# Addition and Multiplication: Spectral Orthogonality and Innovation in the Arithmetic of Integers

Juan Moreno Borrallo[1, *]

[1] *Independent Researcher*

(Dated: January 22, 2026)

**Abstract.** *The arithmetic of the integers is governed by two fundamental operations, addition and multiplication, whose interaction lies at the core of many deep problems in number theory. While multiplication preserves prime factorization in a rigid and conservative manner, addition typically destroys multiplicative structure and generates new prime content.*

*In this work, we develop a unified structural framework that explains this asymmetry through spectral and operator-theoretic principles. By embedding the integers into a Hilbert space, we show that multiplication acts as a diagonal, layer-preserving operator in the prime spectral basis, whereas addition acts as a non-local, mixing operator driven by carry propagation. This spectral incompatibility leads to an arithmetic uncertainty principle, forbidding simultaneous localization in additive and multiplicative bases.*

*Building on this structure, we introduce additive innovation as a quantitative measure of the new prime information created by a sum. We prove that the only obstruction to innovation arises from smoothness and $S$-unit phenomena in the coprime core. Using classical results on smooth numbers, we show that additive innovation is typically large, yielding unconditional abc-type inequalities in density.*

*Finally, we develop an information-theoretic perspective, showing that addition produces entropy across prime scales while multiplication remains information-preserving. These results provide a structural explanation for the sum-product phenomenon and reframe classical problems as manifestations of the intrinsic incompatibility between additive and multiplicative spectral structures.*

*"Entia non sunt multiplicanda praeter necessitatem"*
— Ockham's Razor

*"Padre, Señor del cielo y de la tierra, te doy gracias porque has ocultado todo esto a los sabios y entendidos y se lo has revelado a los que son como niños."*
— Matthew 11:25

* juan.morenoborrallo@gmail.com

## CONTENTS



# I. INTRODUCTION

The integers $\mathbb{Z}$ are equipped with two fundamental operations, addition and multiplication, whose coexistence defines arithmetic. Yet, despite their elementary appearance, the interaction between these two operations remains one of the deepest sources of complexity in number theory. Many central problems—from the abc conjecture [1, 2] to the sum–product phenomenon [3, 4], and from Diophantine equations to questions of pseudorandomness—can be traced to a fundamental tension between additive and multiplicative structure.

From a multiplicative viewpoint, the arithmetic of $\mathbb{Z}$ is remarkably rigid. By the fundamental theorem of arithmetic [5], every positive integer admits a unique prime factorization, and multiplication acts by combining prime exponents in a completely deterministic manner. In this sense, multiplication preserves prime information: no genuinely new prime structure is created by multiplying integers. This rigidity suggests that multiplication acts internally within a fixed prime spectral architecture.

Addition behaves in a strikingly different way. Even when the prime factorizations of two integers $a$ and $b$ are fully known, the prime factorization of their sum $a + b$ is typically unpredictable. In the coprime case, addition produces a number whose prime divisors are entirely disjoint from those of the summands. More generally, except in highly constrained situations, addition introduces new prime factors that are not inherited from either input. While this phenomenon is familiar from elementary examples, it has far-reaching consequences and has lacked a systematic structural formulation.

The purpose of this paper is to provide such a formulation. We propose a conceptual framework in which multiplication is viewed as a *conservative* operation on prime spectra, while addition is viewed as a generically *orthogonal* and information-generating operation. To formalize this asymmetry, we introduce the notion of *additive innovation*, which measures the genuinely new prime content created by a sum beyond that present in the summands.

A central guiding principle of our approach is that this asymmetry is structural rather than accidental. To capture this, we develop an operator-theoretic framework embedding the integers into a Hilbert space. In this setting, multiplication acts as a spectrally local and layer-preserving operator, whereas addition is inherently non-local and mixing. This duality leads naturally to an arithmetic uncertainty principle, preventing simultaneous localization in additive and multiplicative spectral bases.

Our first goal is structural. We identify the precise obstruction to additive innovation and show that all failures of innovation arise from a single mechanism: a smoothness condition on the coprime core of the sum. This places exceptional configurations within the classical theory of $S$-unit equations [6]. From this viewpoint, the abc philosophy acquires a structural interpretation: the rarity of additive relations with abnormally small radical growth is precisely the rarity of innovation collapse.

Our second goal is quantitative. Using classical results on smooth numbers and their distribution [7, 8], together with probabilistic tools in the spirit of Erdős–Kac [9], we show that additive innovation is large for almost all pairs of integers. This yields unconditional abc-type inequalities in density. We further develop an information-theoretic perspective, introducing entropy production as a probabilistic analogue of innovation [10]. Under mild equidistribution assumptions, verified in natural arithmetic models, we show that addition produces entropy across prime scales, while multiplication does not.

As an application, we obtain a spectrum-growth principle providing a structural explanation for the sum–product phenomenon: sets with nontrivial multiplicative diversity cannot remain additively rigid without forcing the appearance of new prime scales. This clarifies additive–multiplicative expansion as a consequence of intrinsic prime orthogonality rather than combinatorial complexity alone.

Finally, we discuss how this framework interfaces with analytic number theory and Diophantine geometry. While we do not address conjectures such as the Riemann Hypothesis or the Birch–Swinnerton-Dyer conjecture directly, we explain how improvements in prime distribution translate into stronger innovation bounds, and how abc-type control of innovation connects naturally with height growth phenomena. The framework developed here isolates with precision the unique structural mechanism by which abc-type inequalities can fail, and shows that this mechanism is rare in a strong quantitative sense.

## II. PRIME SPECTRA AND MULTIPLICATIVE CONSERVATION

Let $\mathbb{P}$ denote the set of all prime numbers. For a positive integer $n \geq 1$, we define its *prime spectrum* as

$$S(n) := \{p \in \mathbb{P} : p \mid n\}.$$

This set captures the qualitative multiplicative structure of $n$, ignoring multiplicities. The fundamental theorem of arithmetic implies that multiplication acts conservatively on prime spectra.

**Proposition II.1** (Multiplicative Conservation). *For any* $a, b \in \mathbb{Z}_{>0}$,

$$S(ab) = S(a) \cup S(b).$$

Multiplication does not introduce genuinely new prime information. In this sense, it is a deterministic and reversible operation at the level of prime support. One may interpret integers as vectors in an infinite coordinate space indexed by primes, with multiplication corresponding to vector addition in the exponent coordinates.

From the perspective of prime spectra, multiplication is a conservative operation: it rearranges and aggregates existing prime directions, but cannot generate new ones. This rigidity reflects a deep structural property of multiplicative arithmetic, which stands in sharp contrast with the behavior of addition. In the next section, we show that addition cannot be described within the same spectral framework, and necessarily acts as a mixing operator across prime directions.

## III. OPERATORIAL FRAMEWORK AND SPECTRAL DYNAMICS

The results established so far describe a sharp asymmetry between addition and multiplication at the level of prime spectra. In this section, we provide a structural explanation for this phenomenon by embedding the integers into an operator-theoretic framework. This perspective clarifies why multiplication acts conservatively on prime structure, while addition is intrinsically mixing and generative.

### A. The arithmetic Hilbert space

Let

$$\mathcal{H} := \mathrm{span}\{\, |n\rangle : n \in \mathbb{Z}_{\geq 1} \,\}$$

be the Hilbert space with orthonormal basis indexed by the positive integers. The basis vector $|n\rangle$ represents the arithmetic state associated with $n$.

By the fundamental theorem of arithmetic, every integer $n$ admits a unique prime-power decomposition

$$n = \prod_{p \in \mathbb{P}} p^{\nu_p(n)},$$

where $\nu_p(n) \in \mathbb{Z}_{\geq 0}$ is the greatest exponent of some prime $p$, and all but finitely many $\nu_p(n)$ vanish. This allows us to identify each basis vector with its exponent profile,

$$|n\rangle \equiv |\nu(n)\rangle := |(\nu_p(n))_{p \in \mathbb{P}}\rangle,$$

so that $\mathcal{H}$ may be viewed as the $\ell^2$-space of finite-support exponent vectors indexed by primes.

### B. Prime number operators and multiplication

For each prime $p$, define the *prime number operator*

$$N_p\, |\nu\rangle := \nu_p\, |\nu\rangle.$$

The family $\{N_p\}_{p \in \mathbb{P}}$ consists of commuting, diagonal operators whose joint spectrum records the prime exponent structure of an integer.

For $m \in \mathbb{Z}_{>0}$, define the *multiplicative shift operator*

$$T_m |n\rangle := |mn\rangle.$$

In exponent coordinates, $T_m$ acts by translation:

$$\nu_p(mn) = \nu_p(n) + \nu_p(m).$$

In particular, for a prime $p$,

$$[N_q, T_p] = \delta_{pq}\, T_p,$$

so multiplication shifts only the $p$-th coordinate and preserves all others. This expresses the spectral locality of multiplication: $T_m$ acts internally within the prime spectral architecture, modifying amplitudes but never creating new prime directions.

This operator-theoretic formulation recovers multiplicative conservation as a structural statement: multiplication preserves the support of the prime spectrum and acts diagonally with respect to the family $\{N_p\}$.

### C. Addition as a non-local operator

Define the *additive shift*

$$S_1 |n\rangle := |n+1\rangle.$$

In contrast to $T_m$, the operator $S_1$ is highly non-local in the prime spectral basis. Its action on exponent coordinates is governed by the valuation jumps

$$\Delta_p(n) := \nu_p(n+1) - \nu_p(n).$$

These jumps are unbounded in both directions:

$$\sup_n \Delta_p(n) = +\infty, \qquad \inf_n \Delta_p(n) = -\infty,$$

as witnessed by the transitions $p^k - 1 \mapsto p^k$ and $p^k \mapsto p^k + 1$. Consequently, $S_1$ does not preserve any finite stratification defined by the operators $\{N_p\}$.

Equivalently, for every finite set of primes $P$ and every bound $K$, the action of $S_1$ mixes states whose truncated prime profiles

$$(\min(\nu_p(n), K))_{p\in P}$$

differ arbitrarily. In this precise sense, addition is an intrinsically non-local and mixing operation in prime spectral coordinates.

## D. Innovation as Spectral Flow

The operator-theoretic framework allows us to interpret additive innovation as a spectral flow.

**Proposition III.1** (Innovation as Commutator Expectation)**.** *Let $\rho$ be a density state representing a uniform distribution over a large interval $[1, N]$. The expected additive innovation at prime $p$ is proportional to the expectation of the commutator between the additive shift and the p-adic number operator:*

$$\mathbb{E}[\nu_p(n+1)] \approx \mathrm{Tr}(\rho[S_1^\dagger N_p S_1])$$

*More precisely, although $S_1$ is not unitary on $\ell^2(\mathbb{Z}_{\geq 1})$ but rather a unilateral isometry, it satisfies*

$$S_1^\dagger S_1 = I.$$

*Consequently,*

$$\Delta_p(n) = \langle n \mid S_1^\dagger N_p S_1 - N_p \mid n\rangle.$$

*The* additive innovation *is defined as the positive part of the spectral flow generated by this non-vanishing commutator.*

## E. Structural interpretation

The operatorial distinction between $T_m$ and $S_1$ provides a conceptual explanation for the asymmetry between multiplication and addition observed at the level of prime spectra.

Multiplication is spectrally conservative because it acts by translations in a fixed prime coordinate system, preserving prime support and internal structure. Addition, by contrast, does not act internally within this coordinate system: it induces large and unavoidable spectral jumps across prime directions. The appearance of new prime factors in sums is therefore not accidental, but a direct manifestation of the non-local and mixing action of $S_1$ with respect to the prime spectral basis.

This operator-theoretic framework isolates the structural origin of additive generativity and multiplicative rigidity. In the following section, we show that this asymmetry reflects a deeper incompatibility between the natural spectral representations of addition and multiplication, leading to a fundamental duality between their respective notions of locality.

# IV. ADDITIVE–MULTIPLICATIVE DUALITY AND SPECTRAL INCOMPATIBILITY

The non-locality of addition established in the prime spectral framework raises a natural question: *is this behavior intrinsic to addition, or merely an artifact of the chosen representation?*

In this section, we show that the observed asymmetry reflects a deeper structural duality. There exist spectral bases in which addition becomes local and transparent, but in such bases multiplication necessarily becomes mixing. *No representation can simultaneously render both operations spectrally local.*

## A. Additive spectral decompositions

Every integer $n \geq 1$ admits a unique binary expansion

$$n = \sum_{k\geq 0} b_k(n)\, 2^k, \qquad b_k(n) \in \{0, 1\}.$$

This expansion induces an alternative labeling of basis states,

$$|n\rangle \equiv |b_0(n), b_1(n), b_2(n), \dots\rangle,$$

which we refer to as an *additive spectral basis*. More generally, one may work in base $q \geq 2$, but the binary case already captures the essential structure.

Define the additive number operators

$$\widetilde{N}_k\,|b\rangle := b_k\,|b\rangle, \qquad k \geq 0,$$

which are commuting diagonal operators whose joint spectrum records the additive digit structure of $n$. In this representation, the integer itself may be recovered as the eigenvalue of the additive Hamiltonian

$$\widetilde{H} := \sum_{k\geq 0} 2^k\,\widetilde{N}_k, \qquad \widetilde{H}|n\rangle = n\,|n\rangle.$$

This additive spectral representation provides a natural framework in which to analyze the locality properties of addition itself, as we now show.

## B. Locality of addition in additive bases

In the additive spectral basis, the action of the additive shift $S_1|n\rangle = |n+1\rangle$ is simple. Adding 1 flips the least significant zero digit to 1 and resets all trailing ones to zero. Define the binary carry length

$$\ell_2(n) := \max\{k \geq 0 :\ b_0(n) = \cdots = b_{k-1}(n) = 1\}.$$

Then the action of $S_1$ modifies only the first $\ell_2(n)+1$ digits of the binary expansion.

This behavior is local in the additive spectral coordinates, up to the carry chain. Moreover, the carry length has a geometric tail:

$$\mathbb{P}(\ell_2(n) \geq k) = 2^{-k},$$

so long carry chains are exponentially rare. Thus, in an additive spectral basis, addition is nearly local, with controlled non-locality arising solely from carries.

## C. Multiplication as a mixing operator in additive bases

In sharp contrast, multiplication is non-local in additive spectral coordinates. Indeed, multiplication by an odd prime $p$ propagates carries across binary digits, so that the output digits of $pn$ depend on input digits at many scales. Consequently, while addition is spectrally simple in additive bases, multiplication is not: for every depth $L$ there exist integers $n$ for which the map

$$n \longmapsto pn$$

changes digits well beyond the first $L$, and hence $T_p$ does not preserve any finite stratification defined by the digit observables $\{\widetilde{N}_k\}$. Equivalently, $T_p$ acts as a genuinely mixing operator on the additive spectrum.

## D. Structural duality

The two spectral representations exhibit a precise inversion of roles:

- In the *prime spectral basis*, multiplication is local and layer-preserving, while addition is non-local and mixing.
- In an *additive spectral basis*, addition is local (up to carries), while multiplication is non-local and mixing.

This duality is structural rather than accidental. It reflects the fact that addition and multiplication generate incompatible decompositions of the integers. Each operation admits a natural spectral basis in which it is simple, but these bases are mutually incompatible.

In particular, there exists no basis in which both the additive shift $S_1$ and the multiplicative shifts $\{T_p\}_p$ are simultaneously local. Any attempt to simplify one operation necessarily complicates the other.

# V. AN ARITHMETIC UNCERTAINTY PRINCIPLE

The incompatibility between additive and multiplicative spectral representations places a fundamental limitation on the simultaneous localization of arithmetic structure. Below, we propose a formalization of this limitation.

Let $A \subset \{1, 2, \dots, N\}$ be a finite set. Define:

- *Additive localization* at depth $L$ by
$$B_L(A) := \Big|\{(b_0(n), \dots, b_{L-1}(n)) : n \in A\}\Big|.$$
- *Multiplicative localization* at resolution $(P, K)$ by
$$V_{P,K}(A) := \Big|\{(\min(\nu_p(n), K))_{p\in P} : n \in A\}\Big|.$$

**Theorem V.1** (Arithmetic uncertainty principle (trivial regime))**.** [1] *There is no family of finite sets $A \subset \{1, \dots, N\}$, with $|A| \to \infty$, such that both*

$$B_L(A) = O(1) \quad \textit{and} \quad V_{P,K}(A) = O(1)$$

*hold uniformly for arbitrarily large choices of $L$, $P$, and $K$.*

*Proof.* Suppose, for the sake of contradiction, that there exists a family of finite sets

$$A_N \subset \{1, 2, \dots, N\}, \qquad |A_N| \to \infty \quad \text{as } N \to \infty,$$

such that the bounds

$$B_L(A_N) = O(1) \qquad \text{and} \qquad V_{P,K}(A_N) = O(1)$$

hold simultaneously and uniformly for arbitrarily large choices of $L$, $P$, and $K$.

Fix $N$ and set $A := A_N$. Choose an integer $L$ such that

$$2^L > N.$$

(Note that such values of $L$ can be taken arbitrarily large.)

Consider the map

$$\Phi_L : \{1, \dots, N\} \longrightarrow \{0,1\}^L, \qquad \Phi_L(n) := \big(b_0(n), \dots, b_{L-1}(n)\big),$$

where $(b_k(n))_{k\geq 0}$ denotes the binary digits of $n$. If $2^L > N$, then for any $m, n \in \{1, \dots, N\}$ one has

$$\Phi_L(m) = \Phi_L(n) \quad \Longrightarrow \quad m = n.$$

Indeed, the binary expansion of any integer $n < N < 2^L$ is completely determined by its first $L$ digits, so $\Phi_L$ is injective on $\{1, \dots, N\}$.

In particular, the restriction $\Phi_L|_A$ is also injective, and therefore

$$B_L(A) = \big|\Phi_L(A)\big| = |A|.$$

However, by hypothesis $B_L(A) = O(1)$ for arbitrarily large $L$, which forces $|A| = O(1)$. This contradicts the assumption that $|A_N| \to \infty$.

Hence, no such family $(A_N)$ can exist. □

Subsequently, no infinite arithmetic configuration can be simultaneously localized in additive and multiplicative spectral coordinates.

## A. An arithmetic uncertainty principle (quantitative form)

The previous result shows that a non-trivial uncertainty principle must necessarily operate below the full additive resolution. We now state and prove a quantitative form capturing this phenomenon.

For an integer $y \geq 3$, let

$$\mathcal{P}(y) := \{p \text{ prime} : \ 3 \leq p \leq y\}$$

(the odd primes up to $y$). Define the binary-digit localization as before:

$$B_L(A) := \Big|\{(b_0(n), \dots, b_{L-1}(n)) : \ n \in A\}\Big|,$$

and define the multiplicative localization at level $y$ by

$$V_y(A) := \Big|\{(1_{p|n})_{p\in\mathcal{P}(y)} : \ n \in A\}\Big|.$$

Thus $V_y(A)$ counts how many distinct divisibility patterns by odd primes $\leq y$ occur inside $A$.

**Theorem V.2** (Quantitative arithmetic uncertainty principle)**.** *There exists an absolute constant $C > 0$ such that the following holds. For every $N \geq 2$, every $A \subset \{1, \dots, N\}$, every $L \geq 1$, and every $y \geq 3$,*

$$|A| \ \leq \ B_L(A)\, V_y(A) \left(\frac{C\,N}{2^L \log y} + 1\right). \tag{V.1}$$

*In particular, whenever $2^L \log y \to \infty$ along a family $N \to \infty$, no family of sets $A \subset \{1, \dots, N\}$ with $|A| \to \infty$ can satisfy*

$$B_L(A) = O(1) \quad \textit{and} \quad V_y(A) = O(1)$$

*uniformly in that regime.*

*Proof.* Set $q := 2^L$.

**Step 1: additive localization ⇒ few residue classes mod $q$.** If two integers $m, n$ have the same $L$ lowest binary digits, then $m \equiv n \pmod q$. Hence the set of $L$-bit patterns attained on $A$ corresponds to a set of residue classes modulo $q$. Therefore there exist residues $r_1, \dots, r_B \in \{0, \dots, q-1\}$, with $B := B_L(A)$, such that

$$A \ \subset \ \bigcup_{j=1}^{B} \{n \leq N : \ n \equiv r_j \pmod q\}. \tag{V.2}$$

[1] The triviality of this regime follows from the injectivity of the map $n \mapsto (b_0(n), \dots, b_{L-1}(n))$ once $2^L > N$, hence the terminology.

**Step 2: multiplicative localization ⇒ few prime-divisibility patterns.** For $n \in \mathbb{Z}_{>0}$ define its $y$-pattern vector

$$\pi_y(n) := (1_{p|n})_{p\in\mathcal{P}(y)} \ \in \ \{0,1\}^{\mathcal{P}(y)}.$$

By definition of $V_y(A)$, the set $\pi_y(A)$ has cardinality $V := V_y(A)$. Thus we may write

$$A \subset \bigcup_{\mathbf{v}\in\pi_y(A)} A(\mathbf{v}), \qquad A(\mathbf{v}) := \{n \in A : \ \pi_y(n) = \mathbf{v}\}. \tag{V.3}$$

**Step 3: bounding one residue class and one pattern (a sieve bound).** Fix a residue $r \pmod q$ and a pattern $\mathbf{v} \in \{0,1\}^{\mathcal{P}(y)}$. Let $S(\mathbf{v}) \subset \mathcal{P}(y)$ be the set of primes with $v_p = 1$ and set

$$d(\mathbf{v}) := \prod_{p\in S(\mathbf{v})} p \quad \text{(an odd integer, possibly } d(\mathbf{v}) = 1\text{).}$$

Any $n$ with $\pi_y(n) = \mathbf{v}$ has the form $n = d(\mathbf{v})\, m$, where $m$ is *not* divisible by any prime in $\mathcal{P}(y) \setminus S(\mathbf{v})$. Moreover, since $d(\mathbf{v})$ is odd, it is invertible modulo $q$, hence the congruence $n \equiv r \pmod q$ is equivalent to

$$m \equiv r\, d(\mathbf{v})^{-1} \pmod q.$$

Therefore,

$$\#\{n \le N : \ n \equiv r \ (\mathrm{mod}\ q), \ \pi_y(n) = \mathbf{v}\} \le \#\left\{m \le N/d(\mathbf{v}) : m \equiv r d(\mathbf{v})^{-1} (\mathrm{mod}\ q), p \nmid m \forall p \in \mathcal{P}(y) \setminus S(\mathbf{v})\right\} \tag{V.4}$$

Now apply the simplest upper-bound sieve in an arithmetic progression: among integers in a fixed residue class modulo $q$ (with $(q,p) = 1$ for all odd primes), the proportion not divisible by a given odd prime $p$ is at most $(1 - 1/p)$. Iterating over primes in $\mathcal{P}(y) \setminus S(\mathbf{v})$ yields

$$\#\left\{m \le X : \ m \equiv a \ (\mathrm{mod}\ q), \ p \nmid m \ \forall p \in \mathcal{P}(y) \setminus S(\mathbf{v})\right\} \le \frac{X}{q} \prod_{p\in\mathcal{P}(y)\setminus S(\mathbf{v})} \left(1 - \frac{1}{p}\right) \ + \ 1, \tag{V.5}$$

valid for all $X \ge 1$ and all residues $a$. With $X = N/d(\mathbf{v})$ we obtain

$$\#\{n \le N : \ n \equiv r \ (\mathrm{mod}\ q), \ \pi_y(n) = \mathbf{v}\} \le \frac{N}{q\, d(\mathbf{v})} \prod_{p\in\mathcal{P}(y)\setminus S(\mathbf{v})} \left(1 - \frac{1}{p}\right) \ + \ 1. \tag{V.6}$$

Rewrite the product by multiplying and dividing by $\prod_{p\in S(\mathbf{v})}(1 - 1/p)$:

$$\frac{1}{d(\mathbf{v})} \prod_{p\in\mathcal{P}(y)\setminus S(\mathbf{v})} \left(1 - \frac{1}{p}\right) = \prod_{p\in\mathcal{P}(y)} \left(1 - \frac{1}{p}\right) \cdot \prod_{p\in S(\mathbf{v})} \frac{1}{p(1 - 1/p)} = \prod_{p\in\mathcal{P}(y)} \left(1 - \frac{1}{p}\right) \cdot \prod_{p\in S(\mathbf{v})} \frac{1}{p-1} \ \le \ \prod_{p\in\mathcal{P}(y)} \left(1 - \frac{1}{p}\right), \tag{V.7}$$

since each factor $1/(p-1) \le 1$ for $p \ge 3$. Thus for every $r$ and every pattern $\mathbf{v}$,

$$\#\{n \le N : \ n \equiv r \ (\mathrm{mod}\ q), \ \pi_y(n) = \mathbf{v}\} \ \le \ \frac{N}{q} \prod_{p\in\mathcal{P}(y)} \left(1 - \frac{1}{p}\right) + 1 \tag{V.8}$$

**Step 4: Mertens bound.** By Mertens' theorem (applied to odd primes), there is an absolute constant $C > 0$ such that

$$\prod_{p\in\mathcal{P}(y)} \left(1 - \frac{1}{p}\right) \ \le \ \frac{C}{\log y} \qquad (y \ge 3).$$

Combining with (V.8) gives

$$\#\{n \le N : \ n \equiv r \ (\mathrm{mod}\ q), \ \pi_y(n) = \mathbf{v}\} \ \le \ \frac{C\,N}{q \log y} + 1. \tag{V.9}$$

**Step 5: sum over residues and patterns.** By (V.2) and (V.3), $A$ is contained in the union over at most $B$ residues and at most $V$ patterns. Therefore,

$$|A| \ \le \ \sum_{j=1}^{B} \ \sum_{\mathbf{v}\in\pi_y(A)} \#\{n \le N : \ n \equiv r_j \ (\mathrm{mod}\ q), \ \pi_y(n) = \mathbf{v}\} \le B\,V \left(\frac{C\,N}{q \log y} + 1\right), \tag{V.10}$$

which is exactly (V.1) since $q = 2^L$. □

**Remark V.3** (How this matches the "uncertainty" intuition)**.** *The inequality* (V.1) *shows a quantitative trade-off: if $A$ occupies only $O(1)$ low-bit configurations (small $B_L(A)$) and only $O(1)$ prime-divisibility configurations up to a large cutoff $y$ (small $V_y(A)$), then $A$ cannot be large unless the scale factor $N/2^L$ is large enough to compensate for the sieving loss $\asymp 1/\log y$.*

**Remark V.4** (Density form and Heisenberg-type interpretation)**.** *Dividing* (V.1) *by $N$, one obtains a bound on the density $\delta(A) := |A|/N$:*

$$\delta(A) \ \le \ B_L(A)\, V_y(A) \left(\frac{C}{2^L \log y} + \frac{1}{N}\right).$$

*In the asymptotic regime $N/(2^L \log y) \gg 1$, the second term becomes negligible, and the inequality takes the form*

$$\delta(A) \ \lesssim \ \frac{B_L(A)\, V_y(A)}{2^L \log y}$$

*This admits a natural "uncertainty principle" interpretation. Localization in additive spectral coordinates at depth $L$ (volume $\sim 2^L$) and localization in multiplicative spectral coordinates up to scale $y$ (volume $\sim \log y$) jointly constrain the achievable density of $A$. From this perspective, the quantity $2^L \log y$ plays the role of an* arithmetic phase space volume, *and the inequality above expresses that the density of $A$ is bounded by the inverse of this volume. Increasing localization in either representation therefore forces a loss of density, reflecting an intrinsic trade-off between additive and multiplicative structure, in the spirit of Heisenberg-type uncertainty principles.*

*The additive constant $+1$ in* (V.1) *is essential to ensure validity at finite scales; the Heisenberg-type interpretation emerges only in the asymptotic regime.*

## B. Operator-theoretic interpretation

In the Hilbert space $\mathcal{H}$, the additive observables $\{\widetilde{N}_k\}$ are diagonal in additive spectral bases, while the multiplicative observables $\{N_p\}$ are diagonal in the prime spectral basis. States that are approximately joint eigenstates of many $\widetilde{N}_k$ cannot simultaneously be approximate joint eigenstates of many $N_p$.

Dynamically, this incompatibility reflects the behavior of the shift operators: $T_m$ is local in prime coordinates, but mixing in additive coordinates, and $S_1$ is local in additive coordinates, but mixing in prime coordinates.

**Remark V.5** (Scope and relation to later results). *The uncertainty principles established in this section impose* static *constraints on the size of a set under simultaneous additive and multiplicative localization. They should be distinguished from the* capacity *and* pair-count *mechanisms developed in later sections, which control the frequency of additive relations and force the appearance of large prime factors in $A+A$. The two viewpoints are complementary rather than redundant.*

### C. Relation to sum–product phenomena

The arithmetic uncertainty principle provides a conceptual explanation for sum–product type results. A set with strong multiplicative structure cannot remain additively rigid under translation without forcing expansion into new additive configurations, and vice versa. This clarifies why additive–multiplicative expansion is a structural necessity rather than a consequence of combinatorial complexity.

In the subsequent sections we quantify this incompatibility more precisely, connecting it to additive innovation, entropy production, and the appearance of large prime factors.

## VI. ADDITIVE INNOVATION AND ITS STRUCTURAL OBSTRUCTION

This section introduces the central quantitative notion of *additive innovation*: the genuinely new prime content created by a sum. We identify the maximally innovative regime (coprime inputs), and show that *all* failures of innovation arise from a precise and rigid structural obstruction, which admits a natural formulation in terms of smoothness and $S$-unit phenomena. This framework provides the arithmetic backbone for the abc-type results developed later in the paper.

### A. Arithmetic orthogonality and strict innovation

Addition behaves in a fundamentally different manner from multiplication. Given $a, b \in \mathbb{Z}_{>0}$ and $c = a + b$, even complete knowledge of the prime spectra of $a$ and $b$ typically gives little information about the prime spectrum of $c$.

The key structural notion governing this phenomenon is coprimality, which we interpret as *arithmetic orthogonality.*

**Definition VI.1** (Arithmetic orthogonality). *Two positive integers $a$ and $b$ are said to be* orthogonal *if* $\gcd(a, b) = 1$.

**Proposition VI.2** (Strict additive orthogonality). *If* $\gcd(a, b) = 1$*, then*

$$S(a+b) \cap \big(S(a) \cup S(b)\big) = \varnothing.$$

*Proof.* Suppose a prime $p$ divides both $a + b$ and $a$. Then $p \mid b$, contradicting $\gcd(a, b) = 1$. The same argument applies if $p \mid b$. □

Thus, for orthogonal pairs, addition produces a number whose prime spectrum is entirely disjoint from those of the summands. In this regime, addition acts as a *maximally innovative* operation.

### B. Innovation set and innovation measure

The preceding discussion motivates a quantitative invariant measuring the new prime content created by a sum.

**Definition VI.3** (Additive innovation set). *For $a, b \in \mathbb{Z}_{>0}$, define*

$$\Delta(a, b) := S(a+b) \setminus \big(S(a) \cup S(b)\big).$$

The set $\Delta(a, b)$ consists exactly of the new prime directions generated by the sum $a + b$.

**Definition VI.4** (Additive innovation measure). *The* additive innovation *of $(a, b)$ is defined by*

$$\mathcal{I}(a, b) := \sum_{p \in \Delta(a,b)} \log p.$$

The logarithmic weight is natural in view of radicals, heights, and Diophantine inequalities.

### C. Maximal innovation for coprime sums

For coprime pairs, Proposition VI.2 yields a simple structural identity: all prime factors of the sum are necessarily new.

**Theorem VI.5** (Maximal innovation for coprime sums). *If* $\gcd(a, b) = 1$*, then*

$$\Delta(a, b) = S(a+b) \text{ and hence } \mathcal{I}(a, b) = \sum_{p \mid (a+b)} \log p = \log \operatorname{rad}(a+b)$$

*In particular,* $\mathcal{I}(a, b) \geq \log 2$.

*Proof.* If $\gcd(a, b) = 1$, then Proposition VI.2 implies $S(a+b) \cap (S(a) \cup S(b)) = \varnothing$, hence $\Delta(a, b) = S(a+b)$. The identity for $\mathcal{I}(a, b)$ follows from the definition. Since $a + b \geq 2$, we have $\operatorname{rad}(a + b) \geq 2$, giving $\mathcal{I}(a, b) \geq \log 2$. □

### D. Normalization and the emergence of obstruction

The coprime case exhibits the strongest possible form of innovation. When $\gcd(a, b) > 1$, innovation may be partially reduced or even vanish altogether. We now show that *all* such failures arise from a single structural mechanism.

Let $a, b \in \mathbb{Z}_{>0}$ and write

$$g = \gcd(a, b), \qquad a = gx, \qquad b = gy, \qquad \gcd(x, y) = 1. \tag{VI.1}$$

Then $c = a + b = g(x + y)$ and

$$S(c) = S(g) \cup S(x + y).$$

Since $\gcd(x, y) = 1$, strict additive orthogonality implies

$$S(x + y) \cap \big(S(x) \cup S(y)\big) = \varnothing.$$

Consequently, any failure of innovation must arise from the containment

$$S(x + y) \subseteq S(g),$$

that is, from the absorption of the prime spectrum of the coprime core $x + y$ into the common factor $g$.

**Definition VI.6** ($S$-smooth collapse). *We say that $(a, b)$ exhibits an $S$-smooth collapse if, with* (VI.1)*, one has*

$$S(x + y) \subseteq S(g).$$

Equivalently,

$$\mathcal{I}(a, b) = \sum_{p \mid (x+y),\, p \nmid g} \log p, \tag{VI.2}$$

so $\mathcal{I}(a, b) = 0$ if and only if $(a, b)$ exhibits an $S$-smooth collapse.

**Remark VI.7** (Collapse or a new prime). *With the normalization $a = gx$, $b = gy$, $\gcd(x, y) = 1$, either the coprime core $x+y$ has all its prime factors contained in $S(g)$, or else it contributes at least one genuinely new prime to the sum. In the latter case,* $\mathcal{I}(a, b) \geq \log 2$.

### E. Innovation deficit and an abc-type inequality with obstruction

The abc conjecture predicts that for coprime $a, b, c$ with $a+b=c$, the radical $\operatorname{rad}(abc)$ is rarely much smaller than $c$. In our framework, this corresponds to the typical largeness of additive innovation. To isolate the precise obstruction, we introduce the following invariant.

**Definition VI.8** (Innovation deficit)**.** *Define the* innovation deficit

$$D(a,b) := \log \operatorname{rad}(x+y) - \mathcal{I}(a,b) = \sum_{\substack{p\mid(x+y)\\ p\mid g}} \log p,$$

*where* (VI.1) *holds.*

Thus $D(a,b)$ measures exactly the portion of the prime spectrum of $x+y$ that is absorbed by the common factor $g$.

**Theorem VI.9** (abc-type inequality with explicit deficit)**.** *Let $a, b \in \mathbb{Z}_{>0}$, $c = a+b$, and let $g = \gcd(a,b)$, $a = gx$, $b = gy$, $\gcd(x,y) = 1$. Then*

$$\log c \ \le\ \log g + \log \operatorname{rad}(g) + \mathcal{I}(a,b) + D(a,b).$$

*Proof.* The identity $c = g(x+y)$ gives $\log c = \log g + \log(x+y)$. Writing

$$\log \operatorname{rad}(x+y) = \mathcal{I}(a,b) + D(a,b)$$

and using $\log(x+y) \le \log \operatorname{rad}(x+y) + \log(x+y/\operatorname{rad}(x+y))$ yields the stated inequality after regrouping terms. □

**Interpretation.** An abc-type bound is obstructed *only* by the deficit $D(a,b)$, that is, by primes of $x+y$ already present in $g$.

### F. A restricted abc principle under bounded deficit

We record a clean conditional formulation making this equivalence explicit.

**Theorem VI.10** (Restricted abc via bounded deficit)**.** *Fix $\varepsilon > 0$. Suppose that for all pairs $(a,b)$ in a class $\mathscr{C}$ one has*

$$D(a,b) \le \varepsilon \log(x+y) + O_\varepsilon(1),$$

*where* (VI.1) *holds. Then for all $(a,b) \in \mathscr{C}$,*

$$\mathcal{I}(a,b) \ge (1-\varepsilon)\log(x+y) - O_\varepsilon(1).$$

*Equivalently,*

$$\log \operatorname{rad}(x+y) \ge (1-\varepsilon)\log(x+y) - O_\varepsilon(1).$$

*Proof.* By definition, $\log \operatorname{rad}(x+y) = \mathcal{I}(a,b) + D(a,b)$, and the claim follows immediately. □

This formulation shows that the abc philosophy is precisely equivalent to controlling how often the coprime core $x+y$ shares too much prime support with the common factor $g$.

### G. Collapse as spectral containment

We now give a complete structural characterization of *innovation collapse*. With the normalization (VI.1), since $\gcd(x,y) = 1$ strict additive orthogonality implies

$$S(x+y) \cap \Big(S(x) \cup S(y)\Big) = \varnothing.$$

Hence the innovation set reduces to the clean expression

$$\Delta(a,b) = S(x+y) \setminus S(g), \tag{VI.3}$$

and collapse can occur only through absorption into $g$.

**Theorem VI.11** (Spectral containment ⇔ innovation collapse)**.** *Let $a, b \in \mathbb{Z}_{>0}$ and write $a = gx$, $b = gy$ with $\gcd(x,y) = 1$. Then the following are equivalent:*

*(i)* $\Delta(a,b) = \varnothing$ *(no additive innovation);*

*(ii)* $S(a+b) \subseteq S(a) \cup S(b)$*;*

*(iii)* $S(x+y) \subseteq S(g)$ *(the coprime core is $S(g)$-smooth).*

*Equivalently,* all *failures of additive innovation arise from the containment of the prime spectrum of $x+y$ inside that of the gcd $g$.*

*Proof.* Since $c = g(x+y)$ we have $S(c) = S(g) \cup S(x+y)$, while $S(a) \cup S(b) = S(g) \cup S(x) \cup S(y)$. Using strict orthogonality for $(x,y)$,

$$S(x+y) \cap (S(x) \cup S(y)) = \varnothing,$$

we obtain

$$\Delta(a,b) = S(c) \setminus (S(a) \cup S(b)) = S(x+y) \setminus S(g).$$

Thus $\Delta(a,b) = \varnothing$ if and only if $S(x+y) \subseteq S(g)$. The equivalence with (ii) is immediate from the definition of $\Delta(a,b)$. □

**Interpretation.** Innovation collapse is a *level-2* phenomenon: it cannot occur at the level of the coprime inputs $(x,y)$ themselves, but only through absorption into a shared multiplicative factor.

### H. $S$-unit reformulation

The containment condition admits a natural reformulation in the language of $S$-integers and $S$-units.

**Definition VI.12** ($S$-integers and $S$-units)**.** *Fix a finite set of primes $S$. An $S$-*integer *is an integer whose prime spectrum is contained in $S$. An $S$-*unit *is a rational number of the form*

$$u = \pm \prod_{p\in S} p^{k_p}, \qquad k_p \in \mathbb{Z}.$$

**Proposition VI.13** ($S$-unit form of innovation collapse)**.** *Let $a = gx$, $b = gy$ with $g = \gcd(a,b)$ and $\gcd(x,y) = 1$. Then $\Delta(a,b) = \varnothing$ if and only if $x+y$ is an $S(g)$-integer. Equivalently, writing $u := x+y$, one has*

$$\frac{x}{u} + \frac{y}{u} = 1$$

*with $\frac{x}{u}$ and $\frac{y}{u}$ being $S(g)$-units.*

*Proof.* By Theorem VI.11, $\Delta(a,b) = \varnothing$ is equivalent to $S(x+y) \subseteq S(g)$, i.e. $x+y$ is an $S(g)$-integer. Setting $u = x+y$ and dividing by $u$ yields the $S$-unit equation. The converse is immediate. □

### I. Finiteness for fixed prime support

A key consequence of the $S$-unit reformulation is finiteness: collapse is not only rigid but arithmetically sparse when the prime support is fixed.

**Corollary VI.14** (Finiteness for fixed prime support)**.** *Fix a finite set of primes $S$. Up to scaling by $g = \gcd(a,b)$, there exist only finitely many coprime pairs $(x,y)$ such that $x+y$ is an $S$-integer.*

In other words, for fixed $S$ the solutions to $x+y=u$ with $u$ an $S$-integer and $\gcd(x,y) = 1$ form only finitely many normalized configurations; scaling recovers the full family $(a,b) = (gx, gy)$.

### J. Relation to the abc/Szpiro philosophy (structural level)

**Remark (abc as rarity of the $S$-unit obstruction).** One standard form of the abc conjecture asserts that for every $\varepsilon > 0$, there exists $K_\varepsilon > 0$ such that for coprime $a, b, c$ with $a + b = c$,

$$c \;\le\; K_\varepsilon \,\mathrm{rad}(abc)^{\,1+\varepsilon}.$$

In the present framework, the obstruction to such inequalities is precisely the phenomenon isolated above: configurations in which the prime support of the coprime core $x + y$ remains trapped inside a bounded set of primes already present in $g$ (equivalently, an $S$-unit/smoothness phenomenon). Thus, the abc/Szpiro philosophy becomes the statement that *severe innovation collapse or severe innovation deficit is exceptionally rare.*

**Remark (Szpiro dictionary, integer model).** At a heuristic level, the usual Szpiro–abc dictionary reads:

- **Height** $\leftrightarrow$ size, modeled here by $\log c$.
- **Conductor** $\leftrightarrow$ prime support, modeled here by $\mathrm{rad}(abc)$.
- **Exceptional families** $\leftrightarrow$ bounded prime support / $S$-unit phenomena, modeled here by $S(x+y) \subseteq S(g)$ and large overlap $S(x + y) \cap S(g)$.

The next sections move beyond structure and quantify typical innovation and the rarity of collapse in probabilistic and density senses.

## VII. TYPICAL LAWS OF ADDITIVE INNOVATION

This section quantifies the *typical* behavior of additive innovation. Section VI isolated the unique structural mechanism by which innovation can fail: absorption of the coprime core $x + y$ into the prime support of the common factor $g = \gcd(a, b)$ (equivalently, an $S$-unit / smoothness phenomenon). We now show that this obstruction is rare in several complementary senses: (i) on average, addition creates about $\log\log N$ genuinely new prime factors; (ii) at small and medium prime scales, innovation obeys a Gaussian law; and (iii) severe failure of innovation is controlled by classical smoothness estimates and therefore occurs on sets of vanishing density.

### A. Mean laws for the number of innovative primes

Recall the innovation set

$$\Delta(a, b) := S(a + b) \setminus \big(S(a) \cup S(b)\big),$$

and define the innovation count

$$\omega_{\mathrm{innov}}(a, b) := |\Delta(a, b)| = \#\{p :\; p \mid (a + b),\; p \nmid a,\; p \nmid b\}.$$

**Theorem VII.1** (First moment law for additive innovation)**.** *Let $(a, b)$ be chosen uniformly at random from $[1, N]^2$. Then*

$$\mathbb{E}\,\omega_{\mathrm{innov}}(a, b) = \log\log N + O(1) \qquad (N \to \infty).$$

*The same estimate holds if the expectation is taken over coprime pairs $(a, b)$ (i.e. conditioned on $\gcd(a, b) = 1$).*

*Proof.* Write

$$\omega_{\mathrm{innov}}(a, b) = \sum_{p \le 2N} Z_p(a, b), \quad Z_p(a, b) := \mathbf{1}_{\{p \mid (a+b),\; p\nmid a,\; p\nmid b\}}$$

By linearity of expectation,

$$\mathbb{E}\,\omega_{\mathrm{innov}}(a, b) = \sum_{p \le 2N} \mathbb{P}(Z_p = 1).$$

Fix a prime $p \le 2N$. We count pairs $(a, b) \in [1, N]^2$ with $a \equiv -b \pmod p$ and $p \nmid a, b$. For each residue class $r \in (\mathbb{Z}/p\mathbb{Z})^\times$, the number of $a \in [1, N]$ with $a \equiv r \pmod p$ is $N/p + O(1)$, and similarly for $b$. Thus

$$\begin{aligned}&\#\{(a, b) \in [1, N]^2 :\; a \equiv -b \pmod p,\; p \nmid a,\; p \nmid b\}\\ &= (p-1)\Big(\frac{N}{p} + O(1)\Big)^2 = \frac{(p-1)N^2}{p^2} + O\Big(\frac{N}{p} + p\Big)\end{aligned} \tag{VII.1}$$

Dividing by $N^2$ gives

$$\mathbb{P}(Z_p = 1) = \frac{p-1}{p^2} + O\Big(\frac{1}{pN} + \frac{p}{N^2}\Big) = \frac{1}{p} + O\Big(\frac{1}{p^2}\Big) + O\Big(\frac{1}{pN}\Big)$$

Summing over $p \le 2N$ and using Mertens' theorem yields

$$\mathbb{E}\,\omega_{\mathrm{innov}}(a, b) = \sum_{p \le 2N} \frac{1}{p} + O(1) = \log\log N + O(1).$$

where we have used that

$$\begin{aligned}\sum_{p \le 2N} \Big(\frac{1}{pN} + \frac{p}{N^2}\Big) &\ll \frac{\log\log N}{N} + \frac{1}{N^2}\sum_{p \le 2N} p\\ &\ll \frac{\log\log N}{N} + 1 \;=\; O(1),\end{aligned} \tag{VII.2}$$

so these error terms contribute only $O(1)$ to the prime sum.

For the coprime-conditioned expectation, note that $\mathbb{P}(\gcd(a, b) = 1) = 6/\pi^2 + o(1)$ and conditioning removes only a constant proportion of pairs; the same modular counting remains valid with the same $O(1)$-level error. ☐

**Interpretation.** A typical integer $n \le N$ has $\omega(n) \sim \log\log N$ distinct prime factors on average. Theorem VII.1 shows that, for a typical pair $(a, b)$, the sum $a+b$ contributes about $\log\log N$ *additional* prime factors not present in either summand. In this sense, addition acts as a robust generator of multiplicative randomness: two prime spectra "collide" and typically produce a third spectrum with essentially fresh support.

This also complements the arithmetic uncertainty principle from Section V: while exact joint localization is forbidden there, the typical laws proved here show that the resulting spread into new multiplicative coordinates is not only necessary, but quantitatively stable across scales.

### B. Gaussian fluctuations at small and medium prime scales

Fix $\theta \in (0, \frac{1}{2})$. Define the truncated innovation count

$$\begin{aligned}\omega^{(\theta)}_{\mathrm{innov}}(a, b) &:= \#\{p \le N^\theta :\; p \mid (a + b),\; p \nmid a,\; p \nmid b\}\\ &= \sum_{p \le N^\theta} Z_p(a, b),\end{aligned} \tag{VII.3}$$

where $Z_p$ is as above.

**Lemma VII.1.1** (Mean and variance of truncated innovation)**.** *Let $(a, b)$ be uniform in $[1, N]^2$ and fix $\theta \in (0, \frac{1}{2})$. Then, as $N \to \infty$,*

$$\mathbb{E}\,\omega^{(\theta)}_{\mathrm{innov}}(a, b) = \sum_{p \le N^\theta} \frac{p-1}{p^2} + O(1) = \log\log\big(N^\theta\big) + O(1),$$

$$\mathrm{Var}\Big(\omega^{(\theta)}_{\mathrm{innov}}(a, b)\Big) = \log\log\big(N^\theta\big) + O(1),$$

*where the implied constants may depend on $\theta$.*

*Proof.* Write $\omega^{(\theta)}_{\text{innov}} = \sum_{p \le N^\theta} Z_p$. The estimate for $\mathbb{E}Z_p$ follows from the same modular counting as in Theorem VII.1:

$$\mathbb{E}Z_p = \mathbb{P}(p \mid a+b,\ p \nmid a,\ p \nmid b) = \frac{p-1}{p^2} + O\Big(\frac{1}{pN}\Big) = \frac{1}{p} + O\Big(\frac{1}{p^2}\Big) + O\Big(\frac{1}{pN}\Big) \tag{VII.4}$$

Summing over $p \le N^\theta$ gives

$$\mathbb{E}\,\omega^{(\theta)}_{\text{innov}} = \sum_{p \le N^\theta} \frac{1}{p} + O(1) = \log\log\big(N^\theta\big) + O(1)$$

For the variance, expand

$$\operatorname{Var}\Big(\sum_{p \le N^\theta} Z_p\Big) = \sum_{p \le N^\theta} \operatorname{Var}(Z_p) + 2 \sum_{p<q \le N^\theta} \operatorname{Cov}(Z_p, Z_q).$$

Since $Z_p$ is an indicator with $\mathbb{E}Z_p \asymp 1/p$, we have

$$\sum_{p \le N^\theta} \operatorname{Var}(Z_p) = \sum_{p \le N^\theta} \mathbb{E}Z_p(1-\mathbb{E}Z_p) = \log\log\big(N^\theta\big) + O(1)$$

To control the covariance, fix distinct primes $p \ne q \le N^\theta$. By the Chinese remainder theorem and the same residue-class counting as above, the joint constraint $(p \mid a+b,\ q \mid a+b)$ and $(p \nmid ab,\ q \nmid ab)$ has asymptotic probability

$$\mathbb{E}(Z_p Z_q) = \frac{(p-1)(q-1)}{p^2 q^2} + O\Big(\frac{1}{N}\Big),$$

and hence

$$\operatorname{Cov}(Z_p, Z_q) = \mathbb{E}(Z_p Z_q) - \mathbb{E}Z_p\,\mathbb{E}Z_q = O\Big(\frac{1}{N}\Big).$$

Therefore

$$\sum_{p<q \le N^\theta} \operatorname{Cov}(Z_p, Z_q) \ll \frac{1}{N}\,\pi(N^\theta)^2 \ll \frac{N^{2\theta}}{N(\log N)^2} = o(1),$$

since $\theta < \frac{1}{2}$. This yields

$$\operatorname{Var}(\omega^{(\theta)}_{\text{innov}}) = \log\log\big(N^\theta\big) + O(1)$$

□

**Theorem VII.2** (Truncated Erdős–Kac law for innovation)**.** *Fix $\theta \in (0, \frac{1}{2})$. Then, as $N \to \infty$,*

$$\frac{\omega^{(\theta)}_{\text{innov}}(a,b) - \log\log\big(N^\theta\big)}{\sqrt{\log\log\big(N^\theta\big)}} \Longrightarrow \mathcal{N}(0,1)$$

*in distribution.*

*Proof.* Apply the classical Erdős–Kac method of moments to the triangular array $\{Z_p\}_{p \le N^\theta}$. For fixed distinct primes $p_1, \ldots, p_k \le N^\theta$, Chinese remainder theory and uniform residue counting give

$$\mathbb{E}\prod_{j=1}^{k} Z_{p_j} = \prod_{j=1}^{k} \frac{p_j - 1}{p_j^2} + O_k\Big(\frac{1}{N}\Big) = \prod_{j=1}^{k} \mathbb{E}Z_{p_j} + O_k\Big(\frac{1}{N}\Big),$$

and the error is negligible since $p_1 \cdots p_k \le N^{k\theta} = o(N)$ when $\theta < \frac{1}{2}$. This verifies the moment convergence to those of a standard Gaussian. □

## C. Smoothness bounds and rarity of collapse

Recall that complete collapse is governed by smoothness of the coprime core: with $a = gx$, $b = gy$, $\gcd(x,y) = 1$, one has

$$\Delta(a,b) = \varnothing \iff S(x+y) \subseteq S(g)$$

by Theorem VI.11 in Section VI. A natural quantitative proxy is the size of the largest prime factor of $x+y$.

Let $P^+(n)$ denote the largest prime divisor of $n$ (with $P^+(1) = 1$), and let

$$\Psi(X,y) := \#\{n \le X :\ P^+(n) \le y\}$$

be the standard smooth-number counting function.

**Lemma VII.2.1** (Pairs with smooth sum)**.** *Let $N \ge 1$ and $y \ge 2$. Then*

$$\#\{(x_1,x_2) \in [1,N]^2 :\ x_1+x_2 \textit{ is } y\textit{-smooth}\} \le N \cdot \Psi(2N, y).$$

*Proof.* For each $n \in [2, 2N]$, the number of representations $n = x_1 + x_2$ with $1 \le x_1, x_2 \le N$ is at most $N$. Summing over $y$-smooth $n \le 2N$ gives the bound. □

**Proposition VII.3** (Smooth sums are rare on power scales)**.** *Fix $\beta \in (0,1)$. Then*

$$\mathbb{P}_{(x_1,x_2)\sim \mathrm{Unif}([1,N]^2)}\big(P^+(x_1+x_2) \le N^\beta\big) \ll \rho\Big(\frac{1}{\beta}\Big) + o(1), \quad (N \to \infty), \tag{VII.5}$$

*where $\rho$ denotes the Dickman function.*

*Proof.* By Lemma VII.2.1,

$$\mathbb{P}\Big(P^+(x_1+x_2) \le N^\beta\Big) \le \frac{N\Psi(2N, N^\beta)}{N^2} = \frac{\Psi(2N, N^\beta)}{N}.$$

The Dickman–de Bruijn estimate gives $\Psi(2N, N^\beta) = (2N)\rho(1/\beta)(1+o(1))$, yielding the claim. □

## D. Large new primes: probabilistic and density formulations

We now translate smoothness rarity into innovation lower bounds for general pairs $a = gx$, $b = gy$ with $\gcd(x,y) = 1$.

**Definition VII.4** (Largest new prime factor)**.** *Let $a = gx$, $b = gy$ with $\gcd(x,y) = 1$. Define*

$$P_{\text{new}}(a,b) := \max\{p :\ p \mid (x+y),\ p \nmid g\},$$

*with the convention $P_{\text{new}}(a,b) = 1$ if the set is empty.*

Then $\mathcal{I}(a,b) \ge \log P_{\text{new}}(a,b)$.

**Theorem VII.5** (Large new prime for almost all coprime cores)**.** *Fix $\beta \in (0,1)$. Let $(x,y)$ be chosen uniformly from $[1,N]^2$, conditioned on $\gcd(x,y) = 1$. Then*

$$\mathbb{P}\Big(P^+(x+y) \le N^\beta\Big) \ll \rho\Big(\frac{1}{\beta}\Big) + o(1).$$

*Equivalently, the number of pairs $(x,y) \in [1,N]^2$ such that $P^+(x+y) \le N^\beta$ is*

$$\ll_\beta N^2\,\rho\Big(\frac{1}{\beta}\Big).$$

*Proof.* Proposition VII.3 gives an upper bound for uniform pairs $(x,y) \in [1,N]^2$. Conditioning on $\gcd(x,y) = 1$ changes probabilities by at most a constant factor, since

$$\mathbb{P}(\gcd(x,y) = 1) = \frac{6}{\pi^2} + o(1).$$

Hence the same upper bound remains valid (up to absolute constants) under the coprimality conditioning, yielding the claim. □

**Corollary VII.6** (Innovation on polynomial prime scales)**.** *Fix* $\beta \in (0,1)$*. Suppose that all prime divisors of* $g$ *are* $\leq N^{\beta}$*. Then, with probability* $1 - O(\rho(1/\beta))$ *over* $(x,y) \in [1,N]^2$ *with* $\gcd(x,y) = 1$*, the sum* $x+y$ *has a prime divisor* $p > N^{\beta}$*, hence*

$$P_{\rm new}(a,b) \;>\; N^{\beta} \qquad \textit{and} \qquad \mathcal{I}(a,b) \;\geq\; \beta \log N.$$

*Proof.* By Theorem VII.5, with high probability there exists a prime $p \mid (x+y)$ with $p > N^{\beta}$. Since all primes dividing $g$ are $\leq N^{\beta}$, such $p$ cannot divide $g$ and is therefore new, contributing at least $\log p \geq \beta \log N$ to $\mathcal{I}(a,b)$. ☐

**Remark VII.7.** *The preceding results show that additive innovation is not merely nonzero or large on average, but overwhelmingly typical in a strong density sense: except on a set of vanishing probability controlled by smoothness statistics, the sum* $x + y$ *produces a genuinely new prime factor on a polynomial scale.*

## E. From logarithmic innovation to many new primes

Let $\omega(n) := \#\{p : \; p \mid n\}$ denote the number of distinct prime divisors. In the coprime regime, $\Delta(x,y) = S(x+y)$ and hence $|\Delta(x,y)| = \omega(x+y)$.

**Lemma VII.7.1** (Few primes with large radical force a high prime power)**.** *Let* $n \geq 2$ *and write* $n = \prod_{i=1}^{r} p_i^{\alpha_i}$ *with* $r = \omega(n)$*. Fix* $\varepsilon \in (0,1)$*. If*

$$\textit{rad}(n) \;\geq\; n^{1-\varepsilon} \quad \textit{and} \quad \omega(n) = r,$$

*then* $n$ *is divisible by a prime power* $p^k$ *with*

$$k \;\geq\; \frac{1}{\varepsilon r}.$$

*Proof.* Identical to the argument given previously: bounding $n/\mathrm{rad}(n)$ forces some $\alpha_i$ to be large. ☐

**Theorem VII.8** (Few innovative primes are rare (explicit power-saving))**.** *Fix an integer* $k \geq 1$*. Then there exists* $\delta_k > 0$ *(e.g.* $\delta_k = \frac{1}{10k}$*) such that the number of coprime pairs* $(x,y) \in [1,N]^2$ *satisfying*

$$|\Delta(x,y)| \leq k$$

*is*

$$\ll_k \; N^{2-\delta_k}.$$

*Proof.* Split into the small-radical class $\mathrm{rad}(x+y) \leq (x+y)^{1-\varepsilon}$, controlled by Theorem X.1 [2], and the complementary class where $\mathrm{rad}(x+y)$ is large but $\omega(x+y) \leq k$, which forces a large prime power divisor by Lemma VII.7.1. Counting such integers $n = x + y \leq 2N$ and then lifting to pairs gives a power-saving bound. Choosing $\varepsilon := \frac{1}{5k}$ yields $\delta_k = \frac{1}{10k}$. ☐

**Corollary VII.9** (Many new primes in the coprime regime)**.** *Fix* $k \geq 1$*. For all but* $\ll_k N^{2-\delta_k}$ *coprime pairs* $(x,y) \in [1,N]^2$*, one has* $|\Delta(x,y)| \geq k+1$*.*

[2] All power-saving exponents appearing in Sections VII–X ultimately derive from Theorem XI.1, which provides a uniform bound of the form $N^{2-\varepsilon/2}$ for the exceptional set of sums with abnormally small radical. No later exponent exceeds this threshold without additional hypotheses

# VIII. ARITHMETIC ENTROPY AND INFORMATION GENERATION BY ADDITION

We develop an information-theoretic formulation of the same phenomenon studied spectrally throughout the paper: multiplication is conservative at the level of prime data, while addition acts as a source of *entropy* in prime divisibility patterns. The key point is that we do not condition on the full integers, but only on their *multiplicative summaries* (prime spectra).

Conceptually, this section serves as a bridge between the spectral theory developed earlier and the probabilistic laws proved later. The operator-theoretic framework explains *why* additive–multiplicative incompatibility must generate new structure, while the statistical results quantify *how much* structure is typically produced. The present entropy formulation explains *why this production is robust*: addition acts as a genuine source of arithmetic information.

## A. Prime-spectrum random variables

Let $U$ be a positive-integer-valued random variable. For a finite set of primes $\mathcal{P}$, define the prime divisibility indicator vector

$$\mathbf{X}_{\mathcal{P}}(U) := \left(\mathbf{1}_{\{p|U\}}\right)_{p\in\mathcal{P}} \in \{0,1\}^{\mathcal{P}}.$$

**Definition VIII.1** (Prime-spectrum entropy on $\mathcal{P}$)**.** *The* $\mathcal{P}$*-spectrum entropy of* $U$ *is*

$$H_{\mathcal{P}}(U) := H\big(\mathbf{X}_{\mathcal{P}}(U)\big),$$

*where* $H(\cdot)$ *denotes Shannon entropy.*

We interpret $H_{\mathcal{P}}(U)$ as the amount of uncertainty in the divisibility pattern of $U$ by primes in $\mathcal{P}$.

## B. Multiplication is information-preserving

Let $A, B$ be independent random variables. For fixed $\mathcal{P}$, multiplication corresponds to a deterministic combination rule on indicators:

$$\mathbf{1}_{AB}(p) = \mathbf{1}_A(p) \vee \mathbf{1}_B(p) \qquad (p \in \mathcal{P}),$$

where $\vee$ is logical OR. Hence $\mathbf{X}_{\mathcal{P}}(AB)$ is a deterministic function of $(\mathbf{X}_{\mathcal{P}}(A), \mathbf{X}_{\mathcal{P}}(B))$, so

$$H_{\mathcal{P}}(AB) \leq H\big(\mathbf{X}_{\mathcal{P}}(A), \mathbf{X}_{\mathcal{P}}(B)\big) = H_{\mathcal{P}}(A) + H_{\mathcal{P}}(B),$$

expressing a form of "conservation": multiplication creates no new prime-spectrum randomness beyond that present in the inputs. From an information-flow perspective, multiplication acts within a closed multiplicative channel: all prime-spectrum information of the output is already encoded in the prime spectra of the inputs. Addition, by contrast, diverts information into new prime directions that are invisible to multiplicative summaries, effectively generating "noise" that emerges as new arithmetic structure.

## C. Addition creates entropy via spectral orthogonality

In the coprime setting, addition forces strict orthogonality of prime spectra: if $\gcd(a,b) = 1$, then $S(a+b)$ is disjoint from $S(a) \cup S(b)$ (Section VI). This suggests that the prime spectrum of $A + B$ should remain significantly unpredictable even after one reveals the prime spectra of $A$ and $B$.

Since $A + B$ is deterministic given $(A, B)$, the relevant notion is not $H(A+B \mid A, B)$, but the uncertainty of *prime data* of $A + B$ given the prime data of the inputs:

$$H\Big(\mathbf{X}_{\mathcal{P}}(A+B) \,\Big|\, \mathbf{X}_{\mathcal{P}}(A), \mathbf{X}_{\mathcal{P}}(B)\Big).$$

**Definition VIII.2** (Additive spectrum entropy production)**.** *Define the $\mathcal{P}$-entropy production of addition by*

$$\mathrm{EP}_{\mathcal{P}}(A,B) := H\Big(\mathbf{X}_{\mathcal{P}}(A+B)\,\Big|\,\mathbf{X}_{\mathcal{P}}(A),\mathbf{X}_{\mathcal{P}}(B)\Big).$$

Heuristic principle: for "typical" large $A, B$, the quantity $\mathrm{EP}_{\mathcal{P}}(A,B)$ is extensive in $|\mathcal{P}|$, reflecting that addition mixes prime directions in a way not recoverable from multiplicative summaries of the inputs. More precisely, one expects extensivity in the following weak sense: under mild mixing or equidistribution assumptions on $(A,B)$ (such as those established in the operator-theoretic framework of Sections III-V), there exists a constant $c > 0$ such that for finite prime sets $\mathcal{P}$,

$$\mathrm{EP}_{\mathcal{P}}(A,B) \;\geq\; c\sum_{p\in\mathcal{P}}\frac{\log p}{p},$$

up to lower-order terms. This matches the Mertens scale $\sum_{p\leq P}\frac{\log p}{p} \sim \log P$ and is consistent with the typical innovation laws proved in Section VII.

**Remark VIII.3** (Entropy production and information deficit)**.** *Since*

$$H(X\mid Y) = H(X) - I(X;Y),$$

*the entropy production* $\mathrm{EP}_{\mathcal{P}}(A,B)$ *can be rewritten as*

$$\mathrm{EP}_{\mathcal{P}}(A,B) = H_{\mathcal{P}}(A{+}B) - I\Big(\mathbf{X}_{\mathcal{P}}(A{+}B);\mathbf{X}_{\mathcal{P}}(A),\mathbf{X}_{\mathcal{P}}(B)\Big).$$

*Thus entropy production measures precisely the portion of the prime-spectrum information of $A+B$ that cannot be inferred from the multiplicative summaries of the inputs, providing an information-theoretic interpretation of additive innovation.*

## D. Primewise entropy production

Lower bounds on $\mathrm{EP}_{\mathcal{P}}$ can be delicate because they require control of correlations across primes. A robust surrogate is a primewise sum of single-prime conditional entropies.

For each prime $p$, define

$$X_p := \mathbf{1}_{\{p|A\}}, \qquad Y_p := \mathbf{1}_{\{p|B\}}, \qquad U_p := \mathbf{1}_{\{p|(A+B)\}}.$$

**Definition VIII.4** (Primewise entropy production)**.** *The primewise entropy production up to $P$ is defined as*

$$\mathrm{PEP}(A,B;P) := \sum_{p\leq P} H\Big(U_p \mid X_p, Y_p\Big).$$

This quantity measures the cumulative unpredictability, prime by prime, of the divisibility of $A+B$ after revealing the multiplicative data of $A$ and $B$ at each prime.

## E. A toy model illustrating the mechanism

Fix a prime $p$. Consider a simplified model where $(A,B)$ is approximately uniform modulo $p$ after conditioning on whether $p$ divides $A$ or $B$, and where the event $A \equiv -B \pmod p$ occurs with probability $\asymp 1/p$. In such a model,

$$\mathbb{P}(p\mid A+B) \approx \frac{1}{p},$$

and the divisibility event $p \mid (A+B)$ retains substantial entropy even after conditioning on $(X_p, Y_p)$. Aggregating over primes suggests extensive entropy production: addition introduces congruence conditions largely unrelated to the multiplicative structure of the inputs.

This primewise uniformity mechanism is the same one underlying the modular counting arguments used in Section VII, now interpreted through an information-theoretic lens.

**Remark VIII.5.** *This toy model is not meant as an additional assumption, but as a simplified reflection of the modular mixing phenomena established earlier in the paper. In particular, the operator-theoretic analysis of additive shifts (Sections III-V) implies that for sufficiently delocalized arithmetic states, the distribution of $A+B$ modulo $p$ approaches uniformity, up to negligible errors.*

The next section turns this heuristic into a rigorous inequality linking entropy production to the deterministic innovation measure $\mathcal{I}(a,b)$.

# IX. ENTROPY–INNOVATION INEQUALITY AND MIXING FROM EQUIDISTRIBUTION

We connect the deterministic innovation measure $\mathcal{I}(a,b)$ to the information-theoretic quantity $\mathrm{EP}_{\mathcal{P}}(A,B)$ from Section VIII. The bridge is obtained by writing $\mathcal{I}$ as a weighted sum of prime-divisibility events and lower bounding entropy production prime-by-prime under a modular mixing hypothesis, which we then verify from equidistribution modulo primes.

## A. Innovation as a weighted sum of prime events

Let $a = gx$, $b = gy$ with $\gcd(x,y) = 1$. For each prime $p$, define the indicator of a *new* prime appearing in the sum:

$$Z_p(a,b) := \mathbf{1}_{\{p|(x+y)\}}\cdot\mathbf{1}_{\{p\nmid g\}}.$$

Then

$$\mathcal{I}(a,b) = \sum_p (\log p)\, Z_p(a,b),$$

where the sum is finite.

Taking expectations over any distribution on pairs $(a,b)$ yields the exact identity

$$\mathbb{E}\,\mathcal{I}(a,b) = \sum_p (\log p)\,\mathbb{P}\Big(p\mid(x+y)\text{ and }p\nmid g\Big). \tag{IX.1}$$

## B. Entropy production conditioned on input spectra

Fix a finite set of primes $\mathcal{P}$. Recall the entropy production (Definition VIII.2)

$$\mathrm{EP}_{\mathcal{P}}(A,B) := H\Big(\mathbf{X}_{\mathcal{P}}(A+B)\,\Big|\,\mathbf{X}_{\mathcal{P}}(A),\mathbf{X}_{\mathcal{P}}(B)\Big).$$

## C. A Bernoulli entropy lower bound

**Lemma IX.0.1** (Bernoulli entropy lower bound)**.** *Let $U \in \{0,1\}$ be a Bernoulli random variable with $\mathbb{P}(U=1) = q$. Then*

$$H(U) = -q\log q - (1-q)\log(1-q) \;\geq\; 2q(1-q),$$

*where the logarithm is natural and the constant 2 is absolute.*

*Proof.* The inequality $-t\log t-(1-t)\log(1-t) \geq 2t(1-t)$ for $t\in[0,1]$ follows from convexity and a Taylor comparison around $t=\frac{1}{2}$. □

### D. Primewise modular mixing

**Hypothesis 1** (Primewise modular mixing). *Let $p$ be a prime. There exists $\delta_p \in (0, 1/2]$ such that*

$$\mathbb{P}\Big(p \mid (A+B) \,\Big|\, \mathbf{X}_{\{p\}}(A), \mathbf{X}_{\{p\}}(B)\Big) \in [\delta_p,\, 1-\delta_p] \quad \textit{almost surely.}$$

### E. Entropy production controls innovation

**Theorem IX.1** (Entropy production controls innovation)**.** *Assume Hypothesis 1 holds for all primes in a finite set $\mathcal{P}$. Then* [3]

$$\mathrm{EP}_{\mathcal{P}}(A, B) \;\geq\; \sum_{p \in \mathcal{P}} 2\delta_p(1-\delta_p) \cdot \mathbb{P}\Big(p \mid (A+B) \text{ and } p \nmid \gcd(A, B)\Big)$$

*In particular, if $\delta_p \geq \delta$ for all $p \in \mathcal{P}$, then*

$$\mathrm{EP}_{\mathcal{P}}(A, B) \;\geq\; 2\delta(1-\delta) \sum_{p \in \mathcal{P}} \mathbb{P}\Big(p \mid (A+B) \text{ and } p \nmid \gcd(A, B)\Big).$$

*Proof.* For each $p \in \mathcal{P}$, let $U_p := \mathbf{1}_{\{p \mid (A+B)\}}$. By the chain rule for conditional entropy and monotonicity under coarsening,

$$\begin{aligned}\mathrm{EP}_{\mathcal{P}}(A, B) &= H\Big((U_p)_{p \in \mathcal{P}} \,\Big|\, \mathbf{X}_{\mathcal{P}}(A), \mathbf{X}_{\mathcal{P}}(B)\Big) \\ &\geq \sum_{p \in \mathcal{P}} H\Big(U_p \,\Big|\, \mathbf{X}_{\mathcal{P}}(A), \mathbf{X}_{\mathcal{P}}(B)\Big) \end{aligned} \tag{IX.2}$$

Since conditioning on more information can only reduce entropy,

$$H\Big(U_p \,\Big|\, \mathbf{X}_{\mathcal{P}}(A), \mathbf{X}_{\mathcal{P}}(B)\Big) \geq H\Big(U_p \,\Big|\, \mathbf{X}_{\{p\}}(A), \mathbf{X}_{\{p\}}(B)\Big).$$

By Hypothesis 1, the conditional success probability of $U_p$ lies in $[\delta_p, 1-\delta_p]$. Applying Lemma IX.0.1 conditionally gives a uniform lower bound $\geq 2\delta_p(1-\delta_p)$ and summing over $p \in \mathcal{P}$ yields the claim. ☐

**Remark IX.2** (Innovation as weighted entropy production)**.** *The connection between deterministic innovation and entropy production can now be made explicit. Recalling that*

$$\mathbb{E}\,\mathcal{I}(a, b) = \sum_p (\log p)\, \mathbb{P}\Big(p \mid (x+y) \text{ and } p \nmid g\Big),$$

*while Theorem IX.1 gives*

$$\begin{aligned}&\mathrm{EP}_{\mathcal{P}}(A, B) \;\geq\; \sum_{p \in \mathcal{P}} c_p\, \mathbb{P}\Big(p \mid (A+B) \text{ and } p \nmid \gcd(A, B)\Big), \\ &c_p := 2\delta_p(1-\delta_p),\end{aligned} \tag{IX.3}$$

*we see that both quantities are driven by the same underlying indicator variables $Z_p$ encoding the appearance of new primes. From this perspective, arithmetic innovation $\mathcal{I}(a, b)$ is a* logarithmically weighted *version of entropy production, while $\mathrm{EP}_{\mathcal{P}}$ captures the unweighted information generated by the same prime events.*

[3] Note that the condition $p \nmid \gcd(A, B)$ ensures that $p$ corresponds to a genuinely new prime factor of the coprime core $x + y$ in the normalization $A = gx$, $B = gy$.

### F. Mixing from equidistribution modulo primes

Heuristically, the mixing mechanism may be visualized as follows: equidistribution of $(x, y)$ mod $p$ forces the sum $x + y$ to sweep uniformly across residue classes modulo $p$. This "scrambling" effect breaks any attempt to predict divisibility of $A + B$ from the multiplicative data of $A$ and $B$, and is the fundamental source of entropy production at each prime.

We now show that Hypothesis 1 follows from standard equidistribution properties and can therefore be verified in natural arithmetic models. Let $(x, y)$ be chosen uniformly from $[1, N]^2$ conditioned on $\gcd(x, y) = 1$. For a fixed prime $p$, classical results imply that the residues ($x$ mod $p$, $y$ mod $p$) are asymptotically uniform over $(\mathbb{Z}/p\mathbb{Z})^2$ minus the diagonal $p \mid x, y$.

**Lemma IX.2.1** (Uniformity mod $p$)**.** *For fixed $p$ and $N \to \infty$,*

$$\mathbb{P}(x + y \equiv 0 \pmod{p} \mid p \nmid x,\ p \nmid y) = \frac{1}{p-1} + o(1).$$

*Proof.* Conditioned on $p \nmid x$ and $p \nmid y$, the residues $x, y$ are asymptotically independent and uniform in $(\mathbb{Z}/p\mathbb{Z})^\times$. Exactly one choice of $y$ modulo $p$ satisfies $x + y \equiv 0$ for each $x$, giving probability $1/(p-1)$. ☐

**Proposition IX.3** (Verified modular mixing)**.** *For $(x, y)$ chosen as above, Hypothesis 1 holds with*

$$\delta_p = \frac{1}{2(p-1)}$$

*for all sufficiently large $N$.*

*Proof.* By Lemma IX.2.1, the conditional probability that $p \mid (x + y)$ is $1/(p-1) + o(1)$. For $p \geq 3$ this lies strictly between 0 and 1, hence inside $[\delta_p, 1-\delta_p]$ for the stated choice when $N$ is large. ☐

**Remark IX.4** (The prime $p = 2$ and parity effects)**.** *For $p = 2$, the mixing behavior depends on the parity classes of $A$ and $B$. If $A$ and $B$ are both odd, then $A + B$ is even with probability one and the conditional entropy at $p = 2$ vanishes. More generally, the equidistribution and mixing phenomena described above are generic for odd primes $p \geq 3$, which are the primary focus throughout the paper. Accordingly, in entropy and innovation estimates we may either exclude $p = 2$ or treat it separately without affecting asymptotic conclusions.*

### G. Entropy growth across primes

**Theorem IX.5** (A logarithmic lower bound for global entropy production under mixing)**.** *Let $\mathcal{P}$ be the set of primes $p \leq P$. Then for $(x, y)$ chosen as above,*

$$\mathrm{EP}_{\mathcal{P}}(A, B) \;\gg\; \sum_{p \leq P} \frac{1}{p} \;\sim\; \log\log P.$$

*Proof.* For each $p \leq P$, Proposition IX.3 gives $\delta_p \asymp 1/p$. Applying Theorem IX.1 and summing over $p \leq P$ yields the claim. ☐

Thus addition produces entropy at a logarithmic rate across prime scales, while multiplication does not, providing an information-theoretic expression of the additive–multiplicative incompatibility.

**Remark IX.6.** *The global entropy production $\mathrm{EP}_{\mathcal{P}}(A, B)$ is sensitive to correlations across primes, whereas $\mathrm{PEP}(A, B; P)$ is a primewise, additive proxy designed to be tractable. In general one has*

$$\mathrm{EP}_{\mathcal{P}}(A, B) \;\leq\; \mathrm{PEP}(A, B; P).$$

**Remark IX.7** (Information-theoretic form of arithmetic uncertainty). *This section provides an information-theoretic validation of the arithmetic uncertainty principle of Section V. If additive and multiplicative data could be simultaneously localized, then $\mathbf{X}_{\mathcal{P}}(A+B)$ would be almost deterministic given $\mathbf{X}_{\mathcal{P}}(A)$ and $\mathbf{X}_{\mathcal{P}}(B)$, forcing entropy production to vanish. The lower bound $\mathrm{EP}_{\mathcal{P}}(A,B) \gg \log\log P$ shows that such joint localization is not only impossible but quantitatively obstructed by inevitable information generation across prime scales.*

Entropy production, additive innovation, and modular equidistribution thus appear as three complementary manifestations of the same underlying arithmetic mixing phenomenon.

## X. UNCONDITIONAL DENSITY THEOREMS FOR THE RADICAL

The preceding sections developed two complementary viewpoints on additive innovation: a structural description of innovation collapse in terms of smoothness and $S$-unit phenomena (Section VI), and a probabilistic–information-theoretic perspective based on entropy production and modular mixing (Sections VIII–IX). In this section, we pass from mechanism to consequence. We establish *unconditional* density results for the radical of a typical sum, showing that abnormally small radical growth is extremely rare. These theorems provide the first fully quantitative incarnation of the abc philosophy in this work: while uniform bounds remain conjectural, abc-type inequalities hold with strong power-saving outside a negligible exceptional set.

The results of this section serve as the deterministic backbone for the capacity and prime-scale forcing mechanisms developed subsequently.

### A. Unconditional abc-type statements in density for the radical

Our first theorem is an unconditional abc-type statement in density for the radical of a typical sum. The second theorem quantifies, in an information-theoretic sense, the entropy generated by addition across prime scales via a primewise entropy functional.

We write $\mathrm{rad}(n) := \prod_{p|n} p$ for the radical of $n$.

**Lemma X.0.1** (Squarefull extraction from small radical). *Let $n \geq 2$ and let $\varepsilon \in (0,1)$. If $\mathrm{rad}(n) \leq n^{1-\varepsilon}$, then $n$ is divisible by a squarefull integer $d \geq n^{\varepsilon}$.*

*Proof.* Write $n = \prod_{p^k \| n} p^k$. Then

$$\frac{n}{\mathrm{rad}(n)} = \prod_{p^k \| n} p^{k-1},$$

which is squarefull [4]. If $\mathrm{rad}(n) \leq n^{1-\varepsilon}$ then $n/\mathrm{rad}(n) \geq n^{\varepsilon}$. □

**Theorem X.1** (Large radical for most sums). *Fix $\varepsilon \in (0,1)$. Then the number of pairs $(x,y) \in [1,N]^2$ such that*

$$\mathrm{rad}(x+y) \;\leq\; (x+y)^{1-\varepsilon}$$

*is $O_\varepsilon(N^{2-\varepsilon/2})$. In particular, for a uniformly random pair $(x,y) \in [1,N]^2$,*

$$\mathbb{P}\Big(\mathrm{rad}(x+y) \leq (x+y)^{1-\varepsilon}\Big) = O_\varepsilon(N^{-\varepsilon/2}).$$

[4] We recall that an integer is squarefull if every prime divisor appears with exponent at least 2 in $n$, equivalently with exponent at least 1 in $n/\mathrm{rad}(n)$

*Proof.* Let $n := x+y$. For each $n \in [2,2N]$, the number of representations $n = x+y$ with $1 \leq x,y \leq N$ is at most $N$. Hence it suffices to bound

$$\#\{n \leq 2N : \ \mathrm{rad}(n) \leq n^{1-\varepsilon}\}.$$

Observe that

$$\frac{n}{\mathrm{rad}(n)} = \prod_{p^k \| n} p^{k-1}.$$

If $\mathrm{rad}(n) \leq n^{1-\varepsilon}$, then

$$\frac{n}{\mathrm{rad}(n)} \geq n^{\varepsilon} \geq (2N)^{\varepsilon}/2^{\varepsilon}$$

Set $Y := N^{\varepsilon}/2^{\varepsilon}$. Then it suffices to bound

$$\#\{n \leq 2N : \ n/\mathrm{rad}(n) \geq Y\}$$

By Lemma X.0.1, $n/\mathrm{rad}(n)$ is a *squarefull* integer (every prime divisor appears with exponent at least 1, coming from $k-1 \geq 1$ whenever $k \geq 2$). In particular, if $n/\mathrm{rad}(n) \geq Y$, then $n$ is divisible by some squarefull integer $d \geq Y$ (namely $d = n/\mathrm{rad}(n)$). Therefore

$$\begin{aligned}&\#\{n \leq 2N : \ n/\mathrm{rad}(n) \geq Y\} \\ &\leq \sum_{\substack{d \geq Y \\ d \text{ squarefull}}} \left\lfloor \frac{2N}{d} \right\rfloor \;\leq\; 2N \sum_{\substack{d \geq Y \\ d \text{ squarefull}}} \frac{1}{d}\end{aligned} \tag{X.1}$$

It remains to bound the tail sum over squarefull integers. A standard parameterization is: $d$ is squarefull iff $d = m^2 s$ where $s$ is squarefree and $s \mid m$. Hence

$$\begin{aligned}\sum_{\substack{d \geq Y \\ d \text{ squarefull}}} \frac{1}{d} &\leq \sum_{m \geq \sqrt{Y}} \sum_{s|m} \frac{1}{m^2 s} \\ &\leq \sum_{m \geq \sqrt{Y}} \frac{\tau(m)}{m^2} \ll \int_{\sqrt{Y}}^{\infty} \frac{\log t}{t^2}\, dt \ll \frac{\log Y}{\sqrt{Y}},\end{aligned} \tag{X.2}$$

where $\tau$ is the divisor function and we used the classical bound $\sum_{m \geq M} \tau(m)/m^2 \ll (\log M)/M$. Therefore,

$$\begin{aligned}\#\{n \leq 2N : \ \mathrm{rad}(n) \leq n^{1-\varepsilon}\} &\ll 2N \cdot \frac{\log Y}{\sqrt{Y}} \\ \ll_\varepsilon N \cdot \frac{\log N}{N^{\varepsilon/2}} &= O_\varepsilon(N^{1-\varepsilon/2} \log N)\end{aligned} \tag{X.3}$$

Multiplying by the at most $N$ representations of each $n$ as $x+y$ yields

$$\begin{aligned}&\#\{(x,y) \in [1,N]^2 : \ \mathrm{rad}(x+y) \leq (x+y)^{1-\varepsilon}\} \\ &= O_\varepsilon(N^{2-\varepsilon/2} \log N).\end{aligned} \tag{X.4}$$

Absorbing the $\log N$ into the implied constant (or keeping it explicitly) proves the claim. □

### B. Primewise entropy production

For a prime $p$, let

$$X_p := \mathbf{1}_{\{p|A\}}, \qquad Y_p := \mathbf{1}_{\{p|B\}}, \qquad U_p := \mathbf{1}_{\{p|(A+B)\}}.$$

Define the *primewise entropy production up to $P$* by

$$\mathrm{PEP}(A,B;P) := \sum_{p \leq P} H\Big(U_p \mid X_p, Y_p\Big),$$

where $H(\cdot)$ is Shannon entropy (natural logarithm).

**Theorem X.2** (Primewise entropy production is logarithmic)**.** *Let $A, B$ be independent and uniformly distributed on $\{1, 2, \dots, N\}$. Fix $P \le N^{1/3}$. Then*

$$\mathbb{E}\,\mathrm{PEP}(A,B;P) \;=\; (1+o(1)) \sum_{p\le P} \left(1-\frac{1}{p}\right)^2 h\left(\frac{1}{p-1}\right),$$

*where $h(t) := -t\log t - (1-t)\log(1-t)$ is the binary entropy function. In particular,*

$$\mathbb{E}\,\mathrm{PEP}(A,B;P) \;\gg\; \sum_{p\le P} \frac{\log p}{p} \;\sim\; \log P.$$

*Proof.* Fix a prime $p \le P$. We analyze $H(U_p \mid X_p, Y_p)$ by conditioning on the three cases for $(X_p, Y_p)$.

**Case 1: $X_p = Y_p = 1$.** Then $p \mid A$ and $p \mid B$, hence $p \mid (A+B)$ deterministically, so $U_p = 1$ and $H(U_p \mid X_p = 1, Y_p = 1) = 0$.

**Case 2: exactly one of $X_p, Y_p$ equals 1.** Say $X_p = 1$ and $Y_p = 0$. Then $A \equiv 0 \pmod p$ and $B \not\equiv 0 \pmod p$, so $A + B \equiv B \not\equiv 0 \pmod p$. Thus $U_p = 0$ deterministically and $H(U_p \mid X_p = 1, Y_p = 0) = 0$ (and similarly for $X_p = 0, Y_p = 1$).

**Case 3: $X_p = Y_p = 0$.** Then $A$ and $B$ are not divisible by $p$. In this case, modulo $p$, the pair $(A \bmod p, B \bmod p)$ is asymptotically uniform on $(\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/p\mathbb{Z})^\times$, with error $O(1/N)$ coming from the boundary of the interval $\{1, \dots, N\}$. Therefore,

$$\begin{aligned}&\mathbb{P}(U_p = 1 \mid X_p = 0, Y_p = 0)\\ &= \mathbb{P}(A + B \equiv 0 \pmod p \mid p \nmid A,\ p \nmid B)\\ &= \frac{1}{p-1} + O\left(\frac{1}{N}\right)\end{aligned} \tag{X.5}$$

Hence,

$$H(U_p \mid X_p = 0, Y_p = 0) = h\left(\frac{1}{p-1}\right) + O\left(\frac{\log p}{N}\right).$$

Now combine the cases using the law of total expectation:

$$\mathbb{E}\,H(U_p \mid X_p, Y_p) = \mathbb{P}(X_p = 0, Y_p = 0)\, H(U_p \mid X_p = 0, Y_p = 0),$$

since all other cases contribute 0. Moreover,

$$\mathbb{P}(X_p = 0, Y_p = 0) = \left(1 - \frac{1}{p}\right)^2 + O\left(\frac{1}{N}\right).$$

Thus

$$\mathbb{E}\,H(U_p \mid X_p, Y_p) = \left(1 - \frac{1}{p}\right)^2 h\left(\frac{1}{p-1}\right) + O\left(\frac{\log p}{N}\right).$$

Summing over $p \le P$ gives the first formula, and the error is

$$\sum_{p\le P} O\left(\frac{\log p}{N}\right) = O\left(\frac{P\log P}{N}\right) = o(1) \quad \text{since } P \le N^{1/3}.$$

Finally, for small $t$ one has $h(t) \asymp t\log(1/t)$. With $t = \frac{1}{p-1} \sim \frac{1}{p}$, we obtain

$$h\left(\frac{1}{p-1}\right) \asymp \frac{\log p}{p}.$$

Therefore,

$$\mathbb{E}\,\mathrm{PEP}(A,B;P) = \sum_{p\le P} \mathbb{E}\,H(U_p \mid X_p, Y_p) \gg \sum_{p\le P} \frac{\log p}{p} \sim \log P,$$

as claimed. □

## C. Consequences: innovation frequency and prime scales (C1–C4)

**Corollary X.3** (C1: innovation occurs for almost all coprime pairs)**.** *Fix $\varepsilon \in (0,1)$. For all but $O_\varepsilon(N^{2-\varepsilon/2})$ coprime pairs $(x,y) \in [1,N]^2$, one has*

$$\mathcal{I}(x,y) \;\ge\; (1-\varepsilon)\log(x+y).$$

*Proof.* By Theorem X.1, for all but $O_\varepsilon(N^{2-\varepsilon/2})$ pairs $(x,y) \in [1,N]^2$ we have

$$\log \mathrm{rad}(x+y) \ge (1-\varepsilon)\log(x+y).$$

If $\gcd(x,y) = 1$, strict additive orthogonality gives $\mathcal{I}(x,y) = \log\mathrm{rad}(x+y)$, hence the claim for all but $O_\varepsilon(N^{2-\varepsilon/2})$ coprime pairs. □

**Corollary X.4** (C2: many innovative pairs force many prime-divisibility events)**.** *Let $A \subset \{1, \dots, N\}$ and let $M$ be the number of pairs $(a,b) \in A^2$ for which $a+b$ has a prime divisor $p \notin \Sigma(A)$. Then for any $P \ge 2$,*

$$\begin{aligned}M \;\le\; &\sum_{p\le P} \#\{(a,b)\in A^2 :\ p \mid (a+b)\}\\ &+\ \#\{(a,b)\in A^2 :\ \exists p > P,\ p \mid (a+b)\}\end{aligned} \tag{X.6}$$

*In particular, if all innovative primes were $\le P$, then*

$$M \;\le\; \sum_{p\le P} \#\{(a,b)\in A^2 :\ p \mid (a+b)\}.$$

*Proof.* Each innovative pair contributes to the sum corresponding to (at least) one prime divisor $p$ of $a+b$. Split according to whether $p \le P$ or $p > P$. □

**Corollary X.5** (C3: capacity bound up to $P$)**.** *For any $A \subset \{1, \dots, N\}$ and $P \ge 2$,*

$$\sum_{p\le P} \#\{(a,b)\in A^2 :\ p \mid (a+b)\} \;\ll\; |A|^2 \log\log P + |A|\frac{P}{\log P}.$$

*Proof.* This is Lemma XI.0.2. □

**Corollary X.6** (C4: large-prime forcing (a PNT-style "prime scale" conclusion))**.** *Let $A \subset \{1, \dots, N\}$ and suppose that at least $M \ge c|A|^2$ pairs $(a,b) \in A^2$ are innovative in the sense that $a+b$ has a prime divisor $p \notin \Sigma(A)$. Let $P \ge 3$. If*

$$c|A|^2 \;>\; C\left(|A|^2 \log\log P \;+\; |A|\frac{P}{\log P}\right)$$

*for a sufficiently large absolute constant $C$, then there exists an innovative pair $(a,b)$ for which $a+b$ has a* new *prime divisor $p > P$.*

*Proof.* If all innovative primes were $\le P$, then by Corollary X.4 and Corollary X.5,

$$M \;\ll\; |A|^2 \log\log P \;+\; |A|\frac{P}{\log P},$$

contradicting the hypothesis when $C$ is large enough. □

# XI. CAPACITY BOUNDS AND PRIME-SCALE FORCING

This section bridges the gap between *frequency* of additive innovation and *prime-scale* consequences. The typical laws of Section VII say that innovation is overwhelmingly common for random pairs. Here we give a deterministic *capacity* mechanism showing that primes $\le P$ can certify only a limited number of additive divisibility constraints of the form $p \mid (a+b)$ inside $A^2$. Consequently, if a set $A \subset \{1, \dots, N\}$ produces many *innovative* sums, then at least one innovative sum must introduce a genuinely *new* prime scale $p > P$.

## A. Prime spectra and innovative sums

For a finite set $A \subset \mathbb{Z}_{>0}$ define its prime spectrum union

$$\Sigma(A) := \bigcup_{a\in A} S(a).$$

A pair $(a,b) \in A^2$ is called *innovative* (relative to $A$) if $a+b$ has a prime divisor $p \notin \Sigma(A)$. Let $M$ denote the number of innovative pairs in $A^2$.

## B. Capacity of a single prime

**Lemma XI.0.1** (Divisibility capacity of a fixed prime)**.** *Let $A \subset \{1,2,\ldots,N\}$ be a finite set and let $p$ be a prime. Then*

$$\#\{(a,b) \in A^2 : \ p \mid (a+b)\} \ \le \ \frac{|A|^2}{p} \ + \ |A|.$$

*Proof.* Partition $A$ into residue classes modulo $p$: let $A_r = \{a \in A : \ a \equiv r \pmod p\}$. Then

$$\#\{(a,b) \in A^2 : \ p \mid (a+b)\} = \sum_{r \pmod p} |A_r| \cdot |A_{-r}|.$$

Using $xy \le \frac{x^2+y^2}{2}$ we have

$$\sum_r |A_r||A_{-r}| \le \frac{1}{2}\sum_r \big(|A_r|^2 + |A_{-r}|^2\big) = \sum_r |A_r|^2.$$

Write $|A_r| = |A|/p + u_r$ with $\sum_r u_r = 0$. Then

$$\sum_r |A_r|^2 = \sum_r \left(\frac{|A|}{p} + u_r\right)^2 = \frac{|A|^2}{p} + \sum_r u_r^2.$$

Moreover, since $u_r = |A_r| - |A|/p$, we have $u_r^2 \le |A_r|^2$ and hence $\sum_r u_r^2 \le \sum_r |A_r|^2$. Combining this with the identity above gives

$$\sum_r |A_r|^2 \le \frac{|A|^2}{p} + \sum_r |A_r|^2,$$

and in particular the standard inequality

$$\sum_r |A_r|^2 \le \frac{\left(\sum_r |A_r|\right)^2}{p} + \max_r |A_r| = \frac{|A|^2}{p} + \max_r |A_r| \le \frac{|A|^2}{p} + |A| \qquad \text{(XI.1)}$$

This yields the claimed bound. □

## C. Total capacity up to a cutoff $P$

**Lemma XI.0.2** (Total additive capacity up to $P$)**.** *Let $A \subset \{1,2,\ldots,N\}$ and let $P \ge 2$. Then*

$$\sum_{p\le P} \#\{(a,b) \in A^2 : \ p \mid (a+b)\} \ \ll \ |A|^2 \sum_{p\le P} \frac{1}{p} \ + \ |A|\pi(P).$$

*In particular,*

$$\sum_{p\le P} \#\{(a,b) \in A^2 : \ p \mid (a+b)\} \ \ll \ |A|^2 \log\log P + |A|\frac{P}{\log P}.$$

*Proof.* Sum Lemma XI.0.1 over primes $p \le P$ to obtain

$$\sum_{p\le P} \#\{(a,b) \in A^2 : \ p \mid (a+b)\} \le |A|^2 \sum_{p\le P} \frac{1}{p} + |A|\pi(P).$$

The first term is $\ll |A|^2 \log\log P$ by Mertens' theorem, and the second is $\ll |A|\, P/\log P$ by the standard bound $\pi(P) \ll P/\log P$. □

## D. Double counting and large-prime forcing

**Lemma XI.0.3** (Double counting of innovative pairs)**.** *Let $A \subset \{1,\ldots,N\}$ and let $M$ be the number of innovative pairs $(a,b) \in A^2$. Then for any $P \ge 2$,*

$$M \ \le \ \sum_{p\le P} \#\{(a,b) \in A^2 : \ p \mid (a+b)\} + \ \#\{(a,b) \in A^2 : \ \exists p > P, \ p \mid (a+b)\} \qquad \text{(XI.2)}$$

*In particular, if every innovative pair has all* new *prime divisors $\le P$, then*

$$M \ \le \ \sum_{p\le P} \#\{(a,b) \in A^2 : \ p \mid (a+b)\}.$$

*Proof.* Each innovative pair $(a,b)$ admits at least one prime $p \notin \Sigma(A)$ dividing $a+b$. Split according to whether such a prime divisor satisfies $p \le P$ or $p > P$, and sum. □

**Theorem XI.1** (Capacity obstruction and forced large primes)**.** *Let $A \subset \{1,\ldots,N\}$ and suppose that at least*

$$M \ \ge \ c\,|A|^2$$

*pairs $(a,b) \in A^2$ are innovative. Fix $P \ge 3$. If*

$$c\,|A|^2 \ > \ C\Big(|A|^2 \log\log P \ + \ |A|\frac{P}{\log P}\Big)$$

*for a sufficiently large absolute constant $C$, then there exists an innovative pair $(a,b)$ for which $a+b$ has a prime divisor $p > P$.*

*Proof.* Assume by contradiction that no innovative sum introduces a prime divisor $> P$. Then Lemma XI.0.3 gives

$$M \ \le \ \sum_{p\le P} \#\{(a,b) \in A^2 : \ p \mid (a+b)\}.$$

Applying Lemma XI.0.2 yields

$$M \ \ll \ |A|^2 \log\log P \ + \ |A|\frac{P}{\log P},$$

contradicting the hypothesis when $C$ is chosen sufficiently large. □

## E. Polynomial-scale spectrum growth and sum–product interpretation

**Corollary XI.2** (Forced spectrum growth at a polynomial scale)**.** *Let $A \subset \{1,\ldots,N\}$ and suppose that a positive proportion of pairs $(a,b) \in A^2$ are innovative. Then there exists $\beta > 0$ and a prime $p$ such that*

$$p \ \gg \ N^{\beta} \qquad \text{and} \qquad p \in \Sigma(A+A) \setminus \Sigma(A).$$

*Proof.* Apply Theorem XI.1 with $P := N^\beta$. For $\beta > 0$ sufficiently small (depending only on the innovation proportion),

$$|A|^2 \log\log P + |A| \frac{P}{\log P} = o(|A|^2),$$

so the hypothesis holds and yields a new prime divisor $p > P = N^\beta$ of some sum $a+b$. Such $p$ lies in $\Sigma(A+A) \setminus \Sigma(A)$. □

**Corollary XI.3** (Additive–multiplicative incompatibility via capacity)**.** *Let $A \subset \{1,\dots,N\}$ be such that $\Sigma(A)$ is contained in primes $\le N^\beta$ for some $\beta < 1$. Then either:*

*(i) $A+A$ contains an integer with a prime divisor $> N^\beta$, or*

*(ii) the proportion of innovative pairs in $A^2$ is $o(1)$.*

*Proof.* If a positive proportion of pairs is innovative, Corollary XI.2 produces a new prime divisor in $\Sigma(A+A)\setminus\Sigma(A)$ at a polynomial scale, hence in particular exceeding $N^\beta$ for the stated $\beta$. Otherwise the proportion of innovative pairs is $o(1)$. □

> A set with nontrivial multiplicative rigidity cannot be additively rich without forcing the appearance of new prime scales.

**Remark XI.4** (Connection with typical innovation)**.** *In many natural situations, the hypothesis "a positive proportion of pairs is innovative" is guaranteed by the typical innovation results of Section VII. For instance, if $A$ contains many coprime pairs and its spectrum $\Sigma(A)$ avoids primes above $N^\beta$, then a positive proportion of sums $a+b$ acquire new primes $> N^\beta$, placing $A$ in the forcing regime of Theorem XI.1.*

**Proposition XI.5** (Coprime rigidity)**.** *Let $A \subset \mathbb{Z}_{>0}$ be a finite set such that $\gcd(a,b) = 1$ for all distinct $a,b \in A$. Then for any $a \ne b$ in $A$,*

$$\mathcal{I}(a,b) = \log \operatorname{rad}(a+b).$$

*In particular, additive innovation cannot vanish uniformly on $A$.*

*Proof.* If $\gcd(a,b)=1$, strict additive orthogonality implies $S(a+b) \cap (S(a)\cup S(b)) = \varnothing$, hence $\mathcal{I}(a,b) = \log\operatorname{rad}(a+b)$. □

**abc philosophy revisited.** In the language of additive innovation, abc-type inequalities express the rarity of configurations where $a+b$ is supported predominantly on primes already present in $a$ and $b$. The capacity bound above shows that *many* such additive relations cannot be certified using only primes $\le P$: if failures of innovation occur with positive density, then some sums must force genuinely new prime scales. Combined with the unconditional density laws for $\operatorname{rad}(\cdot)$, this yields a robust "abc-in-density" mechanism driven by prime-scale forcing.

# XII. A SZPIRO-STYLE VIEW: ABC-TYPE BOUNDS IN DENSITY AND THE MECHANISM OF FAILURE

## A. The abc philosophy and our viewpoint

The *abc conjecture* may be stated in the standard form: for every $\varepsilon > 0$ there exists $K_\varepsilon > 0$ such that for all coprime positive integers $a,b,c$ with $a+b=c$,

$$c \;\le\; K_\varepsilon \operatorname{rad}(abc)^{1+\varepsilon}. \tag{XII.1}$$

Equivalently,

$$\log c \;\le\; (1+\varepsilon)\log\operatorname{rad}(abc) + O_\varepsilon(1). \tag{XII.2}$$

Our framework isolates the *structural obstruction* to such inequalities: writing $a = gx$, $b = gy$ with $g = \gcd(a,b)$ and $\gcd(x,y) = 1$, one has $c = g(x+y)$ and complete innovation failure corresponds to the spectral containment $S(x+y) \subseteq S(g)$ (Theorem VI.11), i.e. an $S$-unit/smoothness phenomenon. Thus, the abc philosophy becomes the statement that *severe innovation deficit is exceptionally rare.*

In this section we record unconditional results which are *Szpiro-style*: they bound the height $\log c$ in terms of a conductor-like quantity $\log\operatorname{rad}(abc)$, not uniformly for all triples (conjectural), but for *almost all* additive relations in a strong density sense. We proceed in three steps:

- Firstly, we record a normalization identity relating $\operatorname{rad}(a+b)$ to innovation and the gcd (Lemma XII.0.1);
- Next, we invoke the density radical-growth theorem (Theorem XII.1) and translate it into a typical-maximal innovation statement in the coprime regime; and
- Finally, we formulate a Szpiro-style height–conductor inequality in density and isolate explicit obstruction terms beyond the coprime case.

## B. A normalization identity: radical vs. innovation vs. gcd

**Lemma XII.0.1** (Innovation decomposition)**.** *Let $a,b \in \mathbb{Z}_{>0}$ and write $a = gx$, $b = gy$ with $g = \gcd(a,b)$ and $\gcd(x,y)=1$. Then*

$$\log\operatorname{rad}(a+b) = \log\operatorname{rad}(g) + \log\operatorname{rad}(x+y) - \log\operatorname{rad}(\gcd(g,x+y)).$$

*Moreover, since $\gcd(x,y) = 1$ implies $S(x+y)\cap(S(x)\cup S(y)) = \varnothing$, we have*

$$\log\operatorname{rad}(x+y) = \mathcal{I}(a,b) + D(a,b),$$
$$D(a,b) := \sum_{\substack{p \mid (x+y) \\ p \mid g}} \log p = \log\operatorname{rad}(\gcd(g,x+y)). \tag{XII.3}$$

*In particular, the only obstruction to $\mathcal{I}(a,b) \approx \log\operatorname{rad}(x+y)$ is the overlap $S(x+y)\cap S(g)$.*

*Proof.* Since $a+b = g(x+y)$, we have

$$S(a+b) = S(g)\cup S(x+y), \qquad \operatorname{rad}(a+b) = \frac{\operatorname{rad}(g)\operatorname{rad}(x+y)}{\operatorname{rad}(\gcd(g,x+y))},$$

which gives the first identity after taking logs. If $\gcd(x,y) = 1$ and $p \mid x$ and $p \mid (x+y)$, then $p \mid y$, a contradiction; similarly for $y$. Hence $S(x+y)\cap(S(x)\cup S(y)) = \varnothing$ and the prime factors of $x+y$ split into those dividing $g$ and those not, yielding $\log\operatorname{rad}(x+y) = \mathcal{I}(a,b) + D(a,b)$. □

## C. abc in density for sums

**Theorem XII.1** (abc in density for sums (radical growth))**.** *Fix $\varepsilon \in (0,1)$. Then the number of pairs $(a,b) \in [1,N]^2$ such that, writing $c = a+b$,*

$$\operatorname{rad}(c) \;\le\; c^{1-\varepsilon}$$

*is $O_\varepsilon(N^{2-\varepsilon/2})$.*

*Proof.* This is exactly Theorem X.1. □

## D. Translation into additive innovation (coprime regime)

Recall that for coprime $x, y$ one has strict additive orthogonality:

$$S(x+y) \cap (S(x) \cup S(y)) = \varnothing.$$

Hence, in this case,

$$\mathcal{I}(x,y) = \log \operatorname{rad}(x+y).$$

**Corollary XII.2** (Innovation is typically maximal)**.** *Fix $\varepsilon \in (0,1)$. Then, for all but $O_\varepsilon(N^{2-\varepsilon/2})$ coprime pairs $(x,y) \in [1,N]^2$,*

$$\mathcal{I}(x,y) \;\ge\; (1-\varepsilon)\log(x+y).$$

*Proof.* By Theorem X.1, for all but $O_\varepsilon(N^{2-\varepsilon/2})$ pairs $(x,y) \in [1,N]^2$ we have

$$\log \operatorname{rad}(x+y) \ge (1-\varepsilon)\log(x+y).$$

When $\gcd(x,y) = 1$, strict additive orthogonality gives $\mathcal{I}(x,y) = \log \operatorname{rad}(x+y)$, proving the claim. □

## E. Szpiro dictionary and a height–conductor inequality in density

We interpret the abc/Szpiro philosophy through the standard heuristic dictionary:

- **Height** $\leftrightarrow$ size, modeled here by $\log c$.
- **Conductor** $\leftrightarrow$ product of primes of bad reduction, modeled here by $\operatorname{rad}(abc)$.
- **Exceptional families** $\leftrightarrow$ bounded prime support ($S$-unit/smoothness phenomena), modeled here by large overlap $S(x+y) \cap S(g)$.

**Lemma XII.2.1** (Radical domination for coprime sums)**.** *If $\gcd(a,b) = 1$ and $c = a+b$, then*

$$\operatorname{rad}(abc) \;\ge\; \operatorname{rad}(c) = \operatorname{rad}(a+b).$$

*Proof.* The radical $\operatorname{rad}(abc) = \operatorname{rad}(ab(a+b))$ contains every prime dividing $a+b$, hence $\operatorname{rad}(abc) \ge \operatorname{rad}(a+b)$. □

**Theorem XII.3** (abc in density: Szpiro-style height–conductor inequality)**.** *Fix $\eta \in (0,1)$. Then the number of coprime pairs $(a,b) \in [1,N]^2$ such that, writing $c = a+b$,*

$$\log c \;>\; \frac{1}{1-\eta}\log \operatorname{rad}(abc)$$

*is $O_\eta(N^{2-\eta/2})$. Equivalently, for a uniformly random coprime pair $(a,b) \in [1,N]^2$,*

$$\mathbb{P}\Big(\log(a+b) \;>\; \frac{1}{1-\eta}\log \operatorname{rad}(ab(a+b))\Big) = O_\eta(N^{-\eta/2}).$$

*Proof.* By Lemma XII.2.1, it suffices to bound $\operatorname{rad}(a+b) < (a+b)^{1-\eta}$, which is exactly Theorem X.1. Restricting to $\gcd(a,b) = 1$ can only decrease the count. □

## F. A $(1+\varepsilon)$-form in density

**Corollary XII.4** (abc in density in $(1+\varepsilon)$ form)**.** *Fix $\varepsilon \in (0,1)$. Then for all but $O_\varepsilon(N^{2-\varepsilon/4})$ coprime pairs $(a,b) \in [1,N]^2$ (with $c = a+b$) one has*

$$\log c \;\le\; (1+\varepsilon)\log \operatorname{rad}(abc).$$

*Proof.* Take $\eta := \varepsilon/2$. For $\varepsilon \in (0,1)$ one has $\frac{1}{1-\eta} \le 1+\varepsilon$. Apply Theorem XII.3. The exceptional set is $O_\varepsilon(N^{2-\varepsilon/4})$. □

## G. Mechanism of failure in density: large squarefull divisors

**Corollary XII.5** (Mechanism of abc failure in density: large squarefull divisors)**.** *Fix $\eta \in (0,1)$. For all but $O_\eta(N^{2-\eta/2})$ pairs $(a,b) \in [1,N]^2$ (with $c = a+b$), the integer $c$ is* not *divisible by any squarefull integer $d \ge N^\eta$.*

*Proof.* If $\operatorname{rad}(c) \le c^{1-\eta}$, then by Lemma X.0.1 the integer $c$ is divisible by a squarefull $d \ge c^\eta \ge N^\eta$. The number of $c \le 2N$ divisible by some squarefull $d \ge N^\eta$ is $O_\eta(N^{1-\eta/2}\log N)$, and each $c$ has at most $N$ representations as $c = a+b$ with $a, b \le N$. Hence the number of exceptional pairs is $O_\eta(N^{2-\eta/2}\log N)$. □

## H. General gcd: a Szpiro-type inequality with explicit obstruction

**Proposition XII.6** (Height bound with explicit obstruction terms)**.** *Let $a, b \in \mathbb{Z}_{>0}$, set $c = a+b$, and write $a = gx$, $b = gy$ with $g = \gcd(a,b)$ and $\gcd(x,y) = 1$. Then*

$$\log c \;\le\; \log g \;+\; \mathcal{I}(a,b) \;+\; D(a,b) \;+\; \Delta(x+y),$$

*where*

$$D(a,b) := \sum_{\substack{p \mid (x+y) \\ p \mid g}} \log p = \log \mathit{rad}(\gcd(g, x+y)),$$

$$\Delta(x+y) := \log\left(\frac{x+y}{\mathit{rad}(x+y)}\right) \tag{XII.4}$$

*Proof.* Since $c = g(x+y)$, we have $\log c = \log g + \log(x+y)$ and

$$\log(x+y) = \log \operatorname{rad}(x+y) + \log\left(\frac{x+y}{\operatorname{rad}(x+y)}\right)$$
$$= \log \operatorname{rad}(x+y) + \Delta(x+y). \tag{XII.5}$$

By Lemma XII.0.1, $\log \operatorname{rad}(x+y) = \mathcal{I}(a,b) + D(a,b)$. □

**Remark XII.7** (Structural vs. arithmetic obstruction)**.** *The term $D(a,b)$ measures* structural *obstruction: pre-existing prime support in $g$ capturing primes of $x+y$ (an $S$-unit/smoothness phenomenon). The term $\Delta(x+y)$ measures* arithmetic *obstruction: atypically small radical growth of the coprime sum $x+y$ (equivalently, the presence of large squarefull divisors). Our density theorem controls $\Delta(x+y)$ for almost all pairs, leaving $D(a,b)$ as the genuine remaining obstruction beyond the coprime regime.*

The results above support a guiding principle: addition generically forces new prime information, while the rare failures are organized by explicit obstruction terms (overlap with the gcd and squarefull defect). In a companion work, we argue that the same structural dichotomy—multiplication as conservative extension and additive accumulation as generative mixing—admits a geometric/variational reading in which physical law emerges from the operational cycle *multiplication → integration → variation.*

# XIII. CONCLUSION AND OUTLOOK

Throughout this work, we have developed a structural theory of *additive innovation* in the arithmetic of integers, centered on the interaction between addition, multiplication, and prime spectra. We have made precise the idea that addition generically destroys multiplicative rigidity, and the results obtained show explicitly that additive–multiplicative compatibility is highly constrained, and that the generic behavior of addition is to generate new prime information across scales.

Several estimates in this paper rely on classical bounds for smooth numbers, equidistribution modulo primes, and divisor sums. These inputs can be quantitatively strengthened under stronger analytic hypotheses. For instance, assuming the Riemann Hypothesis or the Bombieri–Vinogradov theorem, one obtains sharper equidistribution in arithmetic progressions. This leads to improved constants in modular mixing estimates, tighter control of smoothness tails, and enhanced bounds for primewise entropy production. Such refinements extend the forcing results to larger prime cutoffs and improve the quantitative form of innovation-density statements.

Importantly, these analytic assumptions are not required for the qualitative principle established here: *addition generically produces new prime information*, and failures are confined to negligible exceptional sets governed by explicit arithmetic obstructions.

The additive innovation framework is closely aligned with the philosophy of the abc conjecture and related height–conductor inequalities. From this perspective, innovation measures the arithmetic cost of addition in terms of new prime data, while innovation collapse corresponds to rare configurations with abnormally slow radical growth. Moreover, our density results show that abc-type inequalities hold for almost all additive relations, with failures controlled by smoothness and squarefull phenomena. This viewpoint is consistent with the height-growth mechanisms underlying Diophantine geometry, including finiteness results for rational points on curves and the arithmetic behavior of elliptic curves. Thus, while we do not directly address conjectures such as Birch–Swinnerton-Dyer, the structural principle identified here—the incompatibility of additive richness with bounded prime support—fits naturally into the broader landscape of arithmetic geometry.

## A. Final Remark: A geometric and physical outlook

Beyond arithmetic, the results of this paper point toward a more general conceptual picture. Multiplication acts as a conservative operation, preserving prime spectra, while addition acts as a generative operation, producing new structural information. The incompatibility between additive richness and multiplicative rigidity is thus not accidental, but reflects a deeper organizational principle.

In a companion work [11], we explore a geometric and variational interpretation of this dichotomy, in which multiplication corresponds to algebraic extension, integration to additive accumulation, and variation to dynamical response. From this viewpoint, additive innovation acquires a geometric meaning as a form of structural mixing, closely related to entropy production and irreversibility.

This perspective suggests that the arithmetic phenomena studied here are instances of a more universal principle: *structure-preserving operations constrain information, while generative operations force expansion across scales.*


[1] David W. Masser and Joseph Oesterlé. Some conjectures on Fermat curves. *Mathematika*, 33(2):193–208, 1986.

[2] Lucien Szpiro. Discriminant et conducteur des courbes elliptiques. *Astérisque*, 183:7–18, 1990.

[3] György Elekes. On the number of sums and products. *Acta Arithmetica*, 81(4):365–367, 1997.

[4] Terence Tao and Van H. Vu. Additive combinatorics. 2006.

[5] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers.* Oxford University Press, 6 edition, 2008.

[6] Jan-Hendrik Evertse and Kálmán Győry. Unit equations in diophantine number theory. *Cambridge Studies in Advanced Mathematics*, 146, 2015.

[7] Karl Dickman. On the frequency of numbers containing prime factors of a certain relative magnitude. *Arkiv för Matematik, Astronomi och Fysik*, 22A:1–14, 1930.

[8] N. G. de Bruijn. On the number of positive integers $\leq x$ and free of prime factors $> y$. *Nederl. Akad. Wetensch. Proc.*, 54:50–60, 1951.

[9] Paul Erdős and Mark Kac. The gaussian law of errors in the theory of additive number theoretic functions. *American Journal of Mathematics*, 62(1):738–742, 1940.

[10] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* Wiley, 2 edition, 2006.

[11] Juan Moreno Borrallo. From algebraic extension to physical law: Multiplication, integration, and the emergence of variational field reality. ResearchGate preprint, 2025. Available on ResearchGate.