?

The Latent does not add a new level to this hierarchy. It is ORTHOGONAL to it. Instead of asking "in what function space does u live?" (the regularity question), the Latent asks a different question:

> **How many numbers does u contain?**

For an analytic solution ($\sigma > 0$, equivalently $\rho = e^\sigma > 1$), the answer is $N^*(\varepsilon, \rho) = O(\log 1/\varepsilon)$. This is finite. The solution $u(x,t)$ is nominally a function on an infinite-dimensional space, but its actual information content — the number of independent parameters needed to reconstruct it to accuracy ε — is a finite number determined by the PDE itself.

The Latent solution is the trajectory $\Lambda(t) = (c_0(t), \dots, c_{N^*}(t))$ of these parameters. It is not an approximation to u that happens to work well — it is a SUFFICIENT STATISTIC for u, in the sense that no information is lost beyond the certified tolerance ε.

This reframes the classical hierarchy:

Solution concept	Question answered	Information
Classical	Does u exist, pointwise, and is it smooth?	Complete — but existence is hard

Solution concept	Question answered	Information
Mild	Does u satisfy the integral equation?	Complete, derivatives not needed
Weak	Does u satisfy the equation on average?	Partial — uniqueness may fail
Distributional	Does u exist as a generalized function?	Minimal — physics may be lost
Latent	**How many numbers does u require?**	**Certified finite — N^* numbers suffice**

The classical hierarchy asks "does the solution exist in progressively weaker senses?" The Latent asks "given that the solution exists and is analytic, what is its finite essence?" These are complementary, not competing, questions.

The deepest consequence: for an analytic PDE, **the solution IS finite-dimensional**. The PDE lives on an infinite-dimensional function space, but the actual dynamics live on an N^*-dimensional manifold. The infinite-dimensionality is an artifact of the continuum representation, not a feature of the dynamics.

1.3 Two levels of compression: the law and the state

A natural objection: isn't the PDE ITSELF already a "latent"? The Navier–Stokes equations fit in one line — yet they encode every fluid flow that has ever existed or will exist, from a dripping faucet to Jupiter's Great Red Spot. The PDE compresses the infinite complexity of physics into a finite rule.

Yes. There are two distinct compressions at work:

Level	What is compressed	Compressed into	Size
The PDE	All possible behaviors of the physical system	A finite rule: $\partial_t u = F(u)$	A few symbols
The Latent ODE	The entire solution family of the PDE	A finite dynamical system: $\dot{\Lambda} = G(\Lambda)$ on $\mathbb{R}^{N^*}$	$O((N^*)^K)$ coupling coefficients
One Latent trajectory	One specific solution $u(x,t)$ at accuracy ε	A finite state: $\Lambda(t) = (c_0, \ldots, c_{N^*})$	N^* numbers per time step

The PDE is the latent of the **law**. The Latent Λ is the latent of the **state**.

The PDE says: "this is the rule that all solutions must obey." The Latent says: "given this rule and this initial condition, here are the N^* numbers that specify what actually happens."

But a PDE does not have ONE solution — it has a family of solutions, one for each initial condition u_0. Different u_0 produce different solutions $u(x,t)$ and therefore different Latent trajectories $\Lambda(t)$. Does the Latent framework give a separate compression for each solution?

No. The framework is stronger than that. Three levels:

Object	What it captures	Depends on
The PDE: $\partial_t u = F(u)$	The rule	Nothing (fixed)
The Latent ODE: $\dot{c}_k = \sum L_{kj} c_j + \sum B_{k,j_1,j_2} c_{j_1} c_{j_2}$	The entire solution family	Only the PDE
One trajectory: $\Lambda(t) = (c_0(t), \dots, c_{N^*}(t))$	One specific solution	PDE + initial condition $\Lambda(0)$

The Latent ODE — the finite-dimensional ODE on $\mathbb{R}^{N^*}$ derived by projecting the PDE (§3) — is the SAME for all initial conditions. It is determined entirely by the PDE structure (the tensors L_{kj}, B_{k,j_1,j_2}, etc.). Different initial conditions u_0 produce different initial Latent vectors $\Lambda(0) = (c_0(0), \dots, c_{N^*}(0))$, which then evolve under the same ODE to produce different trajectories.

This means:

- **One Latent trajectory** replaces one PDE solution. (Compression of a state.)
- **The Latent ODE** replaces the entire PDE. (Compression of a dynamical system.)

The second statement is the stronger one: the infinite-dimensional PDE $\partial_t u = F(u)$ on a function space is faithfully replaced by a finite-dimensional ODE $\dot{\Lambda} = G(\Lambda)$ on $\mathbb{R}^{N^*}$, for ALL solutions with analyticity radius $\rho \geq \rho_{\min}$. Every analytic solution of the PDE corresponds to a solution of the Latent ODE, and vice versa, up to the certified error ε.

This connects to the concept of **inertial manifolds** (Foias–Sell–Temam, 1988): for certain dissipative PDEs, the long-time dynamics live on a finite-dimensional manifold in function space. The Latent framework provides a quantitative version: the manifold has dimension N^*, the dimension is computable from ρ, and the approximation error is exponentially small. See §6 for the precise relationship.

The deep connection between the two compression levels is the content of this paper: **the simplicity of the law creates compressibility in the state.** The PDE has finite grade K (its right-hand side is a polynomial of degree K) and its solutions have analyticity radius $\rho > 1$. These two structural properties — inherited from the law — are precisely what makes the state finitely compressible, and what makes the Latent ODE a faithful finite-dimensional surrogate.

A random function in L^2 cannot be compressed: it requires $O(\varepsilon^{-d})$ modes for accuracy ε. But a function that was GENERATED by an analytic, finite-grade PDE can be compressed exponentially: $N^* = O(\log 1/\varepsilon)$. The structure of the generating rule propagates through the dynamics and manifests as structure in the solution.

In information-theoretic language: the Kolmogorov complexity of the PDE is small (short description), and this forces the effective Kolmogorov complexity of its solutions to be small (finitely many parameters). The Latent framework makes this implication quantitative: grade K + analyticity ρ $\to N^*(\varepsilon, \rho)$ parameters suffice.

1.4 The Latent map: PDE → Latent ODE

The derivation of the Latent ODE from a PDE is fully constructive. It is a map

$$\mathcal{L} : (\text{PDE},\ \text{basis},\ \varepsilon) \ \longmapsto\ (\text{ODE on } \mathbb{R}^{N^*},\ \text{error} \leq \varepsilon)$$

that takes an analytic PDE and produces a certified finite-dimensional ODE. The map has five steps:

Step 1. Grade decomposition of the operator. Write the right-hand side of the PDE $\partial_t u = F(u)$ as a sum of k-linear operators (its Taylor expansion at $u = 0$):

$$F(u) = A^{(0)} + A^{(1)}[u] + A^{(2)}[u, u] + \cdots + A^{(K)}[\underbrace{u, \ldots, u}_{K}]$$

If F is a polynomial of degree K in u, the sum is finite. If F is transcendental (e.g., e^u, $\sin u$), the sum is infinite but converges with exponential bounds $\|A^{(k)}\|_{\mathrm{op}} \le C_0/\rho_F^k$. The grade K (or effective grade K_{eff}) is the output of this step.

Step 2. Choice of basis. Select an orthonormal basis $\{\varphi_k\}$ for $L^2(\Omega)$. For periodic domains: Fourier modes. For bounded domains: eigenfunctions of the linear part $A^{(1)}$. For irregular domains: wavelets or finite elements. The basis determines the coordinate system; the Latent is basis-dependent (like coordinates on a manifold).

Step 3. Determination of N^*. From the analyticity radius $\rho = e^\sigma$ of the initial data u_0 (or of the solution, if known a priori), compute:

$$N^* = \left\lceil \frac{\log(M \cdot C_d/\varepsilon)}{\log \rho} \right\rceil$$

This is the dimension of the Latent ODE. It depends on ρ (a property of the data/solution) and ε (the desired accuracy), but NOT on any external choice.

Step 4. Computation of coupling tensors. Project each k-linear operator onto the basis to obtain the coupling coefficients of the Latent ODE:

$$\begin{aligned} L_{kj} &= \langle A^{(1)}[\varphi_j], \varphi_k \rangle \qquad \text{(linear coupling)} \\ B_{k,j_1,j_2} &= \langle A^{(2)}[\varphi_{j_1}, \varphi_{j_2}], \varphi_k \rangle \qquad \text{(bilinear coupling)} \\ T^{(m)}_{k,j_1,\ldots,j_m} &= \langle A^{(m)}[\varphi_{j_1}, \ldots, \varphi_{j_m}], \varphi_k \rangle \qquad \text{(grade-}m\text{ coupling)} \end{aligned}$$

These are computable inner products. For the Fourier basis, many have explicit closed forms (e.g., the NS bilinear tensor reduces to $B_{k,j_1,j_2} \propto k \cdot \delta_{k,j_1+j_2}$, the convolution constraint).

Step 5. Assembly of the Latent ODE. The Latent ODE on $\mathbb{R}^{N^*}$ is:

$$\dot{c}_k = \sum_{|j| \le N^*} L_{kj}\, c_j + \sum_{|j_1|,|j_2| \le N^*} B_{k,j_1,j_2}\, c_{j_1} c_{j_2} + \cdots + \sum_{|j_1|,\ldots,|j_K| \le N^*} T^{(K)}_{k,j_1,\ldots,j_K}\, c_{j_1} \cdots c_{j_K}$$

for $|k| \le N^*$. This is a polynomial ODE of degree K on $\mathbb{R}^{N^*}$. Its solutions approximate the PDE solutions with error $\le \varepsilon$ (Theorems 1 and 3).

Worked example 1: Heat equation (grade 1). The PDE $\partial_t u = \kappa \Delta u$ on $\mathbb{T}^d$:

Step	Output
1. Grade decomposition	$A^{(1)}[u] = \kappa\Delta u$. Grade $K = 1$ (linear).
2. Basis	Fourier modes $\varphi_k(x) = e^{ik\cdot x}$
3. N^*	$\lceil \log(MC_d/\varepsilon)/\log\rho_0 \rceil$, where $\rho_0 = e^{\sigma_0}$ is the initial analyticity radius
4. Tensors	$L_{kj} = -\kappa\|k\|^2\delta_{kj}$ (diagonal — no coupling!)
5. Latent ODE	$\dot{c}_k = -\kappa\|k\|^2 c_k$ — each mode decays independently

The Latent ODE decouples completely. Closure error is exactly zero. The analyticity radius GROWS: $\sigma(t) \geq \sigma_0 + \kappa t$, so $N^*(t) \to 0$ — the solution becomes infinitely compressible.

Worked example 2: Viscous Burgers equation (grade 2). The PDE $\partial_t u = \nu\partial_{xx}u - u\,\partial_x u$ on $\mathbb{T}^1$:

Step	Output
1. Grade decomposition	$A^{(1)}[u] = \nu\partial_{xx}u$, $A^{(2)}[u, v] = -\frac{1}{2}(u\partial_x v + v\partial_x u)$. Grade $K = 2$.
2. Basis	Fourier modes $\varphi_k(x) = e^{ikx}$ on $\mathbb{T}^1$
3. N^*	$\lceil \log(M/\varepsilon)/\log\rho \rceil$
4. Tensors	$L_{kj} = -\nu k^2\delta_{kj}$, $B_{k,j_1,j_2} = -\frac{i}{2}(j_1 + j_2)\delta_{k,j_1+j_2}$
5. Latent ODE	$\dot{c}_k = -\nu k^2 c_k - \frac{i}{2}\sum_{j_1+j_2=k}(j_1 + j_2)\, c_{j_1} c_{j_2}$

The bilinear convolution constraint $j_1 + j_2 = k$ gives the ODE its structure. Viscosity (νk^2) damps high modes; the quadratic term transfers energy between modes. For $\nu > 0$, solutions remain analytic for all time ($\sigma(t) > 0$), so the Latent is valid globally. Burgers is the simplest grade-2 PDE — the same structure (linear dissipation + bilinear transfer) appears in Navier–Stokes, where the full derivation is given in the companion paper (Nagy, 2026, NS-Grade, §8.5).

The map $\mathcal{L}$ is deterministic: the same PDE + basis + ε always produce the same Latent ODE. Different PDEs produce different Latent ODEs. The map preserves the structure of F — a grade-K PDE always yields a degree-K polynomial ODE. Sections 2–4 develop the formal construction and prove the four structural theorems.

1.5 Galerkin is a method. The Latent is a property.

The Galerkin method — projecting a PDE onto finitely many basis functions and solving the resulting ODE — has been the workhorse of PDE theory and numerical analysis since Galerkin (1915) and Leray (1934). The distinction between the Latent framework and Galerkin is not mechanical — the projection step is identical. The distinction is conceptual:

Galerkin asks: "How do I approximate this PDE?" You choose a truncation level N, project, solve the ODE, and hope the result is accurate. The quality of the approximation depends on your choice and the specifics of the problem.

The Latent asks: "How many numbers does this PDE actually need?" The answer is $N^*(\varepsilon, \rho) = \lceil \log(C_0/\varepsilon)/\log\rho \rceil$. This is not a choice — it is an intrinsic property of the PDE and its data, determined by the analyticity radius ρ. The N^* modes are not an approximation of the solution; they are a **sufficient statistic** for the solution to accuracy ε.

The analogy is image compression:

- **JPEG**: "I will compress this image to 100 KB." An external choice. The quality depends on whether 100 KB is enough, which you don't know a priori.
- **Rate-distortion theory**: "This image has information content 100 KB at distortion ε." An intrinsic property. The image REQUIRES this many bits, regardless of the compression algorithm.

Galerkin is JPEG — a compression method. The Latent is rate-distortion theory — a representation theorem. The Galerkin method is how you COMPUTE the Latent representation, but N^* exists as a property of the solution independently of how you compute it.

This conceptual difference has four concrete structural consequences:

Property	Standard Galerkin	Latent Galerkin
Truncation level N	Chosen externally, ad hoc	Self-determined: $N^* = \lceil \log(C_0/\varepsilon)/\log\rho \rceil$
Truncation error	Asymptotic, problem-dependent	Certified: $\leq \varepsilon$, exponential in N^*
Mode coupling	All-to-all (N^K terms)	Grade-aware: high-high suppressed by ρ^{-2N^*}
Closure	Uncontrolled (spurious dynamics possible)	Bounded: resolved-unresolved coupling $\leq O(\varepsilon)$
Perturbation $F \to F + G$	Requires new analysis	ρ changes by computable amount
Degradation warning	None — you don't know when N is too small	Built-in: $\rho(t) \to 1$ signals loss of compressibility

The closure row deserves emphasis. In standard Galerkin, the truncated system can LIE: it may have different attractors, different stability properties, and artificial energy pileup at the truncation wavenumber. This is why turbulence simulations require subgrid models — the Galerkin truncation introduces artifacts that must be corrected by hand. In the Latent framework, the truncated system is FAITHFUL: the grade decay guarantees that unresolved modes affect resolved modes at amplitude $O(\varepsilon)$. No subgrid model is needed.

In short: the same machinery, but the Latent UNDERSTANDS what Galerkin merely DOES.

1.6 Two layers: the Latent of u and the grade of F

A precise statement requires distinguishing two related but conceptually distinct structures:

The spectral Latent of the solution u. An analytic function u on a periodic domain $\Omega = \mathbb{T}^d$ has Fourier coefficients satisfying $|\hat{u}_k| \leq Ce^{-\sigma|k|}$, where $\sigma > 0$ is the width of the analyticity strip in the complex plane. Setting $\rho = e^{\sigma}$, this becomes $|\hat{u}_k| \leq C/\rho^{|k|}$. The Latent of u is the vector of these coefficients, truncated at N^* where the tail is below ε.

The grade structure of the operator F. The right-hand side of the PDE $\partial_t u = F(u)$, viewed as a map on function space, has a grade decomposition: $F(u) = A^{(0)} + A^{(1)}[u] + A^{(2)}[u,u] + \cdots$, where $A^{(k)}$ is the k-th Fréchet derivative — the k-body interaction term. If F is a polynomial of degree K in u, then $A^{(k)} = 0$ for $k > K$.

The spectral Latent tells you HOW MANY coefficients to keep (N^*, from the decay rate of u's Fourier modes). The grade structure tells you HOW THEY INTERACT (through at most K-body couplings, from the polynomial structure of F).

1.7 Anticipating the objection: "This is just Foias–Temam"

Foias and Temam (1989) proved that Navier–Stokes solutions with analytic data maintain a positive analyticity radius $\sigma(t) > 0$ (equivalently $\rho(t) = e^{\sigma(t)} > 1$) as long as the solution is smooth. This is the founding result of Gevrey regularity for NS.

The Latent framework goes beyond Foias–Temam in four specific ways:

1. **Explicit N^*.** Foias–Temam proved $\sigma(t) > 0$ but did not compute the number of modes needed for a given accuracy. We give the explicit formula $N^* = \lceil \log(C_0/\varepsilon)/\log \rho \rceil$.
2. **Grade coupling structure.** Foias–Temam did not decompose the Galerkin ODE by interaction order. The Grade Galerkin Theorem (§4) shows that the coupling tensor has exponentially decaying entries, making the truncated ODE provably sparse.
3. **Exponential closure bound.** Foias–Temam did not bound the closure error of the truncated system. Theorem 3 (§3.3) shows the unresolved modes affect the resolved modes at amplitude $O(\varepsilon)$.
4. **Perturbation stability.** Foias–Temam studied NS specifically. The grade framework applies to ANY analytic PDE, and the combined analyticity radius under perturbation is bounded by $\rho_{F+G} \geq \rho_F \rho_G/(\rho_F + \rho_G)$ (Nagy, 2026, Grade Regularity).
5. **Interpolation.** The continuous interpolation between exponential compression ($\rho > 1$) and no compression ($\rho \leq 1$) is not present in Foias–Temam.

1.8 Structure of the paper

Section 2 defines the spectral Latent and the self-determined truncation. Section 3 proves the four structural theorems. Section 4 develops the Grade Galerkin method. Section 5 proves the Interpolation Theorem and identifies the phase transition at $\rho = 1$. Section 6 relates the framework to classical PDE theory (Leray, Foias–Temam, inertial manifolds, determining modes). Section 7 gives worked examples. Section 8 discusses limitations.

1.9 What this paper does NOT claim

- We do not introduce a "new solution concept" in the sense of Schwartz distributions. The Latent solution is a quantitative refinement of the Galerkin projection, not a new mathematical object.
- We do not solve the Navier–Stokes millennium problem. The question "does $\rho(t) > 0$ persist for all time?" remains open.
- We do not claim the framework is useful for non-analytic data. For L^2 or Sobolev data, the exponential compression vanishes and the Latent reduces to standard Galerkin.

---

2. The Spectral Latent

2.1 Spectral coefficients of analytic functions

Let $\Omega = \mathbb{T}^d = [0, 2\pi]^d$ be the d-dimensional torus and $\{\varphi_k\}_{k \in \mathbb{Z}^d}$ the standard Fourier basis. For $u \in L^2(\Omega; \mathbb{R}^m)$, the spectral coefficients are:

$$c_k = \langle u, \varphi_k \rangle_{L^2}, \qquad k \in \mathbb{Z}^d$$

If u is real-analytic on Ω, it extends to a holomorphic function on a strip $\{z \in \mathbb{C}^d : |\mathrm{Im}(z_j)| < \sigma,\ j = 1, \dots, d\}$ for some $\sigma > 0$. The Paley–Wiener theorem gives:

$$|c_k| \leq M\, e^{-\sigma |k|} \tag{SD}$$

where $M = \sup_{|\mathrm{Im}(z)| < \sigma} |u(z)|$ and $|k| = |k_1| + \cdots + |k_d|$.

Definition 1 (Analyticity radius). The analyticity radius of u is $\rho := e^{\sigma}$, where σ is the maximal strip width. The spectral decay (SD) becomes:

$$|c_k| \leq \frac{M}{\rho^{|k|}}$$

2.2 The truncated spectral Latent

Definition 2 (Spectral Latent). The **spectral Latent** of u at accuracy ε is:

$$\Lambda^{N^*}(u) := \{c_k : |k| \leq N^*\} \in \mathbb{R}^{|\mathcal{K}_{N^*}|}$$

where $\mathcal{K}_{N^*} = \{k \in \mathbb{Z}^d : |k| \leq N^*\}$ and:

$$N^*(\varepsilon, \rho) = \left\lceil \frac{\log(M \cdot C_d / \varepsilon)}{\log \rho} \right\rceil$$

with C_d a dimensional constant from the tail sum. The cardinality $|\mathcal{K}_{N^*}| = O((N^*)^d)$.

Theorem 1 (Self-determined sufficiency). The truncated spectral Latent is an ε-sufficient representation:

$$\left\| u - \sum_{|k| \leq N^*} c_k \varphi_k \right\|_{L^2(\Omega)} \leq \varepsilon$$

Proof. The tail satisfies:

$$\sum_{|k| > N^*} |c_k|^2 \leq M^2 \sum_{|k| > N^*} \rho^{-2|k|} \leq M^2 \cdot \frac{C_d\, \rho^{-2N^*}}{1 - \rho^{-2}}$$

By choice of N^*, this is at most ε^2. □

Remark. The dimension $|\mathcal{K}_{N^*}| = O((N^*)^d) = O((\log 1/\varepsilon)^d)$ — polylogarithmic in $1/\varepsilon$. Compare with standard L^2 approximation: $O(\varepsilon^{-d})$ modes needed for the same accuracy without analyticity.

2.3 Time-dependent Latent: the coefficient trajectory

For an evolution PDE $\partial_t u = F(u)$ with solution $u(x,t)$, the spectral coefficients evolve:

$$\dot{c}_k(t) = \langle F(u(\cdot,t)), \varphi_k \rangle, \qquad k \in \mathbb{Z}^d$$

This is a (generally infinite-dimensional) ODE on the sequence space ℓ^2. The **Latent trajectory** is the restriction to the resolved modes:

$$\Lambda(t) = \{c_k(t) : |k| \le N^*\}$$

The central question: does the truncated trajectory faithfully represent the full dynamics?

---

3. The Four Structural Theorems

3.1 Theorem 2: Grade-structured coupling

The structure of the Latent ODE depends on the polynomial degree of F — its **grade**.

Definition 3 (Grade of a PDE). The PDE $\partial_t u = F(u)$ has **grade** K if the right-hand side, viewed as a map $F : X \to X$ on a function space X, is a polynomial of degree K in u:

$$F(u) = A^{(0)} + A^{(1)}[u] + A^{(2)}[u,u] + \cdots + A^{(K)}[\underbrace{u,\ldots,u}_{K}]$$

where $A^{(k)} : X^k \to X$ is a bounded k-linear operator (the k-th Fréchet derivative of F at $u = 0$, divided by $k!$).

PDE	$F(u)$	Grade	Structure
Heat	$\kappa\Delta u$	1	$A^{(1)}[u] = \kappa\Delta u$
Navier–Stokes	$\nu\Delta u - \mathbb{P}(u\cdot\nabla u)$	2	$A^{(1)} = \nu\Delta$, $A^{(2)} = -\mathbb{P}(\cdot\nabla\cdot)$
Euler	$-\mathbb{P}(u\cdot\nabla u)$	2	$A^{(2)} = -\mathbb{P}(\cdot\nabla\cdot)$
Cubic NLS	$i\Delta\psi + i\|\psi\|^2\psi$	3	$A^{(1)} = i\Delta$, $A^{(3)} = i\|\cdot\|^2\cdot$
Boltzmann	$Q(f,f)$	2	$A^{(2)} = Q(\cdot,\cdot)$

Theorem 2 (Grade Galerkin). If the PDE has grade K, the Latent ODE on the resolved modes $\{c_k : |k| \le N^*\}$ has the form:

$$\dot{c}_k = \sum_{|j|\le N^*} L_{kj}\, c_j + \sum_{\substack{|j_1|,|j_2|\le N^* \\ |j_1+j_2|\le N^*}} B_{k,j_1,j_2}\, c_{j_1} c_{j_2} + \cdots + \sum_{|j_1|,\dots,|j_K|\le N^*} T^{(K)}_{k,j_1,\dots,j_K}\, c_{j_1}\cdots c_{j_K}$$

with at most K-body couplings. The interaction tensors are:

$$L_{kj} = \langle A^{(1)}[\varphi_j], \varphi_k\rangle, \quad B_{k,j_1,j_2} = \langle A^{(2)}[\varphi_{j_1}, \varphi_{j_2}], \varphi_k\rangle, \quad \dots$$

Proof. The Galerkin projection of a polynomial is a polynomial of the same degree. Since F is a polynomial of degree K in u, the projected equation involves products of at most K coefficients. □

Remark (Infinite-grade PDEs). For PDEs with transcendental nonlinearities (e.g., $F(u) = e^u$, $\sin(u)$), the grade is infinite. The grade decomposition has infinitely many terms, but with exponential decay: $\|A^{(k)}\|_{\mathrm{op}} \le C_0/\rho_F^k$, where ρ_F is the analyticity radius of F as a map on X. The Grade Galerkin theorem generalizes: the k-body coupling contributes at amplitude $O(1/\rho_F^k)$, so truncating the grade expansion at $K_{\mathrm{eff}} = \lceil \log(C_0/\varepsilon)/\log\rho_F \rceil$ gives an ε-accurate Latent ODE.

3.2 Theorem 3: Exponential closure

The CLOSURE PROBLEM is the central difficulty of Galerkin methods: the truncated system evolves differently from the full system because the unresolved modes ($|k| > N^*$) interact with the resolved modes ($|k| \le N^*$). In standard Galerkin, this closure error is uncontrolled and can introduce spurious dynamics (e.g., spurious energy pileup at the truncation wavenumber in turbulence simulations).

The grade framework bounds this closure error.

Theorem 3 (Exponential closure). Let $c_k^{(N^*)}(t)$ denote the solution of the truncated Latent ODE (modes $|k| \le N^*$) and $c_k(t)$ the exact spectral coefficients. If the exact solution has analyticity radius $\rho(t) \ge \rho_{\min} > 1$ on $[0, T)$, then for all $|k| \le N^*$:

$$|c_k^{(N^*)}(t) - c_k(t)| \le C(T) \cdot \frac{1}{\rho_{\min}^{N^*}} \le C(T)\cdot\varepsilon$$

on $[0, T_\varepsilon)$, where $C(T)$ depends on $\|u\|$, the PDE structure, and T, and $T_\varepsilon \to T$ as $\varepsilon \to 0$.

Proof sketch. The resolved-unresolved coupling in a grade-K PDE enters through terms like $B_{k,j_1,j_2} c_{j_1} c_{j_2}$ where $|j_1|$ or $|j_2| > N^*$. By the spectral decay (SD), each such term is bounded by $|B_{k,j_1,j_2}| \cdot M^2/\rho^{|j_1|+|j_2|}$. Since $|j_1| + |j_2| > N^*$ for at least one index, the total coupling is $O(1/\rho^{N^*})$. Gronwall's inequality propagates this to $O(C(T)/\rho^{N^*})$. □

Why this matters. In standard Galerkin, the truncated system may have qualitatively different dynamics from the true system — different attractors, different stability, energy piling up at the truncation wavenumber. The exponential closure bound guarantees this cannot happen for the Latent truncation: the resolved dynamics track the true dynamics to accuracy ε for times up to T_ε. The truncation is FAITHFUL, not merely a projection.

3.3 Perturbation stability (inherited from Grade Regularity)

The fourth structural property — perturbation stability of ρ under modifications $F \to F + G$ — is proved in the companion paper (Nagy, 2026, Grade Regularity). We state the result:

Theorem (Perturbation Stability — Grade Regularity, Thm 3). If F has analyticity radius ρ_F and G has analyticity radius ρ_G, then $F + G$ has analyticity radius $\rho_{F+G} \geq \rho_F \rho_G / (\rho_F + \rho_G)$.

For the Latent framework: adding a smooth perturbation G to the PDE changes N^* by a bounded amount (since ρ changes by a bounded factor) but preserves the entire structure — grade coupling, exponential closure, self-determined truncation. No part of the analysis needs to be rebuilt.

---

4. The Latent ODE for Grade-2 PDEs

Grade-2 PDEs — Navier–Stokes, Euler, Boltzmann, Einstein in harmonic gauge — are the most important class. We develop the Latent ODE in detail.

4.1 The quadratic Latent ODE

For a grade-2 PDE with linear part L and bilinear part B, the Latent ODE on $\{c_k : |k| \leq N^*\}$ is:

$$\dot{c}_k = \sum_{|j| \leq N^*} L_{kj}\, c_j + \sum_{\substack{j_1 + j_2 = k \\ |j_1|, |j_2| \leq N^*}} B_{k, j_1, j_2}\, c_{j_1}\, c_{j_2} \tag{LO}$$

The convolution constraint $j_1 + j_2 = k$ arises because $\langle (\varphi_{j_1} \cdot \nabla \varphi_{j_2}), \varphi_k \rangle = 0$ unless $j_1 + j_2 = k$ (for Fourier modes). This is the spectral signature of the bilinear interaction.

4.2 Effective sparsity from spectral decay

The bilinear tensor B_{k, j_1, j_2} couples three modes. Not all couplings are dynamically significant. The spectral decay bound gives:

$$|B_{k, j_1, j_2}\, c_{j_1}\, c_{j_2}| \leq |B_{k, j_1, j_2}| \cdot M^2 \cdot \rho^{-(|j_1| + |j_2|)}$$

For high-high interactions ($|j_1|, |j_2| > N^*/2$), the amplitude is $O(\rho^{-N^*}) = O(\varepsilon)$. The dynamically significant interactions are:

- **Low-low** ($|j_1|, |j_2| \leq N^*/2$): fully resolved, $O(1)$ amplitude.
- **Low-high** ($|j_1| \leq N^*/2$, $|j_2| > N^*/2$, or vice versa): partially resolved, amplitude decays as $\rho^{-|j_2|}$.
- **High-high** ($|j_1|, |j_2| > N^*/2$): exponentially suppressed, $O(\varepsilon)$.

This is the grade-theoretic explanation of a well-known phenomenon in turbulence: the energy cascade is carried by the low-high interaction (the sweeping of small eddies by large ones), while the high-high self-interaction is negligible. In the Latent framework, this is not an empirical observation but a THEOREM from the spectral decay.

For Navier–Stokes specifically, this recovers the high-high suppression result: $\|\mathbb{P}(u_{\text{hi}} \cdot \nabla u_{\text{hi}})\|_{L^2} \leq C_0^2/\rho^{2K}$ for modes $|k| > K$ (Nagy, 2026, NS-Grade).

4.3 Computational complexity

Method	Unknowns	Bilinear terms	Scaling
DNS on mesh h	$O(h^{-d})$	$O(h^{-2d})$	Algebraic in $1/h$
Standard Galerkin (N modes)	$O(N^d)$	$O(N^{2d})$	Polynomial in N
Latent Galerkin (N^* modes)	$O((\log 1/\varepsilon)^d)$	$O((\log 1/\varepsilon)^{2d})$	**Polylogarithmic**

The exponential compression from analyticity makes the Latent Galerkin system polylogarithmic in $1/\varepsilon$ — exponentially smaller than standard methods at the same accuracy.

4.4 The radius ODE: when does the Latent persist?

The analyticity radius $\rho(t)$ (equivalently $\sigma(t) = \log \rho(t)$) evolves according to:

$$\frac{d\sigma}{dt} \geq R_+(\sigma, u) - R_-(\sigma, u)$$

where R_+ is the linear restoring (viscous smoothing increases σ) and R_- is the nonlinear straining (grade-2 interactions decrease σ). The Latent persists — meaning N^* remains finite and the truncation remains valid — as long as $\sigma(t) > 0$, i.e., $\rho(t) > 1$.

PDE	R_+ (restoring)	R_- (straining)	Persistence
Heat equation	κ/σ	0	Always ($\sigma \to \infty$)
Navier–Stokes	ν/σ	$C\|u\|_\sigma^2$	Open (millennium problem)
Euler	0	$C\|u\|_\sigma^2$	Shocks possible ($\sigma \to 0$)
Cubic NLS	0	$C\|u\|_\sigma^3$	Dimension-dependent

The Latent framework does not ANSWER the persistence question — it REFORMULATES it as the finite-dimensional question: does the ODE (LO) have a solution whose coefficients maintain exponential decay?

---

5. The Interpolation Theorem

5.1 Three regimes of coefficient decay

The quality of the Latent representation depends on the smoothness of the data through the coefficient decay rate:

Regime A: Analytic data (exponential decay). u extends holomorphically to a strip of width $\sigma > 0$. Coefficients satisfy $|c_k| \leq Me^{-\sigma|k|}$. The Latent has $N^* = O(\sigma^{-1}\log(M/\varepsilon))$ modes — logarithmic in $1/\varepsilon$.

Regime B: Smooth data (algebraic decay). $u \in C^s(\Omega)$ with $s < \infty$. Integration by parts gives $|c_k| \leq C_s\|u\|_{C^s}/|k|^s$. No analyticity strip exists ($\sigma = 0$, $\rho = 1$). The Latent truncation at N^* requires $\sum_{|k|>N^*} C_s^2/|k|^{2s} \leq \varepsilon^2$, giving $N^* = O(\varepsilon^{-d/(2s-d)})$ for $s > d/2$ — polynomial in $1/\varepsilon$.

Regime C: Rough data (no decay). $u \in L^2(\Omega)$ with no smoothness. Coefficients satisfy only $\sum |c_k|^2 < \infty$. The truncation requires $N^* = O(\varepsilon^{-d})$ modes — the worst-case rate. No compression is possible. The Latent reduces to the standard Fourier-Galerkin projection.

Theorem 4 (Spectral interpolation). The transition from Regime A to Regime C is characterized by the coefficient decay rate α defined by $|c_k| \sim e^{-\alpha(|k|)}$:

- $\alpha(n) = \sigma n$ (linear) $\Leftrightarrow$ Regime A. $N^* = O(\log 1/\varepsilon)$.
- $\alpha(n) = s\log n$ (logarithmic) $\Leftrightarrow$ Regime B. $N^* = O(\varepsilon^{-d/(2s-d)})$.
- $\alpha(n) = 0$ (no decay) $\Leftrightarrow$ Regime C. $N^* = O(\varepsilon^{-d})$.

5.2 The phase transition at the analyticity boundary

The transition from Regime A to Regime B is discontinuous in the compression rate:

$$\frac{dN^*}{d\varepsilon} = \begin{cases} O(1/\varepsilon) & \text{Regime A (logarithmic derivative)} \\ O(\varepsilon^{-d/(2s-d)-1}) & \text{Regime B (polynomial derivative)} \end{cases}$$

At $\sigma = 0$ (equivalently $\rho = 1$), the compression rate jumps. This is the **Latent phase transition**: the solution passes from exponentially compressible to only polynomially compressible. In dynamical terms: a solution whose analyticity radius $\rho(t)$ drops to 1 has lost its finite Latent representation. The system transitions from a certified finite-dimensional problem to an infinite-dimensional one.

This phase transition is the grade-theoretic signature of blowup: the solution becomes incompressible precisely when the analyticity strip collapses.

5.3 The Latent interpolation diagram

The three regimes can be organized by their representational efficiency:

	Modes needed	Grade coupling useful?	Closure bounded?	Framework
A ($\rho > 1$)	$O((\log 1/\varepsilon)^d)$	Yes (exponential sparsity)	Yes ($O(\varepsilon)$)	**Latent**
B ($\rho = 1$)	$O(\varepsilon^{-d/(2s-d)})$	Partially (algebraic sparsity)	Weakly	Spectral Galerkin
C ($\rho < 1$)	$O(\varepsilon^{-d})$	No	No	Standard Galerkin

The Latent framework is most powerful in Regime A, progressively weaker in Regime B, and reduces to standard methods in Regime C. This is honest: the framework's power derives from analyticity, and it gracefully degrades as analyticity is lost.

---

6. Relation to Classical PDE Theory

6.1 Leray's construction as the $\rho \to 1$ limit

Leray (1934) proved existence of weak NS solutions by the Galerkin method: project onto N Fourier modes, solve the ODE, extract a weak-$*$ limit as $N \to \infty$. The limiting process discards structural information — the subsequence is non-constructive, and the convergence is only in the weak topology.

The Latent framework, when applicable ($\rho > 1$), replaces Leray's non-constructive limit with a constructive truncation: set $N = N^*(\varepsilon, \rho)$ and solve the ODE ONCE. The convergence is STRONG (L^2 norm, not weak-$*$), and the error is explicitly bounded.

As $\rho \to 1$ (analyticity radius shrinks to zero), $N^* \to \infty$ and the Latent construction degenerates into Leray's: you must take $N \to \infty$ and extract a weak limit. Leray's construction is the limiting case of the Latent at the analyticity boundary.

6.2 Foias–Temam as the precursor

Foias and Temam (1989) proved $\sigma(t) > 0$ for smooth NS solutions — the existence of a Gevrey regularity radius. Their result establishes that Regime A persists as long as the solution is smooth.

The Latent framework takes the Foias–Temam radius $\sigma(t)$ and builds a quantitative theory on it: the explicit N^*, the grade coupling structure, the exponential closure bound, and the perturbation stability. The relationship:

$$\text{Foias–Temam: } \sigma(t) > 0 \qquad \Longrightarrow \qquad \text{Latent: } N^*(t) < \infty \text{ with explicit bound}$$

6.3 Inertial manifolds

The theory of inertial manifolds (Foias, Sell, Temam, 1988) proved that certain dissipative PDEs have finite-dimensional long-time dynamics: the solution converges to a finite-dimensional manifold in function space.

The Latent representation is related but distinct in three ways:

1. **Timing.** Inertial manifolds describe ASYMPTOTIC behavior ($t \to \infty$). The Latent representation is valid IMMEDIATELY, from $t = 0$.
2. **Hypothesis.** Inertial manifolds require a SPECTRAL GAP condition on the linear operator. The Latent requires ANALYTICITY of the data.
3. **Dimension.** The inertial manifold dimension is bounded above but rarely computed explicitly. The Latent dimension N^* is given by an explicit formula.

6.4 Determining modes

Foias and Prodi (1967) proved that for 2D NS, the first N_0 Fourier modes determine the solution: if two solutions agree on modes $|k| \le N_0$ for $t > T_0$, they agree on all modes.

The Latent framework makes this result QUANTITATIVE and IMMEDIATE: - **Quantitative**: $N_0 = N^*(\varepsilon, \rho)$ with explicit formula. - **Immediate**: the agreement holds from $t = 0$ to accuracy ε, not just asymptotically. - **General**: works for any analytic PDE, not just 2D NS.

---

7. Examples

7.1 The heat equation: perfect compression

For $\partial_t u = \kappa \Delta u$ on $\mathbb{T}^d$, the exact solution in Fourier space is:

$$c_k(t) = c_k(0)\, e^{-\kappa |k|^2 t}$$

The analyticity strip width GROWS: for initial data with $|c_k(0)| \le M e^{-\sigma_0 |k|}$, the solution has $|c_k(t)| \le M e^{-(\sigma_0 + \kappa |k| t)|k|}$, so $\sigma(t) \ge \sigma_0$ and the effective decay is super-exponential. The Latent becomes MORE compressed over time.

The Latent ODE decouples completely (grade 1 — linear PDE):

$$\dot{c}_k = -\kappa |k|^2 c_k$$

with explicit solution. No coupling between modes. The closure error is exactly zero.

7.2 Viscous Burgers equation: the simplest grade-2 Latent

The Burgers equation $\partial_t u = \nu \partial_{xx} u - u \partial_x u$ on $\mathbb{T}^1$ is the prototype grade-2 PDE. The Latent ODE is:

$$\dot{c}_k = -\nu k^2 c_k - \frac{i}{2} \sum_{j_1 + j_2 = k} (j_1 + j_2)\, c_{j_1} c_{j_2}$$

The structure is transparent: linear dissipation ($-\nu k^2 c_k$) versus bilinear energy transfer ($\sum c_{j_1} c_{j_2}$). For $\nu > 0$, the dissipation wins at high modes, and $\sigma(t) > 0$ for all time — the Latent representation is globally valid. As $\nu \to 0$, shocks form and $\sigma(t) \to 0$ in finite time — the Latent degrades.

This is the 1D analogue of the Navier–Stokes Latent ODE. The same grade-2 structure (dissipation + convolution-constrained bilinear coupling) appears there, with the added complexity of incompressibility and the Leray projection. The full NS Latent derivation is given in the companion paper (Nagy, 2026, NS-Grade, §8.5).

7.3 Cubic NLS: grade-3 coupling

$$i\partial_t\psi + \Delta\psi \pm |\psi|^2\psi = 0$$

Grade 3: the cubic nonlinearity $|\psi|^2\psi = \bar{\psi}\psi\psi$ produces 3-body mode interactions. The Latent ODE has $O(|\mathcal{K}_{N^*}|^3)$ cubic coupling terms.

The + sign (defocusing) is globally well-posed in 1D and 2D: $\sigma(t) > 0$ persists. The Latent representation is valid for all time.

The − sign (focusing) admits blowup in 2D and above: $\sigma(t) \to 0$ in finite time. The Latent representation degrades — $N^* \to \infty$ — as the solution approaches the singularity. The phase transition at $\sigma = 0$ marks the collapse of finite representability.

7.4 Reaction-diffusion: high-grade coupling

$$\partial_t u_i = d_i \Delta u_i + f_i(u_1, \dots, u_m)$$

If f_i is a polynomial of degree K, the PDE has grade K. The Latent ODE has K-body coupling. For $K = 2$ (mass-action kinetics), the structure mirrors NS. For $K > 2$ (Hill kinetics, cooperative binding), higher-order interactions appear.

The grade hierarchy provides a universal diagnostic: if $d_{\min}/\sigma$ (linear restoring) exceeds the grade-K straining rate, regularity is preserved. Adding new reaction terms of degree $\leq K$ does not change the grade — the Latent ODE structure is unchanged, and only the coupling coefficients adjust.

---

8. Limitations and Honest Scoping

8.1 What the Latent framework provides

- A **self-determined, certified** finite-dimensional encoding of PDE solutions when data is analytic.
- **Grade-aware coupling** that provably sparsifies the truncated ODE.
- **Exponential closure bounds** guaranteeing the truncated system faithfully represents the full dynamics.
- **Perturbation stability**: modifying the PDE changes ρ (and hence N^*) but preserves the structure.
- **Continuous interpolation** from exponential compression (analytic) to no compression (L^2).

8.2 What it does NOT provide

- It does NOT prove global regularity for any specific PDE. The question "$\rho(t) > 0$ for all t?" is the millennium problem.
- It is NOT a numerical method. The framework is theoretical; practical implementation requires efficient computation of the bilinear tensor B_{k,j_1,j_2} and real-time estimation of $\rho(t)$.
- It does NOT help when analyticity is absent. For shock waves, free boundaries, and measure-valued solutions, the compression advantage vanishes entirely.
- It does NOT replace weak solution theory. Distributions remain necessary for non-smooth regimes.

8.3 The honest comparison

Property	Standard Galerkin	Latent Galerkin
Truncation level	External choice	Self-determined by ρ
Truncation error	Uncontrolled	Exponentially bounded
Mode coupling	All-to-all	Grade-sparse (high-high suppressed)
Closure error	May introduce spurious dynamics	$O(\varepsilon)$ (faithful)
Perturbation	Requires new analysis	ρ changes, framework persists
Requires analyticity?	No	**Yes** (for exponential compression)

The last row is the tradeoff. Standard Galerkin works for any L^2 data but provides no structural guarantees. The Latent framework provides all four structural guarantees but requires analytic data. The honest verdict: **use the Latent when the data is smooth enough to justify it; use standard Galerkin when it isn't.**

---

9. Discussion

9.1 The Latent as quantitative Galerkin

The Latent solution is best understood not as a new mathematical object but as the Galerkin method equipped with four quantitative theorems derived from analyticity and the grade structure. The Galerkin method provides the mechanism (project and solve an ODE). The Latent framework provides the control (how many modes, how they couple, how accurate the truncation is, and how perturbations propagate).

This positioning is deliberate. Distributions (Schwartz, 1950) genuinely extended the function concept — they allow operations (differentiation of discontinuous functions) that are impossible in classical analysis. The Latent does not extend the function concept. It compresses it: for analytic functions, the infinite-dimensional object is faithfully captured by finitely many numbers. The novelty is in the CERTIFICATION (the four theorems), not in the object itself.

9.2 Why the distinction matters

The Galerkin-vs-Latent comparison (§1.5) invites a natural objection: "If the projection step is the same, what is new?" Three consequences of the conceptual shift deserve emphasis.

Truthfulness of the truncated system. In standard Galerkin, the truncated ODE is a different dynamical system from the PDE. Its long-time behavior may have artificial attractors, spurious symmetry-breaking, or energy pileup at the truncation wavenumber — pathologies that turbulence modelers spend entire careers correcting with ad hoc subgrid models. In the Latent framework, Theorem 3 (exponential closure) guarantees that the truncated ODE is faithful to the PDE at order $O(\varepsilon)$. Spurious dynamics require the resolved-unresolved coupling to overcome an exponential barrier. The Latent ODE is not an approximation to the PDE — it is the PDE, compressed.

A built-in diagnostic. Galerkin provides no signal when the truncation is inadequate. The numerics may look plausible while being wrong. The Latent framework comes with a built-in warning system: the analyticity radius $\rho(t)$. If $\rho(t) \to 1$ (equivalently $\sigma(t) \to 0$), the compression ratio degrades and $N^*(\varepsilon, \rho)$ grows — signaling that the finite representation is losing fidelity. This is not a heuristic; it follows from Theorem 1. Every Latent computation carries its own quality certificate.

Perturbation stability without rework. Adding a term G to the PDE $\partial_t u = F(u)$ requires, in standard Galerkin, an entirely new convergence analysis (new estimates, new compactness arguments, possibly new function spaces). In the Latent framework, the effect of G is absorbed into a change of analyticity radius: $\rho_{F+G} \geq (\rho_F^{-1} + \rho_G^{-1})^{-1}$. The four structural theorems still hold with the updated ρ. This modularity is what makes the framework universal rather than equation-specific.

9.3 The ρ-persistence conjecture

We conjecture that for physically reasonable dissipative PDEs (energy-dissipating, mass-conserving), the analyticity radius $\sigma(t)$ stays positive for all time.

Evidence: - Heat equation: $\sigma(t) \to \infty$ (parabolic smoothing). - 2D NS: $\sigma(t) > 0$ for all time (the 2D regularity problem is solved). - 3D NS with small data: $\sigma(t) > c > 0$ (classical perturbative result). - Boltzmann near equilibrium: $\sigma(t) > 0$ (perturbative theory).

Counterevidence: - 3D Euler (no viscosity): $\sigma(t) \to 0$ is possible (shocks in compressible; open for incompressible). - Focusing NLS in $d \geq 2$: blowup solutions with $\sigma(t) \to 0$.

The conjecture is STRONGER than the NS millennium problem (which is one instance). A proof would establish global regularity for a broad class of dissipative PDEs simultaneously.

9.4 Open problems

1. **ρ-persistence for NS.** Does $\sigma(t) > 0$ for all time for 3D NS with smooth data? (Equivalent to the millennium problem.)
2. **Optimal basis.** The Latent depends on the basis $\{\varphi_k\}$. Can a data-adapted basis (e.g., empirical eigenfunctions, or wavelet packets) improve N^*?
3. **Practical ρ-estimation.** Can $\sigma(t)$ be estimated from DNS output or experimental data, enabling real-time monitoring of the Latent phase transition?
4. **Non-analytic extensions.** Can the grade coupling and closure bounds be extended to Gevrey ($\sigma > 0$ but sub-exponential decay) or Sobolev data with meaningful (non-trivial) improvement over standard Galerkin?
5. **Stochastic PDEs.** For SPDEs with additive or multiplicative noise, how does noise affect $\sigma(t)$? Does white noise destroy analyticity (likely) or does it average out?
6. **The closure problem in detail.** Theorem 3 bounds the closure error in L^2. Can stronger norms (H^s, L^∞) be controlled? What about long-time behavior (T_ε vs T)?

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

- Galerkin, B. G (1915). "Rods and plates: series in some questions of elastic equilibrium of rods and plates." *Vest. Inzh. Tekh.* Vest. Inzh. Tekh.*, 897-908.
- Leray, J (1934). Sur le mouvement d'un liquide visqueux emplissant l'espace. *Acta Math.*, 193-248. DOI: 10.1007/bf02547354
- Schwartz, L (1950). Théorie des Distributions. *Théorie des Distributions.*.
- Foias, C., Prodi, G (1967). Sur le comportement global des solutions non-stationnaires des équations de Navier–Stokes en dimension 2. *Rend. Sem. Mat. Univ. Padova*, 1-34.
- Foias, C., Sell, G. R., Temam, R (1988). Inertial manifolds for nonlinear evolutionary equations. *J. Differential Equations*, 309-353.
- Foias, C., Temam, R (1989). Gevrey class regularity for the solutions of the Navier-Stokes equations. *J. Funct. Anal.*, 87(2), 359-369. DOI: 10.1016/0022-1236(89)90015-3
- DiPerna, R. J., Lions, P.-L (1989). "On the Cauchy problem for Boltzmann equations: global existence and weak stability." *Annals of Math.* Annals of Math.*, 321-366.
- Grujić, Z., Kukavica, I (1998). "Space analyticity for the Navier–Stokes and related equations with initial data in L^p." *J. Funct. Anal.* J. Funct. Anal.*, 447-466.
- Beale, J.T., Kato, T., Majda, A (1984). Remarks on the breakdown of smooth solutions for the 3-D Euler equations. *Comm. Math. Phys.*, 94(1), 61-66. DOI: 10.1007/bf01212349
- De Lellis, C., Székelyhidi, L. Jr. (2009). The Euler equations as a differential inclusion. *Annals of Mathematics*, 170(3), 1417-1436. DOI: 10.4007/annals.2009.170.1417
- Villani, C (2002). A review of mathematical topics in collisional kinetic theory. *Handbook of Mathematical Fluid Dynamics, Vol. 1*, 71-305.
- Nagy, T. (2026). The Latent: Finite Sufficient Representations of Smooth Systems. *Zenodo*. DOI: 10.5281/zenodo.19101209
- Nagy, T. (2026). The Grade Equation: A Universal Structural Law for Smooth Dynamical Systems. *Working paper.*
- Nagy, T. (2026). Grade Regularity: A Universal Criterion for Strong Solutions of Nonlinear PDEs. *Working paper.*
- Nagy, T. (2026). The Grade Structure of Navier-Stokes: Why Blowup Requires Grade-2 Saturation. *Working paper.*
- Nagy, T. (2026). The Latent Algebra: A Universal Representational Language for Smooth Systems. *Working paper.*