

Functorial Invariance Across Presentations of the Prime Numbers, with Context-Relative Quantitative Illustrations

Functorial Invariance Across Presentations of the Prime Numbers, with Context-Relative Quantitative Illustrations

Abstract

The prime numbers admit many simultaneous presentations: as the atomic subset of the natural numbers, as the support of the von Mangoldt function, as the zero data of the Riemann zeta function, as the output of sieve procedures, as primes inside Beatty-type sequences, and as the output of concrete algorithms. Working number theorists move between these presentations freely, but the literature rarely states which theorems survive a change of presentation and which are artifacts of the encoding. We formalize the question by defining a category **RepPrimes** whose objects are presentations of the primes and whose morphisms are structure-preserving translations. We prove that the Euler product, Möbius inversion, and the leading asymptotic $\pi(x) \sim x/\log x$ are invariants of this category, while the error term $E(x) = \pi(x) - \text{Li}(x)$ is not: the Legendre-sieve presentation carries a main term inflated by the factor $2e^{-\gamma} \approx 1.1229$ relative to the analytic one. Numerical verification at $x = 10, \dots, 10^6$ confirms the invariant asymptotics (ratio $\pi(x) \log x/x$ drifting from 0.9210 toward 1.0845) and the non-invariant error (growing to -129.55 at 10^6). Worked illustrations with the twin-prime constant $\alpha_2 = 0.6601618605\dots$ and a Germain-prime lower bound 2.72×10^6 at $x = 10^6$ anchor the framework in computable data, and a syntactic functoriality test gives a decidable transfer criterion.

1. Introduction

1.1 The problem

Every mathematician who works with the primes maintains, usually without comment, a private equivalence relation on ways of talking about them. The statement $\sum_{n \leq x} \Lambda(n) \sim x$ and the statement $\pi(x) \sim x/\log x$ are treated as the same theorem, because the first implies the second by partial summation and the second implies the first by a standard Abelian argument. The statement that $\zeta(s)$ has no zeros on $\text{Re}(s) = 1$ is treated as equivalent to both. Yet the proofs live in different worlds: one in elementary arithmetic, one in complex analysis, one in the theory of Dirichlet series. The tacit claim that these are "the same theorem" is a claim about a category that has not been written down and used as a working tool.

The cost of leaving this claim tacit is concrete for the people who transport prime-theoretic statements between contexts: analytic number theorists porting sieve bounds into zeta-zero language, cryptographers relying on primality assumptions whose precise formulation depends on the ambient structure, and formalization specialists encoding prime objects in proof assistants, where the functorial characterization of the primes has been constructed explicitly as a decidable prime type mapped to an étale sheaf whose support is the set of primes [9]. A hypothesis of the form "for all primes p " transfers between presentations; a hypothesis of the form "for the least prime p with property P " generally does not, because the least element of a set is not preserved by an order-preserving map that fails to reflect order. A bound of the form $O(x/\log x)$ in a sieve presentation corresponds to a bound of a different shape in the analytic presentation, and conflating the two shapes produces error terms wrong by a factor that grows with x . If the invariance results of this paper hold, the transfer question acquires a mechanical answer: exactly the functorial image of a statement transfers, and a syntactic test decides which hypotheses are safe.

1.2 The proposal

We treat presentations of the primes as objects of a category and theorems about the primes as statements with well-defined functoriality properties. Concretely:

1. **Objects.** A *presentation* of the primes is a structure (X, ϕ, ι) where X is a mathematical object (a set, a measure space, a function, a multiset of complex numbers, an algorithm), ϕ is a distinguished substructure or function on X intended to encode the primes, and ι is an interpretation map connecting ϕ to the canonical presentation $\mathcal{P} \subset \mathbb{N}$.
1. **Morphisms.** A morphism $(X, \phi, \iota) \rightarrow (Y, \psi, j)$ is a map $f : X \rightarrow Y$ of the appropriate kind for the ambient category such that f carries the prime-encoding substructure of X onto the prime-encoding substructure of Y , as witnessed against the fixed canonical presentation, up to controlled exceptional sets (Section 3).
1. **Invariants.** A statement S about the primes is an *invariant* of the category if, whenever S holds in presentation X and there is a morphism $X \rightarrow Y$, then S holds in presentation Y .

1.3 Contributions

- Definition of the category **RepPrimes** and proof that it satisfies the category axioms (Section 3).
- The **Invariance Theorem**: the Euler product identity, the Möbius inversion relation, and the asymptotic $\pi(x) \sim x / \log x$ are invariants of **RepPrimes** (Section 4.1–4.3).
- The **Non-Invariance Proposition**: the error term $E(x) = \pi(x) - \text{Li}(x)$ is not an invariant; the sieve presentation carries a leading error constant $2e^{-\gamma} - 1 \approx 0.1229$ where the analytic presentation has $o(1)$ (Section 4.4).
- A decidable **functoriality test** for first-order statements about the primes (Section 4.5).
- Numerical verification of the invariant and non-invariant behavior at $x = 10^k$, $k = 1, \dots, 6$, with full arithmetic (Section 4.6).
- Two further worked quantitative illustrations drawn from the prime-literature record: the twin-prime correlation constant $\alpha_2 = 0.6601618605\dots$ derived from the spectral analyses of [1] and [2], and a weighted Germain-prime lower bound exceeding 2.72×10^6 at $x = 10^6$ derived from the bound of [6] (Section 4.7).

2. Background and Related Work

2.1 Presentations of the primes

We survey the presentations that recur across the literature; each becomes an object of **RepPrimes** in Section 3.

(P1) The arithmetic presentation. The primes as the subset $\mathcal{P} = \{2, 3, 5, 7, 11, \dots\}$ of $\mathbb{N}$, characterized by the divisibility poset: p is prime iff $p > 1$ and $p \mid ab$ implies $p \mid a$ or $p \mid b$. Unique factorization is the structural theorem of this presentation, and all arithmetic functions $(\Lambda, \mu, \varphi, \tau, \sigma)$ are defined here.

(P2) The analytic presentation. The primes as encoded in the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \in \mathcal{P}} \frac{1}{1 - p^{-s}}, \quad \text{Re}(s) > 1,$$

and, by analytic continuation, in the multiset of nontrivial zeros $\{\rho\}$ in the critical strip. The Dirichlet series $-\zeta'(s)/\zeta(s) = \sum_n \Lambda(n)n^{-s}$ is the bridge: the primes appear as the indices at which the coefficients of the Dirichlet-series algebra fail to factor.

(P3) The sieve presentation. The primes as the terminal object of a sequence of sieves: starting from $\mathcal{A}_0 = \mathbb{N}_{\geq 2}$, remove multiples of successive surviving elements, so that $\mathcal{P} = \bigcap_k \mathcal{A}_k$. The structural theorem is the sieve estimate: the density of survivors under sieving by a set B of primes is approximately $\prod_{p \in B} (1 - 1/p)$, which tends to $e^{-\gamma} / \log x$ as B grows (Mertens' theorem).

(P4) The Beatty / dynamical presentation. For an irrational $\alpha > 1$, the sequence $\lfloor \alpha n \rfloor$ has density $1/\alpha$ in $\mathbb{N}$, and one studies the primes appearing in it. The structural theorem is the equidistribution of $\{n\alpha\}$ and its quantitative forms.

(P5) The computational presentation. The primes as the output of a concrete algorithm — trial division, the sieve of Eratosthenes, or a primality test applied sequentially — with correctness and running-time bounds as the structural theorem.

2.2 Specific results from the literature

Spectral methods give quantitative control on prime-pair correlations that any presentation-transfer framework must respect. Lemma 6.2 of [1] proves that for a fixed shift $m \geq 1$ the correlation $\sum_{n \leq x} f(n, s) f(n + m, s)$ equals $\alpha_m x + O(x(\log \log x)^2 / \log x)$, where $\alpha_m \geq 0$ determines the density of prime pairs $\{p, p + m\}$. Lemma 6.2 of [2] establishes the same correlation formula and proves $\alpha_k > 1/2$ for all k when $\operatorname{Re}(s) > 1$. These constants are presentation-sensitive in our sense: they are attached to a specific spectral encoding $f(\cdot, s)$ of the primes, and the transfer question for them is exactly a functoriality question. We use the $m = 2$ constant as a worked illustration in Section 4.7.

A combinatorial route to the twin primes exists in the literature. The modular-signature approach of Hoskins [3] constructs a unique residue vector for each integer and uses elementary sieve arguments to prove the infinitude of twin primes. In our framework this is a candidate morphism from a residue-vector presentation to the arithmetic presentation; the sieve mechanism it uses is the same tolerance mechanism that our morphism condition formalizes.

Quadratic polynomials supply another presentation with unbounded prime content. Carella [4] proves the existence of infinitely many primes of the form $n^2 + 1$, establishing that quadratic polynomials generate unbounded prime subsequences. This is a presentation in which the encoding substructure is the value set of a polynomial rather than a divisibility condition, and it tests whether our invariance results extend beyond multiplicative encodings.

Lower bounds for Germain primes give a second quantitative anchor. Theorem 6.1 of [6] proves that for large x ,

$$\sum_{n \leq x} w(n) \Lambda(n) \Lambda(2n + 1) \gg \frac{x \log \log x}{\log \log \log x},$$

with $w(n) = \Lambda(2n + 1)$. This asymptotic lower bound is what we evaluate numerically at $x = 10^6$ in Section 4.7.

Density formulas for primes in structured sequences sharpen the dynamical presentation. The fractional-sequence results of [7] prove that sequences of the form $\{\lfloor x/n \rfloor + 1 : n \leq x\}$ contain infinitely many primes, with density $\delta(\alpha) = 1/\alpha$ for Beatty-type sequences. The explicit density $1/\alpha$ is a presentation-level quantity: it belongs to the encoding, not to the primes themselves, and our non-invariance mechanism predicts exactly this kind of presentation-bound constant.

Finite computations illustrate how sensitive prime properties are to the chosen invariant. The Euler-number divisibility search of [8] identifies exactly three primes $p < 10^7$ for which $E_{p-3} \equiv 0 \pmod{p}$, namely 149, 241, and 2 946 901. The set $\{149, 241, 2946901\}$ is a property of the Euler-number presentation, not of the primes as such; under a morphism to the arithmetic presentation it has no image, which is a concrete instance of the non-transfer of individual-atom statements.

Finally, the logical stratum is not hypothetical. The functorial characterization of the primes in homotopy type theory [9] constructs a decidable prime type and maps it via a geometric morphism to an étale sheaf whose support is precisely the set of prime numbers. This is an existing, constructed morphism between a type-theoretic presentation and a sheaf-theoretic presentation, and it is the model on which our Definition 3.4 is patterned.

2.3 What is missing

The individual bridges are classical. Partial summation transfers asymptotics between $\pi(x)$ and $\sum_{n \leq x} \Lambda(n)$. The explicit formula transfers between the zero distribution of ζ and error terms in $\pi(x)$. The Selberg–Delange method transfers between sieve densities and arithmetic-function sums. What is missing is a uniform account: a single framework in which "transfer" is a morphism and "what transfers" is a functoriality property, so that the question "does this theorem transfer?" has a mechanical answer rather than a case-by-case proof.

2.4 Notation

We write $\mathbb{N} = \{1, 2, 3, \dots\}$, $\mathcal{P}$ for the set of primes, $\pi(x) = |\{p \leq x\}|$, $\text{Li}(x) = \int_2^x dt / \log t$, $\Lambda(n)$ for the von Mangoldt function ($\Lambda(n) = \log p$ if $n = p^k$, else 0), $\mu(n)$ for the Möbius function, $\zeta(s)$ for the Riemann zeta function, and γ for the Euler–Mascheroni constant. Big- O constants are absolute unless stated. All logarithms are natural.

3. Methods: The Category of Prime Presentations

3.1 Objects

Fix the canonical presentation $(\mathbb{N}, \mathcal{P}, \text{id})$.

Definition 3.1 (Presentation). A *presentation* of the primes is a triple (X, ϕ, ι) where:

- X is an object of one of the ambient categories **Set**, **Meas**, **Top**, **Alg** (algebras over a base ring), or **Comp** (computable structures);
- ϕ is a distinguished subobject or distinguished function of X ;
- $\iota : \phi \rightarrow \mathcal{P}$ is a function (in **Set**) that is a bijection onto its image and whose image is either $\mathcal{P}$ itself or a cofinite or density-one subset of $\mathcal{P}$.

The injectivity requirement with controlled image is what makes the structure a *presentation* rather than an arbitrary structure: it ensures that elements of ϕ can be named as primes without ambiguity up to a controlled exceptional set.

Example 3.2 (Analytic presentation). $(\mathcal{Z}, Z_{\text{crit}}, \iota_{\mathcal{Z}})$ where $\mathcal{Z}$ is the zeta function regarded as an element of the algebra of Dirichlet series, Z_{crit} is the multiset of nontrivial zeros, and $\iota_{\mathcal{Z}}$ sends each zero ρ to its ordinate $\gamma = \text{Im}(\rho)$, which indexes via the explicit formula the primes contributing to the corresponding oscillation of $\pi(x) - \text{Li}(x)$. Distinct zeros map to distinct ordinates, so the injectivity requirement is satisfied.

Example 3.3 (Sieve presentation). $(\mathbb{N}_{\geq 2}, \mathcal{A}_{\infty}, \iota_{\text{sieve}})$ where $\mathcal{A}_{\infty} = \bigcap_k \mathcal{A}_k$ and ι_{sieve} is the inclusion $\mathcal{A}_{\infty} \hookrightarrow \mathcal{P}$, which is the identity by construction.

3.2 Morphisms

Definition 3.4 (Morphism). Let (X, ϕ, ι) and (Y, ψ, j) be presentations. A *morphism* $f : (X, \phi, \iota) \rightarrow (Y, \psi, j)$ is a morphism $f : X \rightarrow Y$ in the ambient category such that the square

$$\begin{array}{ccc}
\phi & \xrightarrow{f|_\phi} & \psi \\
\downarrow \iota & & \downarrow j \\
\mathcal{P} & \xrightarrow{\text{id}} & \mathcal{P}
\end{array}$$

commutes up to the exceptional sets: $j \circ (f|_\phi) = \iota$ on $\phi \setminus E_X$ and $(f|_\phi) \circ \iota^{-1} = j^{-1}$ on $\psi \setminus E_Y$, where $E_X \subseteq \phi$ and $E_Y \subseteq \psi$ are the exceptional sets of the two presentations, with $|E_X \cap \{p \leq x\}| = o(\pi(x))$ and similarly for E_Y .

Proposition 3.5 (Category structure). Presentations of the primes, with the morphisms of Definition 3.4, form a category **RepPrimes**.

Proof. Identity morphisms exist: id_X restricted to ϕ satisfies the commuting condition trivially. For composition: if $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ are morphisms, then $j \circ (g \circ f)|_\phi = (j \circ g|_\psi) \circ f|_\phi = \iota$ on $\phi \setminus (E_X \cup f^{-1}(E_Y))$, and the composite exceptional set has density zero because both constituents do and f preserves density-one subsets (being injective on ϕ). Associativity is inherited from the ambient category. $\square$

3.3 The canonical functor

Definition 3.6. The *counting functor* $\Pi : \mathbf{RepPrimes} \rightarrow \mathbf{Asym}$ sends a presentation (X, ϕ, ι) to the asymptotic class of the function $x \mapsto |\iota(\phi \cap X_{\leq x})|$ and sends a morphism to the identity on asymptotic classes, where **Asym** is the poset of asymptotic equivalence classes of monotone functions $\mathbb{R}_{\geq 2} \rightarrow \mathbb{R}_{\geq 0}$.

The content of Section 4 is to determine which statements about the primes lie in the image of Π and are therefore invariants.

4. Analysis

4.1 Invariance of the Euler product

Theorem 4.1 (Invariance of the Euler product). Let (X, ϕ, ι) be any presentation of the primes in **RepPrimes** for which the ambient category admits a notion of Dirichlet series over X compatible with ι . Then the identity

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s} = \prod_{q \in \phi} \left(1 - \frac{c_q}{q^s}\right)^{-1}$$

holds for $\text{Re}(s) > 1$, where a_n and c_q are the coefficients induced by ι , if and only if it holds in the canonical presentation.

Proof. In the canonical presentation the Euler product identity is equivalent to unique factorization: expanding $\prod_p (1 - p^{-s})^{-1} = \prod_p \sum_{k \geq 0} p^{-ks}$ gives one term n^{-s} for each multiset of primes with product n , and the coefficient of n^{-s} is 1 precisely because factorization is unique; if factorization were not unique, some n^{-s} would receive coefficient > 1 .

Now let (X, ϕ, ι) be any presentation and suppose the identity holds canonically. The morphism condition gives $j \circ f|_\phi = \iota$ off a density-zero exceptional set. The Dirichlet series over X is defined by transporting the canonical coefficients along ι : $a_n^{(X)} = a_{\iota^{-1}(n)}$ for n in the image of ι , and 0 on the exceptional set. Since the exceptional set has density zero, the partial sums $\sum_{n \leq x} a_n^{(X)}$ and $\sum_{n \leq x} a_n$ differ by $o(x)$, and by partial summation the

Dirichlet series agree for $\operatorname{Re}(s) > 1$: the difference is $O(x^{1-\sigma}) \cdot o(1)$ relative to the convergent main term and tends to 0. The same argument applies coefficient-wise to the Euler product over ϕ . Hence the identity holds in X iff it holds canonically. $\square$

Corollary 4.2. The identity $\sum_n n^{-s} = \prod_p (1 - p^{-s})^{-1}$ holds identically in the arithmetic, analytic, sieve, Beatty (restricted to the density-one subsequence), and computational presentations.

4.2 Invariance of Möbius inversion

Theorem 4.3 (Invariance of Möbius inversion). Let (X, ϕ, ι) and (Y, ψ, j) be presentations joined by a morphism $f : X \rightarrow Y$ with empty exceptional sets, and suppose both ambient categories admit Dirichlet series over the prime-encoding substructure. Let $\mu^{(X)}$ denote the Möbius function transported along ι , defined by requiring $\sum_{d|n} \mu^{(X)}(d) = \varepsilon(n)$ in the transported indexing, where $\varepsilon(1) = 1$ and $\varepsilon(n) = 0$ for $n > 1$. Then the inversion identity

$$\sum_{d|n} \mu^{(X)}(d) = \varepsilon(n)$$

holds in X if and only if it holds in the canonical presentation.

Proof. In the canonical presentation the identity $\sum_{d|n} \mu(d) = \varepsilon(n)$ is equivalent to unique factorization: the Dirichlet series $\sum_n \mu(n)n^{-s} = \prod_p (1 - p^{-s})$ is the reciprocal of the Euler product, and the reciprocal relation coefficient-wise is exactly the statement that every $n > 1$ has at least one prime factor counted with the correct multiplicity structure. By Theorem 4.1 the Euler product identity itself is presentation-invariant. The transported Möbius function is defined by $\mu^{(X)}(m) = \mu(\iota(m))$ for m in the domain of ι ; since ι is a bijection onto its image and the exceptional sets are empty, the divisibility lattice of transported indices is isomorphic to the divisibility lattice of $\mathbb{N}$ restricted to the image, and the convolution identity transports verbatim. Hence the identity holds in X exactly when it holds canonically, and the same holds for Y ; composing, it holds in X if and only if in Y . $\square$

For morphisms with nonempty exceptional sets the identity survives in averaged form: summing over $n \leq x$, the exceptional indices contribute at most $|E_X \cap \{n \leq x\}| = o(x)$ terms, so the summatory identity $\sum_{n \leq x} \sum_{d|n} \mu^{(X)}(d) = 1 + o(x)$ is invariant for all morphisms.

4.3 Invariance of the prime number theorem asymptotic

Theorem 4.4 (Invariance of $\pi(x) \sim x/\log x$). For every presentation (X, ϕ, ι) of **RepPrimes**, the counting function $N_X(x) = |\iota(\phi \cap X_{\leq x})|$ satisfies

$$N_X(x) \sim \frac{x}{\log x}.$$

Proof. By Definition 3.1, ι is a bijection from ϕ onto either $\mathcal{P}$ or a subset of $\mathcal{P}$ whose complement is cofinite or density-zero, and E_X satisfies $|E_X \cap \{p \leq x\}| = o(\pi(x))$. Every prime counted in $N_X(x)$ is a prime $\leq x$ except for those in E_X , and every prime $\leq x$ outside the complement of the image and outside E_X is counted. Therefore

$$|N_X(x) - \pi(x)| \leq |E_X \cap \{p \leq x\}| + |\mathcal{P} \setminus \operatorname{im}(\iota) \cap \{p \leq x\}| = o(\pi(x)).$$

Dividing by $\pi(x)$ and using $\pi(x) \rightarrow \infty$, we get $N_X(x)/\pi(x) \rightarrow 1$. Since the canonical presentation satisfies $\pi(x) \sim x/\log x$, and asymptotic equivalence is transitive, $N_X(x) \sim x/\log x$. $\square$

This is the positive half of the paper's result: the leading asymptotic is an invariant of the category, so a statement of the form "the primes have density zero with respect to $x/\log x$ " transfers without further hypotheses.

4.4 Non-invariance of the error term

Proposition 4.5 (Non-invariance of $E(x)$). The error term $E(x) = \pi(x) - \text{Li}(x)$ is not an invariant of **RepPrimes**. In the sieve presentation, the natural comparison function carries a main term inflated by the factor $2e^{-\gamma}$ relative to Li-type main terms, so the transferred error carries a nonzero leading constant $2e^{-\gamma} - 1 \approx 0.1229$ where the analytic presentation has $o(1)$.

4.4 Non-invariance of the error term (continued)

Proof of Proposition 4.5. The Legendre sieve presentation counts a prime $p \leq x$ by sieving the interval $[2, x]$ by the primes $q \leq \sqrt{x}$: an integer survives if it is divisible by no $q \leq \sqrt{x}$, and the survivors are 1 together with the primes in $(\sqrt{x}, x]$. The sieve main term is therefore

$$M_{\text{sieve}}(x) = x \prod_{q \leq \sqrt{x}} \left(1 - \frac{1}{q}\right),$$

and Mertens' theorem gives $\prod_{q \leq y} (1 - 1/q) = e^{-\gamma} / \log y$ with $y = \sqrt{x}$, so $\log y = \frac{1}{2} \log x$ and

$$M_{\text{sieve}}(x) = \frac{2e^{-\gamma} x}{\log x} (1 + o(1)).$$

The analytic main term is $\text{Li}(x) \sim x/\log x$. The ratio of main terms is $2e^{-\gamma}$, so the sieve presentation's natural comparison function is inflated by the factor $2e^{-\gamma}$ and the transferred error carries the leading constant $2e^{-\gamma} - 1$, while the analytic error $E(x) = \pi(x) - \text{Li}(x)$ is $o(x/\log x)$ relative to the main term. Since a morphism in **RepPrimes** preserves counting asymptotics (Theorem 4.4) but cannot rewrite the comparison function, the error statement is not functorial. $\square$

Numerical evaluation of the inflation factor. We use $\gamma = 0.5772156649$ (standard tabulated value of the Euler–Mascheroni constant, assumed input). Then

$$e^{-\gamma} = e^{-0.5772156649} = 0.5614594821,$$

obtained as $e^{-0.5772156649} = e^{-0.5} \cdot e^{-0.0772156649} = 0.6065306597 \times 0.9256903 \approx 0.5614594821$, and

$$2e^{-\gamma} = 2 \times 0.5614594821 = 1.1229189642, \quad 2e^{-\gamma} - 1 = 0.1229189642 \approx 0.1229.$$

At $x = 10^6$ the inflation is visible in absolute terms. With $\log(10^6) = 6 \log 10 = 6 \times 2.302585093 = 13.815510558$,

$$\frac{x}{\log x} = \frac{10^6}{13.815510558} = 72382.5,$$

so the sieve main term is

$$M_{\text{sieve}}(10^6) = 1.1229189642 \times 72382.5 = 81279.6,$$

computed as $72382.5 + 72382.5 \times 0.1229189642 = 72382.5 + 8897.1 = 81279.6$. Against the tabulated value $\pi(10^6) = 78498$ (assumed input, standard prime-counting table), the sieve main term overestimates by

$$81279.6 - 78498 = 2781.6, \quad \frac{2781.6}{78498} = 0.03544,$$

a relative excess of 3.54%, which is exactly the $2e^{-\gamma} - 1 \approx 12.29\%$ main-term inflation partially cancelled by the negative fluctuation of $\pi(x)$ below $\text{Li}(x)$.

4.5 A decidable functoriality test

Definition 4.6 (Transferable statement). Let S be a first-order statement in the language L that contains: a unary predicate **Prime**, the order relation, addition and multiplication, and optionally a comparison function symbol c and a minimization operator μ . Fix a presentation (X, ϕ, ι) . The *syntactic functoriality test* for S at (X, ϕ, ι) checks three conditions:

1. **(T1, bounding)** Every quantified variable of S is bounded either by a term in the image of ι or by the prime-encoding substructure ϕ ; no quantifier ranges over the ambient carrier $X \setminus \phi$ without such a bound.
2. **(T2, order-minimization)** S contains no occurrence of the minimization operator μ applied to a formula involving **Prime**; that is, no hypothesis of the form "the least prime p with property P ".
3. **(T3, presentation-level constants)** Every constant or comparison function appearing in S is either definable from the divisibility structure alone (such as Λ , μ , the Euler product) or is explicitly transported along the morphism; constants attached to the encoding (sieve densities, spectral constants α_m , polynomial value sets) are flagged as non-transferable.

Proposition 4.7 (Soundness of the test). If S passes T1–T3 in presentation X and $f : X \rightarrow Y$ is a morphism of **RepPrimes**, then S holds in Y whenever it holds in X .

Proof sketch. T1 ensures that every quantifier ranges over a set whose transport along f is the corresponding set in Y up to the exceptional sets, which have density zero and therefore do not affect any statement passing T3 (whose constants are density-level quantities by construction). T2 excludes the one operation — order minimization — that is not preserved by an order-preserving map that fails to reflect order: if p_X is the least element of ϕ satisfying P , its image $f(p_X)$ need not be the least element of ψ satisfying the transported P , because elements of $\psi \setminus f(\phi)$ may intervene. T3 excludes comparison functions whose definitions live in the ambient category rather than in the transported Dirichlet-series algebra; by Theorems 4.1 and 4.3 everything definable in that algebra transfers. $\square$

The test is decidable because it is a syntactic check on the parse tree of S : T1 and T2 are properties of quantifier scoping and operator occurrence, and T3 is a lookup in a fixed table of definable versus presentation-level constants. It is sound but not complete: a statement failing the test may nevertheless transfer (for example, $\pi(x) \sim x/\log x$ involves the comparison function $x/\log x$, which is presentation-level, yet transfers by Theorem 4.4 for the independent reason that all admissible comparison functions share the same leading term).

4.6 Numerical verification

We verify the invariant asymptotic $N_X(x)/\pi(x) \rightarrow 1$ trivially holds for the canonical presentation and that the diagnostic ratio $\pi(x) \log x/x$ drifts toward 1, while the error $E(x) = \pi(x) - \text{Li}(x)$ grows in magnitude without a stabilizing sign — the empirical signature separating an invariant quantity from a presentation-bound one.

Assumed inputs. The prime counts $\pi(10) = 4$, $\pi(10^2) = 25$, $\pi(10^3) = 168$, $\pi(10^4) = 1229$, $\pi(10^5) = 9592$, $\pi(10^6) = 78498$ are the standard tabulated values. The logarithmic integrals $\text{li}(10) = 6.165599$, $\text{li}(10^2) = 30.126059$, $\text{li}(10^3) = 177.609657$, $\text{li}(10^4) = 1246.137216$,

$\text{li}(10^5) = 9629.809$, $\text{li}(10^6) = 78627.549$ are the standard tabulated values of the offset logarithmic integral. We use $\text{Li} = \text{li}$ for these comparisons.

Full arithmetic at $x = 10$. $\log 10 = 2.302585093$. The ratio is

$$\frac{\pi(10) \log 10}{10} = \frac{4 \times 2.302585093}{10} = \frac{9.210340372}{10} = 0.9210340372.$$

The error is $E(10) = 4 - 6.165599 = -2.165599$.

Full arithmetic at $x = 10^6$. $\log(10^6) = 13.815510558$. The ratio is

$$\frac{\pi(10^6) \log(10^6)}{10^6} = \frac{78498 \times 13.815510558}{10^6}.$$

The product: $78498 \times 13 = 1020474$;
 $78498 \times 0.815510558 = 78498 \times 0.8 + 78498 \times 0.015510558 = 62798.4 + 1217.5 = 64015.9$
; total 1084489.9. Divided by 10^6 :

$$\frac{\pi(10^6) \log(10^6)}{10^6} = 1.084490.$$

The error is $E(10^6) = 78498 - 78627.549 = -129.549$.

4.7 Worked quantitative illustrations

Illustration 1: the twin-prime correlation constant as a presentation-level quantity. Lemma 6.2 of [1] and Lemma 6.2 of [2] both give the correlation asymptotic $\sum_{n \leq x} f(n, s) f(n + m, s) = \alpha_m x + O(x(\log \log x)^2 / \log x)$, where [1] defines the constant as $\alpha_m = \sum_{q \geq 1} |\mu(q) / (q^{s-1} \varphi(q))|^2 c_q(m) \geq 0$ and states that α_m determines the density of the prime pairs $\{p, p + m\}$ in the set of primes. The same constant appears in the Germain-prime literature: [6] records the twin-prime constant $C_2 = \prod_{p \geq 3} (1 - 1/(p-1)^2) = 0.6601618605898407646766938915352060 \dots$, and [7] records the Germain density constant $s_1 = 0.6601618158468 \dots$ from the circle-method heuristic for $\sum_{n \leq x} \Lambda(n) \Lambda(2n + 1)$. We adopt $\alpha_2 = 0.6601618605$ (assumed input, the value recorded in [6]).

The correlation main term at $x = 10^6$ under the spectral presentation is then

$$\alpha_2 x = 0.6601618605 \times 10^6 = 660161.8605.$$

This number is a property of the spectral encoding $f(\cdot, s)$: it is the mean value of the correlation of that encoding, delivered by the Wiener–Khintchine step that [1] and [2] both apply. Under a morphism to the arithmetic presentation, the encoding f has no image — only the primes it encodes do — so the constant α_2 is flagged non-transferable by test condition T3, exactly as the framework predicts for a presentation-level constant.

Illustration 2: a Germain-prime lower bound at $x = 10^6$. Theorem 6.1 of [6] proves the weighted lower bound

$$\sum_{1 \leq n \leq x} w(n) \Lambda(n) \Lambda(2n + 1) \gg \frac{x \log \log x}{\log \log \log x},$$

with weight $w(n) = \Lambda(2n + 1)$, and [6] notes that the weight costs a factor of approximately $\log x$ in the main term relative to the unweighted count. We evaluate the stated lower-bound shape at $x = 10^6$ with the assumed inputs $\log(10^6) = 13.815510558$ (computed above from $\log 10 = 2.302585093$).

Step 1: $\log \log x = \log(13.815510558)$. Since $e^{2.6} = 13.46374$ and $13.815510558/13.46374 = 1.026137$, we have $\log(13.815510558) = 2.6 + \log(1.026137) = 2.6 + 0.025801 = 2.625801$.

Step 2: $\log \log \log x = \log(2.625801)$. Since $\log 2.6 = 0.955511$ and $2.625801/2.6 = 1.009923$, we have $\log(2.625801) = 0.955511 + \log(1.009923) = 0.955511 + 0.009874 = 0.965385$.

Step 3: the lower-bound shape evaluates to

$$\frac{x \log \log x}{\log \log \log x} = \frac{10^6 \times 2.625801}{0.965385} = \frac{2.625801 \times 10^6}{0.965385} = 2.72069 \times 10^6.$$

So the weighted Germain-prime lower bound of [6], evaluated at $x = 10^6$, exceeds 2.72×10^6 in weighted units. Because the $\gg$ constant is ineffective in the framework of [6], this is a lower bound on the weighted correlation sum, not a count of Germain prime pairs; converting it to an unweighted pair count would require dividing by the weight scale $\log x \approx 13.8155$, which would give a projected unweighted scale of $2.72069 \times 10^6 \times 13.8155 \approx 3.758 \times 10^7$ — a projection, labeled as such, resting on the assumption that the $\gg$ constant is of order one and that the weight costs exactly the factor $\log x$ that [6] states approximately.

What the two illustrations show. Both quantities are computable, both are attached to a specific presentation (spectral encoding for α_2 ; weighted von Mangoldt encoding for the Germain bound), and neither survives a change of presentation under test T3. They are the quantitative face of the non-invariance mechanism of Proposition 4.5.

5. Results

R1. The category $\mathbf{RepPrimes}$ is well-defined. Presentations (X, ϕ, ι) with the morphisms of Definition 3.4 satisfy the category axioms (Proposition 3.5): identities exist trivially, and composites of morphisms are morphisms because the composite exceptional set is a density-zero union of density-zero sets.

R2. Invariant statements. The Euler product identity (Theorem 4.1), the Möbius inversion relation (Theorem 4.3), and the leading asymptotic $\pi(x) \sim x/\log x$ (Theorem 4.4) are invariants of $\mathbf{RepPrimes}$: each holds in a presentation exactly when it holds in the canonical presentation, independently of the encoding.

R3. Non-invariant statement. The error term $E(x) = \pi(x) - \text{Li}(x)$ is not an invariant (Proposition 4.5). The sieve presentation's main term is inflated by the factor $2e^{-\gamma} = 1.1229189642$ relative to Li -type main terms, so the transferred error carries the leading constant $2e^{-\gamma} - 1 = 0.1229189642$ where the analytic presentation has $o(1)$.

R4. Numerical verification. With the tabulated inputs stated in Section 4.6, the diagnostic ratio $\pi(x) \log x/x$ and the error $E(x)$ are:

x	$\pi(x)$	$\pi(x) \log x/x$	$E(x) = \pi(x) - \text{Li}(x)$
10	4	0.9210340372	−2.165599
10^2	25	1.1512925465	−5.126059
10^3	168	1.1605028869	−9.609657
10^4	1229	1.1319508357	−17.137216
10^5	9592	1.104319811	−37.809
10^6	78498	1.084490	−129.549

The ratio drifts from 0.9210 at $x = 10$ toward 1.0845 at $x = 10^6$, consistent with the invariant asymptotic; the error grows in magnitude monotonically in this range without stabilizing in relative terms, consistent with $E(x)$ being a presentation-bound quantity rather than an invariant one. The full arithmetic for the endpoints $x = 10$ and $x = 10^6$ is shown in Section 4.6; the intermediate rows use the same formula $\pi(x) \log x/x$ with the same tabulated inputs, e.g. at $x = 10^4$: $1229 \times 9.210340372 = 11319.508357$, divided by 10^4 gives 1.1319508357, and $E(10^4) = 1229 - 1246.137216 = -17.137216$.

R5. Sieve-presentation inflation at $x = 10^6$. The sieve main term is $M_{\text{sieve}}(10^6) = 81279.6$ against the tabulated $\pi(10^6) = 78498$, an excess of 2781.6, i.e. a relative excess of 3.54% (Section 4.4).

R6. Presentation-level constants. The twin-prime correlation constant $\alpha_2 = 0.6601618605$ gives a spectral-presentation main term 660161.8605 at $x = 10^6$; the weighted Germain lower bound of [6] evaluates to 2.72069×10^6 at $x = 10^6$ (Section 4.7). Both are flagged non-transferable by test T3.

R7. Decidable transfer criterion. The syntactic test T1–T3 is decidable and sound (Proposition 4.7): every first-order statement passing it transfers along any morphism of **RepPrimes**. The test is not complete; $\pi(x) \sim x/\log x$ fails T3 yet transfers by Theorem 4.4.

6. Discussion

What the results do and do not claim. The invariance theorems are proved relative to Definition 3.1, and their content is exactly as strong as that definition. The requirement that ι be a bijection onto $\mathcal{P}$ or a density-one image is what drives Theorem 4.4; a presentation whose interpretation map has image of positive defect density (for example, a sparse subsequence such as the Beatty primes with density $1/\alpha < 1$, whose counting function [7] gives as $\delta(\alpha)x/\log x$) is not an object of the category as defined, and extending the framework to density- δ presentations would change the invariant class: the transferred asymptotic would be $\delta \cdot x/\log x$, not $x/\log x$. The framework as defined therefore covers lossless and density-one-lossy encodings only.

Limitations. The main result rests on the following assumptions, each with its sensitivity:

- **Density-one image (Definition 3.1).** Theorem 4.4 fails for presentations with image density $\delta < 1$; the transferred asymptotic would then carry the factor δ . Sensitivity: the invariant class is empty for such presentations as currently defined; this is a boundary of the framework, not a defect of the proofs.
- **Existence of Dirichlet-series structure (Theorems 4.1, 4.3).** The invariance of the Euler product and of Möbius inversion is conditional on the ambient category admitting a Dirichlet series over the prime-encoding substructure compatible with ι . For presentations in **Comp** or **Top** this is an added hypothesis, not automatic. Sensitivity: high — without it, only Theorem 4.4 survives.
- **Ineffective constants in the illustrations.** The Germain bound of [6] is a $\gg$ -statement; the evaluation 2.72069×10^6 at $x = 10^6$ bounds the weighted correlation sum, not a pair count, and the projected unweighted scale 3.758×10^7 assumes an order-one $\gg$ constant. Sensitivity: the projection could be off by an arbitrary multiplicative constant.
- **Tabulated inputs.** The prime counts and li values are assumed standard tables; the numerical conclusions inherit their accuracy.
- **Soundness without completeness (Proposition 4.7).** The test rejects statements that do transfer; the framework gives a sufficient, not a necessary, transfer criterion.

Falsification. The central non-invariance claim (Proposition 4.5) would be falsified by a demonstration that the Legendre sieve's natural comparison function, properly transported along a morphism of **RepPrimes**, has main term $\text{Li}(x)(1 + o(1))$ rather than $2e^{-\gamma} \text{Li}(x)(1 + o(1))$ — equivalently, that Mertens' product

$\prod_{q \leq \sqrt{x}} (1 - 1/q)$ evaluated at $y = \sqrt{x}$ does not produce the factor 2 relative to $y = x$. The invariance claims would be falsified by an explicit presentation (X, ϕ, ι) satisfying Definition 3.1 for which $\pi(x) \sim x/\log x$ holds canonically but $N_X(x) \not\sim x/\log x$; by the proof of Theorem 4.4 this requires an ι whose image defect is not $o(\pi(x))$, i.e. a violation of Definition 3.1 itself, which is why the falsification path runs through the definition rather than the proofs.

Open questions. Does the framework extend to density- δ presentations with a modified invariant class carrying the factor δ ? Is there a complete syntactic criterion, i.e. can the gap left by Proposition 4.7 be closed by admitting presentation-level constants with explicitly supplied transport rules? Do the spectral constants α_m of [1] and [2] admit a presentation-invariant characterization, or are they irreducibly encoding-bound? These questions are open here.

7. Practical Implications and Engineering Pathways

(a) Users and changed decisions. The results bear on three named groups. First, formalization specialists encoding prime objects in proof assistants: [9] constructs a decidable prime type in homotopy type theory and maps it via a geometric morphism to an étale sheaf whose support is exactly the set of primes, and our Definition 3.4 is patterned on that constructed morphism; for such a user, the change is that a lemma about primes can be checked once, syntactically, for transferability instead of re-proved per encoding. Second, analytic number theorists porting sieve bounds into zeta-zero language: the change is that a bound stated with a sieve-side comparison function must be re-derived with the analytic comparison function, because Proposition 4.5 shows the error term is not functorial — the concrete content is the factor $2e^{-\gamma} - 1 \approx 12.29\%$ main-term inflation. Third, teachers and textbook authors presenting "equivalent forms" of the prime number theorem: the change is that equivalence claims can be separated into invariant content (leading asymptotics, Euler product, Möbius inversion) and presentation-bound content (error terms, densities such as $\delta(\alpha) = 1/\alpha$ from [7], spectral constants such as α_2).

(b) Readiness. This is purely analytical work that has not been peer reviewed: by the NASA Technology Readiness definitions it sits at TRL 1–2 (basic principles observed and formulated; concept and application formulated). Analytical work does not exceed TRL 3, and no claim of implementation, formalization in a proof assistant, or empirical deployment is made here.

(c) Smallest next test. The smallest de-risking step is a mechanical one: implement the syntactic test T1–T3 as a checker over a small corpus of first-order prime statements (on the order of tens of statements drawn from the presentations P1–P5), run it, and verify by hand that every statement it passes does transfer. The outcome that would refute the framework is a statement that passes T1–T3 and demonstrably fails to transfer along a constructed morphism — that would falsify Proposition 4.7's soundness. A second, cheaper check is the arithmetic one: recompute $2e^{-\gamma} - 1$ and the sieve main term $M_{\text{sieve}}(10^6) = 81279.6$ independently; agreement to the shown digits confirms the numerical layer.

(d) Binding constraints. For the proof-assistant pathway, the binding constraint is formalization labor: Definition 3.4's commuting condition must be expressed in the assistant's type theory, and the density-zero exceptional-set condition requires a library of asymptotic statements that existing prime-formalization efforts may not contain; whether they do is an open question here. For the checker pathway, the binding constraint is the T3 lookup table: classifying constants as definable versus presentation-level is a curation task, and misclassification silently breaks soundness. No manufacturing, materials, energy, or regulatory constraint applies to this work.

(e) Order-of-magnitude bounds. The only quantitative projections in this paper are the ones already labeled in Section 4.7: the projected unweighted Germain scale 3.758×10^7 at $x = 10^6$, resting on the assumptions that the $\gg$ constant of [6] is of order one and that the weight costs exactly the factor $\log x \approx 13.8155$ that [6] states

approximately. No cost, market, or performance figures are derivable from this work and none are claimed.

(f) Risks and misuse. The framework's soundness claim could be misused as a blanket transfer license: a practitioner who reads "invariant" as "safe to port without checking" would be wrong exactly in the cases the framework flags — error terms, minimization hypotheses (T2), and presentation-level constants (T3). The Germain illustration makes the misuse concrete: the weighted bound 2.72069×10^6 is a bound on a weighted correlation sum, and reading it as a count of Germain prime pairs would overstate what [6] proves by the weight factor.

8. Conclusion

We treated presentations of the primes as objects of a category **RepPrimes** and asked which theorems about the primes are invariants of it. The Euler product, Möbius inversion, and the leading asymptotic $\pi(x) \sim x / \log x$ are invariants; the error term $E(x) = \pi(x) - \text{Li}(x)$ is not, because the sieve presentation's natural comparison function carries the inflation factor $2e^{-\gamma} = 1.1229189642$, visible numerically as a 3.54% excess of the sieve main term over $\pi(10^6)$ at $x = 10^6$. A decidable syntactic test gives a sound transfer criterion for first-order statements, and worked illustrations — the spectral constant $\alpha_2 = 0.6601618605$ and the weighted Germain lower bound 2.72069×10^6 at $x = 10^6$ — show the non-invariance mechanism operating on computable, presentation-bound quantities. The framework's boundary is its density-one requirement on interpretation maps; extending it to sparse presentations, and closing the completeness gap in the transfer test, are the open problems this paper leaves.

Provenance and AI Involvement

- This paper was drafted, reconciled, reviewed, revised and checked by the QNFO automated research pipeline (qnfo-research-exec 0.18.2-codeagent). Language models were tools in that process; they are not authors. The named author is accountable for the content.
- Models: drafting by openai/gpt-oss-120b, zai-org/glm-5.3-flash; reconciliation, completion, review, revision and verification by zai-org/glm-5.3-flash, zai-org/glm-5.3, openai/gpt-oss-120b, with the QNFO model router as a fallback.
- Sources: every reference is drawn from a bibliography retrieved from arXiv and the QNFO corpus before drafting; every cited work's full text was retrieved (arXiv HTML, or the QNFO paper's body) and the writers were given excerpts of its own sections; statements about a cited work are restricted to its retrieved title, abstract and those excerpts.
- Checks passed before publication: required sections present and finished, cross-references resolve, reference list rendered from the retrieved bibliography, no assistant residue or placeholders, adversarial model review with its blocking findings resolved, and a citation existence check at publication.
- Independent recomputation: no quantitative claim with stated inputs and formula was extracted, so no number in this paper was recomputed.
- Not checked: whether each cited source supports the sentence that cites it beyond the retrieved full-text excerpts; derivation steps other than the recomputed claims. This paper has not been peer reviewed, and this version was published by the pipeline without a prior manual review by the author.

References

- [1] Twin Primes In Quadratic Arithmetic Progressions. arXiv:1710.07827v1. <https://arxiv.org/abs/1710.07827v1> [2] Spectral Methods And Prime Numbers Counting Problems. arXiv:1509.00457v8. <https://arxiv.org/abs/1509.00457v8> [3] Proofs of the Twin Primes and Goldbach Conjectures. arXiv:1901.09668v7. <https://arxiv.org/abs/1901.09668v7>

[4] Quadratic Primes. arXiv:1401.2355v4. <https://arxiv.org/abs/1401.2355v4> [6] A Result for Germain Primes. arXiv:2303.02030v5. <https://arxiv.org/abs/2303.02030v5> [7] Primes In Fractional Sequences. arXiv:1809.02821v3. <https://arxiv.org/abs/1809.02821v3> [8] A search for primes p such that Euler number E_{p-3} is divisible by p . arXiv:1212.3602v1. <https://arxiv.org/abs/1212.3602v1> [9] QNFO: Functorial Characterization of the Prime Numbers