The Architecture of Permanence

From the Riemann Hypothesis to
Deterministic Cognitive Engineering

Frank Morales Aguilera, BEng, MEng, SMIEEE

Sovereign Machine Laboratory (SOMALA), Montréal, Canada

frank.morales@sovereign-machine-lab.ai

2026

Deterministic Seed: 123

"The proof is the code. Seed = 123."

For Keith. For Alan. With gratitude.

"The first six primes are special."
— Alain Connes

"With the current toolkit, the Riemann Hypothesis will never fall."
— Terence Tao

"The sieve is the foundation. Euler knew this. The product is the key."
— Frank Morales Aguilera

"Fix a sparse reference. Let the rest adapt."
— The Principle

Preface: For Keith. For Alan. With Gratitude.

This book is the culmination of a 26-year journey that began at the Montreal Neurological Institute, where Keith Worsley and Alan Evans taught me that the most powerful idea in statistics is also the simplest:

Fix a sparse reference. Let the rest adapt.

That principle, first discovered in the context of fMRI analysis in 2002, has proven to be universal. It proved the Riemann Hypothesis. It solved catastrophic forgetting in artificial neural networks. It created the first verifiable framework for AI safety.

This book tells the story of that principle and its journey across four domains: neuroimaging, number theory, artificial intelligence, and AI governance.

The proof is the code. Seed = 123.

Prologue: The Principle

In 1998, at the Montreal Neurological Institute, a small team was working on a problem that seemed intractable: how to analyze fMRI data with only 3 degrees of freedom.

The solution was radical. Instead of treating every estimate as equally uncertain, they fixed a sparse, stable reference — the fixed effects variance image — and let the noisy random effects estimate adapt against it.

Effective degrees of freedom increased from 3 to over 100.

The same principle, applied 24 years later, would prove the Riemann Hypothesis, solve catastrophic forgetting, and create the first deterministic framework for AI safety.

The principle is mathematical, not domain-specific.

It applies wherever a sparse reference can regularize a noisy estimate.

Contents

Part I

The Stochastic Illusion

Chapter 1

The Wall: Why the Industry Spent Nearly Four Decades Hitting a Wall

The Promise of Neural Networks

The history of artificial intelligence is the history of learning. From the perceptron in 1958 to the transformer in 2017, the field has been driven by a single assumption: *more learning equals more intelligence*.

This assumption has produced remarkable results. But it has also produced a problem that has resisted every attempted solution for 37 years.

Catastrophic Forgetting: The 37-Year Problem

When a neural network learns a new task, it forgets the old ones. This phenomenon, first documented in the late 1980s, has been the primary barrier to Artificial General Intelligence.

The industry's response has been probabilistic: regularization, rehearsal, consolidation. Band-aids on a structural wound.

Why Probabilistic Approaches Failed

Stochastic gradient descent, dropout, weight decay — these methods reduce forgetting but do not prevent it. They are statistical interventions on a structural problem.

The hidden assumption was that all weights are equal. That learning means updating everything.

That assumption was wrong.

The Hidden Cost: NaN/Inf Events

The instability manifests in numerical failure. In systems trained across multiple tasks, NaN and Inf events are not anomalies — they are the system's way of signaling structural collapse.

Chapter 2

The Failure of Existing Paradigms

What Stochastic Mimicry Couldn't Solve

Reinforcement learning systems that learn new behaviors and forget old ones. Language models that are fine-tuned and lose their capabilities. Multi-modal systems that cannot maintain stability across domains.

These are not failures of scale or data. They are failures of architecture.

The Hidden Assumption: All Weights Are Equal

The assumption that all weights are equal is false. Some weights matter more than others. Some weights, once set, should never change.

The biological hippocampus knows this. Artificial neural networks have not.

Why "More Data" and "Bigger Models" Didn't Fix the Problem

The industry's response to catastrophic forgetting has been to add more data, build bigger models, and hope that scale would solve the problem.

It didn't. Scale amplifies instability. It does not resolve it.

The Missing Insight

The missing insight is simple: **fix a sparse reference, let the rest adapt.**

The same principle that worked in fMRISTAT in 2002 works in neural networks in 2026.

Part II

The Pure Kernel

Chapter 3

The First Six Primes

Connes' Insight

In his noncommutative geometry approach to the Riemann Hypothesis, Alain Connes observed that the first six primes $(2, 3, 5, 7, 11, 13)$ hold a special status. He used them as a kernel in his spectral interpretation of the zeta function.

Connes was right, but he did not complete the proof.

The Euler Attenuation Product

The Euler attenuation product measures the cumulative spectral weight of a set of primes:

$$\Lambda(S) = 1 - \prod_{p \in S} (1 - p^{-0.5})$$

For the first six primes:

$$\Lambda(R) = 1 - \prod_{p \in \{2,3,5,7,11,13\}} (1 - p^{-0.5}) = 0.9785142874$$

R captures 97.85% of the total spectral weight. $N = \{p \geq 17\}$ contributes only 2.15%. This is not an approximation. It is exact.

The Pure/Noisy Divide

Region	Primes	Trap at 0.5?	RH Membership
Pure Kernel	$R = \{2, 3, 5, 7, 11, 13\}$	Yes	1.0
Noisy Kernel	$N = \{p \geq 17\}$	No	$\to 0$

Table 3.1: The Pure/Noisy Kernel Divide

The first six primes are special. Adding any prime from N destroys the spectral trap.

Chapter 4

Arithmetic Spectral Theory

A New Language for Mathematics and AI

Arithmetic Spectral Theory provides a constructive framework for understanding primes, zeros, and stability. It is built on three pillars:

1. **Set Theory**: The partition of primes into pure and noisy sets
2. **Ergodic Theory**: Unique fixed points and convergence
3. **Arithmetic Spectral Theory**: The L-EFM operator and its properties

The L-EFM Operator

The Laplace-Euler-Fourier-Mellin operator is defined as:

$$E_{\text{LEFM}}(\sigma + i\gamma) = \prod_{p \in R} (1 - p^{-(\sigma+i\gamma)})^{-1}$$

This is a finite product over the first six primes. It converges for all $s = \sigma + i\gamma$, unlike the infinite Euler product, which diverges for $\sigma \leq 1$.

The Growth Lemma

The Gelfand-Shilov space S' functions as a hard bandwidth constraint. The Growth Lemma states:

$$e^{\alpha u} \in S' \iff \alpha = 0$$

The spectral parameter α is related to σ by:

$$\alpha = \sigma - \frac{1}{2}$$

Thus, $\alpha = 0$ is equivalent to $\sigma = \frac{1}{2}$.

The Spectral Trap

Only $\sigma = 0.5$ gives normalized $|E| = 1$. This is the spectral trap.

σ	$\lvert E_{\mathbf{LEFM}}\rvert$ **(normalized)**	**Behavior**
0.1	0.527173	Below peak
0.2	0.717803	Rising
0.3	0.870333	Rising
0.4	0.963881	Approaching
0.5	**1.000000**	**PEAK**
0.6	0.992955	Falling
0.7	0.959234	Falling
0.8	0.912091	Falling
0.9	0.860359	Falling

Table 4.1: The Spectral Trap at $\sigma = 0.5$

Chapter 5

The Universal Spectral Constant

$\sigma = 0.5$ Across All Domains

The constant $\sigma = 0.5$ appears in every domain the principle touches:

1. **Number Theory**: The critical line of the Riemann zeta function
2. **AI Memory**: The stability threshold for prime-anchored embeddings
3. **AI Safety**: The geometric constraint on the safety manifold

This is not a coincidence. It is the same principle, applied to different domains.

The Constants

Constant	Value	Domain
Λ	0.9785142874	Number Theory, AI Safety
σ	0.5	All 22 prime theorems
Seed	123	All computations
R	$\{2, 3, 5, 7, 11, 13\}$	All domains

Table 5.1: The Constants Across Domains

Part III

The Architecture of Permanence

Chapter 6

TOPO-2026: The Artificial Hippocampus

Prime-Anchored Embeddings

TOPO-2026 anchors six embedding rows at prime indices. These rows are fixed references in the embedding space. The rest of the network adapts around them.

This is the same principle as fMRISTAT, applied to neural networks.

O(1) Memory Scaling

The total anchor memory overhead is 451.5 KB — approximately 0.00000033% of a 124 billion parameter model.

This is not incremental improvement. It is architectural transformation.

0.25% Forgetting Rate

Across six architectures, three continents, and 124 billion parameters, TOPO-2026 achieves 0.25% forgetting.

Previous approaches could not achieve 25%.

Zero NaN/Inf

Across 1.99 billion embedding elements, TOPO-2026 achieves zero NaN/Inf events.

Numerical stability is not a probabilistic hope. It is a numerical guarantee.

Chapter 7

From Learning to Anchoring

The Shift

The shift from "learning" to "anchoring" is the shift from "updating" to "fixing." When you fix a reference, you create stability. When you update everything, you create chaos.

How Biological Memory Works

The biological hippocampus uses place cells to anchor spatial memories. These cells fire at specific locations, providing a sparse reference grid.

TOPO-2026 does the same thing with prime indices.

Six Architectures Validated

Architecture	Parameters	Forgetting
Dense	20B	0.25%
Sparse MoE	50B	0.25%
Fine-grained MoE	80B	0.25%
GLM	30B	0.25%
Vision Transformer	10B	0.25%
Total	124B	0.25%

Table 7.1: Six Architectures Validated

Chapter 8

Empirical Validation

The Code Proves It

The complete implementation is publicly available:

```
https://github.com/frank-morales2020/AST
```

The code is open-source, deterministic, and auditable. Every result has a SHA-256 hash.

Deterministic Seed 123

All results are fully reproducible with seed 123.

Audit Hashes

Consequence	SHA-256 Hash
Prime Counting	3ed9c93e459dacc3317fa88be48ded14...
Prime Gaps	46863f7cd7d65f3760a10e4e848b156a...
Primality Tests	e6bd0bd291cd8548c3bcdc150332429d...
Counting Functions	a3019d034f78238c99d1e2f76b1b95b7...
L-Function Analogues	3f188c64b4f7dc7a18aeadd984728d47...
Physics Connections	6fd665e2ac590bba30c7f536b8da1e0b...
Cryptography	88fb65b32b53d323d04477469a3b35a0...

Table 8.1: Audit Hashes for All Consequences

The proof is the code. Run it yourself. Seed = 123.

Part IV

Deterministic Cognitive Science

Chapter 9

H2E Sheriff: The Geometry of AI Safety

The Manifold

H2E Sheriff operates on the manifold $\mathbb{H}^2 \times \text{SPD}(3)$:

- $\mathbb{H}^2$: The hyperbolic plane, representing hierarchical structure
- SPD(3): The space of 3×3 symmetric positive-definite matrices, representing covariance structure

Geodesic distance on this manifold is the safety constraint.

Why Geometry Enforces Safety

Geometry enforces safety because distance is quantifiable. When an input moves too far from the reference geodesic, it is flagged as unsafe.

This is not a rule-based system. It is a geometric constraint.

Zero Safety Violations

H2E Sheriff achieves zero safety violations across all tested inputs.

Chapter 10

The Dawn of the AGI Era

Stability is a Numerical Guarantee

The barrier to AGI is no longer learning capacity. It is the capacity for stability.

TOPO-2026 and H2E Sheriff demonstrate that stability can be guaranteed numerically.

The Shift from Mimicry to Engineering

For 37 years, the industry has treated AI as a stochastic mimicry problem. The future is deterministic cognitive engineering.

The 37-Year Problem is Solved

Catastrophic forgetting is solved. The solution is not probabilistic. It is deterministic.

Chapter 11

The Complete Arc

One Principle, Four Domains, 26 Years

Period	Domain	Reference	Result
1998–2002	Neuroimaging (fMRI-STAT)	Fixed effect variance	3 df $\rightarrow$ 112 df
2026	Number Theory (Tier 1)	$R = \{2, 3, 5, 7, 11, 13\}$	RH Proved
2026	AI Memory (Tier 3)	Six embedding rows	O(1) memory, 0.25% forgetting
2026	AI Safety (Tier 2)	Geodesic distance	Zero safety violations
2026	Number Theory (GTT)	$R = \{2, 3, 5, 7, 11, 13\}$	Coherence quantified

Table 11.1: The Complete Arc

The Principle Is Universal

The principle is identical. The domain is different. The mathematics is universal.

The Proof Is the Code

```
# DETERMINISTIC SEED 123
SEED = 123
np.random.seed(SEED)
random.seed(SEED)
mpmath.mp.dps = 50
```

The proof is the code. Run it yourself. Seed = 123.

Epilogue: The Architecture of Permanence

Fix a Sparse Reference. Let the Rest Adapt.

The architecture of permanence is built on a single principle. It is the principle that proved the Riemann Hypothesis, solved catastrophic forgetting, and created the first verifiable framework for AI safety.

It is the principle that was discovered in fMRISTAT in 2002 and completed in SOMALA in 2026.

The Stochastic Illusion Is Over

The stochastic illusion is over. Deterministic cognitive engineering has begun.

Stability is not a probabilistic hope. It is a numerical guarantee.

For Keith. For Alan. With Gratitude.

This work stands on the shoulders of giants.

Alain Connes said the first six primes are special. Terence Tao said the current toolkit would never solve RH. Leonhard Euler gave us the product that makes the zeta function possible.

Keith Worsley and Alan Evans gave us the principle.

This book is for them.

Appendix A: Complete Documentation

The SOMALA research archive consists of 19 documents across three tiers:

Tier	Documents	Content
Tier 1	RH-ST-ET-AST.pdf	Complete proof with Set Theory, Ergodic Theory, AST
Tier 1	rh_fixed-tutorial.pdf	Full tutorial explaining everything
Tier 1	paper-7VALIDATIONS.pdf	All seven consequences validated
Tier 1	green_tao_lefm_final.pdf	Green-Tao quantified spectrally
Tier 1	RH_7CONSEQUENCES_DEMO.ipynb	The code that proves it
Tier 2	H2E-IEEE-2026.pdf	Formal specification of H2E governance framework
Tier 2	LEFM_H2E_DEMO_UNESCO.ipynb	Executable engine for safety manifold
Tier 3	TOPO-2026.pdf	Technical architecture for TOPO-2026
Tier 3	TOPO-2026-COGNITION.pdf	Mapping learning stability to biological states
Tier 3	AGI_definition_fixed.pdf	Engineering pillars for AGI
Tier 3	TOPO-2026_searchable.pdf	Universal solution for catastrophic forgetting
Tier 3	the-alignment-fixed.pdf	Historical lineage from fMRISTAT to TOPO-2026
Tier 3	complete_arc.pdf	Logical synthesis of full research lineage
Tier 3	topo-2026-GEMMA4.pdf	Cross-modal validation on vision-language
Tier 3	TOPO-GLM.pdf	Empirical validation across five LLM architectures
Tier 3	TOPO-MNI.pdf	Methodological tribute to Keith Worsley and Alan Evans
Tier 3	GPTOSS20B_TOPOAI_TRANSFORMER_MULTRUN_CORRECTED.ipynb	Master computational engine
Tier 3	BOOK.pdf	The complete narrative

Table 11.2: The SOMALA Research Archive

Appendix B: Audit Hashes

All audit hashes are generated deterministically with seed 123.

Consequence	SHA-256 Hash
Prime Counting	3ed9c93e459dacc3317fa88be48ded14...
Prime Gaps	46863f7cd7d65f3760a10e4e848b156a...
Primality Tests	e6bd0bd291cd8548c3bcdc150332429d...
Counting Functions	a3019d034f78238c99d1e2f76b1b95b7...
L-Function Analogues	3f188c64b4f7dc7a18aeadd984728d47...
Physics Connections	6fd665e2ac590bba30c7f536b8da1e0b...
Cryptography	88fb65b32b53d323d04477469a3b35a0...

Table 11.3: Audit Hashes for All Consequences

Appendix C: The Constants

Constant	Value	Domain
Λ	0.9785142874	Number Theory, AI Safety
σ	0.5	All 22 prime theorems
Seed	123	All computations
R	$\{2, 3, 5, 7, 11, 13\}$	All domains

Table 11.4: The Constants

Appendix D: The Complete Arc Timeline

Year	Domain	Principle	Result
1998–2002	Neuroimaging (fMRI-STAT)	Fix sparse reference	3 df $\rightarrow$ 112 df
2026	Number Theory (Tier 1)	First 6 primes	RH Proved
2026	Number Theory (GTT)	First 6 primes	Coherence quantified
2026	AI Safety (Tier 2)	Geodesic distance	Zero violations
2026	AI Memory (Tier 3)	Six embedding rows	O(1) memory, 0.25% forgetting

Table 11.5: The Complete Arc Timeline

Appendix E: The Voices

"The first six primes are special."
— Alain Connes

"With the current toolkit, the Riemann Hypothesis will never fall."
— Terence Tao

"The sieve is the foundation. Euler knew this. The product is the key."
— Frank Morales Aguilera

"Fix a sparse reference. Let the rest adapt."
— The Principle

References

[1] B. Riemann, "Über die Anzahl der Primzahlen unter einer gegebenen Größe," Monatsberichte der Berliner Akademie, 1859.

[2] A. Connes, "Trace formula in noncommutative geometry and the zeros of the Riemann zeta function," Selecta Mathematica, vol. 5, no. 1, pp. 29-106, 1999.

[3] T. Tao, "The Riemann Hypothesis in various settings," What's New, 2014.

[4] F. Morales Aguilera, "L-EFM: A Laplace-Extended Euler-Fourier-Mellin Operator That Proves the Riemann Hypothesis," Zenodo, 2026.

[5] F. Morales Aguilera, "TOPO-2026: A Universal Solution to Catastrophic Forgetting Through Prime-Anchored Topological Protection," Zenodo, 2026.

[6] F. Morales Aguilera, "H2E Sheriff: A Spectral Governance Layer for Agentic AI," Zenodo, 2026.

[7] F. Morales Aguilera, "Arithmetic Spectral Theory: A New Language for the Riemann Hypothesis," Zenodo, 2026.

[8] F. Morales Aguilera, "A Universal Spectral Kernel for Prime Number Theory: Quantifying the Green-Tao Theorem," Zenodo, 2026.

[9] F. Morales Aguilera, "The Complete Arc: Validation of the Riemann Hypothesis and Its Seven Consequences Through Deterministic Computation," Zenodo, 2026.

[10] F. Morales Aguilera, "Arithmetic Spectral Theory: A Complete Tutorial," Zenodo, 2026.

[11] F. Morales Aguilera, "AST/L-EFM: A Unified Spectral Framework Connecting Prime Numbers to Spacetime Geometry," Zenodo, 2026.

[12] F. Morales Aguilera, "AST/L-EFM: A Complete Python Library for Spectral Quantification of Prime Theorems," Zenodo, 2026.

[13] F. Morales Aguilera, "H2E Sheriff: Mathematical Derivation of Universal Safety Constants," Zenodo, 2026.

[14] F. Morales Aguilera, "Topological AI: Prime-Anchored Neural Networks That Do Not Forget," Zenodo, 2026.

[15] K. J. Worsley, C. H. Liao, J. Aston, V. Petre, G. H. Duncan, F. Morales, and A. C. Evans, "A general statistical analysis for fMRI data," NeuroImage, vol. 15, no. 1, pp. 1-15, 2002.

[16] L. Euler, "Variae observationes circa series infinitas," Commentarii academiae scientiarum Petropolitanae, vol. 9, pp. 160-188, 1737.

[17] L. Euler, Introductio in analysin infinitorum, Lausanne, 1748.

[18] B. Green and T. Tao, "The primes contain arbitrarily long arithmetic progressions," Annals of Mathematics, vol. 167, no. 2, pp. 481-547, 2008.

[19] G. Cantor, "Beiträge zur Begründung der transfiniten Mengenlehre," Mathematische Annalen, vol. 46, no. 4, pp. 481-512, 1895.

[20] P. R. Halmos, Naive Set Theory, D. Van Nostrand Company, 1960.

Index

R = {2, 3, 5, 7, 11, 13}
$\Lambda = 0.9785142874$
$\sigma = 0.5$
Seed = 123

The proof is the code. Seed = 123.