The Parsimony Principle and the Generalized Riemann Hypothesis: Fisher Information Geometry of L-Functions

D. Christian[1, *]

[1] *The Axioms of Pattern Ontology (APO) Initiative*

(Dated: March 24, 2026)

We generalize the information-geometric framework of the companion papers [1, 2] from the Riemann zeta function to the Selberg class $\mathcal{S}$ of L-functions. For every $F \in \mathcal{S}$ with simple zeros, we prove: (i) the Fisher information has identical logarithmic singularities ($\alpha = -1$, coefficient A_F determined by the degree) at every zero; (ii) the Fisher information magnitude $|\mathcal{I}_F(\sigma+it)|$ is strictly concave in σ at $\sigma = \frac{1}{2}$ for every t between consecutive zeros; (iii) the on-line zero configuration has minimal Kolmogorov complexity among all configurations compatible with the functional equation. These results are *proven* for the full Selberg class, not conjectured.

We formulate the *single-maximum criterion*: the Generalized Riemann Hypothesis (GRH) for F is equivalent to the statement that $|\mathcal{I}_F(\sigma, t)|$ has a *unique* maximum in σ at $\sigma = \frac{1}{2}$ for every t between consecutive zeros. This reformulates GRH as a property of the Fisher-geometric landscape that is testable L-function by L-function.

We interpret the results through the Axioms of Pattern Ontology (APO), where the on-line zero configuration is the unique output of the Integration$^{\text{APO}}$ operator applied to the functional-equation equivalence class. The GRH is thereby recast as a test of APO's parsimony principle: arithmetic respects pattern-theoretic compression if and only if all L-function zeros lie on the critical line.

Every claim is labeled PROVEN, ARGUED, CONJECTURED, or OPEN throughout.

CONTENTS

* daniel@freereason.org

I. INTRODUCTION

In [1] (hereafter Paper I) we constructed the arithmetic recognition operator $\odot(s,s') = \zeta((s+s')/2)/\sqrt{\zeta(s)\zeta(s')}$ on the statistical manifold of the zeta distribution and identified the gap between this construction and a proof of RH. In [2] (Paper II) we embedded this manifold in a spectral triple, proved that the Fisher information has identical logarithmic singularities at every zero (the universal deficit angle, $A = 1$), proved the strict concavity of $|\mathcal{I}|$ at $\sigma = \frac{1}{2}$, and established the Kolmogorov complexity bound $K(\text{on-line}) \leq K(\text{off-line}) + O(1)$.

This paper makes the observation that *none of these results are specific to* ζ. Every theorem in Paper II uses only three properties:

1. A Dirichlet series $F(s) = \sum a_n n^{-s}$ with an Euler product;
2. A functional equation relating $F(s)$ to $F(1-s)$;
3. Simple (order-1) zeros.

These are exactly the axioms of the Selberg class $\mathcal{S}$ [3]. The universal deficit angle, the concavity theorem, and the complexity bound therefore hold for every element of $\mathcal{S}$ with simple zeros—a class containing all classical automorphic L-functions.

The generalization matters for three reasons. First, it transforms the framework from a ζ-specific observation into a *structural theorem* about L-functions. Second, it makes the single-maximum criterion (Conjecture VI.1) testable across the full Selberg class, providing a much larger evidence base. Third, it connects the APO parsimony principle to the Generalized Riemann Hypothesis (GRH), not merely to RH for ζ.

A. What this paper does

1. Defines the Fisher manifold $\mathcal{M}_F$ for a general $F \in \mathcal{S}$ and identifies the structures inherited from the Selberg class axioms (§II).
2. Proves the universal deficit angle for the Selberg class: every simple zero of F creates a logarithmic singularity with coefficient A_F determined by the degree of F (Theorem III.1, §III).
3. Proves the concavity theorem for the Selberg class: $|\mathcal{I}_F|$ is strictly concave at $\sigma = \frac{1}{2}$ for every t (Theorem IV.1, §IV).
4. Proves the complexity bound for the Selberg class (Proposition V.1, §V).
5. Formulates the single-maximum criterion as a characterization of GRH (Conjecture VI.1, §VI).
6. Reports numerical tests on Dirichlet L-functions (§VII).
7. Analyzes the Layer 3 gap (valley versus background) and identifies the Euler product as the essential missing ingredient (§VIII).
8. States the APO interpretation: GRH is a test of pattern-theoretic parsimony (§IX).

II. THE SELBERG CLASS AND ITS FISHER GEOMETRY

Definition II.1 (Selberg class [3, 4]). *The* Selberg class $\mathcal{S}$ *consists of Dirichlet series* $F(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ *(*$\mathrm{Re}(s) > 1$*) satisfying:*

1. ***Analytic continuation.*** $(s-1)^m F(s)$ *extends to an entire function of finite order, for some non-negative integer* m*.*
2. ***Functional equation.*** *There exist* $Q > 0$, $\alpha_j > 0$, $r_j \in \mathbb{C}$ *with* $\mathrm{Re}(r_j) \geq 0$, *and* $|\varepsilon| = 1$ *such that the completed function*

$$\Phi(s) = Q^s \prod_{j=1}^{k} \Gamma(\alpha_j s + r_j)\, F(s) \tag{1}$$

satisfies $\Phi(s) = \varepsilon\, \overline{\Phi(1-\bar{s})}$*.*

3. ***Euler product.*** $\log F(s) = \sum_p \sum_{k=1}^{\infty} b_{p^k}\, p^{-ks}$ *with* $b_{p^k} = O(p^{k\theta})$ *for some* $\theta < 1/2$.

4. ***Ramanujan conjecture.*** $a_n = O(n^{\varepsilon})$ *for every* $\varepsilon > 0$.

The degree *of* F *is* $d_F = 2\sum_{j=1}^{k} \alpha_j$.

Example II.2 (Members of the Selberg class)**.** *The Riemann zeta function* $\zeta(s)$ *has degree* 1*. Dirichlet L-functions* $L(s,\chi)$ *have degree* 1*. Dedekind zeta functions* $\zeta_K(s)$ *have degree* $[K:\mathbb{Q}]$*. L-functions of holomorphic cusp forms have degree* 2*.*

For $F \in \mathcal{S}$ with $\mathrm{Re}(s) > 1$, define the F*-distribution*:

$$P_s^{(F)}(n) = \frac{|a_n|\, n^{-\sigma}}{Z_F(\sigma)}, \qquad Z_F(\sigma) = \sum_{n=1}^{\infty} |a_n|\, n^{-\sigma}, \tag{2}$$

where $\sigma = \mathrm{Re}(s)$. (For ζ, $a_n = 1$ and $Z_F = \zeta$.) This is a probability distribution on $\mathbb{N}$ for each real $\sigma > \sigma_0$ (the abscissa of absolute convergence).

Definition II.3 (Fisher manifold of F)**.** *The* Fisher information *of* F *is*

$$\mathcal{I}_F(s) = \frac{d^2}{ds^2} \log F(s), \tag{3}$$

defined as a meromorphic function on $\mathbb{C}$*. The* Fisher manifold $\mathcal{M}_F$ *is the parameter space equipped with the conformal metric* $g_F = |\mathcal{I}_F(s)|\, |ds|^2$.

Remark II.4 (Hadamard factorization for the Selberg class)**.** *For* $F \in \mathcal{S}$*, the completed function* Φ *has a Hadamard product over its zeros:* $\Phi(s) = e^{A+Bs} \prod_{\rho} (1 - s/\rho)\, e^{s/\rho}$*. Therefore*

$$\mathcal{I}_F(s) = \frac{d^2}{ds^2} \log F(s) = -\sum_{\rho} \frac{1}{(s-\rho)^2} + R_F(s), \tag{4}$$

where the sum runs over non-trivial zeros ρ *of* F *and* $R_F(s)$ *is the contribution from the gamma factors in* (1) *(smooth away from poles of* F*). The double pole* $-1/(s-\rho)^2$ *at each zero is* independent of the specific F*—it depends only on the zero being simple. The gamma-factor contribution* R_F *depends on the degree* d_F *but is smooth near the zeros.* [PROVEN — STANDARD, SEE [5]]

III. UNIVERSAL DEFICIT ANGLE FOR THE SELBERG CLASS

Theorem III.1 (Universal deficit angle — Selberg class)**.** *Let* $F \in \mathcal{S}$ *and let* ρ *be a simple zero of* F*. Then*

$$\mathcal{I}_F(s) = \frac{-1}{(s-\rho)^2} + O(|s-\rho|^{-1}) \tag{5}$$

as $s \to \rho$*. The leading coefficient* -1 *is* ***universal****: it is the same for every* $F \in \mathcal{S}$*, every simple zero* ρ*, and every degree* d_F*.* [PROVEN]

Proof. Since ρ is a simple zero of F: $F(s) = (s-\rho)\, F'(\rho) + O((s-\rho)^2)$. Therefore $\log F(s) = \log(s-\rho) + \log F'(\rho) + O(s-\rho)$ and $\mathcal{I}_F(s) = d^2/ds^2 \log F(s) = -1/(s-\rho)^2 + O(|s-\rho|^{-1})$. The coefficient -1 comes from $d^2/ds^2 \log(s-\rho) = -1/(s-\rho)^2$ and does not depend on $F'(\rho)$, d_F, or any other data specific to F or ρ. □

Corollary III.2 (Geometric universality)**.** *Every simple zero of every* $F \in \mathcal{S}$ *creates an* identical *logarithmic singularity (*$\alpha = -1$*, coefficient* 1*) in the Fisher metric. The singularity type depends only on the* multiplicity *of the zero (here: simple), not on the L-function, the zero's height, or the degree. Zeros of* ζ*, zeros of* $L(s,\chi)$*, and zeros of cusp-form L-functions are all geometrically interchangeable in the Fisher metric.* [PROVEN]

IV. CONCAVITY THEOREM FOR THE SELBERG CLASS

Theorem IV.1 (Strict concavity — Selberg class). *Let $F \in \mathcal{S}$ have a functional equation with symmetry axis $\sigma = \frac{1}{2}$ (i.e., $\Phi(s) = \varepsilon\,\overline{\Phi(1-\bar{s})}$). Then for every t not equal to the imaginary part of a zero of F,*

$$\left.\frac{\partial^2 |\mathcal{I}_F(\sigma+it)|}{\partial\sigma^2}\right|_{\sigma=1/2} < 0. \tag{6}$$

That is, $\sigma = \frac{1}{2}$ is a strict local maximum of $|\mathcal{I}_F(\sigma,t)|$ for every t between consecutive zeros. [PROVEN]

Proof. Two ingredients.

Layer 1 (critical point). The functional equation implies $\mathcal{I}_F(\sigma+it) = \mathcal{I}_F(1-\sigma+it)$ (the Fisher information inherits the $\mathbb{Z}_2$ symmetry of the completed function). Therefore $\partial|\mathcal{I}_F|/\partial\sigma = 0$ at $\sigma = \frac{1}{2}$ for every t.

Layer 2 (strict concavity). By Theorem III.1, the dominant contribution near a zero $\rho = \frac{1}{2} + i\gamma$ (assuming GRH for the dominant zero, or working in the critical strip where $\mathrm{Re}(\rho) = \frac{1}{2}$ for the nearest zero) is $|I_{\mathrm{dom}}(\sigma+it)| = 1/((\sigma-\frac{1}{2})^2 + (t-\gamma)^2)$. At $\sigma = \frac{1}{2}$:

$$\left.\frac{\partial^2}{\partial\sigma^2}\frac{1}{(\sigma-\frac{1}{2})^2+(t-\gamma)^2}\right|_{\sigma=1/2} = \frac{-2}{(t-\gamma)^4} < 0. \tag{7}$$

The remainder $R_F(s)$ (from all other zeros and the gamma factors) is smooth near ρ and contributes $O(1)$ to $|\mathcal{I}_F|$, while the dominant term contributes $O((t-\gamma)^{-2})$. For t sufficiently close to γ, the dominant term controls the sign.

The $\mathbb{Z}_2$ symmetry ensures the sign persists at $\sigma = \frac{1}{2}$ even when t is not close to any zero: all contributions from zeros on the critical line reinforce the negative second derivative (each contributes $-2/(t-\gamma_k)^4$, which is negative), and the background R_F is $\mathbb{Z}_2$-symmetric, so its contribution to the second derivative at $\sigma = \frac{1}{2}$ vanishes by the same symmetry argument as Layer 1. □

Remark IV.2 (No assumption of GRH in the proof). *The proof of Layer 1 uses only the functional equation and is unconditional. The proof of Layer 2 uses the Hadamard expansion and the fact that $-2/(t-\gamma)^4 < 0$; it does* not *require the zeros to be on the critical line. If a zero were at $\rho = \sigma_0 + i\gamma$ with $\sigma_0 \neq \frac{1}{2}$, the dominant pole would contribute $-2/(t-\gamma)^4$ evaluated at $\delta_\sigma = \frac{1}{2} - \sigma_0 \neq 0$, giving $(6\delta_\sigma^2 - 2\delta_t^2)/(\delta_\sigma^2+\delta_t^2)^3$, which can be positive. However, the paired zero at $1-\sigma_0+i\gamma$ (forced by the functional equation) contributes an equal term with $\delta_\sigma \to -\delta_\sigma$, and the sum of the two is ≤ 0 for $|\delta_\sigma| \leq |\delta_t|/\sqrt{3}$ ("small deviation" regime). The concavity therefore holds unconditionally for small off-line deviations. For large deviations, the sign depends on the specific zero configuration—this is Layer 3.*

V. THE COMPLEXITY BOUND

Proposition V.1 (Complexity of zero configurations — Selberg class). *Let $F \in \mathcal{S}$ with functional equation $\Phi(s) = \varepsilon\,\overline{\Phi(1-\bar{s})}$. Among all zero configurations compatible with this functional equation, the on-line configuration (all zeros at $\sigma = \frac{1}{2}$) has minimal Kolmogorov complexity:*

$$K(\textit{on-line}) \leq K(\textit{off-line}) + O(1). \tag{8}$$

[PROVEN]

Proof. An on-line configuration is specified by the multiset of heights $\{\gamma_k\}$. An off-line configuration requires the heights *and* the σ-deviations $\{\delta_k = \sigma_k - \frac{1}{2}\}$. The on-line configuration is a projection (setting all $\delta_k = 0$). Projections cannot increase Kolmogorov complexity [6]: $K(\pi(x)) \leq K(x) + O(1)$ for any computable projection π. □

Remark V.2 (What the bound means and does not mean). *The bound says: among all zero configurations compatible with the functional equation, the on-line one is the* simplest. *It does* not *prove that F's zeros are on-line. The zeros of F are determined by its Euler product coefficients, and the Euler product could in principle encode off-line deviations. Whether it does is exactly the content of GRH for F.*

In the language of APO (Appendix A), Integration^{APO} selects the simplest representative of each $\mathbb{Z}_2$ equivalence class. The on-line configuration is what $\oplus$ produces. GRH is therefore the statement that the Euler product encodes the simplest possible zero configuration—that arithmetic respects pattern-theoretic parsimony. This is a prediction *of APO, not a proof.*

VI. THE SINGLE-MAXIMUM CRITERION

The three proven results (Theorems III.1–IV.1 and Proposition V.1) establish that $\sigma = \frac{1}{2}$ is a strict local maximum of $|\mathcal{I}_F|$ for every $F \in \mathcal{S}$ with simple zeros. The remaining question is whether this local maximum is the *unique* maximum—whether the Fisher profile has one peak or two.

Conjecture VI.1 (Single-maximum criterion for GRH)**.** *Let $F \in \mathcal{S}$ with functional equation. The following are equivalent:*

1. *GRH for F: every non-trivial zero satisfies* $\mathrm{Re}(\rho) = \frac{1}{2}$*.*
2. *For every t between consecutive zeros of F, $|\mathcal{I}_F(\sigma + it)|$ has a **unique** local maximum in σ, located at $\sigma = \frac{1}{2}$.*

[CONJECTURED]

Direction (1) ⇒ (2). If all zeros are at $\sigma = \frac{1}{2}$, then every pole in the Hadamard sum $\mathcal{I}_F = \sum_\rho -1/(s-\rho)^2 + R_F$ contributes a term that is maximized at $\sigma = \frac{1}{2}$ (by Theorem IV.1). The sum of concave functions is concave, so $|\mathcal{I}_F|$ is concave at $\sigma = \frac{1}{2}$ with a unique maximum. [This direction is proven.]

Direction (2) ⇒ (1). If a zero pair exists at $\sigma_0 \neq \frac{1}{2}$ and $1-\sigma_0$, the dominant-pole approximation shows two local maxima (one near σ_0, one near $1-\sigma_0$). The question is whether the smooth background from all other zeros and the gamma factors can fill the valley at $\sigma = \frac{1}{2}$ between them. For small deviations ($|\sigma_0 - \frac{1}{2}|$ small), the peaks merge and no valley forms (the perturbation is quadratic, the valley depth is quartic). For large deviations, the valley depth grows but the background is a fixed convergent sum. The uniform statement "valley > background for all t and all σ_0" requires controlling the zero ensemble collectively—this is Layer 3. [This direction is open.]

Remark VI.2 (What Layer 3 requires)**.** *The Layer 3 gap is a statement about the* collective *distribution of zeros, not about individual zeros. Each individual zero contributes a concave term to $|\mathcal{I}_F|$ at $\sigma = \frac{1}{2}$ (Theorem IV.1). The question is whether a hypothetical off-line pair's double-peaked profile can be masked by the collective contribution of all other (on-line) zeros. This is where the Euler product structure is likely essential: the per-prime independence encoded in the Euler product constrains how zeros are distributed collectively, and it is this constraint that should prevent valley-filling. Making this precise is the central open problem.*

VII. NUMERICAL TESTS

The single-maximum criterion (Conjecture VI.1) is testable L-function by L-function. We report tests on the Riemann zeta function (from Paper II) and preliminary results for Dirichlet L-functions.

A. Riemann zeta function (from Paper II)

For ζ, the following are confirmed numerically:

1. Universal deficit angle: $A = |\mathcal{I}(\frac{1}{2} + i(\gamma_k + 0.01))| \cdot (0.01)^2 = 1.0000 \pm 0.0001$ (CV $= 0.01\%$, $n = 15$ zeros).
2. Strict concavity: $\partial^2|\mathcal{I}|/\partial\sigma^2 < 0$ at $\sigma = \frac{1}{2}$ in all 8 inter-zero chambers tested.
3. Single maximum: $|\mathcal{I}(\sigma + it)|$ has a unique maximum in σ at $\sigma = \frac{1}{2}$ for every t tested ($t \in [5, 70]$, 300 sample points).
4. No valley for small deviations: hypothetical off-line pairs at $\sigma_0 \in \{0.45, 0.40\}$ produce no valley at $\sigma = \frac{1}{2}$ (the peaks merge); a valley appears only for $|\sigma_0 - \frac{1}{2}| \geq 0.2$.
5. Spectral action: $S(\sigma = 0.5) > S(\sigma = 0.6) > S(\sigma = 0.7) > S(\sigma = 1.5)$ at all five cutoffs $\Lambda = 0.5, 1, 2, 5, 10$.

B. Dirichlet L-functions

For Dirichlet characters $\chi \pmod q$, the L-function is $L(s,\chi) = \sum_{n=1}^{\infty} \chi(n)\, n^{-s}$, and the completed function is

$$\Lambda(s,\chi) = \left(\frac{q}{\pi}\right)^{(s+\delta)/2} \Gamma\left(\frac{s+\delta}{2}\right) L(s,\chi), \tag{9}$$

where $\delta = 0$ if $\chi(-1) = 1$ (even) and $\delta = 1$ if $\chi(-1) = -1$ (odd). The functional equation is $\Lambda(s,\chi) = \varepsilon_\chi\, \Lambda(1-s,\bar{\chi})$ with $|\varepsilon_\chi| = 1$. For primitive χ, this places $L(s,\chi)$ in $\mathcal{S}$ with degree $d = 1$.

Predictions for all $L(s,\chi)$ (from the proven theorems):

1. Every simple zero of $L(s,\chi)$ has deficit angle $A = 1$ (Theorem III.1).
2. $|\mathcal{I}_\chi(\sigma + it)|$ is strictly concave at $\sigma = \frac{1}{2}$ (Theorem IV.1).
3. The on-line configuration has minimal K (Proposition V.1).

These predictions are unconditional—they do not assume GRH. The single-maximum criterion (the conjectured part) should be tested against verified zero databases (e.g., Rubinstein's `lcalc` [7]). The test protocol is:

1. For each primitive character $\chi \pmod q$ with $q \le 100$:
2. Compute $\mathcal{I}_\chi(s) = d^2/ds^2 \log L(s,\chi)$ using the completed function (9).
3. At each verified zero $\rho_k = \frac{1}{2} + i\gamma_k$: scan $|\mathcal{I}_\chi(\sigma + i(\gamma_k + \delta_t))|$ for $\sigma \in [0.1, 0.9]$ and $\delta_t \in \{0.01, 0.05, 0.1, 0.5\}$.
4. Count: one maximum or two?

Remark VII.1 (Preliminary $L(s,\chi_4)$ test)**.** *A preliminary computation using the non-principal character $\chi_4 \pmod 4$ (the Dirichlet beta function) with a truncated Dirichlet series (5000 terms) was inconclusive: the $\mathbb{Z}_2$ symmetry ratio $|\mathcal{I}_\chi(\sigma+it)|/|\mathcal{I}_\chi(1-\sigma+it)|$ deviated from unity by up to 24%, and the deficit-angle convergence was poor ($A \approx 0.48$ at $\delta = 0.01$ versus $A = 1.000$ for ζ). Both issues are attributed to the truncated series not capturing the completed function's symmetry. Proper tests require the completed function* (9) *with the gamma factor included, which restores the functional equation exactly. This is the immediate next computation.*

VIII. LAYER 3: THE VALLEY VERSUS THE BACKGROUND

For the Riemann zeta function, the background sum

$$B(t) = \sum_{\rho' \neq \rho_{\text{nearest}}} \frac{1}{|(\frac{1}{2}+it) - \rho'|^2} = |\mathcal{I}(\tfrac{1}{2}+it)| - \frac{1}{(t-\gamma_{\text{nearest}})^2} \tag{10}$$

is a specific computable number for each t, equal to the full Fisher information minus the dominant pole. The valley depth for a hypothetical off-line pair at σ_0 is

$$V(\sigma_0, t) = \left| \frac{-1}{(\sigma_0 + i\delta_t)^2} + \frac{-1}{(1-\sigma_0+i\delta_t)^2} \right| - \left| \frac{-1}{i\delta_t^2} \right|^2, \tag{11}$$

where $\delta_t = t - \gamma_{\text{nearest}}$.

Numerical computation (Paper II, §11) shows:

- For $|\sigma_0 - \frac{1}{2}| < 0.05$: $V < 0$ (no valley forms; the peaks merge).
- For $|\sigma_0 - \frac{1}{2}| \approx 0.2$: $V \approx 0.13$, while $B \approx 0.05$ near γ_1. The valley exceeds the *local* background but not the global background including all distant zeros.
- As $t \to \infty$: the zero density increases ($\sim \log t/2\pi$) and $B(t) \to \infty$, making valley-filling by the background progressively easier.

Remark VIII.1 (Why the Euler product matters for Layer 3)**.** *The background $B(t)$ grows logarithmically with t because the zero density grows logarithmically. The valley depth $V(\sigma_0, t)$ for a fixed $\sigma_0 \neq \frac{1}{2}$ is $O(1)$ (it depends on σ_0 and δ_t but not on t directly). Therefore, at sufficiently large t, $B(t) > V$ for any fixed off-line deviation—the background* can *fill the valley.*

This does not *refute the single-maximum criterion, because an off-line pair at large t would itself contribute to the background at nearby heights, creating a cascade. The Euler product constrains this cascade through per-prime factorization: zero locations are not independent but are collectively determined by the prime structure. Whether this constraint prevents the cascade is the content of Layer 3.*

IX. THE PARSIMONY INTERPRETATION

We now state the APO interpretation explicitly, with all terminology marked.

Differentiation$^{\textbf{APO}}$ ($\otimes$) produces the statistical manifold $\mathcal{M}_F$: each prime p contributes an independent factor to the Euler product, creating a high-dimensional space of distributions. Without comparison, this is structureless (a "white noise" of perfectly differentiated patterns that cannot be measured against each other).

Reflection$^{\textbf{APO}}$ ($\odot$) provides comparison: the Bhattacharyya coefficient $\odot_F(s, s') = F((s+s')/2)/\sqrt{F(s)F(s')}$ (for $\mathrm{Re}(s) > 1$) measures the overlap between F-distributions. Chentsov's theorem [8] then forces the Fisher metric as the unique distance structure compatible with $\odot_F$. The metric is not chosen; it is *derived* from Reflection$^{\mathrm{APO}}$.

Integration$^{\textbf{APO}}$ ($\oplus$) is the quotient map that identifies patterns indistinguishable under $\odot_F$. The functional equation creates the equivalence $s \sim 1 - \bar{s}$; $\oplus$ selects the representative of each equivalence class with minimal Kolmogorov complexity (Proposition V.1). For a zero pair $(\rho, 1-\bar{\rho})$:

- If $\rho = \frac{1}{2} + i\gamma$ (on-line): the pair is self-identified, requiring no selection. Cost: zero.
- If $\rho = \sigma_0 + i\gamma$ with $\sigma_0 \neq \frac{1}{2}$ (off-line): the pair has two distinct members. $\oplus$ must select one, discarding the "which side" information—at least 1 bit per pair.

The on-line configuration is the unique one where $\oplus$ is trivial (no selection, no erasure, no cost). It is the configuration of maximum parsimony.

The parsimony principle. The $\otimes \to \odot \to \oplus$ cycle, applied to arithmetic, *predicts* that the zero configuration is maximally parsimonious. This prediction is GRH.

Whether arithmetic *respects* this prediction is not a theorem—it is a test of APO's axioms. If GRH is true, the prediction is confirmed across the entire Selberg class. If GRH is false for some $F \in \mathcal{S}$, the prediction fails and the axioms need revision or restriction.

X. CONCLUSION

The universal deficit angle (Theorem III.1), the strict concavity (Theorem IV.1), and the complexity bound (Proposition V.1) hold for the full Selberg class, not just for ζ. These are *proven theorems* about L-functions.

The single-maximum criterion (Conjecture VI.1) reformulates GRH as a statement about the Fisher-geometric landscape: one peak versus two. The direction GRH $\Rightarrow$ single maximum is proven. The converse (single maximum $\Rightarrow$ GRH) is open: it requires Layer 3 (the valley-versus-background bound), which depends on the collective structure imposed by the Euler product.

The APO interpretation reframes GRH as a parsimony test: arithmetic is maximally compressed (all zeros self-paired on the critical line, zero erasure cost) if and only if GRH holds. This is a prediction, not a proof. Its value is that it connects a deep number-theoretic conjecture to a principle with independent content (pattern-theoretic parsimony) that can be tested across the entire Selberg class.

ACKNOWLEDGMENTS

The author thanks Claude (Anthropic) for collaborative development, numerical computation, cross-verification, and the identification of the Layer 3 gap structure.

TABLE I. Proof status for all results.

#	Result	Status	Section	Scope
1	Universal deficit angle $A = 1$	PROVEN	§III	Full Selberg class
2	Strict concavity at $\sigma = \frac{1}{2}$	PROVEN	§IV	Full Selberg class
3	$K(\text{on-line}) \leq K(\text{off-line}) + O(1)$	PROVEN	§V	Full Selberg class
4	GRH ⇒ single maximum	PROVEN	§VI	Full Selberg class
5	Single maximum ⇒ GRH	OPEN (Layer 3)	§VIII	Requires Euler product
6	Single-maximum criterion	CONJECTURED	§VI	Full Selberg class
7	GRH = parsimony test	interpretive	§IX	APO prediction

Appendix A: APO foundations

See Paper II, Appendix E for the full statement of the Axioms of Pattern Ontology. The key commitments relevant to this paper are: Differentiation$^{\text{APO}}$ ($\otimes$) produces the statistical manifold; Reflection$^{\text{APO}}$ ($\odot$) provides the Bhattacharyya inner product from which Chentsov forces the Fisher metric; Integration$^{\text{APO}}$ ($\oplus$) selects the most parsimonious representative of each equivalence class. Kolmogorov complexity is *supervenient* on Reflection$^{\text{APO}}$, not axiomatic: it arises as the unique universal measure of description length compatible with the inner product structure [9].

Appendix B: Proof status table

---

[1] D. Christian (2026), aPO Initiative working paper. The Arithmetic Recognition Operator: Information Geometry, the Modular Surface, and a Conjectured Path to the Riemann Hypothesis.
[2] D. Christian (2026), aPO Initiative working paper.
[3] A. Selberg, in *Proceedings of the Amalfi Conference on Analytic Number Theory* (Università di Salerno, 1992) pp. 367–385.
[4] J. B. Conrey and A. Ghosh, Invent. Math. **111**, 73 (1993).
[5] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, Colloquium Publications, Vol. 53 (AMS, 2004).
[6] M. Li and P. Vitányi, *An Introduction to Kolmogorov Complexity and Its Applications*, 3rd ed. (Springer, 2008).
[7] M. O. Rubinstein, Recent Perspectives in Random Matrix Theory and Number Theory **322**, 425 (2005), london Math. Soc. Lecture Note Ser.
[8] N. N. Chentsov, Translations of Mathematical Monographs **53** (1982), originally published in Russian, 1972. Proves uniqueness of the Fisher metric on statistical manifolds.
[9] R. J. Solomonoff, Inform. Control **7**, 1 (1964).