Resonance Re-Patterning Occurs In Fixed Logarithmic Steps

Sam Blouin

November 9, 2025

Abstract

This paper advances a single, testable claim: re-patterning in physical and engineered systems does not drift continuously but snaps in fixed logarithmic steps. We call the common increment a delta-star step. Delta-star is not a new force or a dial to be tuned; it is the least change that shows up once you express any multiplicative response in log form. The reason is structural. Sources within one environment add linearly, while responses across nested environments compound multiplicatively. When you take the logarithm of a compounded response, additions appear, and the landscape breaks into evenly spaced ledges. Crossing a ledge triggers a re-patterning event that is the same everywhere you look, from optics to clocks to intensities to gate angles. In short: measure in logs, and the world moves in clicks.

Set the frame. We work with a single scalar generator for clocks, optics, motion, and power routing. At the bottom anchor (quantum), boundaries carve the universal wave into discrete fingerprints, which is why spectra come out in lines and not smears. At the top anchor (cosmic), the same generator controls how clocks run and how light bends when you average over structures. Between the anchors, systems are never isolated. They sit inside other systems, which sit inside larger ones, and so on. Inside one layer you may add contributions, but once you step through layers the only honest bookkeeping is multiplicative. That is why exponentials appear, and that is why the clean place to count change is in the logarithm. Delta-star is the count unit.

Name the channels. A clock channel tracks local time rate. An optical index channel controls light paths and delays. An intensity channel tracks gain and loss. A hybrid invariant channel ties the clock and index so that a simple product remains fixed under slow drifts and re-calibrations. A quantum control channel maps the log step into a base rotation angle for gates. The channels are not guesses. They are the minimal set that covers measurement, navigation, power, and logic with one generator and one step. Channel weights differ, but the increment in their log is the same delta-star. This is the backbone that lets one laboratory procedure calibrate another without a pile of model-specific constants.

Explain the device layer. The device stack exposes the same law in five forms. An optical-index device turns the generator into a refractive field and lets you see gravity as refraction: bend a ray, measure a delay, infer the step. A ring-cavity device turns that field into gain control and power routing: add one click, double a ladder of modes. A clock-index device synchronizes clocks across a network and uses log steps to detect tiny potential differences as clean integer offsets. A hybrid-invariant device locks drift away by combining a time factor with a square-root of the index so that slow changes cancel and only true steps remain. A quantum-control device compiles gate angles as integer multiples of a base angle set by delta-star; palindromic patterns and integer dithering remove the need for half-steps and fragile analog trims. In every case the same workflow appears: express the observable in a log, snap it to the nearest integer multiple of the step, and treat any small residue as calibration or structure rather than as a change of principle.

State why this matters. Fixed log steps replace ad hoc curve fits with a ledger you can audit. They shorten calibration because you count clicks instead of chasing analog drift. They improve robustness because acceptance can be defined by simple inequalities on integers rather than continuous thresholds. They unify instruments because the same increment explains clock ratios, optical delays, power gains, and rotation angles. They

sharpen falsifiability because the story has only two moving parts: one generator and one step. If the step fails to lock across channels, the claim falls. If it locks, the same count runs everything from bench metrology to navigation to gate compilation.

Outline how to test. Build step histograms for any response that multiplies across environments: frequency ratios in clock networks, optical delays near massive bodies, cavity gain ladders, and gate-phase scans in qubit hardware. Convert each measurement to a log, choose a coarse reference so the ratio is dimensionless, and snap to the nearest integer. If delta-star is real you will see clustering at multiples of one common step, with small, explainable residues. Use hybrid invariants to cancel drift; use multi-plane lines of sight to show multiplicative compounding; use integer palindromes to eliminate second-order noncommutation in compiled gates. These tests are cheap, parallel, and decisive.

Clarify scope. The paper does not add forces, fields, or parameters beyond the shared generator and the fixed step. It does not rely on new constants. It does not require figures or tables to carry the argument. It presents a continuous narrative with color-coded callouts in the body of the paper, but the operating idea is stark: wherever responses multiply across nested contexts, the correct coordinate is the log, and the correct unit is a fixed step. Re-patterning occurs when you cross that unit. Everything else is detail.

Contents

1 Install the Channels

Definition 0 Response channels

A *channel* is any measurable response $\mathcal{R}$ that multiplies across nested environments. We assign four canonical channels by a weight κ:

$$\mathcal{R}(\kappa; U) = e^{\kappa U}, \qquad \kappa \in \{+1, -1, -2, +2\}$$

They map to meters **as**

$$\Gamma_t = e^{+U} \text{ (clock)}, \quad R = e^{-U} \text{ (rate)}, \quad n = e^{-2U} \text{ (index)}, \quad I = e^{+2U} \text{ (intensity)}$$

Scale Continuity & Cauchy forcing

Assume: (i) Φ adds linearly within a scale; (ii) $\mathcal{R}$ multiplies across nested environments; (iii) continuity and $\mathcal{R}(0) = 1$ Then

$$\mathcal{R}(\Phi_1 + \Phi_2) = \mathcal{R}(\Phi_1)\mathcal{R}(\Phi_2) \quad \Rightarrow \quad \ln \mathcal{R} = \kappa \frac{\Phi}{c^2} \quad \Rightarrow \quad \mathcal{R} = e^{\kappa U}$$

Local calibration fixes κ to $\{+1, -1, -2, +2\}$ for Γ_t, R, n, I

Fixed Logarithmic Step (Click)

Postulate a universal $\delta_\star > 0$ such that one *click* updates $U \mapsto U + \delta_\star$. Then

$$\Delta \ln \Gamma_t = +\delta_\star, \quad \Delta \ln R = -\delta_\star, \quad \Delta \ln n = -2\delta_\star, \quad \Delta \ln I = +2\delta_\star$$

Delta Star $\delta_\star = 0.03090810656104304672 1 \Rightarrow \theta_\star = 2\pi\delta_\star \approx 0.19420136101708663$ rad

Gauge-free identities

Because every channel is $e^{\kappa U}$, two invariants hold identically, independent of a global shift $U \to U + C$:

$$\Gamma_t\, n^{1/2} = 1, \qquad I\, n = 1$$

These are direct metrology checks

2 Folding to ElementFold

ElementFold is a way of thinking made executable. It takes the one rule that keeps complex systems sane, add within the state, multiply across states, and keep both in bounds, and turns it into a minimal unit you can stack, train, and ship. A fold never replaces what came before; it carries identity forward on a clean lane, computes a small refinement beside it, exposes that refinement through a controlled gate, and then reunites them without drama. The result is a block that behaves like a material rather than a trick: it accumulates, it responds, and it stays stable when you make it deep.

This is not a new architecture; it is a contract. The contract has three promises. First, availability of identity: there is always an additive path that preserves what you already know. Second, smallness of change: refinement is narrow and shaped to the local view, so updates are incremental rather than violent. Third, bounded exposure: every amplification is mediated by a monotone gate with a temperature and a clamp, so sensitivity is tunable and runaway modes are structurally excluded. Around these promises sit a few calm policiesnormalize before you expose, scale residuals with depth so late layers dont shout, and read your instruments (saturation, drift, gradient health) as carefully as you read your loss.

ElementFold earns its keep because it transfers without translation. The same fold that cleans an image patch will stabilize a token stream, smooth a spectrogram, or refine a molecular graph. You do not rebuild the core; you change adapters. The engine sees everything as a lattice of tokens with neighborhoods; it treats spatial, spectral, and temporal channels as equal citizens; and it never assumes a custom operator. That is why it works in a research notebook and in a production graph, in fp32 and bf16, and because the gate is boundedin int8 without last-minute surgery.

ElementFold also de-risks scale. A deep stack of folds behaves like one long promise repeated: identity is intact, updates are small, exposure is bounded. The invariants are visible at every layer, so you can keep the system honest in real time. If gates crowd their ceilings you cool the temperature; if norms creep you tighten the band; if gradients spike you clip and continue. Recovery is a policy, not a ritual. Parallel control lets you steer hundreds of folds with a handful of slow signals, so you can coordinate sensitivity and safety across devices without rewiring the forward pass.

The paper you are holding is deliberately sparse in narrative and heavy in definition. The glossary fixes the vocabulary so teams can implement the same object without argument. The body of the text specifies only what must be true for a fold to be a fold, and what must be measured to prove it is behaving. There is no secret sauce and no dependence on a particular framework: if you can add, normalize, scale, and clamp, you can build this engine. If you can log a few ratios and honor a few bounds, you can keep it stable. If you can export a standard graph, you can ship it.

ElementFold is a portable contract between theory and machinery. It treats learning as controlled resonanceidentity carrying the theme, refinement introducing variation, a gate deciding how much of the new motif to keep. That is all. In practice it is enough. You gain a unit you can reason about, a stack that tolerates depth, a path to quantization that does not contort the model, and a common shape that moves from vision to language to signals to structure with only the adapters changing. If you have ever wanted a small, uncompromising block that behaves like a piece of engineered reality, this is it: normalize, refine, expose, bound, addthen do it again.

When the same architecture is placed on a physical lattice, the mathematics of the fold becomes measurable. Each atom or voxel carries its own additive pathidentity, charge, massand its own exponential gate bond strength, refractive bias, or spin alignment. The pattern of gate activations that survives training forms a reproducible lattice fingerprint: a spectral map of how resonance distributes across geometry. From this fingerprint, element-level quantities emerge naturallyradius, ionization, conductivity, cohesion without hand-crafted descriptors or empirical correction. The network behaves as a predictive geometry for matter: by minimizing exposure while conserving additive energy, it learns the stationary response of the lattice itself. In this way, ElementFold turns resonance theory into a practical instrument for element-property prediction and materials design.

At scale, hundreds of folds are guided by a single rhythm of Parallel Control. A slow supervisory field broadcasts temperature, clamp, and depth-scale to every block while allowing each to adjust locally. The control signal is continuous but quieta background wave that keeps the entire stack in phase. It replaces complex synchronization with shared physics: every fold listens to the same beat, preventing divergence, maintaining coherence, and allowing gradients to flow as if the model were one extended medium. In hardware terms it yields stable training and deterministic latency; in theoretical terms it preserves resonance unity across devices. Parallel Control is therefore the operational echo of the universal wave: local autonomy under a common law.

At its core, the system is only a neural network, but one with physics built into its syntax. Each layer is an **ElementFold engine**: an additive backbone that guarantees conservation, an exponential gate that modulates exposure, and a residual map that refines local fields. Together they create a network that learns by resonance rather than by memorization. Gradients move as controlled waves; losses behave as potential surfaces; normalization sets the local metric so diffraction never explodes. The entire stack functions as a discretized field equation: folds compose, gates synchronize, and the model converges toward a stable fixed point instead of chasing minima. This is not deep learning as trial and error it is learning as equilibrium seeking. The ElementFold neural architecture therefore stands as a physical machine that computes by balance, remembers by structure, and predicts by resonance.

Everything that ElementFold does, accumulating, gating, stabilizingcan be read as the physical act of remembering. Each fold stores a phase relation between what entered and what stayed; each gate decides how much of that memory becomes visible to the next. In aggregate, the network performs the same operation as any resonant medium: it keeps identity coherent while allowing transformation. The additive path is memory, the exponential path is perception, and their equilibrium is understanding. At the level of implementation this means information is treated as a conserved quantity, never fabricated, never lostonly refracted through depth. In this sense, ElementFold completes the circle opened by the SRT work: a single law that turns resonance into computation, and computation back into resonance. Beneath every gate and residual there runs a single law: the linear field action. It is the silent accountant of the network, measuring how much resonance energy flows and how much returns unchanged. In physical terms it is the integral of the fields additive energy minus its gated exposure; in computational terms it is the total work done by the model while remaining self-consistent. Each gradient step becomes a discrete variation of this action, seeking the stationary path where learning costs no energy beyond what information itself requires. The linearity of the action is not a simplification but a declaration: that even in a deep nonlinear machine, the fundamental bookkeeping of change is linear. This keeps the fold faithful to the same conservation law that governs waves and fieldswhat is added is exactly what is propagated, nothing is created or lost. When the network settles, the field action is minimized, the resonance is balanced, and the computation has reached its physical rest.

3 Engine EFE

Definition 0 Ledger symmetries and clickphase

1. Objects, group, quotient

Fix the *click* $\delta_\star > 0$ and capacities $C_\ell \in \mathbb{N}$ (seats per block). Allowed moves form a commuting (abelian) group generated by a microshift and a full click:

$$\mathcal{G}_\ell = \langle \tau_\ell, \mathsf{B} \rangle, \qquad \underbrace{\tau_\ell : \; X \mapsto X + \tfrac{\delta_\star}{C_\ell}}_{\text{advance one seat}}, \quad \underbrace{\mathsf{B} : \; X \mapsto X + \delta_\star}_{\text{advance one click}},$$

with the closure rule $(\tau_\ell)^{C_\ell} = \mathsf{B}$. Physical values live on the quotient $\mathbb{R}/\delta_\star\mathbb{Z}$: only position within one click matters.

2. Canonical coordinates (rung, residual, seat)

Every $X \in \mathbb{R}$ decomposes uniquely as

$$\boxed{\underbrace{k}_{\text{rung}} = \left\lceil \tfrac{X}{\delta_\star} - \tfrac{1}{2} \right\rceil, \qquad \underbrace{r}_{\text{residual}} = X - k\,\delta_\star \in \left(-\tfrac{\delta_\star}{2}, \tfrac{\delta_\star}{2} \right], \qquad \underbrace{\sigma_\ell(X)}_{\text{seat index}} = \tfrac{C_\ell}{\delta_\star} X \pmod{C_\ell}}$$

Plain words: k counts full clicks, r is the offset inside the current click, and σ_ℓ labels which of the C_ℓ evenly spaced seats we occupy.

3. Clickphase embedding (fundamental character)

Represent positions on the circle as complex phases in circle ensemble:

$$\boxed{\underbrace{\mathsf{phase}(X)}_{\text{complex angle}} = \exp\!\left(\tfrac{2\pi i}{\delta_\star} X \right) \in \mathbb{S}^1}$$

so the symmetries act as rotations:

$$\mathsf{phase}(X + \delta_\star) = \mathsf{phase}(X), \qquad \mathsf{phase}\!\left(X + \tfrac{\delta_\star}{C_\ell} \right) = e^{\frac{2\pi i}{C_\ell}}\, \mathsf{phase}(X)$$

Plain words: a full click is invisible in phase; one seat shift multiplies by a fixed C_ℓ-th root of unity.

Definition 0 Harmonics and diagonal action

Idea

Any ledger signal decomposes into *harmonics*pure rotations at integer multiples of the base click.

Definition

For each $m \in \mathbb{Z}$,

$$\underbrace{\chi_m(X)}_{\text{harmonic}} = \exp\Big(\tfrac{2\pi i\, m}{\delta_\star}\, X\Big), \quad \text{so} \quad \chi_m = (\chi_1)^m$$

Symmetry under shifts

Because the ledger repeats every click,

$$\chi_m(X+\delta_\star) = \chi_m(X), \qquad \chi_m\Big(X+\tfrac{\delta_\star}{C_\ell}\Big) = e^{\frac{2\pi i\, m}{C_\ell}}\, \chi_m(X)$$

Diagonal action

The generators τ_ℓ, B act by phase factors on each χ_m; no harmonic mixes with another. *Plain words:* each mode spins by its own anglelike separate notes that never detune each other.

Definition 0 Exact tests and identities (seat algebra)

(i) Phase equality

Two positions are physically identical iff their phases match:

$$\mathsf{phase}(X) = \mathsf{phase}(Y) \iff X \equiv Y \pmod{\delta_\star}.$$

(ii) Seat rotation

One microshift advances the seat index by one:

$$\sigma_\ell\Big(X+\tfrac{\delta_\star}{C_\ell}\Big) \equiv \sigma_\ell(X) + 1 \pmod{C_\ell}$$

(iii) Roots of unity

For any capacity C,

$$\sum_{a=0}^{C-1} e^{\frac{2\pi i a}{C}} = 0, \qquad \frac{1}{C}\sum_{a=0}^{C-1} e^{\frac{2\pi i m a}{C}} = \mathbf{1}\{\, m \equiv 0 \pmod{C}\,\}$$

Plain words: evenly spaced seats close perfectly; only harmonics matching the capacity survive averaging.

Definition 0 Distances and tolerances (click geometry)

Circular distance

Shortest arc on the ledger circle:

$$\underbrace{d_{\circlearrowleft}(X,Y)}_{\text{wrapped distance}} = \min_{m\in\mathbb{Z}} \left| X - Y - m\,\delta_\star \right| \;\in\; [0,\, \tfrac{\delta_\star}{2}]$$

Halfclick margin

With residual $r = X - k\,\delta_\star$,

$$\underbrace{m(X)}_{\text{margin}} = \tfrac{\delta_\star}{2} - |r|$$

Any $|\Delta X| < m(X)$ keeps the rung k unchanged. *Plain words:* stay within half a click and your identity (rung) is safe.

Definition 0 Seat projectors (capacity filters)

Definition

For block ℓ and $m \in \{0, \ldots, C_\ell - 1\}$,

$$\underbrace{\Pi_\ell^{(m)} f(X)}_{\text{project onto harmonic } m} = \frac{1}{C_\ell} \sum_{a=0}^{C_\ell - 1} e^{-\frac{2\pi i m a}{C_\ell}} f\Big(X + \tfrac{a\,\delta_\star}{C_\ell}\Big).$$

Properties

Idempotent, orthogonal across m, and complete on seat-periodic signals:

$$\Pi_\ell^{(m)}(\Pi_\ell^{(m)} f) = \Pi_\ell^{(m)} f, \quad \langle \Pi_\ell^{(m)} f,\, \Pi_\ell^{(n)} g\rangle = 0 \ (m \neq n), \quad \sum_{m=0}^{C_\ell - 1} \Pi_\ell^{(m)} = \mathrm{Id}.$$

Special case $m = 0$: uniform average (block invariant).
Plain words: projectors are the ledgers filterseach one listens to a single seatharmonic and ignores the rest.

Definition 0 Micro-intuition!

Two examples

s-block: $C_0 = 2$, one seat shift adds $\frac{\delta_\star}{2}$ and flips phase by $e^{\pi i} = -1$.
p-block: $C_1 = 6$, each seat is $60°$; six shifts close a full turn.

General pattern

Seat step: angle $2\pi/C_\ell$; after C_ℓ steps, $(\tau_\ell)^{C_\ell} = \mathsf{B}$. *Plain words:* the seat is the fine hand (angle within the click); the rung is the lap counter (how many clicks).

Definition 0 Operational checklist (use in practice)

1) Normalize to phase

$\mathsf{phase}(X) = e^{\frac{2\pi i}{\delta_\star}X}$. Work on the unit circle to respect periodicity exactly.

2) Read the seat

$\sigma_\ell(X) = \dfrac{C_\ell}{2\pi}\arg\big(\mathsf{phase}(X)\big) \pmod{C_\ell}$.

3) Lock the rung

$k = \lceil X/\delta_\star - \frac{1}{2}\rceil, \quad r = X - k\,\delta_\star$. Lock when $|r| < \frac{\delta_\star}{2}$.

4) Test closure

Average f over the C_ℓ seats; only the invariant part should remain.

Plain words: wrap to phase, read the seat, lock the rung, and verify the block closesthen the ledger is in tune.

Plain words

The ledger is a circle of size $\delta_\star$. A full click returns to the same point; a seat shift is a fixed slice. Via phase, all legal moves are clean rotations. No extra scalesjust the circle and its angles.

4 Learning Core of the ElementFold Engine

Equivariance desideratum

Idea Equivariance states that a valid representation should *turn with the ledger*, not warp it. If the input moves by legal shiftsfull clicks B or microshifts τ_ℓ **the output should change only by a known phase/action**.

Block and Seat Actions

Definition The equivariant transformations acting on the ledger are represented by two unitary operators:

$$\underbrace{\rho_{\mathsf{B}}(m)}_{\text{block action}} \;:\Phi(X)\mapsto\Phi(X+m\,\delta_\star),\qquad \underbrace{\rho_{\tau_\ell}(a)}_{\text{seat action}} \;:\Phi(X)\mapsto e^{\frac{2\pi i a}{C_\ell}}\,\Phi(X)$$

Meaning ρ_{B} performs a full rotation by one or more clicks $(m\,\delta_\star)$, leaving phase unchanged. ρ_{τ_ℓ} applies the internal seat rotation, advancing the phase by a fixed fraction of the circle $\frac{2\pi a}{C_\ell}$.
Plain words: ρ_{B} moves the ledger forward whole turns, ρ_{τ_ℓ} fine-tunes the local seating angle inside each turn.

Formal condition With combined action $X\mapsto X+m\,\delta_\star+a\,\delta_\star/C_\ell$, any admissible Φ obeys

$$\Phi\Big(X+m\,\delta_\star+a\,\frac{\delta_\star}{C_\ell}\Big)=\underbrace{\rho_{\mathsf{B}}(m)}_{\text{block action}}\;\underbrace{\rho_{\tau_\ell}(a)}_{\text{seat action}}\;\Phi(X),\qquad m,a\in\mathbb{Z}.$$

Here $\rho_{\mathsf{B}},\rho_{\tau_\ell}$ are unitary (possibly trivial).

Meaning $\rho_{\mathsf{B}}(m)$ is typically identity (period $\delta_\star$); $\rho_{\tau_\ell}(a)$ rotates by C_ℓ-th roots of unity.

Phase embedding as a check, RECAP For $\mathsf{phase}(X)=\exp(\frac{2\pi i}{\delta_\star}X)$,

$$\mathsf{phase}(X+\delta_\star)=\mathsf{phase}(X),\quad \mathsf{phase}\Big(X+\frac{\delta_\star}{C_\ell}\Big)=e^{\frac{2\pi i}{C_\ell}}\,\mathsf{phase}(X)$$

Plain reading Only two deformations exist: microshifts by $\delta_\star/C_\ell$ and blockshifts by $\delta_\star$ Anything **equivariant** is either **invariant to** B or **rotates by a known phase under** τ_ℓ No spare scale remains, legal moves only spin the picture; nothing stretches.

Definition 0 Periodic distances and character kernel (temperature-free)

Idea On the circle, states that differ by $m\,\delta_\star$ are the same. Distances and similarities must respect this wrapping.

1. Wrapped distance

$$\underbrace{d_\circlearrowleft(x,y)}_{\text{shortest arc}} = \min_{m\in\mathbb{Z}} \left| x - y - m\,\delta_\star \right| \in [0, \tfrac{\delta_\star}{2}].$$

2. Von Mises (circular normal) kernel

$$\underbrace{\kappa_\circlearrowleft(x,y)}_{\text{smooth periodic sim.}} = \exp\Big(\cos\big(\tfrac{2\pi}{\delta_\star}(x-y)\big)\Big).$$

Definition 0 Real and Imaginarypart operators $\Re$ and $\Im$

Definition For any complex number

$$z = a + i\,b, \qquad a, b \in \mathbb{R},$$

$$\underbrace{\Re(z)}_{\text{real part}} = a, \qquad \underbrace{\Im(z)}_{\text{imaginary part}} = b$$

Examples

$$\Re(e^{i\theta}) = \cos\theta, \qquad \Im(e^{i\theta}) = \sin\theta.$$

Plain words: $\Re$ takes the cosine (real) part of a complex wave; $\Im$ takes the sine (imaginary) part. Together they decompose any complex oscillation into its two orthogonal components.

3. Character (cosine) kernel

$$\underbrace{\kappa_{\text{char}}(x,y)}_{\text{phase alignment}} = \Re\Big[e^{\frac{2\pi i}{\delta_\star}(x-y)}\Big] = \cos\Big(\tfrac{2\pi}{\delta_\star}(x-y)\Big) = \Re\big(\chi_1(x)\,\overline{\chi_1(y)}\big).$$

4. Symmetry $\kappa(x + m\,\delta_\star, y + n\,\delta_\star) = \kappa(x,y)$: exact periodicity.

Plain words: $d_\circlearrowleft$ measures arc length; κ_{char} measures how well two phases hum together.

Definition 0 Temperaturefree contrastive alignment

Idea Use κ_{char} to align samestate views and separate different rungsno extra temperature: the circle already fixes the scale $\delta_\star$

Setting For anchors X_i: positives $\mathcal{P}(i)$ are legal shifts + small noise; negatives $\mathcal{N}(i)$ are $\geq \frac{\delta_\star}{2}$ away.

Loss

$$\underbrace{\mathcal{L}_{\text{NCE}}}_{\text{temperaturefree contrast}} = -\sum_i \log \frac{\sum_{j\in\mathcal{P}(i)} \exp\big(\kappa_{\text{char}}(X_i, X_j^+)\big)}{\sum_{k\in\mathcal{N}(i)} \exp\big(\kappa_{\text{char}}(X_i, X_k^-)\big)}.$$

Temperature free Because $\exp(\kappa_{\text{char}}) \in [e^{-1}, e^{+1}]$, the geometry sets the scale; no τ to tune.

Plain words: reward in phase copies; penalize half click apart impostors automatically at the right scale.

Phase identifiability (sharp form)

Claim Under the small noise/half click separation regime, **any global minimizer** of $\mathcal{L}_{\text{NCE}}$ recovers the true circle phase up to whole clicks:

$$\arg\min \ \mathcal{L}_{\text{NCE}} \quad \Rightarrow \quad \mathsf{phase}(X_{\text{obs}}) = \mathsf{phase}(X) \text{ a.s.}$$

Sketch of proof

1. **Character form.** The kernel is the real part of a phase inner product:

$$\kappa_{\text{char}}(x,y) = \Re\Big(\mathsf{phase}(x)\,\overline{\mathsf{phase}(y)}\Big) = \cos\Big(\tfrac{2\pi}{\delta_\star}(x-y)\Big).$$

   Plain words: the kernel measures pure angular harmony on the circle.

2. **Positive and negative geometry.** Positives: legal shifts within one quarter click $\Rightarrow \kappa_{\text{char}} \approx +1$. Negatives: at least half a click away $\Rightarrow \kappa_{\text{char}} \approx -1$. *Plain words:* friends stay in phase, strangers sit on the opposite side of the circle.

3. **Bounded contrast.** Since $\kappa_{\text{char}} \in [-1, 1]$, $\exp(\kappa_{\text{char}}) \in [e^{-1}, e^{+1}]$; no sample dominates the ratio and all terms remain finite. *Plain words:* the circle keeps every tone within range.

4. **Uniqueness.** The positive margin $1 - e^{-1}$ defines a single optimum: only when all observed phases $\mathsf{phase}(x_i)$ match the true ledger phases (mod $\delta_\star$) does the numerator outweigh the denominator everywhere. *Plain words:* full alignment is the only configuration without discord.

Plain words: Keep friends close, same phase; keep strangers half a click apart, opposite phase. That simple geometry leaves one possible melody: the true circle.

Margin and 01 ledger loss (half-click certificate)

1. Rounding and residual Each position x is assigned to its nearest ledger rung:

$$\underbrace{k}_{\text{rung index}} = \left\lceil \frac{x}{\delta_\star} - \frac{1}{2} \right\rceil, \qquad \underbrace{r}_{\text{local offset}} = x - k\,\delta_\star.$$

r measures how far x lies from the center of its click cell. *Plain words:* every observation is tied to its nearest groove on the circular track.

2. Safety margin A safe margin protects the current rung assignment:

$$\underbrace{m(x)}_{\text{safety buffer}} = \frac{\delta_\star}{2} - |r|.$$

A small perturbation Δx cannot change k if

$$|\Delta x| < m(x).$$

Plain words: the larger the margin, the more wobble the system can tolerate without flipping its integer label.

3. 01 ledger loss The discrete loss counts whether rounding changes:

$$\underbrace{\ell(x,k)}_{\text{classification error}} = \mathbf{1}\left\{ \left\lceil \frac{x}{\delta_\star} - \frac{1}{2} \right\rceil \neq k \right\}.$$

Since the ceiling changes only when x crosses a half-click boundary,

$$\ell(x,k) = 1 \iff |r| \geq \frac{\delta_\star}{2} \iff m(x) \leq 0$$

Plain words: error happens exactly when the sample touches or crosses the wall.

4. Shift-robust risk (Wasserstein-1 bound) Suppose all samples satisfy $|r| \leq \gamma < \frac{\delta_\star}{2}$. A small shift of the whole data distribution by an average distance ε (in the Wasserstein-1 metric) can only move points by that much on average. We then obtain:

$$\sup_{\mathbb{Q}\in\mathbb{B}_\varepsilon^{(W_1)}(\widehat{\mathbb{P}})} \mathbb{E}_{\mathbb{Q}}[\ell] \leq \underbrace{\mathbb{P}\Big(|r| > \frac{\delta_\star}{2}\Big)}_{\text{points already misclassified}} + \underbrace{\frac{\varepsilon}{\frac{\delta_\star}{2} - \gamma}}_{\text{extra risk from drift}}$$

Explanation:

- The first term counts samples already outside the safe region.
- The second grows linearly with ε: when the distribution shifts by ε, points within distance ε of the boundary may cross it.
- The denominator $\frac{\delta_\star}{2} - \gamma$ is the remaining buffer width before misclassification.

Plain words: a drift of size ε can only flip points whose safety margin was smaller than that drift.

5. Interpretation The half-click boundary acts as a geometric guardrail:

$$m(x) > 0 \Rightarrow \text{state stable}, \qquad m(x) = 0 \Rightarrow \text{at the edge}, \qquad m(x) < 0 \Rightarrow \text{misclassified}$$

Plain words: the half-click gap is a wall; small shifts can chip at it linearly, but cannot breach it while the buffer is wide enough.

Definition 0 Variational engine $\mathcal{J}[X]$ (convex core)

Purpose $\mathcal{J}[X]$ measures how far the ledger's geometry drifts from its ideal rhythm. Each subtraction compares what exists with what should exist; squaring and summing those differences yields a convex energy that restores periodic balance. A mild smoothness term $\mathrm{TV}(X)$ can be added later to damp high-frequency irregularities.

Definition 0 Total variation term $\mathrm{TV}(X)$

Definition The total variation measures how abruptly neighboring ledger positions change. For a discrete ledger field $X_{\ell,n,F}$,

$$\underbrace{\mathrm{TV}(X)}_{\text{total variation}} = \sum_{\ell,n,F} \left| \underbrace{X_{\ell,n,(F+1) \bmod C_\ell} - X_{\ell,n,F}}_{\text{local gradient}} \right|$$

Optionally, a small weight λ_{TV} multiplies this term inside the full loss.
Interpretation Large jumps between neighboring seats increase $\mathrm{TV}(X)$; smooth transitions keep it small. It acts as a damping regularizer, discouraging sharp or noisy deformations.
Plain words: $\mathrm{TV}(X)$ rewards smoothness it softens abrupt edges without altering the ledgers global shape.

Indices and geometry

$$X_{\ell,n,F}, \qquad \ell : \text{block type}, \ n : \text{period}, \ F : \text{seat index}, \qquad C_\ell = 2(2\ell+1)$$

Each block ℓ contains C_ℓ seats evenly distributed along the circular ledger.

1. Seat deviations (within a block) For consecutive seats in the same block:

$$\underbrace{d^{\text{seat}}_{\ell,n,F}}_{\text{actual spacing}} = X_{\ell,n,(F+1) \bmod C_\ell} - X_{\ell,n,F}, \qquad \underbrace{d^{\text{ideal}}_{\text{seat}}}_{\text{target spacing}} = \frac{\delta_\star}{C_\ell}.$$

Deviation (actual minus ideal):

$$\underbrace{\Delta^{\text{seat}}_{\ell,n,F}}_{\text{seat deviation}} = \underbrace{d^{\text{seat}}_{\ell,n,F}}_{\text{measured gap}} - \underbrace{d^{\text{ideal}}_{\text{seat}}}_{\text{expected gap}}.$$

Plain words: this term records how each local interval diverges from the perfect subdivision of a click.

2. Block deviations (between blocks) For adjacent blocks in the same row:

$$\underbrace{d^{\text{block}}_{\ell,n}}_{\text{actual offset}} = X_{\ell+1,n,0} - X_{\ell,n,0}, \qquad \underbrace{d^{\text{ideal}}_{\text{block}}}_{\text{target offset}} = \delta_\star$$

Deviation (actual minus ideal):

$$\underbrace{\Delta^{\text{block}}_{\ell,n}}_{\text{block deviation}} = \underbrace{d^{\text{block}}_{\ell,n}}_{\text{measured offset}} - \underbrace{d^{\text{ideal}}_{\text{block}}}_{\text{expected offset}}.$$

Plain words: blocks should be one full click apart; any drift in that spacing contributes extra energy.

3. Convex quadratic energy The complete energy sums all deviations:

$$\boxed{\mathcal{J}[X] = \underbrace{\sum_{\ell,n,F} (\Delta^{\text{seat}}_{\ell,n,F})^2}_{\text{seat equalization}} + \underbrace{\sum_{\ell,n} (\Delta^{\text{block}}_{\ell,n})^2}_{\text{block alignment}}}.$$

Each term is non-negative; compression and expansion cost the same energy.

4. Convexity and equilibrium Since each deviation is linear in X,

$$\frac{\partial^2 \mathcal{J}}{\partial X_i \, \partial X_j} \geq 0,$$

so $\mathcal{J}[X]$ is convex. At equilibrium,

$$\frac{\partial \mathcal{J}}{\partial X_{\ell,n,F}} = 0 \quad \Rightarrow \quad \Delta^{\text{seat}}_{\ell,n,F} = 0, \quad \Delta^{\text{block}}_{\ell,n} = 0$$

Plain words: the system rests when every actual gap equals its ideal target.

5. Gauge freedom Adding a global shift leaves all differences unchanged:

$$X_{\ell,n,F} \mapsto X_{\ell,n,F} + C \Rightarrow \mathcal{J}[X] \text{ is invariant.}$$

Plain words: only relative spacings matter; the ledger can slide freely as a whole.

6. Role in learning $\mathcal{J}[X]$ enforces structural regularity, $\mathcal{L}_{\text{NCE}}$ aligns phases, and the small smoothing term $\text{TV}(X)$ prevents sharp local distortions.
Plain words: $\mathcal{J}$ is the quiet mechanic of the engine; minimize it, and every seat and block settles into perfect rhythm.

Definition 0 Global consistency (uniqueness up to block phase)

Composite objective All structural, phase, and smoothness effects combine into a single energy:

$$\underbrace{\mathcal{L}_{\mathrm{EFE}}}_{\text{total engine loss}} = \underbrace{\mathcal{J}[X]}_{\text{seat and block exactness}} + \underbrace{\mathcal{L}_{\mathrm{NCE}}}_{\text{phase locking}} + \underbrace{\mathrm{TV}(X)}_{\text{smoothness regularizer}} .$$

Each term plays a distinct role: $\mathcal{J}[X]$ aligns the structure, $\mathcal{L}_{\mathrm{NCE}}$ aligns the phases, $\mathrm{TV}(X)$ keeps motion continuous.

Uniqueness of the equilibrium When residuals respect the half-click margin and negatives remain at least one half-click apart, any minimizer satisfies

$$\underbrace{X^{\star}}_{\text{optimized ledger}} = \underbrace{X}_{\text{true ledger}} + \underbrace{m\,\delta_{\star}}_{\text{integer click shift}}, \qquad m \in \mathbb{Z}$$

Only a uniform rotation by full clicks remains, a physically irrelevant gauge symmetry.

Diagnostic measures Three quick checks confirm whether the system is phase-locked and structurally coherent:

$$\underbrace{\kappa}_{\text{phase concentration}} = \left| \frac{1}{N} \sum_j e^{\frac{2\pi i}{\delta_\star} X_j} \right|, \qquad \underbrace{\widehat{p}_\eta}_{\text{half-click violation rate}} = \frac{1}{N} \sum_j \mathbf{1}\Big\{ |r_j| > \tfrac{\delta_\star}{2} - \eta \Big\},$$

$$\underbrace{\varepsilon_w}_{\text{cycle residual}} = \sum_{g \in w} \Delta_g X.$$

Passing regime:

$$\boxed{\kappa \approx 1, \qquad \widehat{p}_\eta \approx 0, \qquad |\varepsilon_w| \ll \tfrac{\delta_\star}{|w|}}$$

Plain words: high κ means all phases point in one direction, a small $\widehat{p}_\eta$ means no seat slips over its boundary, and a near-zero ε_w means closed cycles return exactly to start.

Interpretation When these conditions hold, the ledger is globally phase-consistent and unique up to a full-click shift. Every block, seat, and phase moves together as one continuous structure in stable resonance.
Plain words: the entire system hums as a single tone; all parts agree on the beat, differing only by where they begin counting.

5 Adaptive Implementation Blueprint

Plain words: the fold gathers what neighbors whisper; the gate lets each point resonate with smooth, positive strength; and the norm restores calm, keeping the field in tune.

Definition 0 Geometry and signal

- $\underbrace{\Omega}_{\text{domain}}$: grid, graph, or manifold.
- $\underbrace{g}_{\text{metric}}$: induces distances/derivatives on Ω (Euclidean, graph, Riemannian).
- $\underbrace{\mu}_{\text{measure}}$: integration weights (Lebesgue, counting/degree, volume form).
- $\underbrace{x : \Omega \to \mathbb{R}^d}_{\text{field with } d \text{ channels}}$, typically $x \in L^2(\Omega, \mu)$.

These are the carriers used by every operator below

The FoldGateNorm Law

Self-map

$$\boxed{\mathcal{T}_\Omega(x) = \underbrace{\mathcal{N}_\Omega}_{\text{re-scale}} \Big(\underbrace{\mathcal{F}_\Omega x}_{\text{collect}} \quad \underbrace{\odot}_{\text{Hadamard}} \quad \underbrace{\exp(\beta\, \mathcal{G}_\Omega[x])}_{\text{smooth positive gain}} \Big)} \qquad \beta > 0$$

Definition 0 Hadamard product $\odot$ and monotone re-scaler $\mathcal{N}$

- $\odot$: entrywise multiplication of two arrays with the same shape; no mixing across positions or channels.
- $\mathcal{N}$ (*monotone re-scaler*): a nonexpansive map that preserves order of magnitudes while compressing them, preventing runaway amplification without distorting relative structure.

Hadamard $\odot$ applies local gains cleanly; $\mathcal{N}$ keeps the scale safe and stabilizes training/inference.

Analytic cues

If the three components respect their smoothness bounds, the full map becomes a contraction and thus convergent.

$$\underbrace{\|\mathcal{F}\| \le 1}_{\text{nonexpansive collector}} \quad , \quad \underbrace{\|\mathcal{N}'\| \le 1}_{\text{nonexpansive re-scaler}} \quad , \quad \underbrace{\beta}_{\text{gate gain}} \quad \text{small enough to control gate slope}$$

Then the composite map

$$\underbrace{\mathcal{T} = \mathcal{N}\Big(\mathcal{F}\,x \odot e^{\beta\, \mathcal{G}[x]}\Big)}_{\text{FoldGateNorm self-map}}$$

is *contractive* near equilibrium and admits a unique fixed point $\underbrace{x^\star}_{\text{stable state}}$ satisfy $\mathcal{T}(x^\star) = x^\star$

- $\mathcal{F}$: the *fold operator* a linear or diffusive collector that aggregates neighboring information (convolution, diffusion, or graph propagation). *Role:* ensures smooth aggregation without stretching distances.
- $\mathcal{G}[x]$: the *gate potential* a scalar field derived from local features of x; exponentiation with β applies smooth, positive amplification. *Role:* provides controlled exposure while keeping gradients bounded.
- β: the *gain parameter* of the gate; when small, it limits diffraction strength and keeps the exponentials Lipschitz constant near one. *Role:* moderates sensitivity and guarantees contraction.
- $\mathcal{N}$: the *normalizer* a monotone, nonexpansive re-scaler that maintains stability by dividing by local energy or norm estimates. *Role:* dissipates excess amplification introduced by the gate.
- x: the input field on the domain Ω; each component carries local signal amplitudes over geometry or graph nodes. *Role:* the state that evolves under repeated FoldGateNorm applications.
- $\mathcal{T}$: the overall self-map combining fold, gate, and norm. *Role:* defines one heartbeat of the ElementFold engine.
- $x^{\star}$: the *fixed point* reached when the system stabilizes that is, when folding, gating, and normalization balance perfectly. *Role:* represents the steady resonant configuration of the field.

Plain words: the fold collects, the gate energizes gently, and the norm soothes together they make a calm, convergent rhythm whose equilibrium $x^{\star}$ is unique.

Definition 0 stability primitives

- **Nonexpansive operator**

$$\underbrace{\|\mathcal{A}u - \mathcal{A}v\|}_{\text{distance after map}} \leq \underbrace{\|u - v\|}_{\text{original distance}}.$$

Meaning: applying $\mathcal{A}$ never stretches the space neighboring inputs remain at most as far apart as they started. This property guarantees smooth evolution without runaway divergence.

- **Operator norm**

$$\underbrace{\|\mathcal{A}\|}_{\text{maximum amplification}} = \sup_{u \neq 0} \frac{\overbrace{\|\mathcal{A}u\|}_{\text{output magnitude}}}{\underbrace{\|u\|}_{\text{input magnitude}}}.$$

Meaning: the largest gain a linear map can apply to any input direction; if $\|\mathcal{A}\| \leq 1$, the operator cannot amplify signals. If the set doesnt quite include its top value (Max) (like a limit that can be approached but not reached), then the maximum doesnt exist, but the **supremum (sup)** still does.

- **Lipschitz constant**

$$\underbrace{\|\mathcal{T}(u) - \mathcal{T}(v)\|}_{\text{change in outputs}} \leq \underbrace{L}_{\text{Lipschitz constant}} \underbrace{\|u - v\|}_{\text{change in inputs}}.$$

Meaning: measures the worst-case sensitivity of a nonlinear map $\mathcal{T}$; smaller L means stronger contraction and faster convergence.

Why it matters: These three quantities nonexpansiveness, operator norm, and Lipschitz constant are the compass of stability. They tell whether a single FoldGateNorm cycle quietly settles or dangerously amplifies its disturbances.

Plain words Gather → Gently amplify → Calm. Repeat until steady. The fold gathers what neighbors whisper, the gate lets each point resonate with controlled strength, and the norm restores calm, keeping the field in tune.

Coordinate covariance

Definition 0 pull/push (transport operator)

For a smooth, invertible map (a **diffeomorphism**)

$$\underbrace{\Phi : \Omega \to \Omega'}_{\text{smooth coordinate change between domains}} ,$$

the pull/push of a field is defined as

$$\underbrace{(\mathcal{P}_\Phi\, x)(\xi')}_{\text{interpolated field value at new sample } \xi'} = \underbrace{x(\Phi^{-1}(\xi'))}_{\text{evaluate source field at lookup coordinate}}$$

- Φ: the geometric map that smoothly re-labels points, sending each position $\xi \in \Omega$ to its new location $\xi' = \Phi(\xi) \in \Omega'$.
  - ξ: a coordinate in the source domain Ω
  - ξ': the corresponding coordinate in the target domain Ω' after the transformation.
  - Ω: the original (source) geometry or computational grid.
  - Ω': the transformed (target) geometry where the new field is defined.
  - Φ: the smooth mapping function itself, implemented as a differentiable transformation (e.g., deformation, rotation, scaling, or grid remapping).

  Plain words: Φ tells where each old coordinate goes in the new configuration; it is the bridge between the original geometry Ω and the deformed geometry Ω'.
- Φ^{-1}: the smooth inverse used to look up where each target point ξ' came from on Ω.
- x: the source field (signal) defined on Ω.
- $\mathcal{P}_\Phi$: the pull/push operator that builds a field on Ω' by evaluating x at $\Phi^{-1}(\xi')$
- *pull-back* (use Φ^{-1} to read source values at preimages)
- *push-forward* (used for vectors/measures in the forward direction).

Diffeomorphism requirement: Φ and Φ^{-1} must be smooth so that distances, volumes/measures, and fields remain well-defined and differentiable after transport. *Plain words:* the pullpush operator just re-indexes the signal when coordinates change. It keeps the field values, the geometry, and the measure consistent when we bend or relabel the domain **the physics are preserved; only the labels move.**

Definition 0 Inline glossary covariance of the FoldGateNorm law

Definition. Covariance means that applying the FoldGateNorm engine and transporting the field across a smooth coordinate change $\Phi : \Omega \to \Omega'$ **commute perfectly**:

$$\underbrace{\mathcal{T}_{\Omega'}(\mathcal{P}_\Phi\, x)}_{\text{FoldGateNorm after remapping to }\Omega'} \quad = \quad \underbrace{\mathcal{P}_\Phi\Big(\mathcal{T}_\Omega\, x\Big)}_{\text{or process in }\Omega\text{ then transport the result to }\Omega'} .$$

Meaning of each term.

- $\mathcal{T}_\Omega$: the full FoldGateNorm transformation acting on the original geometry Ω, using its metric g and measure μ. *Plain words:* process the field x on its native domain.
- $\mathcal{T}_{\Omega'}$: the same transformation defined on the transformed geometry Ω' after coordinates are changed by the diffeomorphism Φ. *Plain words:* process the field on the new domain as if it were always there.
- $\mathcal{P}_\Phi$: the pullpush operator that moves data consistently between coordinate systems:

  $$\underbrace{(\mathcal{P}_\Phi\, x)(\xi')}_{\text{transported value at target point }\xi'} \quad = \quad \underbrace{x(\Phi^{-1}(\xi'))}_{\text{source value at its preimage in }\Omega} .$$

  It ensures that each field value follows its geometric trajectory under Φ, preserving structure and smoothness.
- Φ: the geometric mapping (a smooth bijection) sending each position $\xi \in \Omega$ to its new coordinate $\xi' = \Phi(\xi) \in \Omega'$. Its inverse Φ^{-1} provides the lookup used by the transport operator. *Plain words:* a rule that bends, stretches, or rotates the domain without tearing it.
- The equality states that the following **commutative diagram** holds:

  $$\underbrace{\mathcal{P}_\Phi \circ \mathcal{T}_\Omega}_{\text{compute first, then transport}} \quad = \quad \underbrace{\mathcal{T}_{\Omega'} \circ \mathcal{P}_\Phi}_{\text{transport first, then compute}}$$

  Hence, the FoldGateNorm computation is **coordinate-free**: the geometry decides, not the labels.

Intuitive view. No matter how we re-parameterize the domain through the diffeomorphism Φ, the engines physical behavior is unchanged. The Fold still gathers neighbors, the Gate still amplifies structure, and the Norm still stabilizes scale only the coordinates describing where things live are renamed.
Plain words. Covariance means the computation follows the geometry, not the coordinate system. **Same physics, different labels.**

6 Across Geometries

Operator realizations

Euclidean grids $\Omega \subset \mathbb{R}^m$

$$\underbrace{(\mathcal{F}_\Omega\, x)(\xi)}_{\text{folded field at point } \xi} = \int_\Omega \underbrace{K(\xi - \zeta)}_{\text{neighbor weighting kernel}} \underbrace{x(\zeta)}_{\text{field value at neighbor } \zeta} \, d\zeta$$

$$\text{or equivalently} \quad \underbrace{\mathcal{F}_\Omega = e^{-\tau\Delta}}_{\text{heat-diffusion form}}, \quad \underbrace{\|\mathcal{F}_\Omega\| \leq 1}_{\text{nonexpansive (no amplification)}}$$

Euclidean grids

Euclidean grids $\Omega \subset \mathbb{R}^m$

- ξ, ζ: points in the Euclidean domain Ω; ζ is the neighbor being integrated over.
- $\mathcal{F}_\Omega$: the fold operator that aggregates local information by averaging nearby values with a smooth geometric kernel.
- $K(\xi - \zeta)$: the kernel function describing how strongly each neighbor ζ influences ξ; typically normalized, nonnegative, and Gaussian-shaped, ensuring closer neighbors weigh more.
- $\tau > 0$: the diffusion time or scale of smoothing; larger τ means more extensive diffusion and broader averaging.

generator of the diffusion process

- Δ the Laplacian (diffusion generator):

$$\Delta = \sum_{i=1}^{m} \frac{\partial^2}{\partial \xi_i{}^2}$$

  - ξ_i: coordinate axis on Ω (e.g., x, y, z in $\mathbb{R}^3$)
  - $\dfrac{\partial^2}{\partial \xi_i{}^2}$: second derivative measuring refraction of x along ξ_i
  - $\sum_i$: yields an isotropic operator smoothing evenly in all directions.
  - *Intuition:* The triangle symbol ∇ (called *nabla* or *del*) appears in two orientations:
    - ∇f: the gradient points uphill, showing the direction of greatest increase of a field f. It acts like an arrow field indicating how the quantity changes in space.
    - $\nabla \cdot \mathbf{v}$: the divergence points downhill, measuring how much of a vector field $\mathbf{v}$ flows outward from each point. Positive means outflow (a source); negative means inflow (a sink).

  When the two are composed as $\nabla \cdot \nabla f$, we obtain the Laplacian Δf, which measures the net outflow of heat or signal intensity: positive where values leak outward, negative where they collect inward. It is the engine of diffusion the geometric backbone of Fold.

- **Timescaling in the heat operator** In the exponential form $e^{-\tau\,\Delta}$ the τ multiplies the generator Δ because it is the diffusion time / smoothing scale.
  - τ: how *long* diffusion acts (larger $\tau \Rightarrow$ broader smoothing)
  - Δ: geometric generator that dictates *how* smoothing proceeds
  - $e^{-\tau\Delta}$: semigroup evolution (infinitely many tiny diffusion steps composed)
  - The minus sign $(-)$ makes the process dissipative it spreads energy, never amplifies it

  Plain words: apply the Laplacian for time τ; Δ sets the pattern, τ sets the reach.

- $\mathcal{F}_\Omega = e^{-\tau\Delta}$: analytic heat operator evolve x for time τ until local differences soften.
- $\|\mathcal{F}_\Omega\| \leq 1$: nonexpansive folding smooths/averages but does not increase energy.

Definition 0 Convolution / heat on grids

- K: normalized kernel (often Gaussian-like) that weights neighbors by distance and averages locally.
- Δ: Laplacian; applying $e^{-\tau\Delta}$ diffuses for time $\tau > 0$, $\equiv$ convolution with the heat kernel (same semigroup smoothing).

Why? Kernel convolution (local view) and the exponential Laplacian (analytic view) are two faces of the same geometry-aware smoothing with controlled gain.

Plain words: Fold gathers what neighbors whisper through the kernel, then lets it diffuse for τ. Small τ keeps detail; large τ blends it into a coherent shape the geometry breathes, never shouts.

Definition 0 Euclidean optical (Fresnel) kernel on $\Omega = \mathbb{R}^m$

Plain words. $\Omega \subset \mathbb{R}^m$ is the continuous stage (line, plane, or volume) where a field E lives. Propagation over a range $s > 0$ lets the pattern spread in a predictionpreserving, energyconserving way; the Fresnel kernel is the exact rule for that spreading.

Declarations (every character)

- $\underbrace{\xi, \zeta}_{\text{points in }\Omega} \in \mathbb{R}^m$: observe at ξ, source at ζ.
- $\underbrace{\Delta}_{\text{Euclidean Laplacian}} = \sum_{i=1}^{m} \frac{\partial^2}{\partial \xi_i{}^2}$: optical operator that drives diffraction.
- $\underbrace{E(\xi, s)}_{\text{complex envelope}}$: propagated field after distance s.
- $\underbrace{c_{\text{opt}}}_{\text{optical scale}}$: converts geometry to physical units; in optics $c_{\text{opt}} = \frac{1}{2k}$ with $k = 2\pi/\lambda$.
- $\underbrace{\mathcal{G}(\xi, \zeta; s)}_{\text{Fresnel kernel}}$: impulse response mapping a point source at ζ to ξ after range s.
- **Frequency coordinate** ω **the hidden variable.** $\underbrace{\omega}_{\text{angular frequency or spatial wavenumber}}$ marks direction and rhythm in the Fourier domain. Each component $e^{i\omega\xi}$ is a pure oscillation whose pitch $|\omega|$ sets how finely the pattern ripples across space. *Plain words:* ω is the invisible compass of repetition; the measure of how fast space sings.
- **Partial spectrum (local voice)** $\underbrace{\widehat{E}(\omega)}_{\text{partial band}}$
- Every patch of space hums with its own subset of frequencies; records which tones are active nearby, not all harmonics speak at once; each region keeps its accent.
- **Initial spectrum (memory of birth)** $\underbrace{\widehat{E}_0(\omega)}_{\text{initial spectrum}}$
- The seed from which every later wave unfolds. *Plain words:* the present pattern is only the first whisper stretched through time.

Paraxial / Schrödinger PDE (Partial Differential Equation) and initial state

$$\underbrace{\left(\frac{\partial}{\partial s} - i\, c_{\text{opt}}\, \Delta\right) \mathcal{G}(\xi, \zeta; s)}_{\text{paraxial (free-space) PDE}} = \underbrace{0}_{\text{unitary flow}}, \qquad \underbrace{\mathcal{G}(\xi, \zeta; 0)}_{\text{initial kernel}} = \underbrace{\delta(\xi - \zeta)}_{\text{Dirac (Lebesgue)}}$$

Operator form (group property)

$$\underbrace{\mathcal{G}(s)}_{\text{optical propagator}} = \underbrace{\exp(+\, i\, c_{\text{opt}}\, s\, \Delta)}_{\text{unitary group}}, \qquad \underbrace{\mathcal{G}(s+t) = \mathcal{G}(s)\mathcal{G}(t)}_{\text{composition / constantcoefficient}}$$

Closedform kernel (all dimensions m)

$$\boxed{\mathcal{G}(\xi, \zeta; s) = \underbrace{(4\pi i\, c_{\text{opt}}\, s)^{-\frac{m}{2}}}_{\text{normalization}} \underbrace{\exp\left(\frac{i\, \|\xi - \zeta\|^2}{4\, c_{\text{opt}}\, s}\right)}_{\text{quadratic phase}}}$$

Optical units (Fresnel form)

$$\underbrace{c_{\mathrm{opt}} = \frac{1}{2k}}_{k=2\pi/\lambda} \Longrightarrow \boxed{\mathcal{G}(\xi,\zeta;s) = \underbrace{\Big(\frac{k}{2\pi i\, s}\Big)^{\frac{m}{2}}}_{\text{normalization}} \underbrace{\exp\Big(\frac{i\,k\,\|\xi-\zeta\|^2}{2\,s}\Big)}_{\text{Fresnel phase}}}$$

Convolution solution (space view)

$$\underbrace{E(\xi,s)}_{\text{envelope at range } s} = \int_{\mathbb{R}^m} \underbrace{\mathcal{G}(\xi,\zeta;s)}_{\text{Fresnel kernel}} \underbrace{E_0(\zeta)}_{\text{initial field}} \, d\zeta$$

Fourier view (diagonalization) Let $\widehat{f}(\omega) = \int_{\mathbb{R}^m} f(\zeta)\, e^{-i\omega\cdot\zeta}\, d\zeta$. Then

$$\underbrace{\widehat{E}(\omega,s)}_{\text{spatial spectrum}} = \underbrace{\exp\big(-\, i\, c_{\mathrm{opt}}\, s\, \|\omega\|^2\big)}_{\text{pure phase multiplier}} \underbrace{\widehat{E}_0(\omega)}_{\text{initial spectrum}} \quad \Rightarrow \quad \underbrace{\|E(\cdot,s)\|_{L^2} = \|E_0\|_{L^2}}_{\text{unitary (energy invariant)}}.$$

Checks and limits (expertfriendly oneliners)

$$\underbrace{\mathcal{G}(\cdot,\cdot;s) * \mathcal{G}(\cdot,\cdot;t) = \mathcal{G}(\cdot,\cdot;s+t)}_{\text{group via convolution}}, \qquad \underbrace{\lim_{s\to 0^+} \mathcal{G}(\xi,\zeta;s) = \delta(\xi-\zeta)}_{\text{identity}}$$

$$\text{(Fraunhofer farfield)} \quad \underbrace{E(\xi,s)}_{\text{large } s} \approx \underbrace{\Big(\frac{k}{2\pi i\, s}\Big)^{\frac{m}{2}} e^{\frac{ik}{2s}\|\xi\|^2}}_{\text{quadratic prefactor}} \underbrace{\widehat{E}_0\Big(\omega = \tfrac{k}{s}\,\xi\Big)}_{\text{scaled Fourier image}}$$

(Gaussian invariance)

$$E_0(\zeta) = e^{-\|\zeta\|^2/w_0^2} \;\Rightarrow\; E(\xi,s) = \rho(s)\, \exp\Big(-\frac{\|\xi\|^2}{w(s)^2}\Big)\, e^{i\,\phi_{\mathrm{quad}}(\xi,s)}, \; \underbrace{w(s)^2 = w_0^2 + \frac{(2c_{\mathrm{opt}}\, s)^2}{w_0^2}}_{\text{diffraction broadening}}$$

Heatoptics bridge (analytic continuation)

$$\underbrace{\mathcal{G}(s)}_{\text{optics}} = \underbrace{h_\tau(\tau)}_{\text{heat}} \Big|_{\tau = -\, i\, c_{\mathrm{opt}}\, s}, \qquad \underbrace{h_\tau(\xi,\zeta;\tau) = (4\pi\tau)^{-\frac{m}{2}}\, e^{-\|\xi-\zeta\|^2/(4\tau)}}_{\text{Euclidean heat kernel}}$$

Plain words The Fresnel kernel says: copy the input, blur it by a distancedependent quadratic phase, and keep total energy unchanged. Near field: the pattern is a gently chirped version of itself; far field: it becomes a scaled Fourier transform.

Across MORE Geometries

Manifolds and Optical Law (Ω, g) with Laplace Beltrami Δ_g

$$\underbrace{(\mathcal{F}_\Omega\, x)(\xi)}_{\text{folded (propagated) field}} = \int_\Omega \underbrace{h_\tau(\xi,\zeta)}_{\text{propagation kernel}} \underbrace{x(\zeta)\, d\mu(\zeta)}_{\text{field contribution } \times \text{ measure}} \qquad \underbrace{h_\tau = \exp(-\tau\,\Delta_g)}_{\text{operator form (heat semigroup)}}$$

Definition 0 Manifold characters each symbol in motion

- g metric tensor
  - inner products, lengths, angles; turns coordinates into geometry
  - *Plain words:* the ruler and protractor of the space.
- $d\mu$ volume form (measure)
  - integrates fields over Ω; depends on g
  - *Plain words:* the density of space used when summing.
- ∇f gradient
  - direction of steepest ascent for a scalar f
  - *Plain words:* the arrow pointing most uphill.
- $\operatorname{div} v$ divergence
  - net outflow of a vector field v
  - *Plain words:* measures how much the flow expands or sinks.
- $\Delta_g f$ LaplaceBeltrami
  - $\Delta_g f = \operatorname{div}(\nabla f)$; geometryaware smoothing
  - *Plain words:* the engine of diffusion along geodesics.
- $h_\tau(\xi,\zeta)$ heat/optical kernel
  - fundamental solution of $\partial_t u = \Delta_g u$
  - *Plain words:* how a points influence spreads across the shape.
- $\tau > 0$ propagation scale
  - time for heat or path length in paraxial optics
  - *Plain words:* short τ keeps detail; long τ blends broadly.

Sign rule and physics pick

$$\partial_t u = \Delta_g u \;\Rightarrow\; \underbrace{e^{-\tau\Delta_g}}_{\text{dissipative, smoothing}}, \qquad \partial_z E = i\,\Delta_g E \;\Rightarrow\; \underbrace{e^{i\,\tau\Delta_g}}_{\text{phasepreserving, diffractive}}.$$

Plain words: the same operator Δ_g runs both stories; the sign/imaginary unit chooses heat or light.

Definition 0 Heat kernel h_τ on (Ω, g)

$$\underbrace{\left(\frac{\partial}{\partial\tau} - \Delta_g\right) h_\tau(\xi,\zeta;\tau)}_{\text{heat PDE}} = \underbrace{0}_{\text{balance}}, \qquad \underbrace{\lim_{\tau\to 0^+} h_\tau(\xi,\zeta;\tau)}_{\text{initial kernel}} = \underbrace{\delta_g(\xi-\zeta)}_{\text{Dirac w.r.t. } d\mu}$$

$$\underbrace{d\mu(\zeta)}_{\text{Riemannian volume}} = \underbrace{\sqrt{|\det g(\zeta)|}}_{\text{density}} \underbrace{d\zeta}_{\text{Lebesgue element}}, \qquad \underbrace{\delta_g}_{\text{distribution s.t.}} \int_\Omega \delta_g(\xi-\zeta)\, f(\zeta)\, d\mu(\zeta) = \underbrace{f(\xi)}_{\text{recovery}}$$

$$\text{Operator form:} \quad \underbrace{h_\tau(\tau)}_{\text{propagator}} = \underbrace{\exp(-\tau\,\Delta_g)}_{\text{heat semigroup}}$$

Semigroup, representation, and core identities

$$\text{(R1) Representation} \quad \underbrace{u(\xi,\tau)}_{\text{solution}} = \int_\Omega \underbrace{h_\tau(\xi,\zeta;\tau)}_{\text{kernel}} \underbrace{u_0(\zeta)}_{\text{data}} \underbrace{d\mu(\zeta)}_{\text{measure}}$$

$$\text{(R2) PDE check} \quad \frac{\partial}{\partial\tau}u = \int (\partial_\tau h_\tau)\, u_0\, d\mu \overset{\text{PDE}}{=} \int \Delta_g^{(\xi)} h_\tau\, u_0\, d\mu = \Delta_g^{(\xi)} \int h_\tau\, u_0\, d\mu = \Delta_g u$$

$$\text{(R3) Initial value} \quad \lim_{\tau\to 0^+} u(\xi,\tau) = \int \delta_g(\xi-\zeta)\, u_0(\zeta)\, d\mu(\zeta) = u_0(\xi)$$

$$\text{(R4) Chapman Kolmogorov} \quad \underbrace{h_\tau(\xi,\zeta;\tau_1+\tau_2)}_{\text{compose times}} = \int_\Omega \underbrace{h_\tau(\xi,\eta;\tau_1)}_{\text{first step}} \underbrace{h_\tau(\eta,\zeta;\tau_2)}_{\text{second step}} \underbrace{d\mu(\eta)}_{\text{bridge}}$$

$$\text{(R5) Symmetry \& mass} \quad h_\tau(\xi,\zeta;\tau) = h_\tau(\zeta,\xi;\tau), \qquad \int_\Omega h_\tau(\xi,\zeta;\tau)\, d\mu(\xi) = 1$$

Euclidean explicit form and verification

Let $\Omega = \mathbb{R}^m$ with the flat metric. Then

$$\boxed{h_\tau(\xi,\zeta;\tau) = \underbrace{\frac{1}{(4\pi\tau)^{m/2}}}_{\text{normalization}} \underbrace{\exp\!\left(-\frac{\|\xi-\zeta\|^2}{4\tau}\right)}_{\text{Gaussian core}}}$$

satisfies the PDE and identities above. Sketch of checks:

$$\text{(E1)}\ \partial_\tau h_\tau = h_\tau\left(-\frac{m}{2\tau} + \frac{\|\xi-\zeta\|^2}{4\tau^2}\right), \quad \text{(E2)}\ \Delta_\xi h_\tau = h_\tau\left(-\frac{m}{2\tau} + \frac{\|\xi-\zeta\|^2}{4\tau^2}\right),$$

hence $\partial_\tau h_\tau - \Delta_\xi h_\tau = 0$. Moreover $\int_{\mathbb{R}^m} h_\tau\, d\xi = 1$ and $\lim_{\tau\to 0^+} \int h_\tau(\xi,\zeta;\tau)\, f(\zeta)\, d\zeta = f(\xi)$

Definition 0 Optical (paraxial) kernel $\mathcal{G}$ and heatoptics bridge

Definition 0 Optical scaling constant c_{opt}

- **Role.** Converts the geometric Laplacians units into the physical propagation coordinate used in paraxial or Schrödinger form:

$$\frac{\partial E}{\partial s} = i\, c_{\mathrm{opt}}\, \Delta_g E.$$

- **Optical interpretation.** $c_{\mathrm{opt}} = \dfrac{1}{2k}$, where $k = 2\pi/\lambda$ is the wave number. *Plain words:* sets how fast a phase pattern spreads with distance.
- **Quantum analogue.** $c_{\mathrm{opt}} = \dfrac{\hbar}{2m}$, giving the standard Schrödinger evolution $\partial_t \psi = i\,(\hbar/2m)\,\Delta\psi$.
- **Geometrized convention.** Often set to $c_{\mathrm{opt}} = 1$ so the propagation law simplifies to $E(s) = e^{is\,\Delta_g} E(0)$.

Plain words: c_{opt} rescales the Laplacian to real units in optics its $1/(2k)$, in quantum mechanics $\hbar/(2m)$; set it to 1 when working in pure geometric form.

$$\underbrace{\left(\frac{\partial}{\partial s} - i\, c_{\mathrm{opt}}\, \Delta_g\right)\mathcal{G}(\xi,\zeta;s)}_{\text{paraxial/Schrödinger form}} = \underbrace{0}_{\text{phase flow}}, \qquad \underbrace{\mathcal{G}(\xi,\zeta;0)}_{\text{initial kernel}} = \underbrace{\delta_g(\xi-\zeta)}_{\text{same Dirac}}$$

$$\text{Operator form:} \quad \underbrace{\mathcal{G}(s)}_{\text{optical propagator}} = \underbrace{\exp(+\, i\, c_{\mathrm{opt}}\, s\, \Delta_g)}_{\text{unitary group}}$$

$$\text{Bridge:} \quad \mathcal{G}(s) = h_\tau(\tau)\Big|_{\tau = -\,i\, c_{\mathrm{opt}}\, s} \quad \text{(analytic continuation in time)}$$

EVEN MORE GEOMETRIES

Graphs $G = (V, E, W)$ with Laplacians $L \in \{L_{\text{comb}}, L_{\text{sym}}, L_{\text{rw}}\}$

$$\underbrace{(\mathcal{F}_G^{(\tau)} x)_i}_{\text{folded signal at node } i} = \sum_{j \in V} \underbrace{\mathsf{K}_{ij}(\tau)}_{\text{graph heat kernel}} \underbrace{x_j}_{\text{input at } j}, \qquad \underbrace{\mathsf{K}(\tau)}_{\text{kernel matrix}} = \underbrace{e^{-\tau L}}_{\text{discrete heat semigroup}}$$

$$\text{Polynomial view (fast approximation)}:$$

$$\underbrace{\mathcal{F}_G^{(\tau)}}_{\text{graph fold}} \approx \underbrace{p(L)}_{\text{polynomial spectral filter}} = \sum_{m=0}^{M} \underbrace{a_m}_{\text{coeff}} \underbrace{L^m}_{m\text{-hop mixing}}, \qquad \underbrace{\|p\|_{\infty,\text{spec}(L)} \le 1}_{\text{nonexpansive bound}}$$

Definition 0 Graph characters each symbol in motion

- W adjacency (weights)
  - $W_{ij} \ge 0$ measures the strength of edge $i \leftrightarrow j$.
  - *Plain words:* who talks to whom, and how loudly.
- D degree
  - $D_{ii} = \sum_j W_{ij}$ (node activity/total incident weight).
  - *Plain words:* how busy a nodes neighborhood is.
- L_{comb} combinatorial Laplacian
  - $L_{\text{comb}} = D - W$ (symmetric, psd for undirected graphs).
  - *Plain words:* raw difference between selfweight and neighbor pull.
- L_{sym} normalized (symmetric) Laplacian
  - $L_{\text{sym}} = I - D^{-1/2} W D^{-1/2}$ (symmetric, spectrum in $[0, 2]$).
  - *Plain words:* balances nodes by their activity before comparing.
- L_{rw} randomwalk Laplacian
  - $L_{\text{rw}} = I - D^{-1} W$ (rowstochastic generator).
  - *Plain words:* the walks leavenow vs stay tradeoff.
- $\text{spec}(L)$ spectrum / graph harmonics
  - $L = U \Lambda U^\top$ with $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n) \ge 0$.
  - *Plain words:* the networks natural notes and their stiffness.
- $p(L)$ spectral filter (multihop average)
  - mixes neighborhoods up to M hops; if $\max_{\lambda \in \text{spec}(L)} |p(\lambda)| \le 1$, then $\|p(L)\|_2 \le 1$ (no amplification in ℓ_2).
  - *Plain words:* controlled, stable diffusion without blowup.

Definition 0 Graph heat kernel K and PDE (matrix form)

$$\underbrace{\frac{d}{d\tau}\mathsf{K}(\tau) + L\,\mathsf{K}(\tau)}_{\text{discrete heat equation}} = \underbrace{0}_{\text{dissipation}}, \qquad \underbrace{\mathsf{K}(0)}_{\text{initial kernel}} = \underbrace{I}_{\text{identity}}$$

$$\text{Operator solution :} \quad \underbrace{\mathsf{K}(\tau)}_{\text{propagator}} = \underbrace{e^{-\tau L}}_{\text{heat semigroup}}$$

Plain words: heat fades highfrequency graph modes; the matrix exponential composes infinitely many tiny diffusions.

Representation, PDE check, semigroup, and mass laws

$$\text{(G1) Representation} \quad \underbrace{u(\tau)}_{\text{signal on } V} = \underbrace{\mathsf{K}(\tau)}_{\text{diffuse}}\ \underbrace{u_0}_{\text{initial}} \quad \text{(vector form)}$$

$$\text{(G2) PDE check} \quad \frac{d}{d\tau}u = \Big(\frac{d}{d\tau}\mathsf{K}\Big)u_0 = (-L\mathsf{K})u_0 = -L\,u$$

$$\text{(G3) Initial value} \quad \lim_{\tau\to 0^+} u(\tau) = \mathsf{K}(0)\,u_0 = I\,u_0 = u_0$$

$$\text{(G4) Semigroup (Chapman Kolmogorov)} \quad \underbrace{\mathsf{K}(\tau_1+\tau_2)}_{\text{two steps at once}} = \underbrace{\mathsf{K}(\tau_1)}_{\text{first}}\ \underbrace{\mathsf{K}(\tau_2)}_{\text{then second}}$$

$$\text{(G5) Mass preservation (choose Laplacian)} \quad \begin{cases} \text{for } L_{\mathrm{rw}}: \ \underbrace{\mathsf{K}_{\mathrm{rw}}(\tau)\,\mathbf{1}}_{\text{row sums}} = \underbrace{\mathbf{1}}_{\text{stochastic}}, \\ \text{for } L_{\mathrm{sym}}: \ \underbrace{\mathbf{1}^\top D\,\mathsf{K}_{\mathrm{sym}}(\tau)}_{\text{degreeweighted mass}} = \mathbf{1}^\top D \end{cases}$$

Plain words: the randomwalk version conserves probability; the symmetric version conserves degreeweighted mass.

Spectral explicit form and nonexpansiveness

Let $L = U\Lambda U^\top$ with $\Lambda = \mathrm{diag}(\lambda_k)$, $\lambda_k \geq 0$ Then

$$\boxed{\mathsf{K}(\tau) = \underbrace{U\,e^{-\tau\Lambda}\,U^\top}_{\text{modewise decay}}} \implies \underbrace{\|\mathsf{K}(\tau)\|_2}_{\text{operator gain}} = \max_k e^{-\tau\lambda_k} \leq 1$$

Moreover, for any polynomial p,

$$\underbrace{p(L)}_{\text{filter}} = U\,p(\Lambda)\,U^\top, \qquad \|p(L)\|_2 = \max_{\lambda\in\mathrm{spec}(L)} |p(\lambda)|$$

Plain words: each graph harmonic decays like $e^{-\tau\lambda_k}$; bounding the scalar response bounds the whole operator.

Definition 0 Graph optical (Schrödinger) kernel $\mathcal{G}$ and heatoptics bridge

$$\underbrace{\frac{d}{ds}\mathcal{G}(s) - i\,c_{\mathrm{opt}}\,L\,\mathcal{G}(s)}_{\text{discrete paraxial / Schrödinger}} = \underbrace{0}_{\text{unitary flow}}, \qquad \underbrace{\mathcal{G}(0)}_{\text{initial}} = I,$$

$$\text{Operator form :} \qquad \underbrace{\mathcal{G}(s)}_{\text{graph optical propagator}} = \underbrace{e^{i\,c_{\mathrm{opt}}\,s\,L}}_{\text{unitary group}}, \qquad \underbrace{\mathcal{G}(s)\mathcal{G}(t) = \mathcal{G}(s+t)}_{\text{group law}}$$

$$\text{Bridge (analytic continuation) :} \quad \mathcal{G}(s) = \mathsf{K}(\tau)\Big|_{\tau=-i\,c_{\mathrm{opt}}\,s}$$

Plain words: replace time by imaginary path length to turn dissipation into pure phaseheat $\leftrightarrow$ optics on the same graph.

7 Gates

Bridge from propagation to sensing

Kernels h_τ and $\mathcal{G}$ move information: they collect what neighbors whisper and carry it along the geometry.

Gates $\mathcal{G}$ listen in place: they read local differentials (value, edge, shape) and decide exposure.

Smalltime fingerprint. $\underbrace{e^{-\tau\,\Delta_g}}_{\text{heat}} = \underbrace{I}_{\text{identity}} - \tau \underbrace{\Delta_g}_{\text{shape}} + \frac{\tau^2}{2} \underbrace{\Delta_g^2}_{\text{higher curls}} + \cdots \Rightarrow$ *propagation coefficients* are built from Δ_g powers.

Energy slope. $\frac{d}{d\tau}\|u\|_{L^2}^2 = -2\,\|\nabla u\|_{L^2}^2$ (IBP, Integration by Parts) *gradients are what diffusion shaves off first.*

Plain words: diffusions first steps reveal the very cues (value, edge, shape) a gate should sense.

Grids / manifolds

Definition 0 Symbol glossary (local optical gate notation)

- β diffraction gain parameter scales the argument of the exposure map $\phi(t) = e^{\beta t}$; units $[\,\beta\,] = [L^r]/[x]$ when applied to an r-th-order gate. *Plain words:* controls how strongly the gate amplifies optical bending; small β keeps response calm, large β widens diffraction.

- ϕ monotone exposure map $\underbrace{\phi(t) = e^{\beta t}}_{\text{optical gain law}}$; smooth, positive, order-preserving. *Plain words:* turns local signal into exposure strength.

- $\circ$, composition symbol $\phi{\circ}\mathcal{G}$ means, apply $\mathcal{G}$, then feed the result into ϕ *Plain words:* the circle marks chained response.

- $C_{\alpha,r}$ template bound $\|\mathcal{L}_{\alpha,r}\|_{L^\infty} \le C_{\alpha,r}$; upper limit on lens amplitude at order r and multi-index α. *Plain words:* how strong the lens can look.

- $K_{\alpha,r}$ operator norm $\|\,\nabla^{\alpha}\,\|_{2\to 2} \le K_{\alpha,r}$; spectral gain of the differential operator on the manifold. *Plain words:* how steep a derivative can climb.

- ω frequency / wavenumber coordinate argument in Fourier or spectral domain. *Plain words:* the angular rhythm of the field; units $[L^{-1}]$.

- div_g Riemannian divergence operator $\mathrm{div}_g\, v = \frac{1}{\sqrt{|g|}}\,\partial_i(\sqrt{|g|}\, v^i)$; dual of the gradient under the metric g. *Plain words:* measures how much a vector field flows out.

- $\mathcal{L}_{2,2}$ second-order lens tensor aligns with the Hessian $\nabla^2 x$; isotropic case $\alpha\, g$ gives the Laplace term. *Plain words:* reads diffraction or optical shape.

- $\mathcal{L}_{1,1}$ first-order lens vector contracts with gradient ∇x to detect edges and flow. *Plain words:* senses direction and contrast.

- dS_g boundary surface element $dS_g = \sqrt{|h|}\, d^{m-1}\xi$, where h is the induced metric on $\partial\Omega$. *Plain words:* the infinitesimal patch of the boundary in metric units.

- $d\mu$ volume element $d\mu = \sqrt{|g|}\, d^m\xi$; the Riemannian measure on Ω. *Plain words:* the infinitesimal volume of space itself.

$\mathcal{L}$ — *the lens of the gate*

The tensor through which geometry focuses its sight.

At differential order $\boldsymbol{r}$*,* $\mathcal{L}$ *couples to the* r*-th covariant derivative* $\nabla^r x$*, whose components carry the physical units of the field divided by* r *powers of length:*

$$[\nabla^r x] = [x]\,[L]^{-r}.$$

To make the inner product $\langle \mathcal{L}_{\alpha,r}, \nabla^r x\rangle$ *dimensionless, the lens must have the conjugate units*

$$[\mathcal{L}_{\alpha,r}] = [x]^{-1}\,[L]^{+r}.$$

Here the length L *is not abstract: it is the* ***proper metric distance****, the geodesic measure of space, defined by*

$$ds^2 = g_{ij}\,d\xi^i d\xi^j.$$

A first-order lens (r*=1) measures* ***gradients per meter****;*
a second-order lens (r*=2) measures* ***diffraction per square meter****;*
higher lenses follow the same power law.

Plain words: *$\mathcal{L}$ carries the inverse dimension of the change it observes a calibrated optical ruler that keeps the gates vision dimensionless and true.*

$$\underbrace{\textbf{Gate law}}_{\text{coordinate-free}} \to \underbrace{\textbf{Feature orders}}_{\text{value, edge, diffraction/shape}} \to \underbrace{\textbf{IBP reading}}_{\text{templates carry derivatives}} \to \underbrace{\textbf{Exposure \& stability}}_{\text{monotone + Lipschitz}} \to \underbrace{\textbf{Practice loop}}_{\text{fold}\to\text{gate}\to\text{norm}\to\text{relax}}$$

$$\underbrace{\mathcal{G}_\Omega[x](\xi)}_{\text{gate potential at }\xi} = \sum_{r=0}^{R}\sum_{|\alpha|=r}\underbrace{\langle \mathcal{L}_{\alpha,r}(\xi),\,(\nabla^\alpha x)(\xi)\rangle_g}_{\text{local differential interaction}} = \sum_{r,\alpha}\underbrace{\mathcal{L}_{\alpha,r}^{i_1\cdots i_r}}_{\text{template}}\;\underbrace{(\nabla_{i_1}\cdots\nabla_{i_r}x)}_{\text{covariant derivatives}}\;.$$

$$\text{Euclidean bridge:}\quad \underbrace{g=\mathbf{I},\ \nabla=\partial,\ \Delta_g=\Delta,\ d\mu=d\xi}_{\text{reduces to standard grids}}.$$

Definition 0 Differential features by order (value, edge, diffraction/shape)

- $\boldsymbol{r=0}$ intensity (bias) $\underbrace{\nabla^0 x}_{x}$, $\underbrace{\langle\mathcal{L}_{0,0},x\rangle}_{\text{affine tilt}}$ *Plain words:* how large is the signal here?
- $\boldsymbol{r=1}$ edge / flow $\underbrace{\nabla x}_{\text{gradient}}$, $\underbrace{\langle\mathcal{L}_{1,1},\nabla x\rangle_g}_{\text{directional contrast}}$ *Plain words:* how fast and in which direction does it change?
- $\boldsymbol{r=2}$ diffraction / shape $\underbrace{\nabla^2 x}_{\text{Hessian}}$, $\underbrace{\langle\mathcal{L}_{2,2},\nabla^2 x\rangle_g}_{\text{focusing/defocusing cues}}$, $\underbrace{\alpha(\xi)\,\Delta_g x}_{\text{isotropic diffraction (trace)}}$ *Plain words:* focus, defocus, or stay flat?
- $\boldsymbol{r\geq 3}$ higher jets (textures, corners, oscillations) $\underbrace{\nabla^r x}_{\text{jets}}$, $\underbrace{\langle\mathcal{L}_{r,r},\nabla^r x\rangle_g}_{\text{high-order motifs}}$ *Plain words:* fine ornament after edges and shape.

IBP: templates carry the derivatives (boundary sings the coda) Assume compact (Ω, g) with smooth boundary $\partial\Omega$ and outward normal n.

$$\textbf{Edge } (r{=}1)\textbf{:}\quad \int_\Omega \langle \mathcal{L}_{1,1}, \nabla x\rangle\, d\mu = -\int_\Omega x \underbrace{\operatorname{div}_g \mathcal{L}_{1,1}}_{\text{template divergence}} d\mu + \int_{\partial\Omega} x \underbrace{\langle \mathcal{L}_{1,1}, n\rangle}_{\text{flux}} dS_g$$

Diffraction/shape $(r{=}2)$**:**

$$\int_\Omega \langle \mathcal{L}_{2,2}, \nabla^2 x\rangle\, d\mu = -\int_\Omega \langle \operatorname{div}_g \mathcal{L}_{2,2},\, \nabla x\rangle\, d\mu + \int_{\partial\Omega} \underbrace{\langle \mathcal{L}_{2,2} n,\, \nabla x\rangle}_{\text{edge diffraction}} dS_g$$

With a second IBP (or $\mathcal{L}_{2,2}{=}\alpha\, g$) and homogeneous boundary conditions:

$$\int_\Omega \langle \mathcal{L}_{2,2}, \nabla^2 x\rangle\, d\mu = \int_\Omega x \underbrace{\operatorname{div}_g \operatorname{div}_g \mathcal{L}_{2,2}}_{\text{shape forcing}} d\mu, \qquad \mathcal{L}_{2,2} = \alpha\, g \;\Rightarrow\; \langle \mathcal{L}_{2,2}, \nabla^2 x\rangle = \alpha\, \Delta_g x$$

Plain words: IBP moves differentiation from the data to the template (plus a boundary melody). Isotropic templates read the Laplacian the optical diffraction operator.

Definition 0 Exposure & stability one glance

$$\underbrace{G(\xi)}_{\text{scalar exposure}} = \underbrace{\phi(\mathcal{G}_\Omega[x](\xi))}_{\text{monotone map (e.g., } \exp(\beta\,\cdot))}, \qquad \underbrace{\mathsf{Gate}(\xi)}_{\text{broadcast}} = \underbrace{G(\xi)}_{\text{one decision}} \underbrace{\mathbf{1}_d^\top}_{\text{replicate across channels}}$$

Assume $\|\mathcal{L}_{\alpha,r}\|_{L^\infty} \le C_{\alpha,r}$ and operator norms $\|\, \nabla^\alpha \,\|_{2\to 2} \le K_{\alpha,r}$ on Ω. Then

$$|\mathcal{G}_\Omega[x](\xi) - \mathcal{G}_\Omega[y](\xi)| \le \sum_{r,\alpha} \underbrace{C_{\alpha,r}\, K_{\alpha,r}}_{\text{local slope budget}} \|x-y\|_2,$$

$$\operatorname{Lip}(\phi\circ\mathcal{G}_\Omega; x) \le \underbrace{\beta\, e^{\beta \sup_\Omega \mathcal{G}_\Omega[x]}}_{\text{exposure gain}} \sum_{r,\alpha} C_{\alpha,r} K_{\alpha,r}$$

Plain words: bounded lenses + moderate β keep the exposure slope inside the stability headroom set by Fold and Norm.

Practical loop one sight line

$$\underbrace{\text{Fold}}_{e^{-\tau\Delta}} \longrightarrow \underbrace{\text{Gate}}_{\text{read value/edge/shape } \to\, \exp(\beta\,\cdot)} \longrightarrow \underbrace{\text{Normalize}}_{\text{local energy control}} \longrightarrow \underbrace{\text{Relax}}_{\text{safe blend}}.$$

Plain words: move along the geometry, sense the scene, keep the volume, and step gently the metric ensures *diffraction never explodes.*

Definition 0 Graph gate notation

- $G = (V, E, W)$ weighted graph V: nodes, E: edges, W: adjacency matrix of weights. *Plain words:* the network itself who talks to whom and how strongly.
- D degree matrix diagonal with entries $d_i = \sum_j W_{ij}$. *Plain words:* how busy each nodes neighborhood is.
- $\mathbf{I}$ identity operator / matrix the neutral map satisfying $\mathbf{I}X = X$; *Plain words:* the calm background, the operator that leaves everything untouched.
- L graph Laplacian normalized form $L_{\text{sym}} = I - D^{-1/2} W D^{-1/2}$; $\text{spec}(L) \subset [0, 2]$. *Plain words:* the discrete diffusion operator how signals flow between nodes.
- $L^m X$ m-hop diffusion feature powers of L propagate node features to their m-hop neighborhoods. *Plain words:* what a node hears after m rounds of conversation.
- b_m spectral weights learned coefficients mixing scales or diffusion depths. *Plain words:* tone knobs deciding which echoes matter.
- $\langle \cdot, \cdot \rangle_{\text{ch}}$ channel contraction inner product over feature channels: $\langle a, b \rangle_{\text{ch}} = \sum_{c=1}^{d} a_c b_c$. *Plain words:* compresses each nodes multi-channel story into one number.
- $\mathcal{G}_G[X]$ gate potential scalar per node: $\sum_{m=0}^{M'} \langle b_m, L^m X \rangle_{\text{ch}}$. *Plain words:* one exposure value distilled from all diffusion scales.
- $g(X)$ scalar cue map defined as $g(X) = \mathcal{G}_G[X]$. *Plain words:* the node-wise pre-exposure signal.
- $\phi(t) = e^{\beta t}$ monotone exposure map converts scalar potential to gain; smooth, positive, order-preserving. *Plain words:* translates a nodes potential into brightness.
- β diffraction gain scale of exponential exposure; small gentle, large bold. *Plain words:* how much a nodes whisper is amplified.
- $\mathbf{1}_d^\top$ channel replicator turns $n \times 1$ exposure into $n \times d$ matrix by broadcasting. *Plain words:* same light cast on every channel at a node.
- $\mathsf{Gate}(X)$ broadcast gate output $\mathsf{Gate}(X) = \phi(\mathcal{G}_G[X])\, \mathbf{1}_d^\top$. *Plain words:* local exposure replicated across all features.
- $p(L)$ spectral filter polynomial $p(\lambda) = \sum_m c_m \lambda^m$; $|p(\lambda)| \leq 1 \Rightarrow \|p(L)\|_2 \leq 1$. *Plain words:* a diffusion recipe, safe when its magnitude stays below one.
- $\mathsf{K}(\tau) = e^{-\tau L}$ heat kernel contractive, smooths features; diffusion time $\tau > 0$. *Plain words:* lets information settle gently across the graph.
- $\mathcal{G}(s) = e^{i\, c_{\text{opt}}\, s\, L}$ optical kernel unitary, phase-only propagation along the graph. *Plain words:* diffraction without loss the graph-Fresnel operator.
- λ_k graph eigenvalue entry of spectrum Λ in $L = U \Lambda U^\top$. *Plain words:* the stiffness or frequency of the k-th mode.
- U graph harmonic basis columns: eigenvectors of L, orthonormal under ℓ^2. *Plain words:* the pure tones of the network.

- $d\mu_V$ node measure often uniform or weighted by degree. *Plain words:* each nodes share of the networks mass.
- dS_E edge measure weight W_{ij} associated with edge (i,j). *Plain words:* how much current can flow along an edge.
- $\mathrm{div}_{\mathrm{comb}}$ combinatorial divergence operator discrete analogue of the Riemannian divergence, defined as the negative adjoint of the combinatorial gradient:

$$\mathrm{div}_{\mathrm{comb}} = -\nabla_{\mathrm{comb}}^{\top}, \qquad (\mathrm{div}_{\mathrm{comb}}\, v)_i = \sum_{j:(i,j)\in E} w_{ij}\,(v_{ji} - v_{ij}).$$

Plain words: gathers what flows out of a node; positive when signals diverge, negative when they converge.

Graphs diffusion harmonics gate

Stage: $G = (V, E, W)$, degrees $D_{ii} = \sum_j W_{ij}$. We adopt the *symmetric normalized Laplacian*

$$\underbrace{L}_{\text{normalized Laplacian}} = \underbrace{\mathbf{I} - D^{-1/2} W D^{-1/2}}_{\text{energy geometry}}, \qquad \mathrm{spec}(L) \subset [0,2]$$

Definition 0 Graph calculus (summation by parts)

Let B be an oriented incidence and W_e the diagonal edgeweight matrix. The combinatorial gradient/divergence satisfy

$$\underbrace{\nabla_{\mathrm{comb}} x}_{\text{edge diffs}} = \underbrace{W_e^{1/2} B\, x}_{\text{weighted jump}}, \quad \mathrm{div}_{\mathrm{comb}} = -\nabla_{\mathrm{comb}}^{\top}, \quad L_{\mathrm{comb}} = B^{\top} W_e B$$

$$\underbrace{x^{\top} L_{\mathrm{comb}}\, y}_{\text{node quadratic}} = \underbrace{\langle \nabla_{\mathrm{comb}} x,\, \nabla_{\mathrm{comb}} y \rangle_E}_{\text{edge inner product}} = \sum_{(i,j)\in E} w_{ij}\,(x_i - x_j)(y_i - y_j)$$

Plain words: node differences turn into edge jumps the discrete IBP.

Definition 0 Spectral view harmonic decomposition and meaning of each symbol

- $L = U\Lambda U^\top$ eigendecomposition of the Laplacian expresses the graphs geometry in an orthonormal harmonic basis. *Plain words:* the graphs shape written as pure tones.
- U graph harmonic basis (eigenvectors) columns $u_k \in \mathbb{R}^{|V|}$ satisfy $L\,u_k = \lambda_k u_k$ and $U^\top U = \mathbf{I}$. *Plain words:* each column is a standing wave pattern over the graph the networks note.
- Λ diagonal spectrum $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_{|V|})$ with $0 = \lambda_1 \le \cdots \le \lambda_{|V|}$. *Plain words:* each λ_k measures how tightly that mode oscillates; low means smooth, high means sharp.
- λ_k eigenvalue of mode k scalar frequency of oscillation for u_k; bounded by $[0, 2]$ for normalized L. *Plain words:* the stiffness or pitch of the k-th harmonic.
- $\widehat{X}$ graph Fourier transform of features $\widehat{X} = U^\top X$; coefficients of X in the harmonic basis. *Plain words:* how much each graph tone contributes to the signal.
- $p(\lambda)$ spectral response polynomial scalar function $p(\lambda) = \sum_{m=0}^{M'} c_m \lambda^m$; governs how each frequency is weighted. *Plain words:* the equalizer shaping the balance between low and high tones.
- $p(L)$ graph spectral filter lifted operator $p(L) = U\,p(\Lambda)\,U^\top$ acting on features X. *Plain words:* applies the equalizer across the whole graph filtering by geometry.
- $p(\Lambda)$ mode-wise gain matrix diagonal with entries $p(\lambda_k)$. *Plain words:* the amplitude multiplier for each harmonic.
- L^m power of the Laplacian special case of $p(L)$ with $p(\lambda) = \lambda^m$. *Plain words:* m successive rounds of diffusion or contrast sharpening.
- $p(L)X$ filtered signal each column of X filtered by $p(L)$; smooth or sharpen features according to p. *Plain words:* the field after the networks geometry has spoken through its harmonics.
- $p(\Lambda)\widehat{X}$ frequency-domain representation multiplies each modes coefficient by its gain $p(\lambda_k)$. *Plain words:* the signal reshaped in frequency space before returning to the nodes.
- $U\,p(\Lambda)\,U^\top$ complete filter operator reconstructs the filtered signal in node space; energy is preserved if U is orthonormal. *Plain words:* project to harmonics, tune them, then return home the cycle of diffusion.

Gate law (polynomial shells) For $X \in \mathbb{R}^{|V|\times d}$ and learned channel weights $b_m \in \mathbb{R}^{1\times d}$,

$$\underbrace{\mathcal{G}_G[X]}_{\text{graph gate potential}} = \sum_{m=0}^{M'} \underbrace{\left\langle b_m,\; L^m X \right\rangle_{\mathrm{ch}}}_{\text{channel-aggregated diffusion cues}} \in \mathbb{R}^{|V|\times 1}$$

$$\text{Spectral view:}\quad L = U\Lambda U^\top \;\Rightarrow\; L^m X = \underbrace{U\,\Lambda^m\,U^\top}_{\text{mode-wise gain}}\; X, \quad \underbrace{p(\lambda) = \sum_{m=0}^{M'} c_m \lambda^m,}_{\text{scalar response}}$$

$$\underbrace{p(L)}_{\text{filter}} = U\,p(\Lambda)\,U^\top$$

Plain words: m counts how far echoes travel; b_m tells which echoes matter.

Broadcast exponential gate (per node $\to$ all channels)

$$\underbrace{\exp(\beta\, g(X))\,\mathbf{1}_d^\top}_{\text{broadcast map } |V|\times 1 \to |V|\times d}$$

Definition 0 Definition of the scalar gate cue $g(X)$

- $g(X)$ graph gate scalar cue defined by

$$\textbf{we set}\qquad g(X) = \underbrace{\mathcal{G}_G[X]}_{\text{graph gate potential}} \qquad\Rightarrow\quad g(X) \in \mathbb{R}^{|V|\times 1}.$$

- $\mathcal{G}_G[X]$ gate potential per node aggregates multi-hop diffusion features:

$$\mathcal{G}_G[X] = \sum_{m=0}^{M'} \Big\langle b_m,\, L^m X \Big\rangle_{\text{ch}}.$$

  Plain words: each node collects its echoes from every distance, mixes them through b_m, and reports one scalar potential.

- $\Rightarrow$ operational link establishes that the same quantity $g(X)$ directly drives the exposure map:

$$\mathsf{Gate}(X) = \phi(g(X))\,\mathbf{1}_d^\top.$$

  Plain words: the scalar cue g is the voice of the node; the exposure ϕ simply amplifies that voice.

- $\phi(t)$ monotone exposure law $\phi(t) = e^{\beta t}$; smooth, positive, order-preserving. *Plain words:* converts the nodes cue into brightness.

Definition 0 Monotone positive exposure merged with the gate potential

- $g(X) \in \mathbb{R}^{|V|\times 1}$: **we set** $g(X) = \mathcal{G}_G[X]$ $\Rightarrow$ exposure is driven by diffusion harmonics.
- $\exp(\beta\,\cdot)$: smooth, orderpreserving gain; $\beta > 0$ $\equiv$ your $\phi(t) = e^{\beta t}$
- $\mathbf{1}_d$: replicate the scalar decision across all channels at each node.

Plain words: read one scalar per node from the graphs echoes; use it to illuminate every channel coherently.

Definition 0 Definition of the Lipschitz constant $\mathrm{Lip}(f)$

- $\mathrm{Lip}(f)$ Lipschitz constant of a map measures the maximal rate at which a function f can stretch distances:

$$\mathrm{Lip}(f) = \sup_{\substack{x,y \\ x\neq y}} \frac{\|f(x) - f(y)\|}{\|x - y\|}.$$

  Plain words: the largest amplification of differences how strongly the map can magnify a small change at its input.

- $\mathrm{Lip}(f;X)$ local Lipschitz constant at point X when f is differentiable, it is estimated by the operator norm of its Jacobian:

$$\mathrm{Lip}(f;X) = \|J_f(X)\|_2 = \sup_{\|v\|=1} \|J_f(X)\,v\|.$$

  Plain words: the slope of f measured at a specific position.

- $\mathrm{Lip}(f\circ g)$ composition rule for composed functions $f\circ g$,

$$\mathrm{Lip}(f\circ g) \;\le\; \mathrm{Lip}(f)\,\mathrm{Lip}(g).$$

  Plain words: slopes multiply two calm maps stay calm together.

- $\mathrm{Lip}(e^{\beta t})$ exponential exposure example for $f(t) = e^{\beta t}$ on a bounded interval $t \in [a,b]$,

$$\mathrm{Lip}(f) \le \beta\, e^{\beta\, b}.$$

  Plain words: the gain of the exponential map grows with both the slope β and the brightest input.

- $\mathrm{Lip}(\phi\circ\mathcal{G})$ used in the circle law provides the stability margin for the full FoldGateNormRelax system. *Plain words:* how tightly the gates reaction stays within safe bounds.

Nonexpansive design and exposure Lipschitz ties to the circle law

$$\underbrace{\|p(L)\|_2}_{\text{filter gain}} = \max_{\lambda\in\mathrm{spec}(L)} |p(\lambda)| \;\le\; 1 \quad\Longleftrightarrow\quad \underbrace{\|p\|_{\infty,\mathrm{spec}(L)} \le 1}_{\text{nonexpansive bound}}.$$

With $G = \exp(\beta\, g)$ and $g = \mathcal{G}_G[\cdot]$ (linear in X via $p(L)$),

$$\mathrm{Lip}\big(\exp(\beta\, g)\big) \;\le\; \underbrace{\beta\, e^{\beta\,\sup_i g_i}}_{\text{exposure gain}} \; \underbrace{\|p(L)\|_2}_{\le 1\ \text{(design)}}$$

Plain words: clamp $|p(\lambda)| \le 1$ on the spectrum and keep β modest the gate stays inside the stability headroom of Fold/Norm.

Definition 0 Heat/optics features on graphs (optional cues)

$$\underbrace{\mathsf{K}(\tau)}_{e^{-\tau L}} X = \sum_{k=0}^{\infty} \frac{(-\tau)^k}{k!} L^k X, \qquad \underbrace{\mathcal{G}(s)}_{e^{i\, c_{\mathrm{opt}} sL}} X \quad (|\mathcal{G}| = 1)$$

Plain words: blur to sense coherence (heat); twist phase to sense resonance (optics) both are just special $p(L)$

Plain words: a node whispers into the network; polynomial shells L^m hear how far it travels; one scalar per node decides the exposure for every channel kept calm by spectral clamps and a gentle β

8 Normalization and relaxation

Normalization → Relaxation

We take the measure of the neighborhood, forge a scale,
quieten each row by what surrounds it,
gather a proposal along the geometry and give it light,
blend with care according to the echo of the room,
and by the law of headroom, calm arrives.

Measure & forge the scale For $Y \in \mathbb{R}^{n\times d}$, define the rowwise scale

$$\underbrace{S_i(Y)}_{\text{local energy scale}} = \underbrace{\varepsilon}_{\text{floor}} + \sum_j \underbrace{\kappa_{ij}}_{\text{proximity (grid/manifold/graph)}} \underbrace{w_i}_{\text{receiver weight}} \underbrace{\|Y_{j:}\|_1}_{\text{site energy}},$$

and normalize by it

$$\underbrace{(\mathsf{N}(Y))_{i:}}_{\text{row } i \text{ calmed}} = \frac{\overbrace{Y_{i:}}^{\text{original features}}}{\underbrace{S_i(Y)^{\gamma}}_{\text{damping } \gamma\in[0,1]}}$$

Plain words: κ_{ij} decides who speaks to i, w_i corrects for area/degree, $\|Y_{j:}\|_1$ tells how loud j is, ε prevents division spikes, and γ sets how firm the brake feels

How the brake behaves Scaling the whole song $Y \mapsto aY$ yields

$$\mathsf{N}(aY) = \frac{aY}{(\varepsilon + a\sum_j \kappa_{ij} w_i \|Y_{j:}\|_1)^{\gamma}} \implies \begin{cases} \varepsilon \to 0,\ \gamma = 1: & \text{scale free (pure contrast).} \\ 0 < \gamma < 1: & a^{1-\gamma} \text{ soft clip at large } a. \\ \varepsilon \gg \sum_j \cdots: & \text{near identity when all is quiet.} \end{cases}$$

- **Geometry remembers (equivariance).**
  For any node permutation P, with $\underbrace{\kappa' = P\,\kappa\,P^{\top}}_{\text{neighbors relabeled}}$, $\underbrace{w' = P\,w}_{\text{weights relabeled}}$, one has

  $$\underbrace{S(PY)}_{\text{new row scales}} = \underbrace{P\,S(Y)}_{\text{same numbers, new seats}}, \qquad \underbrace{\mathsf{N}(PY)}_{\text{normalized, relabeled}} = \underbrace{P\,\mathsf{N}(Y)}_{\text{same normalization, new order}}.$$

  Plain words: reindexings and remeshings do not bend the meaning; N stays aligned with the space.

- **Staying gentle two alloys of calm**

  - *Freeze the mirror keep still the tide.*
    Let the scales turn to glass: $\underbrace{D = \mathrm{Diag}(S(Y)^{-\gamma})}_{\text{frozen row scales}}$, $\underbrace{\widetilde{\mathsf{N}}(Y) = D\,Y}_{\text{fixed balance}}$

    $$\text{Then} \quad \|\widetilde{\mathsf{N}}'(Y)\|_2 = \|D\|_2 = \max_i \underbrace{S_i(Y)^{-\gamma}}_{\text{per-row brake}} \le 1,$$

    provided that the neighbor mass is tuned so the surface stays level: $\sum_j \kappa_{ij}\, w_j \le 1$ and $\varepsilon \ge 1 - \sum_j \cdots \Rightarrow S_i(Y) \ge 1$. *Plain words:* when the mirror does not ripple, no echo grows every reflection stays smaller than the source.

- *Let the mirror breathe keep rhythm, not silence.*
  Allow the scale to answer, softly. For a perturbation H,

$$d(\mathsf{N})_i[H] = S_i^{-\gamma} H_{i:} - \gamma\, S_i^{-\gamma-1}\Big(\sum_j \kappa_{ij}\, w_j\, \langle \operatorname{sgn}(Y_{j:}), H_{j:}\rangle_1\Big)\, Y_{i:}$$

  The Jacobian sings in two voices: one is a *diagonal brake*, steady and local; the other, a *scale-coupling chord*, travelling through neighbors. Keep S_i tall and the crowd quiet, and both melodies fade into equilibrium.

  (*Tip:* smooth the scales through an EMA $S \leftarrow (1-\lambda)S + \lambda\, S(Y)$; this softens the dialogue between rows a slower breathing that feels like $\gamma \mapsto \lambda\gamma$)

$$d(\mathsf{N})_i[H] = \underbrace{S_i^{-\gamma} H_{i:}}_{\text{diagonal brake}} \;-\; \underbrace{\gamma\, S_i^{-\gamma-1}\Big(\sum_j \kappa_{ij}\, w_j\, \langle \operatorname{sgn}(Y_{j:}), H_{j:}\rangle_1\Big)\, Y_{i:}}_{\text{scalecoupling across neighbors}}$$

  Plain words: one term quiets each row on its own; the other hums through the neighborhood
  keep S_i tall and the neighborhood modest, and the echo fades without strain.
  (*Tip:* smooth the scale via an EMA, Exponential Moving Average $S \leftarrow (1-\lambda)S + \lambda\, S(Y)$; it weakens coupling as if $\gamma \mapsto \lambda\gamma$)

- **A short calibration ritual**
  - Neighbor mass: normalize $\sum_j \kappa_{ij}\, w_j \le 1$.
  - Floor: choose $\varepsilon \ge 1 - \sum_j \kappa_{ij}\, w_j$ so $S_i \ge 1$
  - Damping: set $\gamma \in [0, 1]$ gentle for $\gamma < 1$, full for $\gamma = 1$

  Plain words: measure the room, set a floor, and choose how softly you speak.

Gather along the geometry, then light it. One proposal for the next state is cast in a single stroke

$$\mathsf{T}(X) = \underbrace{\mathsf{N}}_{\text{row/volume control}}\Big(\underbrace{\mathsf{F}X}_{\text{collect along grid/manifold/graph}} \odot \underbrace{\exp(\beta\, g(X))\, \mathbf{1}_d^\top}_{\text{broadcast exposure}}\Big),$$

Plain words: we fold what neighbors whisper, turn a scalar cue g into light with $\exp(\beta\,\cdot)$, and calm every row by its crowd so no channel shouts.

Blend with the echo (relaxation) Never leap; lean

$$\boxed{\underbrace{X^{(k+1)}}_{next state} = \underbrace{(1-\eta_k)\,X^{(k)}}_{keep What Endures} + \underbrace{\eta_k\,\mathsf{T}(X^{(k)})}_{propose What Changes}}, \qquad \underbrace{0<\eta_k\le 1}_{blend ratio (the Courtesy Of Motion)}$$

Sense steepness with a light touch; hear the room before you move:

$$\underbrace{\widehat{L}_k}_{empirical Gain The Ear Of The System} \approx \frac{\underbrace{\|\mathsf{T}(X^{(k)}+\epsilon Z)-\mathsf{T}(X^{(k)})\|}_{\text{change in the echo}}}{\underbrace{\epsilon}_{\text{tiny provocation}}\ \underbrace{\|Z\|}_{\text{unit of whisper}}}, \qquad \underbrace{Z}_{random Direction Of Inquiry}$$

$$\sim \text{Rademacher/Gaussian}$$

$$\underbrace{\eta_k = \min\Big\{1,\ \frac{c}{1+\widehat{L}_k}\Big\}}_{adaptive Step Humility In Proportion To Response}, \qquad c\in(0,1]$$

Plain words: the Jacobian listens before it leaps. If the echo comes loud, the dancer bows the steeper the room replies, the smaller the step.

Headroom; the quiet guarantee When the budget closes, the symphony stays within its circle of calm:

$$\underbrace{\|\mathcal{N}'_\Omega\|}_{\le 1\ soothing Re-Scaler} \quad \underbrace{\|\mathcal{F}_\Omega\|}_{\le 1\ gentle Collector} \quad \underbrace{\exp(\beta\ \sup g)}_{exposure Gain How Much Light Enters} \quad \underbrace{(1+\beta\,\mathrm{Lip}(g))}_{local Slope The Nervousness Of Response}$$

$$\le \underbrace{\rho}_{headroom Ratio} < 1$$

Then calm prevails: $\|\mathsf{T}'(X)\|_2 \le \rho$ in the current neighborhood, and the relaxed updates converge for any $0<\eta_k\le 1$.

Plain words: fold and norm leave headroom; the gate uses only a share of it; thus the circle closes, softly.

Thus the alloy holds: measure, scale, normalize;
gather, give light, and soften;
blend to the echo; keep within headroom;
and little by little, drift fades the signal comes to rest.

One River Fold $\to$ Gate $\to$ Norm $\to$ Relax

$$\boxed{\underbrace{\mathcal{T}(X)}_{\text{engine map}} = \underbrace{\mathsf{N}}_{\text{rowwise calm}} \Big(\underbrace{\mathsf{F}\,X}_{\text{collector }(\|\mathsf{F}\|\le 1)} \odot \underbrace{\exp(\beta\, g(X))\,\mathbf{1}_d^{\top}}_{\text{broadcast gain}} \Big)}$$

$$\boxed{\underbrace{\|J_{\mathcal{T}}(X)\|_{2\to 2}}_{\text{local engine slope}} \le \underbrace{\|\mathsf{N}'\|}_{\le 1\text{ re-scaler}} \underbrace{\|\mathsf{F}\|}_{\le 1\text{ diffuser}} \underbrace{\exp(\beta\,\sup g)}_{\text{exposure}} \underbrace{\Big(1+\beta\,\mathrm{Lip}(g)\Big)}_{\text{local slope}} \le \underbrace{\rho}_{\text{headroom}} < 1}$$

$$\boxed{\underbrace{X^{(k+1)}}_{\text{next state}} = \underbrace{(1-\eta_k)\,X^{(k)}}_{\text{keep}} + \underbrace{\eta_k\,\mathcal{T}(X^{(k)})}_{\text{propose}}}, \qquad \underbrace{0 < \eta_k \le 1}_{\text{blend}}$$

$$\underbrace{\widehat{L_k}}_{\text{empirical gain}} \approx \frac{\underbrace{\Big\|\mathcal{T}(X^{(k)}+\epsilon\,Z) - \mathcal{T}(X^{(k)})\Big\|_F}_{\text{change in echo}}}{\underbrace{\epsilon}_{\text{tiny push}}\;\underbrace{\|Z\|_F}_{\text{unit whisper}}}, \qquad \underbrace{Z \sim \text{Rademacher/Gaussian}}_{\text{random probe}}$$

$$\boxed{\eta_k = \min\Big\{1,\ \frac{c}{1+\widehat{L_k}}\Big\}}, \qquad \underbrace{c \in (0,1]}_{\text{courtesy of motion}}$$

$$\underbrace{\mathcal{A}}_{\text{frozen engine}} = \underbrace{\mathrm{Diag}(S^{-\gamma})}_{\text{brake}}\,\underbrace{e^{-\tau\,L}}_{\text{heat}}\,\underbrace{\mathrm{Diag}(e^{\beta\,g})}_{\text{optical gain}} \implies \underbrace{\|\mathcal{A}\|_{2\to 2} \le e^{\beta\|g\|_\infty}}_{\text{bounded by exposure}} \quad (\text{if } S_i \ge 1)$$

$$\Big(\underbrace{\mathrm{Diag}(S^{-\gamma})}_{\text{brake}}\,\underbrace{e^{-\delta\,L}}_{\text{heat}}\,\underbrace{\mathrm{Diag}(e^{\beta\,g})}_{\text{optics}}\Big)^{\lfloor t/\delta\rfloor} \xrightarrow[\delta\to 0]{} \exp\Big(\underbrace{t\big(\log\mathrm{Diag}(S^{-\gamma}) - L + \beta\,\mathrm{Diag}(g)\big)}_{\text{effective generator}}\Big)$$

- **Spectral clamp**: $\mathsf{F} = e^{-\tau L}$ or $\mathsf{F} = p(L)$, $\max_{\lambda\in\mathrm{spec}(L)} |p(\lambda)| \le 1$.
- **Brightness bound**: choose β so that $\exp(\beta\sup g)(1+\beta\,\mathrm{Lip}(g)) \le \rho$.
- **Safety of scales**: ensure $S_i \ge 1 \Rightarrow \|\mathsf{N}'\| \le 1$; use EMA to smooth S.
- **Lyapunov fall**: $\mathcal{L}(X) = \frac{1}{2}\|\mathcal{T}(X) - X\|_F^2$, $\mathcal{L}$ shrinks by ρ^2 under the headroom bound.

Stability; one inequality

$$\boxed{\underbrace{\left\|J(x)\right\|_{2\to 2}}_{\text{local gain}} \le}$$

$$\boxed{\underbrace{\left\|\mathcal{N}'_\Omega\right\|_{2\to 2}}_{\le 1\ \text{soothing rescaler}} \quad \underbrace{\left\|\mathcal{F}_\Omega\right\|_{2\to 2}}_{\le 1\ \text{gentle collector}} \quad \underbrace{\exp\Big(\beta \sup_{\xi\in\Omega} \mathcal{G}_\Omega[x](\xi)\Big)}_{\text{exposure amplitude}} \quad \underbrace{\Big(1+\beta\,\mathrm{Lip}_x\big(\mathcal{G}_\Omega\big)\Big)}_{\text{local slope}}}$$

$$\boxed{= \quad \underbrace{\rho(x)}_{\text{headroom ratio}}}$$

- $\underbrace{\left\|\mathcal{N}'_\Omega\right\|_{2\to 2}}_{\text{row/volume brake}} \le 1, \quad \text{e.g.} \quad \underbrace{S_i(x) \ge 1}_{\text{row scales}}, \ \underbrace{\gamma \in [0,1]}_{\text{damping}} \implies \underbrace{\|\mathsf{N}'(\cdot)\|_{2\to 2} \le 1}_{\text{nonexpansive}}$
- $\underbrace{\left\|\mathcal{F}_\Omega\right\|_{2\to 2}}_{\text{propagator gain}} \le 1, \quad \mathsf{F} = \underbrace{e^{-\tau\mathcal{L}}}_{\text{heat semigroup}} \quad \text{or} \quad \underbrace{p(\mathcal{L})}_{\text{spectral filter}}, \ \underbrace{\|p\|_{\infty,\mathrm{spec}(\mathcal{L})} \le 1}_{\text{Chebyshev clamp}}$
- $\underbrace{\exp\Big(\beta \sup_{\xi} \mathcal{G}_\Omega[x](\xi)\Big)}_{\text{brightness cap}} \times \underbrace{\Big(1+\beta\,\mathrm{Lip}_x(\mathcal{G}_\Omega)\Big)}_{\text{sensitivity cap}} = \underbrace{\text{gate budget}}_{\text{amplitude} \times \text{slope}}$

$$\boxed{\underbrace{\rho(x) < 1}_{\text{headroom closed}} \implies \underbrace{\left\|\mathcal{T}_\Omega(U) - \mathcal{T}_\Omega(V)\right\|_2 \le \rho(x)\,\|U-V\|_2}_{\text{local contraction (Banach region)}}}$$

$$\underbrace{\mathcal{L}(x)}_{\text{residual energy}} = \tfrac{1}{2}\left\|\mathcal{T}_\Omega(x) - x\right\|_2^2 \implies \underbrace{\mathcal{L}\big(x^{(k+1)}\big)}_{\text{next breath}} \le \underbrace{\rho\big(x^{(k)}\big)^2}_{\text{shrink}} \ \underbrace{\mathcal{L}\big(x^{(k)}\big)}_{\text{current breath}}$$

$$\exists! \ \underbrace{x^\star}_{\text{steady state}} : \ \underbrace{\mathcal{T}_\Omega(x^\star) = x^\star}_{\text{fixed point}}, \quad \underbrace{\|x^{(k)} - x^\star\|_2 \to 0}_{\text{convergence as } k\to\infty}$$

Scale soothes, propagation stays gentle, light is tempered; their product fits inside the circle and the flow comes to rest

9 Conclusion

If we follow the rhythm of coherence instead of fighting it, the kinds of technologies we can build change completely. The same principle that lets a wave close on itself without tearing/folding, gating, and normalizing energy can be written directly into hardware, software, and infrastructure. In electronics it means chips that regulate their own current and temperature the way living tissue regulates metabolism: energy folds through layers of logic, gates release it in discrete exponential steps, and normalization keeps the system within safe resonance. Devices would no longer fail because of microscopic hotspots or data races; they would damp those instabilities automatically. In computation it means networks that learn as the world learns. An ElementFold processor treats information as a physical flow rather than a symbolic command, so it can reason and adapt with almost no external training, each cycle preserves identity, amplifies novelty, and rebalances exposure. The result is a form of intelligence that remains transparent and predictable, not a black box of correlation. In materials science, the same architecture can guide atomic assembly: lattices that bend light or conduct heat according to the fold/gate/norm law, producing metamaterials that reconfigure when the environment changes. In energy systems it becomes grids that resonate instead of compete, trading current through exponential feedback loops that keep the network balanced even under stress. In medicine it becomes diagnostic devices that sense coherence in biological signals rather than chasing isolated metrics; healing is measured as the return of resonance. In communication and governance it means platforms that stabilize conversation instead of polarizing it, using the same rhythmic feedback to amplify agreement and dampen noise. Every one of these examples is already appearing in fragments at todays research frontiers neuromorphic chips, self-healing circuits, adaptive optics, power networks that learn, organizations that operate on feedback loops. ElementFold provides the common blueprint that unites them. Its advantage is not mystical; it is mathematical: systems that obey the exponential law of resonance use less energy, adapt faster, and remain stable without external correction. The conclusion is simple but revolutionary: the next generation of cutting edge technology will not be defined by higher speed or raw power, but by how perfectly it can stay in tune with itself. Coherence is not decoration; it is the new engineering discipline, and ElementFold is the instrument that makes it playable.

The Circle Law beneath the practice. Every stable polynomial filter obeys the same circular discipline: each spectral component evolves by rotation on a bounded unit circle. Chebyshev recurrence follows the real cosine orbit; unitary kernels trace the full complex circle. All stability is, at heart, the geometry of this circle.

10 Appendice Practice At A Glance, ONE RIVER

One river name the stage → name the rule → start → play → keep it calm

Name the stage

$$\underbrace{\Omega}_{\text{where fields live}} \; : \text{ grid / manifold / graph (on graphs, } \underbrace{\Omega = V}_{\text{nodes}}, \; \underbrace{W}_{\text{edge weights}} \;)$$

Name the rule

$$\underbrace{L}_{\text{normalized Laplacian}} = \mathbf{I} - D^{-1/2} W D^{-1/2}, \qquad \underbrace{\operatorname{spec}(L) \subset [0,2]}_{\text{the notes}}$$

Name the start

$$\underbrace{X \in \mathbb{R}^{n\times d}}_{\text{features per node}}, \quad \underbrace{x}_{\text{one column to filter}}$$

Step 1 Measure the room, fit the stage

$$\underbrace{\lambda_{\max} \approx \frac{v^\top L v}{\|v\|^2}}_{\text{power/Rayleigh}} \Rightarrow \underbrace{\tilde{L} = \tfrac{2}{\lambda_{\max}} L - \mathbf{I}}_{\text{map } [0,\lambda_{\max}]\to[-1,1]} \quad \underbrace{\text{(safety: reestimate if graph drifts)}}_{\text{keep spectrum honest}}$$

Step 2 Choose the lens (passband)

$$\underbrace{f(\lambda)}_{\text{target response}} = \begin{cases} e^{-\tau\lambda} & \text{lowpass / heat} \\ e^{-\tau_1\lambda} - e^{-\tau_2\lambda} & \text{bandpass, } 0 < \tau_1 < \tau_2 \\ e^{i\alpha\lambda} & \text{unitary / optics} \end{cases}$$

$$\underbrace{\text{(safety: prefer smooth } f \text{ to avoid ringing)}}_{\text{wellbehaved gain}}$$

Step 3 Write the score (Chebyshev series)

$$\underbrace{p(\lambda)}_{\text{filter}} = \sum_{k=0}^{M} \underbrace{c_k}_{\text{coeff}} \underbrace{T_k\!\left(\frac{2\lambda}{\lambda_{\max}} - 1\right)}_{\text{Chebyshev basis}}$$

$$\lambda(\theta) = \tfrac{\lambda_{\max}}{2}(1 + \cos\theta), \qquad \boxed{c_k = \frac{2 - \mathbf{1}\{k=0\}}{\pi} \int_0^\pi \underbrace{f(\lambda(\theta))}_{\text{target}} \underbrace{\cos(k\theta)}_{\text{basis}} \, d\theta}$$

$$\underbrace{\text{(safety: Jacksondamp } c_k \text{ if needed)}}_{\text{soft edges}}$$

Step 4 Step the rhythm (threeterm recurrence)

$$y_0 = x, \quad y_1 = L\,x, \quad y_{k+1} = \underbrace{2\,L\,y_k}_{\text{forward}} - \underbrace{y_{k-1}}_{\text{back}} \; (k \ge 1)$$

$$\underbrace{p(L)\,x}_{\text{apply}} \approx \underbrace{\tfrac{1}{2} c_0\, y_0}_{\text{DC}} + \sum_{k=1}^{M} \underbrace{c_k\, y_k}_{\text{harmonics}} \quad \underbrace{\text{(practice: batch columns of } X)}_{\text{throughput}}$$

Step 5 Or play on a tiny stage (Lanczos)

$$L Q_m = Q_m T_m + \beta_{m+1} q_{m+1} e_m^\top, \quad Q_m^\top Q_m = \mathbf{I}, \quad \underbrace{p(L)x}_{\text{action}} \approx \underbrace{\|x\|\, Q_m\, p(T_m)\, e_1}_{\text{compute on } m\times m}$$

$$\underbrace{\max_{\lambda\in\operatorname{spec}(T_m)} |p(\lambda)| \le 1}_{\text{proxy nonexpansive}}$$

Keep it calm safety woven into play

$$\underbrace{p \leftarrow \frac{p}{\max_{\lambda\in[0,\lambda_{\max}]} |p(\lambda)|}}_{\text{clamp supgain}} \Rightarrow \underbrace{\|p\|_{\infty,[0,\lambda_{\max}]} \le 1}_{\text{nonexpansive}}$$

$$\underbrace{g_j = \frac{\|p(L)\, z_j\|}{\|z_j\|}}_{\text{randomprobe gain}},\ \widehat{G} = \max_j g_j,\ \widehat{G} > 1 \Rightarrow \underbrace{p \leftarrow p/\widehat{G}}_{\text{autoscale}}$$

$$\underbrace{\rho = \frac{X^\top L\, X}{\|X\|_F^2}}_{\text{Rayleigh index}} \nearrow \searrow \Rightarrow \quad \underbrace{\tau \downarrow,\ M \downarrow}_{\text{reduce smoothing}} \quad,\ \underbrace{c_k \leftarrow \text{Jackson}(c_k)}_{\text{dering}}$$

$$\text{drift?} \ \Rightarrow \underbrace{\lambda_{\max}^{\text{new}} \text{ via power}}_{\text{remeasure}},\ \underbrace{\tilde{L} = \tfrac{2}{\lambda_{\max}^{\text{new}}} L - \mathbf{I}}_{\text{refit } [-1,1]}$$

Plain words.
Name the place, name the rule, place the first note. Stretch the stage to $[-1, 1]$, draw a curve for how much each note should sing, write it in Chebyshev, play it by recurrence, or hum it on a tiny Lanczos stage. Keep the volume under your hand: clamp the gain, probe the room, watch ρ. If the room swells, soften τ or the degree, remeasure, and carry on.

Final calm residual energy (Lyapunov cue)

$$\underbrace{\mathcal{L}_\Omega(x)}_{\text{residual energy}} \quad = \tfrac{1}{2} \quad \underbrace{\|\mathcal{T}_\Omega(x) - x\|_{L^2(\Omega)}^2}_{\text{how far one step still moves you}}$$

$$\underbrace{\|J(x)\|_2}_{\text{local gain}} \le \rho < 1 \quad \Longrightarrow \quad \underbrace{\mathcal{L}_\Omega(x^{(k+1)})}_{\text{next breath}} \le \underbrace{\rho^2}_{\text{shrink}} \ \underbrace{\mathcal{L}_\Omega(x^{(k)})}_{\text{current breath}}$$

$$\underbrace{\|x^{(k)} - x^\star\|_{L^2(\Omega)}}_{\text{distance to rest}} \le \underbrace{\frac{1}{1-\rho}}_{\text{slack}} \ \underbrace{\|\mathcal{T}_\Omega(x^{(k)}) - x^{(k)}\|_{L^2}}_{\text{residual (what you already measure)}}$$

$$\underbrace{k_\varepsilon}_{\text{steps to tolerance}} \quad \ge \quad \underbrace{\left\lceil \frac{\ln(\varepsilon \,/\, \mathcal{L}_\Omega(x^{(0)}))}{\ln(\rho^2)} \right\rceil}_{\text{geometric decay rule}}$$

How to read it (everyday use)

- One number to watch: $\mathcal{L}_\Omega(x^{(k)}) = \frac{1}{2}\|\mathcal{T}(x^{(k)}) - x^{(k)}\|_{L^2}^2$. It *must* go down each step when the map is contractive.
- Stop rule: stop when $\sqrt{2\,\mathcal{L}_\Omega(x^{(k)})} \le \varepsilon$, or when its relative drop falls below a tiny threshold.
- If it stalls or rises: cool the gate (reduce β), increase normalization strength, or shorten the relaxation steprestore $\rho < 1$

Plain words: a single gauge falls, steadily; the field exhales, the motions fade, and the state comes to rest at $x^\star$.