Studies of some polynomials with possible applications in physics

Orges Leka

December 11, 2025

Abstract

We propose a geometric and spectral model for the natural numbers derived from the recursive polynomials $f_n(x)$ satisfying $f_n(2) = n$. The prime-step recursion induces integer-valued coefficients $\alpha(n)$ and $C_n(k)$ in the expansion of $\log f_n(x)$, yielding an injective feature embedding $\Phi(n) = (\alpha(n), C_n) \in \mathbb{R} \oplus \ell^2$ with a canonical inner product. This turns $\mathbb{N}$ into a discrete "particle geometry" inside an infinite-dimensional Hilbert space.

Motivated by the primon gas paradigm, we interpret $E_n := \log n$ as an energy (and, in units $c = 1$, a mass) assigned to an integer-atom n, with primes as elementary constituents. Internal structure is probed by prime-restricted Gram matrices $G_{\mathcal{P}(n)}$ and by purely prime-set Grams G_P, whose ordered eigenvalues are interpreted as arithmetic "energy levels." We introduce a radius observable based on distances in Φ-space and identify $\omega(n)$ as a natural dimension parameter of an integer-atom.

Dyson-style diagnostics applied to random large prime-set atoms suggest a robust transition from small-k finite-size behavior toward $\beta = 1$ universality consistent with GOE/LOE statistics, in agreement with the real symmetric nature of the Gram construction. This framework offers a concrete bridge between classical zeta-thermodynamic analogies and random-matrix heuristics, while providing new geometric observables for "atomic" complexity in $\mathbb{N}$.

Contents

1 Introduction

The analogy between prime numbers and elementary particles has long served as a productive metaphor connecting analytic number theory with statistical mechanics. In the primon gas picture, primes are the fundamental excitations and composite integers correspond to multi-particle states. A canonical thermodynamic assignment is

$$E_n := \log n,$$

so that the Boltzmann weight $\exp(-sE_n) = n^{-s}$ yields the partition function

$$Z(s) = \sum_{n\geq 1} e^{-sE_n} = \zeta(s),$$

and the normalized Gibbs law becomes the Zipf-type distribution

$$\mathbb{P}_s(X = n) = \frac{1/n^s}{\zeta(s)}.$$

This viewpoint is aesthetically aligned with Dyson's broader expectation: systems with many effective degrees of freedom should display universal random-matrix statistics, while simpler systems often exhibit Poisson-like behavior.

To make such a philosophy testable on $\mathbb{N}$, one needs a concrete notion of geometry and an associated spectral object. In this work the required structure is induced by a recursion for polynomials $f_n(x) \in \mathbb{Z}[x]$ with the anchor identity $f_n(2) = n$. The prime-step expansion of $\log f_n(x)$ produces two families of invariants: a completely additive scalar $\alpha(n)$ and a finitely supported vector $C_n = (C_n(k))_{k\geq 1}$. This leads to a canonical feature map

$$\Phi(n) := (\alpha(n), C_n) \in \mathcal{H} := \mathbb{R}\oplus\ell^2, \qquad \langle \Phi(i), \Phi(j)\rangle = \alpha(i)\alpha(j)+\sum_{k\geq 1} C_i(k)C_j(k).$$

Because composites linearize in terms of prime features within this embedding, Φ furnishes an arithmetic analogue of "atomic composition" in a Hilbert space.

We then define spectral observables by restricting the Gram form to selected prime sets. For an integer-atom n with prime divisors $\mathcal{P}(n)$, the matrix

$$G_{\mathcal{P}(n)} = (\langle \Phi(p), \Phi(q)\rangle)_{p,q\in\mathcal{P}(n)}$$

provides a small internal spectrum whose positive eigenvalues are interpreted as arithmetic energy levels of n. To model large-complexity limits without constructing n itself, we also study direct prime-set atoms G_P for randomly chosen P of size k.

Finally, we introduce two geometric observables compatible with the physical storyline. First, a radius measuring extension in Φ-space,

$$\mathrm{radius}(n) := \max\{\|\Phi(n) - \Phi(p)\| : \; p \mid n, \; p \text{ prime}\},$$

so that prime are point-like ($\mathrm{radius}(p) = 0$) while composites, have nontrivial spatial extension. Second, a dimension parameter

$$\dim(n) := \omega(n),$$

the number of distinct prime constituents. By Erdős–Kac heuristics, most integers have small $\omega(n)$, suggesting that "typical" integer-atoms are low-dimensional, whereas rare integers with large $\omega(n)$ form the natural arena for random-matrix universality tests.

2 A recursion for $f_n(x)$ and the induced "derivative" $n' = f_n'(2)$

2.1 Definition of f_n

We define polynomials $f_n(x) \in \mathbb{Z}[x]$ recursively by

$$\begin{aligned} f_1(x) &= 1, \\ f_2(x) &= x, \\ f_p(x) &= 1 + f_{p-1}(x) \qquad \text{for primes } p, \\ f_n(x) &= \prod_{p \mid n} f_p(x)^{v_p(n)} \qquad \text{for composite } n, \end{aligned}$$

where $v_p(n)$ is the p-adic valuation.

This is precisely the recursion encoded in `ff(n,x)` in the code snippet.

2.2 Basic evaluation at $x = 2$

Lemma 1. *For all $n \geq 1$ we have $f_n(2) = n$.*

Proof. We proceed by strong induction on n.

For $n = 1, 2$ the statement is immediate. Let $n \geq 3$.

If $n = p$ is prime, then

$$f_p(2) = 1 + f_{p-1}(2) = 1 + (p - 1) = p$$

using the induction hypothesis.

If n is composite, then by definition

$$f_n(2) = \prod_{p|n} f_p(2)^{v_p(n)} = \prod_{p|n} p^{v_p(n)} = n,$$

again by induction. □

2.3 Definition of n' and the correct recursion

Definition 1. *Define*

$$n' := f_n'(2).$$

The code proposes a recursive function `der(n)`. The prime case is correct, but for composites the code uses a *product*. The correct formula for n' is instead a *sum*. We now prove the right recursion.

Proposition 1. *The values $n' = f_n'(2)$ satisfy:*

$$\begin{aligned} 1' &= 0, \\ 2' &= 1, \\ p' &= (p-1)' \qquad \textit{for primes } p, \\ n' &= n \sum_{p|n} v_p(n) \frac{p'}{p} \qquad \textit{for composite } n. \end{aligned}$$

Proof. The initial values are immediate from $f_1(x) = 1$ and $f_2(x) = x$.

If p is prime, $f_p(x) = 1 + f_{p-1}(x)$, hence

$$p' = f_p'(2) = f_{p-1}'(2) = (p-1)'.$$

If n is composite, write

$$f_n(x) = \prod_{p|n} f_p(x)^{v_p(n)}.$$

Taking a logarithmic derivative in x gives

$$\frac{f_n'(x)}{f_n(x)} = \sum_{p|n} v_p(n) \frac{f_p'(x)}{f_p(x)}.$$

Evaluating at $x = 2$ and using Lemma 1 yields

$$\frac{n'}{n} = \sum_{p|n} v_p(n) \frac{p'}{p},$$

which is equivalent to the stated formula. □

Remark 1. *Thus, if one wants* `der(n)` *to equal* $f_n'(2)$*, the composite branch in the code should be implemented as* a sum*:*

```
der(n) = n * sum( valuation(n,p) * der(p) / p for p in prime_divisors(n) ).
```

3 The squarefree Dirichlet series $z(x,s)$ and the ratio $c(s) = a(s)/b(s)$

3.1 Definition of $z(x,s)$

Definition 2. *Let* $\mathrm{rad}(n)$ *denote the radical of* n*. Define*

$$z(x,s) := \sum_{\mathrm{rad}(n)=n} \frac{1}{f_n(x)^s}.$$

For $\Re(s) > 1$ *and* x *in a neighborhood of* 2 *this formal Dirichlet series admits the Euler product*

$$z(x,s) = \prod_p \left(1 + f_p(x)^{-s}\right),$$

since the sum ranges over squarefree integers.

At $x = 2$, Lemma 1 implies

$$z(2,s) = \sum_{\mathrm{rad}(n)=n} \frac{1}{n^s} = \prod_p (1 + p^{-s}) = \frac{\zeta(s)}{\zeta(2s)}, \qquad \Re(s) > 1. \tag{1}$$

3.2 Logarithmic derivative in x

Differentiate the Euler product logarithmically with respect to x. Formally,

$$\begin{aligned} \frac{\partial_x z(x,s)}{z(x,s)} &= \sum_p \frac{\partial_x (1 + f_p(x)^{-s})}{1 + f_p(x)^{-s}} \\ &= \sum_p \frac{-s f_p'(x)\, f_p(x)^{-s-1}}{1 + f_p(x)^{-s}}. \end{aligned} \tag{2}$$

Evaluating at $x = 2$ gives

$$\frac{z_x'(2,s)}{z(2,s)} = -s \sum_p \frac{p'\, p^{-s-1}}{1 + p^{-s}}. \tag{3}$$

3.3 The series $a(s)$ and $b(s)$

Definition 3. *For $\Re(s) > 1$, define*

$$a(s) := \sum_{p} \frac{p' \, p^{-s-1}}{1 + p^{-s}},$$

$$b(s) := \sum_{\mathrm{rad}(n)=n} \frac{n'}{n^{s+1}}.$$

On the other hand, differentiating $z(x,s)$ termwise in the squarefree sum yields

$$z'_x(2,s) = \sum_{\mathrm{rad}(n)=n} \frac{-s \, n'}{n^{s+1}} = -s \, b(s).$$

Combining this with (3) gives, for $\Re(s) > 1$,

$$\frac{-s \, b(s)}{z(2,s)} = -s \, a(s) \quad \Longrightarrow \quad z(2,s) = \frac{b(s)}{a(s)}. \tag{4}$$

Definition 4. *Set*

$$c(s) := \frac{a(s)}{b(s)} \qquad (\Re(s) > 1).$$

Using (1) and (4), we obtain

$$c(s) = \frac{\zeta(2s)}{\zeta(s)}, \qquad \Re(s) > 1. \tag{5}$$

4 Meromorphic continuation of $c(s)$

Proposition 2. *The function $c(s)$ admits a meromorphic continuation to all of $\mathbb{C}$ by*

$$c(s) := \frac{\zeta(2s)}{\zeta(s)}. \tag{6}$$

This continuation agrees with the Dirichlet-series-defined ratio $a(s)/b(s)$ on $\Re(s) > 1$.

Proof. Equation (5) identifies the ratio $a(s)/b(s)$ with $\zeta(2s)/\zeta(s)$ on a nonempty domain. The right-hand side is meromorphic on $\mathbb{C}$, hence defines a meromorphic continuation. □

5 The "doubling formula" and its iterates

5.1 Doubling

Theorem 1 (Doubling identity)**.** *For all $s \in \mathbb{C}$ where both sides are defined,*

$$\zeta(2s) = \zeta(s)\, c(s).$$

Proof. This is immediate from the definition (6). $\square$

5.2 Iteration

A naive statement of the form

$$\zeta(2^k s) = \zeta(s)^k\, c(s)$$

is generally *incorrect*. The correct iterated identity is the telescoping product:

Proposition 3. *For $k \geq 1$,*

$$\zeta(2^k s) = \zeta(s) \prod_{j=0}^{k-1} c(2^j s),$$

whenever all factors are defined.

Proof. Using $\zeta(2t) = \zeta(t)c(t)$ repeatedly with $t = s, 2s, \ldots, 2^{k-1}s$ yields

$$\zeta(2s) = \zeta(s)c(s), \quad \zeta(4s) = \zeta(2s)c(2s), \ \ldots, \ \zeta(2^k s) = \zeta(2^{k-1}s)c(2^{k-1}s).$$

Multiplying these relations telescopes to the stated formula. $\square$

6 Zeros of ζ and zeros of c

From $c(s) = \zeta(2s)/\zeta(s)$ we get the precise relationship between zeros.

Proposition 4. *Let $w \in \mathbb{C}$.*

1. *If $\zeta(w) = 0$ and $\zeta(w/2) \neq 0, \infty$, then*

$$c\Big(\frac{w}{2}\Big) = 0.$$

2. *Conversely, if $c(w/2) = 0$ and $\zeta(w/2) \neq 0, \infty$, then*

$$\zeta(w) = 0.$$

Proof. We have

$$c\Big(\frac{w}{2}\Big) = \frac{\zeta(w)}{\zeta(w/2)}.$$

If the denominator is finite and nonzero, then $c(w/2) = 0$ is equivalent to $\zeta(w) = 0$. □

Remark 2. *Thus, the equivalence*

$$\zeta(w) = 0 \quad \Longleftrightarrow \quad c(w/2) = 0$$

is valid with the natural caveat *that one must exclude the exceptional case where $\zeta(w/2) = 0$ (or a pole). In particular, zeros of ζ at w typically manifest as zeros of c at $w/2$, unless a zero of ζ already occurs at the halved point.*

7 Summary

We introduced the recursively defined polynomials $f_n(x)$ and the induced arithmetic data

$$n' := f_n'(2).$$

We proved the correct recursion

$$n' = n \sum_{p|n} v_p(n) \frac{p'}{p},$$

and showed that, on $\Re(s) > 1$, the ratio

$$c(s) := \frac{a(s)}{b(s)}$$

coincides with $\zeta(2s)/\zeta(s)$. This provides a canonical meromorphic continuation

$$c(s) = \frac{\zeta(2s)}{\zeta(s)}$$

to all of $\mathbb{C}$. The "doubling" identity $\zeta(2s) = \zeta(s)c(s)$ is then tautological, and its correct iteration is

$$\zeta(2^k s) = \zeta(s) \prod_{j=0}^{k-1} c(2^j s).$$

Finally, zeros of ζ at w correspond to zeros of c at $w/2$ provided $\zeta(w/2)$ does not vanish (and is finite).

8 A recursive "log-free on the right" expansion for $\log f_n(x)$

8.1 The polynomials $f_n(x)$

We define polynomials $f_n(x) \in \mathbb{Z}[x]$ by the recursion

$$\begin{aligned} f_1(x) &= 1, \\ f_2(x) &= x, \\ f_p(x) &= 1 + f_{p-1}(x) \qquad \text{for primes } p \geq 3, \\ f_n(x) &= \prod_{p \mid n} f_p(x)^{v_p(n)} \qquad \text{for composite } n, \end{aligned}$$

where $v_p(n)$ denotes the p-adic valuation. (The prime step is intentionally stated only for $p \geq 3$, since the base value $f_2(x) = x$ does not match the rule $f_2 = 1 + f_1$.)

8.2 The power-series block

For any $k \geq 1$ we introduce the formal power series

$$\mathrm{PS}(k; x) := \log\Big(1 + \frac{1}{f_k(x)}\Big) = \sum_{m \geq 1} (-1)^{m+1} \frac{1}{m} \frac{1}{f_k(x)^m}. \tag{7}$$

Analytically, (7) holds whenever $|f_k(x)| > 1$, but we will use it as a formal expansion rule.

8.3 A recursive expansion operator

Define a formal expression $E(n)$ recursively by:

$$E(1) := 0, \tag{8}$$

$$E(2) := \log x, \tag{9}$$

$$E(p) := E(p-1) + \mathrm{PS}(p-1; x) \qquad (p \geq 3 \text{ prime}), \tag{10}$$

$$E(n) := \sum_{p \mid n} v_p(n)\, E(p) \qquad (n \text{ composite}). \tag{11}$$

Lemma 2. *For every $n \geq 1$ we have*

$$E(n) = \log f_n(x).$$

Proof. We proceed by induction on n.

The cases $n = 1, 2$ follow from (8)–(9). If $n = p \geq 3$ is prime, then by definition $f_p = 1 + f_{p-1}$, hence

$$\log f_p = \log f_{p-1} + \log\Big(1 + \frac{1}{f_{p-1}}\Big) = E(p-1) + \mathrm{PS}(p-1; x) = E(p).$$

If n is composite, then $f_n = \prod_{p|n} f_p^{v_p(n)}$, so

$$\log f_n = \sum_{p|n} v_p(n) \log f_p = \sum_{p|n} v_p(n) E(p) = E(n).$$

□

8.4 Coefficient tree and a "log-free" right-hand side

We now expand $E(n)$ in the basis $\{\log x\}$ and the blocks $\mathrm{PS}(k; x)$. Write formally

$$E(n) = \alpha(n)\, \log x + \sum_{k\geq 1} C_n(k)\, \mathrm{PS}(k; x), \tag{12}$$

with integer coefficients $\alpha(n), C_n(k)$.

From the recursion (8)–(11) one immediately reads:

$$\begin{aligned} &\alpha(1) = 0, \qquad \alpha(2) = 1, \\ &\alpha(p) = \alpha(p-1) \qquad (p \geq 3 \text{ prime}), \\ &\alpha(n) = \sum_{p|n} v_p(n)\, \alpha(p) \qquad (n \text{ composite}), \end{aligned}$$

and

$$\begin{aligned} &C_1(k) = 0, \qquad C_2(k) = 0, \\ &C_p(k) = C_{p-1}(k) + \delta_{k,p-1} \qquad (p \geq 3 \text{ prime}), \\ &C_n(k) = \sum_{p|n} v_p(n)\, C_p(k) \qquad (n \text{ composite}). \end{aligned}$$

Now substitute (7) into (12) and *factor out* the coefficients $(-1)^{m+1}\frac{1}{m}$ from the PS-sums. This yields the desired "log-free right-hand side" representation:

Proposition 5. *For every $n \geq 1$,*

$$\log f_n(x) = \alpha(n)\, \log x + \sum_{m\geq 1} (-1)^{m+1} \frac{1}{m} \sum_{k\geq 1} C_n(k)\, \frac{1}{f_k(x)^m}. \tag{13}$$

Proof. Combine Lemma 2 with (12) and (7), then exchange the finite sum over k with the formal series in m. $\square$

8.5 Small examples

The first values of $f_n(x)$ are

$$f_1 = 1, \quad f_2 = x, \quad f_3 = x + 1, \quad f_4 = x^2,$$
$$f_5 = x^2 + 1, \quad f_6 = x(x + 1), \quad f_7 = x^2 + x + 1, \quad f_8 = x^3.$$

Using the recursion for $E(n)$:

$$\log f_3 = \log x + \mathrm{PS}(2; x) = \log x + \sum_{m \geq 1} (-1)^{m+1} \frac{1}{m} \frac{1}{f_2(x)^m},$$

$$\log f_5 = 2 \log x + \mathrm{PS}(4; x) = 2 \log x + \sum_{m \geq 1} (-1)^{m+1} \frac{1}{m} \frac{1}{f_4(x)^m},$$

$$\begin{aligned}\log f_7 &= 2 \log x + \mathrm{PS}(2; x) + \mathrm{PS}(6; x) \\ &= 2 \log x + \sum_{m \geq 1} (-1)^{m+1} \frac{1}{m} \left(\frac{1}{f_2(x)^m} + \frac{1}{f_6(x)^m} \right).\end{aligned}$$

Two slightly larger prime examples illustrate the requested "factored" form.

Example $p = 11$. Since $10 = 2 \cdot 5$ and $f_{10} = f_2 f_5 = x(1 + x^2) = x + x^3$,

$$\log f_{11} = 3 \log x + \mathrm{PS}(4; x) + \mathrm{PS}(10; x) = 3 \log x + \sum_{m \geq 1} (-1)^{m+1} \frac{1}{m} \left(\frac{1}{f_4(x)^m} + \frac{1}{f_{10}(x)^m} \right).$$

Example $p = 13$. Since $12 = 2^2 \cdot 3$ and $f_{12} = f_2^2 f_3 = x^2(x + 1) = x^3 + x^2$,

$$\log f_{13} = 3 \log x + \mathrm{PS}(2; x) + \mathrm{PS}(12; x) = 3 \log x + \sum_{m \geq 1} (-1)^{m+1} \frac{1}{m} \left(\frac{1}{f_2(x)^m} + \frac{1}{f_{12}(x)^m} \right).$$

Remark 3. *Formula (13) packages the entire "prime-step tree" into the integer multiplicities $C_n(k)$. In this sense, the right-hand side is "log-free": the only remaining logarithm is the base term $\alpha(n) \log x$, while all prime-step corrections appear as explicit series in $1/f_k(x)^m$ with the common scalar factor $(-1)^{m+1} \frac{1}{m}$.*

9 Examples

$$\log f_n(x) = \alpha(n)\log x + \sum_{m\geq 1}(-1)^{m+1}\frac{1}{m}\sum_k \frac{C_n(k)}{f_k(x)^m}.$$

n	$\alpha(n)$	$\sum_k \frac{C_n(k)}{f_k(x)^m}$
1	0	0
2	1	0
3	1	$\frac{1}{f_2(x)^m}$
4	2	0
5	2	$\frac{1}{f_4(x)^m}$
6	2	$\frac{1}{f_2(x)^m}$
7	2	$\frac{1}{f_2(x)^m} + \frac{1}{f_6(x)^m}$
8	3	0
9	2	$2\frac{1}{f_2(x)^m}$
10	3	$\frac{1}{f_4(x)^m}$
11	3	$\frac{1}{f_4(x)^m} + \frac{1}{f_{10}(x)^m}$
12	3	$\frac{1}{f_2(x)^m}$
13	3	$\frac{1}{f_2(x)^m} + \frac{1}{f_{12}(x)^m}$
14	3	$\frac{1}{f_2(x)^m} + \frac{1}{f_6(x)^m}$
15	3	$\frac{1}{f_2(x)^m} + \frac{1}{f_4(x)^m}$
16	4	0
17	4	$\frac{1}{f_{16}(x)^m}$
18	3	$2\frac{1}{f_2(x)^m}$
19	3	$2\frac{1}{f_2(x)^m} + \frac{1}{f_{18}(x)^m}$
20	4	$\frac{1}{f_4(x)^m}$

Table 1: Coefficients $\alpha(n)$ and inner sums $\sum_k C_n(k)/f_k(x)^m$ in the expansion $\log f_n(x) = \alpha(n)\log x + \sum_{m\geq 1}(-1)^{m+1}\frac{1}{m}\sum_k C_n(k)/f_k(x)^m$.

Remark 4. *The function $\alpha(n)$ is completely additive by construction.*

n	$\alpha(n)$	$\sum_k \frac{C_n(k)}{f_k(x)^m}$
21	3	$2\frac{1}{f_2(x)^m} + \frac{1}{f_6(x)^m}$
22	4	$\frac{1}{f_4(x)^m} + \frac{1}{f_{10}(x)^m}$
23	4	$\frac{1}{f_4(x)^m} + \frac{1}{f_{10}(x)^m} + \frac{1}{f_{22}(x)^m}$
24	4	$\frac{1}{f_2(x)^m}$
25	4	$2\frac{1}{f_4(x)^m}$
26	4	$\frac{1}{f_2(x)^m} + \frac{1}{f_{12}(x)^m}$
27	3	$3\frac{1}{f_2(x)^m}$
28	4	$\frac{1}{f_2(x)^m} + \frac{1}{f_6(x)^m}$
29	4	$\frac{1}{f_2(x)^m} + \frac{1}{f_6(x)^m} + \frac{1}{f_{28}(x)^m}$
30	4	$\frac{1}{f_2(x)^m} + \frac{1}{f_4(x)^m}$
31	4	$\frac{1}{f_2(x)^m} + \frac{1}{f_4(x)^m} + \frac{1}{f_{30}(x)^m}$
32	5	0
33	4	$\frac{1}{f_2(x)^m} + \frac{1}{f_4(x)^m} + \frac{1}{f_{10}(x)^m}$
34	5	$\frac{1}{f_{16}(x)^m}$
35	4	$\frac{1}{f_2(x)^m} + \frac{1}{f_4(x)^m} + \frac{1}{f_6(x)^m}$
36	4	$2\frac{1}{f_2(x)^m}$
37	4	$2\frac{1}{f_2(x)^m} + \frac{1}{f_{36}(x)^m}$
38	4	$2\frac{1}{f_2(x)^m} + \frac{1}{f_{18}(x)^m}$
39	4	$2\frac{1}{f_2(x)^m} + \frac{1}{f_{12}(x)^m}$
40	5	$\frac{1}{f_4(x)^m}$

Table 2: Coefficients $\alpha(n)$ and inner sums $\sum_k C_n(k)/f_k(x)^m$ in the expansion $\log f_n(x) = \alpha(n)\log x + \sum_{m\geq 1}(-1)^{m+1}\frac{1}{m}\sum_k C_n(k)/f_k(x)^m$.

10 Identification of $\alpha(n)$ with OEIS A064415

10.1 A recursion matching A064415

The OEIS entry A064415 lists several equivalent characterizations. In particular, it states that the function $a(n)$ is completely additive, with $a(1) = 0$, $a(2) = 1$, and for $n > 2$:

$$a(n) = \sum_{p|n} a(p-1), \tag{14}$$

where the sum is over primes dividing n with multiplicity. (Equivalently $a(n) = A003434(n) - (n \bmod 2)$, etc.) This recursion is consistent with the prime rule $a(p) = a(p-1)$ for odd primes.

Proposition 6. *The function $\alpha(n)$ satisfies the OEIS recursion* (14).

Proof. Let $n > 2$. If n is composite, then

$$\alpha(n) = \sum_{p|n} v_p(n)\,\alpha(p).$$

For any odd prime p, we have $\alpha(p) = \alpha(p-1)$. Hence

$$\alpha(n) = \sum_{p|n} v_p(n)\,\alpha(p-1),$$

which is precisely (14) with multiplicity. $\square$

10.2 Equality of sequences

Theorem 2. *For all $n \geq 1$,*

$$\alpha(n) = A064415(n).$$

Proof. Both sequences are determined uniquely by $a(1) = 0$, $a(2) = 1$, complete additivity, and the recursion $a(p) = a(p-1)$ for odd primes (or equivalently the divisor-sum form (14)). These properties hold for $\alpha(n)$, and the OEIS entry states them for A064415. Thus the two functions coincide. $\square$

Remark 5. *Independently, one may observe that $\alpha(n) = \deg f_n(x)$ because $f_p(x) = 1 + f_{p-1}(x)$ preserves degree for $p \geq 3$ and $f_n(x)$ is multiplicative in n. The MathOverflow discussion notes that these degrees match A064415.*

11 Exponentiation and the representation of $f_n(x)$

Corollary 1. *Formally (and analytically whenever the series converges),*

$$f_n(x) = x^{\alpha(n)} \exp\!\left(\sum_{m \geq 1} (-1)^{m+1} \frac{1}{m} \sum_{k \geq 1} C_n(k)\, \frac{1}{f_k(x)^m} \right). \tag{15}$$

Proof. Exponentiate the "log-free" equation for $f_n(x)$. $\square$

12 Specialization at $x = 2$

Using Lemma 1, we have $f_k(2) = k$ for all k. Thus (15) yields:

Theorem 3. *For every $n \geq 1$,*

$$n = 2^{\alpha(n)} \exp\left(\sum_{m \geq 1} (-1)^{m+1} \frac{1}{m} \sum_{k \geq 1} C_n(k) \, \frac{1}{k^m} \right). \tag{16}$$

Proof. Evaluate (15) at $x = 2$ and use $f_n(2) = n$. □

Remark 6. *By Theorem 2, the exponent of* 2 *in* (16) *is precisely A064415*(n)*. Thus the representation* (16) *decomposes each integer n into a canonical power of* 2 *times an exponential correction encoded by the prime-step coefficient tree $C_n(k)$.*

13 Concluding comments

The recursion defining $f_n(x)$ produces two parallel structures:

- a completely additive "degree-like" invariant $\alpha(n)$,
- and a refinement encoded by the coefficients $C_n(k)$ measuring how often prime-step corrections $\mathrm{PS}(k; x)$ appear when expanding $\log f_n(x)$ down to the base index 2.

The identity (16) at $x = 2$ makes this structure explicit at the level of integers.

14 A canonical feature embedding via the coefficients $C_n(k)$

14.1 Setup

Let $f_n(x)$ be defined by

$$\begin{aligned} &f_1(x) = 1, \qquad f_2(x) = x, \\ &f_p(x) = 1 + f_{p-1}(x) \qquad (p \geq 3 \text{ prime}), \\ &f_n(x) = \prod_{p \mid n} f_p(x)^{v_p(n)} \qquad (n \text{ composite}). \end{aligned}$$

Define

$$\mathrm{PS}(k;x) := \log\Bigl(1 + \frac{1}{f_k(x)}\Bigr) = \sum_{m\geq 1} (-1)^{m+1} \frac{1}{m} \frac{1}{f_k(x)^m}.$$

As before, define the expansion operator $E(n)$ by

$$\begin{aligned} &E(1) = 0, \qquad E(2) = \log x, \\ &E(p) = E(p-1) + \mathrm{PS}(p-1;x) \qquad (p \geq 3 \text{ prime}), \\ &E(n) = \sum_{p\mid n} v_p(n)\, E(p) \qquad (n \text{ composite}). \end{aligned}$$

Then one has $E(n) = \log f_n(x)$ for all $n \geq 1$.

We write uniquely

$$\log f_n(x) = \alpha(n) \log x + \sum_{k\geq 1} C_n(k)\, \mathrm{PS}(k;x), \tag{17}$$

where $\alpha(n) \in \mathbb{Z}_{\geq 0}$ and only finitely many $C_n(k)$ are nonzero.

14.2 Evaluation at $x = 2$

A key structural property is

$$f_n(2) = n \qquad (n \geq 1),$$

hence

$$\mathrm{PS}(k;2) = \log\Bigl(1 + \frac{1}{k}\Bigr).$$

Evaluating (17) at $x = 2$ yields the identity

$$\log n = \alpha(n) \log 2 + \sum_{k\geq 1} C_n(k)\, \log\Bigl(1 + \frac{1}{k}\Bigr). \tag{18}$$

Equivalently, exponentiating,

$$n = 2^{\alpha(n)} \prod_{k\geq 1} \Bigl(1 + \frac{1}{k}\Bigr)^{C_n(k)}. \tag{19}$$

Since only finitely many $C_n(k) \neq 0$, the product is finite.

14.3 Injectivity of the feature map

Theorem 4. *The assignment*

$$n \longmapsto \big(\alpha(n), (C_n(k))_{k\geq 1}\big)$$

is injective on $\mathbb{N}$.

Proof. Assume that two integers $n, m \geq 1$ satisfy

$$\alpha(n) = \alpha(m) \quad \text{and} \quad C_n(k) = C_m(k) \text{ for all } k \geq 1.$$

Then the right-hand sides of (17) coincide as formal expressions, hence $\log f_n(x) = \log f_m(x)$. Evaluating at $x = 2$ gives $\log n = \log m$, so $n = m$. □

Remark 7. *The theorem shows that the finite-support vector* $(C_n(k))_{k\geq 1}$*, together with the scalar* $\alpha(n)$*, forms a complete invariant for* n *within this framework.*

14.4 A Hilbert space viewpoint

Let

$$\ell^2_{\text{fin}} := \{(c_k)_{k\geq 1} : \ c_k \in \mathbb{R}, \ c_k = 0 \text{ for all but finitely many } k\} \subset \ell^2.$$

Define the Hilbert space

$$\mathcal{H} := \mathbb{R} \oplus \ell^2, \qquad \langle (a, c), (a', c') \rangle = aa' + \sum_{k\geq 1} c_k c'_k.$$

Then for each n we may define the feature embedding

$$\Phi(n) := \big(\alpha(n), (C_n(k))_{k\geq 1}\big) \in \mathcal{H}. \tag{20}$$

Because $(C_n(k))$ has finite support, $\Phi(n) \in \mathbb{R} \oplus \ell^2_{\text{fin}}$.

By Theorem 4, Φ is an injective (canonical) embedding of $\mathbb{N}$ into the Hilbert space $\mathcal{H}$. Identity (18) shows that the linear functional

$$L(a, (c_k)) := a \log 2 + \sum_{k\geq 1} c_k \log\Big(1 + \frac{1}{k}\Big)$$

recovers $\log n$ from the feature vector:

$$L(\Phi(n)) = \log n.$$

Remark 8. *This provides a concrete "feature geometry" on the integers induced by the prime-step recursion of* $f_n(x)$*. One may study distances* $\|\Phi(n) - \Phi(m)\|$ *or angles in* $\mathcal{H}$ *as quantitative measures of similarity of prime-step structure.*

15 A concrete feature geometry on the integers

15.1 The ambient Hilbert space

Recall the coefficient expansion

$$\log f_n(x) = \alpha(n)\log x + \sum_{k\geq 1} C_n(k)\,\mathrm{PS}(k;x), \qquad \mathrm{PS}(k;x) = \log\Big(1 + \frac{1}{f_k(x)}\Big),$$

with $\alpha(n) \in \mathbb{Z}_{\geq 0}$ and only finitely many $C_n(k) \neq 0$. We define the feature map

$$\Phi(n) := \big(\alpha(n), (C_n(k))_{k\geq 1}\big).$$

Let

$$\ell^2_{\mathrm{fin}} = \{(c_k)_{k\geq 1} : \ c_k = 0 \text{ for all but finitely many } k\} \subset \ell^2,$$

and consider the Hilbert space

$$\mathcal{H} := \mathbb{R} \oplus \ell^2, \qquad \langle (a,c),(a',c')\rangle_{\mathcal{H}} = aa' + \sum_{k\geq 1} c_k c'_k.$$

Since each C_n has finite support, we have $\Phi(n) \in \mathbb{R} \oplus \ell^2_{\mathrm{fin}}$.

15.2 Distance and angle

The induced norm and distance are

$$\|\Phi(n)\|^2 = \alpha(n)^2 + \sum_{k\geq 1} C_n(k)^2, \qquad d(n,m) := \|\Phi(n) - \Phi(m)\|.$$

Explicitly,

$$d(n,m)^2 = \big(\alpha(n) - \alpha(m)\big)^2 + \sum_{k\geq 1}\big(C_n(k) - C_m(k)\big)^2.$$

Thus the distance decomposes into two contributions:

- a *base discrepancy* measured by $\alpha(n) - \alpha(m)$,
- a *prime-step discrepancy* measured coefficientwise by $C_n(k) - C_m(k)$.

Because all C-vectors have finite support, the sum is finite and $d(n, m)$ is well-defined.

Whenever $\Phi(n)$ and $\Phi(m)$ are nonzero, we may also form the angle

$$\cos\theta(n, m) = \frac{\langle \Phi(n), \Phi(m) \rangle_{\mathcal{H}}}{\|\Phi(n)\| \, \|\Phi(m)\|}.$$

A small angle (equivalently large cosine) indicates that n and m share a similar pattern of prime-step corrections, in the sense that their coefficients $C_n(k)$ and $C_m(k)$ point in nearly the same direction in ℓ^2.

15.3 Interpretation via the $x = 2$ identity

Evaluating at $x = 2$ yields

$$\log n = \alpha(n) \log 2 + \sum_{k \ge 1} C_n(k) \, \log\Bigl(1 + \frac{1}{k}\Bigr).$$

Define the linear functional

$$L : \mathcal{H} \to \mathbb{R}, \qquad L\bigl(a, (c_k)\bigr) := a \log 2 + \sum_{k \ge 1} c_k \log\Bigl(1 + \frac{1}{k}\Bigr).$$

Then

$$L(\Phi(n)) = \log n.$$

Hence the feature embedding is not merely geometric: it is *arithmetically complete* in the sense that the scalar observable $\log n$ is a linear readout of the feature vector.

15.4 Nearest-neighbor intuition

The geometry suggests a natural "nearest-neighbor" heuristic: integers n and m with small $d(n, m)$ should share

1. similar values of $\alpha(\cdot)$ (i.e., similar A064415 values),
2. and similar collections of prime-step ancestors, encoded by the support and multiplicities of $C_n(k)$ and $C_m(k)$.

In particular, if n and m share many of the same prime-step indices k with comparable multiplicities, then the ℓ^2 part of $d(n, m)$ will be small.

15.5 Summary

The map

$$\Phi : \mathbb{N} \to \mathcal{H}, \qquad \Phi(n) = \big(\alpha(n), (C_n(k))_{k\geq 1}\big),$$

provides a canonical embedding of the integers into a Hilbert space. The induced distance and angle offer quantitative measures of similarity between the prime-step structures underlying $f_n(x)$, while the identity at $x = 2$ shows that standard arithmetic data ($\log n$) can be recovered as a linear functional of these features.

16 The rank of the Gram matrix and the prime counting function

16.1 Feature vectors and Gram matrices

Recall the coefficient expansion

$$\log f_n(x) = \alpha(n) \log x + \sum_{k\geq 1} C_n(k)\, \mathrm{PS}(k; x), \qquad \mathrm{PS}(k; x) = \log\Big(1 + \frac{1}{f_k(x)}\Big), \tag{21}$$

where $\alpha(n) \in \mathbb{Z}_{\geq 0}$ and only finitely many $C_n(k)$ are nonzero. We define the feature vector

$$\Phi(n) := \big(\alpha(n), (C_n(k))_{k\geq 1}\big) \in \mathbb{R} \oplus \ell^2.$$

We equip $\mathcal{H} := \mathbb{R} \oplus \ell^2$ with the standard inner product

$$\langle (a, c), (a', c') \rangle_{\mathcal{H}} = aa' + \sum_{k\geq 1} c_k c'_k.$$

For $n \geq 1$ let

$$\mathcal{V}_n := \mathrm{span}\{\Phi(1), \dots, \Phi(n)\} \subset \mathcal{H}.$$

The $n \times n$ Gram matrix G_n is defined by

$$(G_n)_{ij} := \langle \Phi(i), \Phi(j) \rangle_{\mathcal{H}} \qquad (1 \leq i, j \leq n).$$

It is a standard fact that

$$\mathrm{rank}(G_n) = \dim(\mathcal{V}_n). \tag{22}$$

16.2 Recursive structure of the coefficients

The coefficients $(\alpha(n), C_n(k))$ satisfy the following recursions:

$$\alpha(1) = 0, \qquad \alpha(2) = 1, \tag{23}$$
$$\alpha(p) = \alpha(p-1) \qquad (p \geq 3 \text{ prime}), \tag{24}$$
$$\alpha(n) = \sum_{p|n} v_p(n)\, \alpha(p) \qquad (n \text{ composite}), \tag{25}$$

and

$$C_1(k) = 0, \qquad C_2(k) = 0, \tag{26}$$
$$C_p(k) = C_{p-1}(k) + \delta_{k,p-1} \qquad (p \geq 3 \text{ prime}), \tag{27}$$
$$C_n(k) = \sum_{p|n} v_p(n)\, C_p(k) \qquad (n \text{ composite}). \tag{28}$$

16.3 Composites do not create new directions

Lemma 3. *If n is composite, then*

$$\Phi(n) = \sum_{p|n} v_p(n)\, \Phi(p).$$

Proof. The first component follows directly from (25). For the second component, fix $k \geq 1$ and use (28):

$$C_n(k) = \sum_{p|n} v_p(n)\, C_p(k).$$

Thus both components of $\Phi(n)$ equal the corresponding linear combination of the components of $\Phi(p)$ for primes $p \mid n$. □

Corollary 2. *For every $n \geq 1$,*

$$\mathcal{V}_n = \operatorname{span}\{\Phi(p) : p \leq n,\ p \text{ prime}\}.$$

Proof. The inclusion "⊃" is obvious. For "⊂", repeatedly apply Lemma 3 to replace each composite $\Phi(m)$ with a linear combination of prime feature vectors with indices $\leq m \leq n$. □

16.4 A uniqueness coordinate for each odd prime

The key observation is that each odd prime p introduces a genuinely new coordinate in the C-part of $\Phi(p)$.

Lemma 4. *Let $p \geq 3$ be prime. Then*

$$C_p(p-1) = 1, \qquad C_m(p-1) = 0 \text{ for all } 1 \leq m < p.$$

In particular, for any other prime $q \neq p$ with $q \leq n$,

$$C_q(p-1) = 0.$$

Proof. The identity $C_p(p-1) = 1$ is immediate from (27):

$$C_p(p-1) = C_{p-1}(p-1) + \delta_{p-1,p-1} = C_{p-1}(p-1) + 1.$$

It therefore suffices to show that $C_{p-1}(p-1) = 0$. We prove the stronger statement that $C_m(p-1) = 0$ for all $m < p$ by induction on m.

For $m = 1, 2$, this follows from (26). Assume $3 \leq m < p$, and that $C_r(p-1) = 0$ for all $r < m$.

Case 1: m is prime. Since $m < p$, we have $m - 1 \neq p - 1$, hence $\delta_{p-1,m-1} = 0$. Thus (27) gives

$$C_m(p-1) = C_{m-1}(p-1) + \delta_{p-1,m-1} = C_{m-1}(p-1) = 0$$

by the induction hypothesis.

Case 2: m is composite. Then (28) yields

$$C_m(p-1) = \sum_{q|m} v_q(m)\, C_q(p-1).$$

All primes $q \mid m$ satisfy $q \leq m < p$, so $C_q(p-1) = 0$ by the induction hypothesis. Hence $C_m(p-1) = 0$.

This completes the induction and proves the claim. □

16.5 Linear independence of the prime feature vectors

Lemma 5. *The set of vectors $\{\Phi(p) : p \leq n,\ p \text{ prime}\}$ is linearly independent.*

Proof. We separate the prime 2 from the odd primes.

First note that $\Phi(2) = (\alpha(2), (0, 0, \dots)) = (1, 0)$ is nonzero, hence provides at least one independent direction.

Now consider a linear relation among prime feature vectors:

$$\lambda_2 \Phi(2) + \sum_{\substack{p \le n \\ p \ge 3 \text{ prime}}} \lambda_p \Phi(p) = 0.$$

Fix an odd prime $p_0 \le n$. Look at the coordinate $k = p_0 - 1$ in the ℓ^2–component. By Lemma 4,

$$C_{p_0}(p_0 - 1) = 1, \quad C_p(p_0 - 1) = 0 \ (p \ne p_0), \quad C_2(p_0 - 1) = 0.$$

Therefore the $k = p_0 - 1$ coordinate of the above linear combination equals λ_{p_0}, which must be zero. Since this holds for every odd prime $p_0 \le n$, all $\lambda_p = 0$ for $p \ge 3$.

The remaining relation is $\lambda_2 \Phi(2) = 0$, so $\lambda_2 = 0$. Thus the prime feature vectors are linearly independent. □

16.6 Rank equals the prime counting function

Theorem 5. *For every $n \ge 1$,*

$$\dim(\mathcal{V}_n) = \pi(n),$$

where $\pi(n)$ denotes the number of primes $\le n$. Consequently,

$$\operatorname{rank}(G_n) = \pi(n).$$

Proof. By Corollary 2,

$$\mathcal{V}_n = \operatorname{span}\{\Phi(p) : p \le n, \ p \text{ prime}\}.$$

By Lemma 5, this spanning set is linearly independent. Hence it is a basis of $\mathcal{V}_n$. Therefore

$$\dim(\mathcal{V}_n) = \#\{p \le n : \ p \text{ prime}\} = \pi(n).$$

Finally, (22) yields $\operatorname{rank}(G_n) = \pi(n)$. □

Corollary 3. *For $n \ge 2$,*

$$\operatorname{rank}(G_n) - \operatorname{rank}(G_{n-1}) = \begin{cases} 1, & n \text{ prime}, \\ 0, & n \text{ composite}. \end{cases}$$

Proof. This is immediate from Theorem 5 and the identity $\pi(n) - \pi(n-1) = 1$ if and only if n is prime. □

16.7 Interpretation

Theorem 5 shows that the first n feature vectors $\Phi(1), \ldots, \Phi(n)$ generate exactly $\pi(n)$ independent directions. The C–coordinates $k = p - 1$ act as "signature features" for odd primes p, while the prime 2 is detected by the α–component. In this sense, the Gram-rank profile of the feature embedding is rigidly controlled by the distribution of primes.

17 Linear functional and continuity of L

We briefly recall the relevant notions and then check that the functional

$$L(a, (c_k)_{k\geq 1}) := a \log 2 + \sum_{k\geq 1} \frac{1}{k}\, c_k \, \log\Bigl(1 + \frac{1}{k}\Bigr)$$

is continuous in the natural Hilbert-space topology used for the feature vectors.

17.1 What is a linear functional?

Let V be a vector space over $\mathbb{R}$ or $\mathbb{C}$. A *linear functional* on V is a linear map

$$L : V \to \mathbb{R} \quad (\text{or } \mathbb{C}),$$

i.e.

$$L(x + y) = L(x) + L(y), \qquad L(\lambda x) = \lambda L(x).$$

In this setting one works with a feature space of the form

$$H := \mathbb{R} \oplus \ell^2,$$

with inner product

$$\langle (a, c), (b, d)\rangle := ab + \sum_{k\geq 1} c_k d_k,$$

and norm $\|(a, c)\|_H^2 = a^2 + \sum_{k\geq 1} c_k^2$.

17.2 Continuity in which sense?

On a normed space $(V, \|\cdot\|)$, a linear functional L is called *continuous* (or *bounded*) if there exists $C > 0$ such that

$$|L(x)| \leq C\|x\| \quad \text{for all } x \in V.$$

For Hilbert spaces this is the standard notion of continuity.

17.3 Continuity of L on $H = \mathbb{R} \oplus \ell^2$

Define the coefficient vector

$$w := \Big(\log 2,\ (w_k)_{k\geq 1}\Big), \qquad w_k := \frac{1}{k}\log\Big(1+\frac{1}{k}\Big).$$

Then L can be written as the inner product

$$L(a,c) = \langle (a,c), w\rangle_H.$$

Hence L is continuous if and only if $w \in H$.

We verify $w \in H$: for large k,

$$\log\Big(1+\frac{1}{k}\Big) = \frac{1}{k} + O\Big(\frac{1}{k^2}\Big),$$

so

$$w_k = \frac{1}{k}\log\Big(1+\frac{1}{k}\Big) = \frac{1}{k^2} + O\Big(\frac{1}{k^3}\Big).$$

Thus $w_k^2 = O(k^{-4})$, and therefore

$$\sum_{k\geq 1} w_k^2 < \infty.$$

Consequently $w \in \mathbb{R} \oplus \ell^2$, and by Cauchy–Schwarz,

$$|L(a,c)| = |\langle (a,c), w\rangle| \leq \|(a,c)\|_H \, \|w\|_H.$$

So L is a bounded (hence continuous) linear functional on H.

17.4 Interpretation via Riesz representation

By the Riesz representation theorem, every continuous linear functional on a Hilbert space is given by inner product with a unique vector. Here that representing vector is exactly w, and the operator norm is

$$\|L\| = \|w\|_H.$$

17.5 Remark

If one changed the ambient topology (e.g. replaced ℓ^2 by ℓ^1 or by the product topology), the continuity question would have a different answer. The above statement is the natural one for the Gram-kernel setup, since the Gram matrix arises from the ℓ^2-type inner product on features.

18 Extension to $\mathbb{R}$ via Continued Fractions

Let $r > 0$ have (canonical) regular continued fraction expansion

$$r = [a_0; a_1, a_2, \dots], \qquad a_0 \in \mathbb{Z}_{\geq 0},\ a_k \in \mathbb{N}\ (k \geq 1),$$

with the standard uniqueness convention for rationals (last digit ≥ 2). Define

$$f_r(x) := [f_{a_0}(x); f_{a_1}(x), f_{a_2}(x), \dots], \tag{29}$$

interpreted as the limit of the convergents

$$f_{r,k}(x) := [f_{a_0}(x); f_{a_1}(x), \dots, f_{a_k}(x)].$$

18.1 Invariants

Anchor at $x = 2$. Using $f_n(2) = 2$ termwise in (29),

$$f_r(2) = r. \tag{30}$$

This is the fundamental property justifying the extension.

Finite complexity for rationals. If $r \in \mathbb{Q}_{>0}$, its continued fraction is finite. Hence $f_r(x)$ is obtained from finitely many $f_{a_j}(x)$ by the operations of addition and reciprocal, and the syntactic "depth" matches the length of the continued fraction (equivalently, the Euclidean algorithm complexity).

18.2 Structural Changes

From polynomials to rational/limit functions.

- For $n \in \mathbb{N}$, $f_n(x) \in \mathbb{Z}[x]$ is a polynomial.
- For $r \in \mathbb{Q}_{>0}$, $f_r(x) \in \mathbb{Q}(x)$ is a rational function.
- For $r \in \mathbb{R} \setminus \mathbb{Q}$, $f_r(x)$ is an infinite functional continued fraction, so it is naturally treated as a limit of rational functions on its convergence domain.

Loss of simple multiplicativity in the index. The identity

$$f_{mn}(x) = f_m(x) f_n(x) \qquad (m, n \in \mathbb{N})$$

does not extend in a natural way to arbitrary real r via continued fractions. The global symmetry of the continued-fraction construction is more closely aligned with fractional-linear transformations (the $\mathrm{GL}(2, \mathbb{Z})$-framework) than with the multiplicative semigroup of indices.

18.3 Emergent Phenomena

Functional recursion in r. Writing $r = a_0 + \frac{1}{r_1}$ with $r_1 = [a_1; a_2, \dots]$ yields

$$f_r(x) = f_{a_0}(x) + \frac{1}{f_{r_1}(x)}, \tag{31}$$

which introduces a dynamical viewpoint reminiscent of continued-fraction maps.

Convergence regimes. For $x = 2$ the coefficients are the positive integers a_k, so convergence is automatic. For $x > 2$ the values $f_{a_k}(x)$ typically grow rapidly, suggesting very fast convergence of (29). For $x < 2$ and in complex x, nontrivial analytic boundaries may appear, defining an "arithmetic stability region" for the functional continued fraction.

19 A Derived Notion of "Arithmetic Derivative" on $\mathbb{R}$

Let $r > 0$ have a (canonical) regular continued fraction expansion

$$r = [a_0; a_1, a_2, \dots], \qquad a_0 \in \mathbb{Z}_{\geq 0},\ a_k \in \mathbb{N}\ (k \geq 1).$$

Assume that a real-indexed extension $r \mapsto f_r(x)$ has been defined via continued-fraction substitution,

$$f_r(x) := [f_{a_0}(x); f_{a_1}(x), f_{a_2}(x), \dots],$$

and that the convergents

$$f_{r,k}(x) := [f_{a_0}(x); \dots, f_{a_k}(x)]$$

are differentiable at $x = 2$.

Define the tails

$$r_k = [a_k; a_{k+1}, \dots], \qquad r_k(x) := f_{r_k}(x),$$

so that in particular $r_0 = r$ and $r_k(2) = r_k$. The continued-fraction recursion becomes the functional identity

$$r_k(x) = f_{a_k}(x) + \frac{1}{r_{k+1}(x)}. \tag{32}$$

Differentiating (32) and evaluating at $x = 2$ yields formally

$$r'_k = a'_k - \frac{1}{r_{k+1}^2}\, r'_{k+1}, \tag{33}$$

where

$$a'_k := f'_{a_k}(2)$$

is the integer arithmetic derivative of the partial quotient a_k.

Iterating (33) by substituting r'_{k+1} into the equation for r'_k produces an explicit alternating expansion for the derivative of $r = r_0$:

$$\begin{aligned} r' &= a'_0 - \frac{1}{r_1^2}\left(a'_1 - \frac{1}{r_2^2} r'_2\right) \\ &= a'_0 - \frac{a'_1}{r_1^2} + \frac{a'_2}{(r_1 r_2)^2} - \frac{a'_3}{(r_1 r_2 r_3)^2} + \cdots . \end{aligned}$$

This suggests the compact formal identity

$$\boxed{r' = f'_r(2) = \sum_{k=0}^{\infty} (-1)^k \frac{a'_k}{\left(\prod_{j=1}^{k} r_j\right)^2}}\,. \tag{34}$$

Informal convergence discussion. The product $\prod_{j=1}^{k} r_j$ grows at least geometrically and is closely related to the denominators q_k of the convergents of r. Thus (34) is expected to converge under mild growth assumptions on the integer derivatives a'_k (e.g. excluding superexponential growth). Conceptually, the formula links *local arithmetic data* (the derivatives a'_k of the partial quotients) with *global diophantine properties* of r encoded by the tails r_j.

20 Extension to $\mathbb{Z}$ and $\mathbb{C}$ in the Index

To support complex indices, extend first from $\mathbb{N}$ to $\mathbb{Z}$ by

$$f_0(x) := 0, \qquad f_{-r}(x) := -f_r(x) \quad (r > 0). \tag{35}$$

Assuming the real-index map $r \mapsto f_r(x)$ has been defined (for instance via continued fractions), define for $z = a + bi \in \mathbb{C}$ with $a, b \in \mathbb{R}$

$$f_{a+bi}(x) := f_a(x) + i\, f_b(x). \tag{36}$$

This complexification is canonical with respect to the decomposition $\mathbb{C} \cong \mathbb{R} \oplus i\mathbb{R}$.

20.1 Properties Preserved by the Complex Extension

Complex anchor at $x = 2$. From the real anchor $f_r(2) = r$ and (36),

$$f_{a+bi}(2) = f_a(2) + i f_b(2) = a + bi = z.$$

Oddness in the index. By (35) and (36),

$$f_{-z}(x) = f_{-a-bi}(x) = f_{-a}(x) + i f_{-b}(x) = -f_a(x) - i f_b(x) = -f_z(x).$$

Conjugation compatibility (for real x). If $x \in \mathbb{R}$, then $f_a(x)$ and $f_b(x)$ are real-valued, hence

$$\overline{f_z(x)} = \overline{f_a(x) + i f_b(x)} = f_a(x) - i f_b(x) = f_{\overline{z}}(x).$$

20.2 The Induced Arithmetic Derivative on $\mathbb{C}$

The definition (36) induces a natural notion of arithmetic derivative for complex indices by differentiating with respect to x and anchoring at $x = 2$.

Definition 5 (Complex Arithmetic Derivative)**.** *For $z = a + bi \in \mathbb{C}$, define*

$$z' := f_z'(2).$$

By linearity of differentiation and (36),

$$z' = \frac{d}{dx}\bigg|_{x=2} \big(f_a(x) + i f_b(x)\big) = f_a'(2) + i f_b'(2), \tag{37}$$

so the induced derivative acts component-wise:

$$\boxed{(a + bi)' = a' + ib'}.$$

Basic properties.

1. **Consistency with the real case.** If $b = 0$, then $z' = a'$, recovering the derivative on $\mathbb{R}$.

2. **Compatibility with conjugation.** Using (37),

$$(\overline{z})' = (a - bi)' = a' - ib' = \overline{a' + ib'} = \overline{z'}.$$

Remark 9. *This induced derivative treats $\mathbb{C}$ primarily as the real vector space $\mathbb{R} \oplus i\mathbb{R}$. In contrast to the classical arithmetic derivative on Gaussian integers (which is designed to satisfy a Leibniz rule), the present construction does not enforce multiplicativity in the index, and therefore does not generally imply*

$$(zw)' = z'w + zw' \quad \textit{for arbitrary } z, w \in \mathbb{C}.$$

20.3 What This Extension Does *Not* Add

The rule (36) is a canonical packaging of the real theory into complex parameters. It does not, by itself, impose a new multiplicative or prime-like arithmetic on $\mathbb{C}$-indices:

$$f_{z+w}(x) \neq f_z(x) + f_w(x), \qquad f_{zw}(x) \neq f_z(x) f_w(x)$$

in general, unless such identities already hold at the real-index level.

21 Summary and Outlook

The continued-fraction extension (29) shifts the organizing principle from prime-factor multiplicativity to the additive/diophantine structure of continued-fraction digits. The anchor identity $f_r(2) = r$ is retained, while the index-arithmetic becomes governed by functional continued-fraction dynamics.

The complex-index extension (36) is canonical and consistent, preserving the anchor and symmetries such as oddness and conjugation compatibility, but it is primarily a linear complexification of the real-index theory.

N	$r = \pi(N)$	$\det G_{\mathcal{P}(N)}$	Vol $= \sqrt{\det}$ (num)
2	1	1	1
3	2	1	1
5	3	1	1
7	4	1	1
11	5	1	1
13	6	1	1
17	7	1	1
19	8	1	1
23	9	1	1
29	10	1	1
31	11	1	1
37	12	1	1
41	13	1	1
43	14	1	1
47	15	1	1

Table 3: Gram determinants and volumes of the prime-spanned parallelotopes.

22 Unimodularity of the Prime Gram Determinant

We work in the Hilbert space

$$H = \mathbb{R}e_\alpha \oplus \ell^2(\mathbb{N}),$$

equipped with the orthonormal basis

$$\{e_\alpha\} \cup \{e_k\}_{k\geq 1}.$$

For each $n \geq 1$ define the feature vector

$$\Phi(n) = \alpha(n)e_\alpha + \sum_{k\geq 1} C_n(k)\, e_k,$$

where $\alpha(n) \in \mathbb{Z}$ and $C_n(k) \in \mathbb{Z}$ are given by the recursions:

- $\alpha(1) = 0$, $\alpha(2) = 1$;
- if $p \geq 3$ is prime, then $\alpha(p) = \alpha(p-1)$;
- if n is composite, then
$$\alpha(n) = \sum_{p|n} v_p(n)\, \alpha(p);$$
- $C_1 = 0$, $C_2 = 0$;
- if $p \geq 3$ is prime, then
$$C_p = C_{p-1} + \delta_{p-1};$$
- if n is composite, then
$$C_n = \sum_{p|n} v_p(n)\, C_p.$$

The inner product on H is

$$\langle \Phi(i), \Phi(j) \rangle = \alpha(i)\alpha(j) + \sum_{k\geq 1} C_i(k)C_j(k).$$

For $N \geq 2$ let

$$\mathcal{P}(N) = \{p_1 < \cdots < p_r\}$$

be the set of primes $\leq N$ with $p_1 = 2$, and let

$$G_{\mathcal{P}(N)} = \big(\langle \Phi(p_i), \Phi(p_j) \rangle\big)_{i,j=1}^{r}$$

be the Gram matrix of the prime feature vectors.

22.1 Support and new-coordinate lemmas

Lemma 6 (Prime-support lemma)**.** *For every* $n \geq 1$*,*

$$\mathrm{supp}(C_n) \subseteq \{\, q-1 : \ q \leq n \ \textit{prime}, \ q \geq 3 \,\}.$$

Proof. We argue by induction on n. For $n = 1, 2$ the claim is immediate since $C_1 = C_2 = 0$. If $n = p \geq 3$ is prime, then

$$C_p = C_{p-1} + \delta_{p-1},$$

so the support of C_p is contained in the union of the support of C_{p-1} and the singleton $\{p-1\}$. If n is composite, then

$$C_n = \sum_{p|n} v_p(n) C_p,$$

so $\mathrm{supp}(C_n)$ is contained in a union of prime supports with primes $p \leq n$. □

Lemma 7 (New coordinate per prime)**.** *Let* $p \geq 3$ *be prime. Then*

$$C_p(p-1) = 1 \quad \textit{and} \quad C_m(p-1) = 0 \textit{ for all } m < p.$$

Proof. By the prime-support lemma, for any $m < p$ the support of C_m is contained in $\{q - 1 : \ q \leq m \text{ prime}\}$, hence cannot contain $p - 1$. Thus $C_{p-1}(p-1) = 0$. Using

$$C_p = C_{p-1} + \delta_{p-1}$$

gives $C_p(p-1) = 1$. □

22.2 Determinant computation

Define the r-element orthonormal set

$$B_N := \{e_\alpha\} \cup \{e_{p_i-1} : \ i = 2, \ldots, r\}.$$

By the prime-support lemma, every $\Phi(p_i)$ with $p_i \leq N$ lies in $\mathrm{span}(B_N)$. Let $A_N \in \mathbb{Z}^{r\times r}$ be the coordinate matrix of $\{\Phi(p_i)\}_{i=1}^{r}$ with respect to B_N (rows are the coordinates).

Lemma 8. *The matrix* A_N *is lower triangular with diagonal entries equal to* 1*. Hence* $\det(A_N) = 1$.

Proof. For $p_1 = 2$, we have $\Phi(2) = e_\alpha$, so the first row of A_N is $(1, 0, \dots, 0)$. For $i \geq 2$, the column corresponding to $e_{p_i - 1}$ records the coefficient $C_{p_j}(p_i - 1)$ in row j. By the new-coordinate lemma, this coefficient equals 1 when $j = i$ and equals 0 for all $j < i$. Therefore all entries above the diagonal in these columns vanish and the diagonal entries are 1. □

Proposition 7 (Prime Gram determinant)**.** *For every $N \geq 2$,*

$$\det\bigl(G_{\mathcal{P}(N)}\bigr) = 1.$$

Consequently the Euclidean volume of the parallelotope spanned by $\{\Phi(p) : p \leq N\}$ equals 1.

Proof. Since B_N is orthonormal and A_N records the coordinates of the prime vectors in that basis, we have

$$G_{\mathcal{P}(N)} = A_N A_N^T.$$

Thus

$$\det(G_{\mathcal{P}(N)}) = \det(A_N)^2 = 1.$$

The volume statement follows from the standard identity $\mathrm{Vol}^2 = \det(\mathrm{Gram})$. □

Remark 10. *The result reflects a strong "unimodularity" property of the prime feature family: each new prime $p \geq 3$ introduces a genuinely new orthonormal coordinate e_{p-1} with coefficient 1 that cannot occur earlier. This makes the coordinate matrix of the prime vectors triangular with unit diagonal, forcing the Gram determinant to be identically 1.*

23 Local lattice counts in the Φ-geometry and comparison with the v_p-encoding

In this section we use, without proof, the degree estimate

$$\frac{\log n}{\log 3} \;\leq\; \alpha(n) = \deg f_n(x) \;\leq\; \frac{\log n}{\log 2}, \qquad n \geq 2. \tag{38}$$

Recall that

$$\Phi(n) = (\alpha(n), C_n), \qquad \|\Phi(n)\|^2 = \alpha(n)^2 + \sum_k C_n(k)^2,$$

and the induced metric satisfies

$$\|\Phi(m) - \Phi(n)\|^2 = (\alpha(m) - \alpha(n))^2 + \|C_m - C_n\|^2 \;\geq\; (\alpha(m) - \alpha(n))^2.$$

23.1 A coarse multiplicative window

Fix $n \in \mathbb{N}$ and $r > 0$ and define the Φ-ball

$$B_\Phi(n, r) := \{m \in \mathbb{N} : \ \|\Phi(m) - \Phi(n)\| \le r\}.$$

From the inequality above we obtain the necessary condition

$$m \in B_\Phi(n, r) \quad \Longrightarrow \quad |\alpha(m) - \alpha(n)| \le r. \tag{39}$$

Using (38) as a logarithmic proxy for α, one is led to the heuristic interpretation

$$\alpha(n) \approx \log_3 n, \qquad \alpha(m) \approx \log_3 m,$$

so that (39) suggests

$$|\log_3 m - \log_3 n| \lesssim r,$$

i.e. a multiplicative window

$$n\, 3^{-r} \ \lesssim \ m \ \lesssim \ n\, 3^{r}. \tag{40}$$

While (40) is deliberately coarse (it ignores the essential $\|C_m - C_n\|$ constraint), it already captures the correct *scaling* in n for fixed r.

23.2 The sinh-bound and an $O(n)$ estimate

The window (40) implies the counting upper bound

$$\begin{aligned} \#B_\Phi(n, r) \ &\lesssim \ \#\big([n3^{-r}, \, n3^{r}] \cap \mathbb{N}\big) \\ &\le \ n(3^{r} - 3^{-r}) + 1 = 2n \, \sinh(r \log 3) + 1. \end{aligned} \tag{41}$$

Hence, for each fixed radius r,

$$\boxed{\#B_\Phi(n, r) = O(n) \qquad (n \to \infty).} \tag{42}$$

In practice, the C-component often makes the ball substantially smaller than (41) predicts; the estimate (41) should be viewed as an α-only envelope.

23.3 Contrast with the pure v_p-encoding

Consider the classical prime-exponent encoding

$$n \longmapsto v(n) := (v_p(n))_{p \text{ prime}} \in \mathbb{Z}^{(\mathcal{P})},$$

equipped, for instance, with the unweighted ℓ^2-metric

$$\|v(m) - v(n)\|^2 = \sum_p (v_p(m) - v_p(n))^2.$$

In this geometry, if $r \geq 1$ and q is any prime not dividing n, then

$$m = nq \quad \Longrightarrow \quad \|v(m) - v(n)\| = 1 \leq r.$$

Since there are infinitely many primes q, every ball with radius $r \geq 1$ contains infinitely many lattice points. Thus:

$$\boxed{\text{In the pure } v_p\text{-geometry, local balls are typically infinite.}} \tag{43}$$

23.4 Adding a size coordinate

The finiteness phenomenon in the Φ-geometry is driven by the size-sensitive component $\alpha(n) \sim \log n$. To obtain a similar "properness" for the v_p-encoding, one would need to augment it by an additional size coordinate, e.g.

$$n \longmapsto (\log n,\ v(n)),$$

so that any metric ball automatically restricts m to a finite multiplicative interval around n. In this sense, $\alpha(n)$ plays the role of a built-in logarithmic anchor in the Φ-framework.

Summary. Assuming (38), the necessary condition $|\alpha(m) - \alpha(n)| \leq r$ suggests a multiplicative window of radius 3^r, yielding the coarse counting estimate

$$\#B_\Phi(n, r) \ \lesssim\ 2n \sinh(r \log 3) + 1,$$

and in particular $\#B_\Phi(n, r) = O(n)$ for fixed r. This sharply contrasts with the unweighted v_p-encoding, where local balls are infinite unless one adds a logarithmic size coordinate.

24 From the "Primon Gas" to a Φ-Geometry of Integers

24.1 Primes as particles and ζ as a partition function

A longstanding theme in analytic number theory and mathematical physics is the analogy between prime numbers and fundamental particles. In the so-called *primon gas* picture, the primes act as elementary excitations and composite integers arise as multi-particle states. This viewpoint is closely tied to the formal identity

$$\zeta(s) = \sum_{n\geq 1} \frac{1}{n^s} = \sum_{n\geq 1} \exp(-sE_n), \tag{44}$$

where one interprets

$$E_n := \log n \tag{45}$$

as an "energy" of the "atom" n. Setting $c = 1$ suggests an energy–mass identification in natural units, so that E_n may be read as a mass scale for the integer-state n.

From this we obtain a canonical Gibbs law:

$$\mathbb{P}_s(X = n) := \frac{e^{-sE_n}}{\sum_{m\geq 1} e^{-sE_m}} = \frac{1/n^s}{\zeta(s)}. \tag{46}$$

Thus the "primon gas" connects directly to a Zipf-type distribution on $\mathbb{N}$.

24.2 Montgomery–Odlyzko–Dyson: atomic spectra and random matrices

The Montgomery pair-correlation conjecture and Odlyzko's computations indicate that the local statistics of the high-lying nontrivial zeros of $\zeta(s)$ match those of the Gaussian Unitary Ensemble (GUE). Dyson's broader philosophy suggests that complicated many-body atomic spectra tend to fall into universal random-matrix classes, while simpler or more integrable systems display Poisson-like statistics.

In this narrative, "large atoms" with many effective degrees of freedom should exhibit random-matrix behavior, whereas "small atoms" should be closer to integrable limits. A key challenge is to equip integers with a geometric and spectral structure that makes this principle testable.

24.3 A Φ-feature embedding induced by the polynomials $f_n(x)$

We consider a family of polynomials $f_n(x) \in \mathbb{Z}[x]$ defined recursively so that $f_n(2) = n$. This recursion induces a "prime-step" expansion of $\log f_n(x)$ in the form

$$\log f_n(x) = \alpha(n) \log x + \sum_{k\geq 1} C_n(k)\, \mathrm{PS}(k; x), \tag{47}$$

where $\alpha(n) \in \mathbb{Z}_{\geq 0}$ is completely additive and $C_n(k) \in \mathbb{Z}$ has finite support. Evaluating at $x = 2$ gives the exact identity

$$\log n = \alpha(n) \log 2 + \sum_{k \geq 1} C_n(k) \log\Big(1 + \frac{1}{k}\Big). \tag{48}$$

This motivates the feature map

$$\Phi(n) := (\alpha(n), C_n) \in \mathbb{R} \oplus \ell^2, \tag{49}$$

together with the Euclidean inner product

$$\langle \Phi(i), \Phi(j) \rangle = \alpha(i)\alpha(j) + \sum_{k \geq 1} C_i(k) C_j(k). \tag{50}$$

In this geometry, primes act as extremal "building blocks" in the sense that composite feature vectors are linear combinations of prime feature vectors with indices dividing the composite.

24.4 "Atomic" spectra from prime subspaces

We now propose a concrete spectral model for the integer-atom n.

Let

$$\mathcal{P}(n) := \{p : \ p \mid n, \ p \text{ prime}\} \tag{51}$$

and define the prime-restricted Gram matrix

$$G_{\mathcal{P}(n)} := \big(\langle \Phi(p), \Phi(q) \rangle\big)_{p,q \in \mathcal{P}(n)}. \tag{52}$$

We interpret the *sorted eigenvalues* of $G_{\mathcal{P}(n)}$ as *internal energy levels* of the integer-atom n. Removing a prime factor p corresponds to

$$n' := \frac{n}{p}, \qquad E_{n'} = \log(n') < \log(n) = E_n, \tag{53}$$

so the mass/energy decreases as one removes a constituent "particle".

24.5 Radius and dimension of an integer-atom

We introduce a coarse geometric notion of size. Define

$$\text{radius}(n) := \max\big\{ \ \|\Phi(n) - \Phi(p)\| \ : \ p \mid n, \ p \text{ prime} \ \big\}. \tag{54}$$

A prime $n = p$ then satisfies

$$\text{radius}(p) = 0, \tag{55}$$

so a prime is "point-like" with no spatial extension in this model. A composite integer with

$$\omega(n) \geq 2 \tag{56}$$

has strictly positive radius and therefore an "extended" structure.

We further propose the intrinsic dimension

$$\dim(n) := \omega(n), \tag{57}$$

the number of distinct prime divisors. Hence, in this cartoon, an integer with $\omega(n) = 1$ behaves like a one-particle excitation (including prime powers), while $\omega(n) = 2, 3, \ldots$ behaves like a multi-particle atom with increasing internal complexity.

By the Erdős–Kac heuristic, most integers have a small value of $\omega(n)$, with the bulk of mass concentrated around relatively low dimensions. In this sense, "most atoms" in $\mathbb{N}$ are low-dimensional, even though exceptional integers with large $\omega(n)$ exist and may constitute the natural arena where random-matrix universality emerges.

24.6 Empirical GOE crossover in random prime ensembles

To test the Dyson philosophy in this arithmetic setting, one may form *random prime-set atoms*: choose a set P of k primes at random from a range, build

$$G_P := \big(\langle \Phi(p_i), \Phi(p_j) \rangle\big)_{p_i, p_j \in P}, \tag{58}$$

and study the spacing-ratio statistic

$$r_i = \frac{\min(s_i, s_{i-1})}{\max(s_i, s_{i-1})}, \tag{59}$$

computed from the ordered eigenvalues of G_P.

Numerically, the mean ratio $\langle r \rangle$ displays a clear size-dependent crossover: for very small k, finite-size effects dominate; for intermediate k one sees level repulsion above Poisson; and for large k the statistics approach the $\beta = 1$ universality class associated to real symmetric ensembles. In one representative set of experiments,

$$k = 5: \ \langle r \rangle \approx 0.27, \tag{60}$$

$$k = 50: \ \langle r \rangle \approx 0.47, \tag{61}$$

$$k = 500: \ \langle r \rangle \approx 0.51, \tag{62}$$

and for

$$k = 400 \text{ random primes} \le 5000, \quad \langle r \rangle \approx 0.531, \tag{63}$$

consistent with an approach to the GOE/LOE benchmark. This behavior matches the structural expectation that the *real, symmetric* Gram construction naturally belongs to the $\beta = 1$ symmetry class.

Summary. Starting from the primon-gas idea $E_n = \log n$, we propose a Φ-feature geometry derived from the polynomials $f_n(x)$. Primes are point-like constituents, composites acquire a geometric extension, and $\omega(n)$ serves as a natural dimension parameter of integer-atoms. Spectra of prime-restricted Gram matrices $G_{\mathcal{P}(n)}$ offer a concrete "atomic" diagnostic. Empirically, random large prime-set Gram spectra exhibit a crossover toward $\beta = 1$ universality (GOE/LOE), aligning with Dyson's expectation that complex many-constituent systems tend to random-matrix statistics, while leaving open a principled path toward unitary symmetry and the Montgomery–Odlyzko regime.

25 Conclusion

We developed a Hilbert-space geometry of the integers derived from the prime-step recursion of the polynomials $f_n(x)$. The induced coefficients $\alpha(n)$ and $C_n(k)$ define a canonical and injective feature embedding $\Phi : \mathbb{N} \to \mathbb{R} \oplus \ell^2$, with an explicit inner product that produces well-defined Gram matrices for arithmetically meaningful subsets of $\mathbb{N}$. Within this framework, primes emerge as elementary, point-like constituents: they carry zero radius in the proposed metric, while composites acquire nontrivial spatial extension.

Interpreting $E_n = \log n$ as energy and mass (in units $c = 1$) connects the present geometry to zeta-thermodynamic models and to a Zipf-type Gibbs law on $\mathbb{N}$. The prime-restricted Gram spectra $G_{\mathcal{P}(n)}$ provide a concrete notion of internal "energy levels" of an integer-atom, while direct prime-set atoms G_P allow us to probe large-k regimes without constructing n. The observed spacing-ratio crossover—from small-k finite-size behavior toward statistics consistent with $\beta = 1$ universality—is structurally natural for a real symmetric Gram construction and aligns with a Dyson-style complexity narrative: increasing the number of constituents drives the system away from Poisson-like simplicity toward random-matrix rigidity.

Several directions remain open. First, a refined notion of mass or inertial content within Φ-space could combine the scalar energy E_n with geometric invariants such as $\|\Phi(n)\|$, radius(n), or local lattice counts in Φ-balls. Second, the relationship between this real $\beta = 1$ regime and the Montgomery–Odlyzko GUE paradigm suggests a symmetry question: to access $\beta = 2$ universality one may need a genuinely complex or phase-sensitive extension of the feature space rather than a purely real Gram kernel. Finally, the rarity of large-$\omega(n)$ integers highlights a natural dichotomy: most arithmetic atoms are low-dimensional and potentially "integrable," while exceptional high-ω atoms offer a plausible laboratory for universality phenomena at the interface of number theory and spectral physics.

26 Appendix

26.1 A Riemannian Φ–geometry on integer "atoms"

26.1.1 Prime–restricted Gram matrices and observables

Let $\Phi : \mathbb{N} \to \mathcal{H}$ be the feature embedding constructed in the previous sections, with $\mathcal{H}$ a real Hilbert space and

$$\langle \Phi(i), \Phi(j) \rangle = \alpha(i)\alpha(j) + \sum_{k \geq 1} C_i(k) C_j(k).$$

For a natural number $n \geq 2$, denote by

$$\mathcal{P}(n) := \{p : \ p \mid n, \ p \text{ prime}\}$$

its set of distinct prime divisors and write

$$\mathcal{P}(n) = \{p_1, \dots, p_r\}, \qquad r = \omega(n).$$

Definition 6 (Prime–restricted Gram matrix and observable)**.** *The* prime Gram matrix *of the integer–atom n is*

$$G_{\mathcal{P}}(n) := \big(\langle \Phi(p_i), \Phi(p_j) \rangle \big)_{1 \leq i,j \leq r} \in \mathbb{R}^{r \times r}.$$

We know that $G_{\mathcal{P}}(n)$ is positive definite, so that the matrix logarithm and square root are well–defined. The associated observable *at n is*

$$A_{\mathcal{P}}(n) := \log G_{\mathcal{P}}(n),$$

viewed as a Hermitian "Hamiltonian" encoding the internal spectrum of n.

In this picture, the eigenvalues of $G_{\mathcal{P}}(n)$ play the role of energy levels of the integer–atom n, and the eigenvalues of $A_{\mathcal{P}}(n)$ are their logarithms.

26.1.2 A local metric at a base atom

We regard the set of admissible prime Gram matrices as a submanifold of the symmetric space of positive definite matrices. Fix a basepoint $r \in \mathbb{N}$. The observable $A_{\mathcal{P}}(r)$ is then an invertible Hermitian matrix.

Definition 7 (Local metric at a base atom)**.** *For integers m, n we define the bilinear form*

$$g_r(m,n) := \operatorname{tr}\Big(A_{\mathcal{P}}(r)^{-1}A_{\mathcal{P}}(m)\,A_{\mathcal{P}}(r)^{-1}A_{\mathcal{P}}(n)\Big).$$

Formally, g_r plays the role of a Riemannian metric on the space of integer–atoms, pulled back from the canonical trace metric on the space of observables. When $m = n = r$ we have

$$g_r(r,r) = \operatorname{tr}(\mathbf{1}) = \omega(r),$$

the rank of the prime Gram matrix of r.

26.1.3 Geodesic distance via the affine–invariant metric

On the manifold $\mathcal{M}$ of symmetric positive definite matrices, the *affine–invariant Riemannian metric* is the unique Riemannian structure invariant under congruence transformations $X \mapsto SXS^{\mathsf{T}}$, $S \in \mathrm{GL}(r, \mathbb{R})$. Its geodesic distance between $X, Y \in \mathcal{M}$ is given by

$$d_{\mathrm{AI}}(X,Y) = \Big\|\log\big(X^{-1/2}YX^{-1/2}\big)\Big\|_{\mathrm{F}} = \sqrt{\operatorname{tr}\Big(\log\big(X^{-1/2}YX^{-1/2}\big)^2\Big)},$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm.

We now use this metric on the prime Gram matrices of integer–atoms.

Definition 8 (Affine–invariant geodesic distance between atoms)**.** *For integers m, n we define*

$$\mathrm{dist}_{\mathrm{geo}}(m,n) := d_{\mathrm{AI}}\big(G_{\mathcal{P}}(m), G_{\mathcal{P}}(n)\big) = \sqrt{\operatorname{tr}\Big(\log\big(G_{\mathcal{P}}(m)^{-1/2}G_{\mathcal{P}}(n)G_{\mathcal{P}}(m)^{-1/2}\big)^2\Big)}.$$

By construction, $\mathrm{dist}_{\mathrm{geo}}$ is a true metric: it is symmetric, nonnegative, vanishes exactly when $G_{\mathcal{P}}(m) = G_{\mathcal{P}}(n)$, and satisfies the triangle inequality. Moreover, it is invariant under simultaneous congruence transformations of all Gram matrices, reflecting the natural symmetry of the underlying feature space.

26.1.4 The one–dimensional prime–power sector

To understand the geometry in the simplest possible setting, consider the "one–dimensional" sector consisting of integers with exactly one distinct prime factor. Let

$$a = p^x, \quad b = q^y, \quad c = r^z,$$

with p, q, r prime and $x, y, z \geq 1$, so that

$$\omega(a) = \omega(b) = \omega(c) = 1.$$

Then

$$\mathcal{P}(a) = \{p\}, \quad \mathcal{P}(b) = \{q\}, \quad \mathcal{P}(c) = \{r\},$$

and each prime Gram matrix is 1×1:

$$G_{\mathcal{P}}(a) = [\,\|\Phi(p)\|^2\,], \quad G_{\mathcal{P}}(b) = [\,\|\Phi(q)\|^2\,], \quad G_{\mathcal{P}}(c) = [\,\|\Phi(r)\|^2\,].$$

The associated observables are scalars:

$$A_{\mathcal{P}}(a) = \log G_{\mathcal{P}}(a) = 2\log\|\Phi(p)\|,$$

and similarly for b, c.

Proposition 8 (Local metric in the prime–power sector)**.** *In the one–dimensional prime–power sector, the metric at $a = p^x$ acting on $b = q^y$ and $c = r^z$ is*

$$g_a(b, c) = \frac{\log\|\Phi(q)\|\ \log\|\Phi(r)\|}{\big(\log\|\Phi(p)\|\big)^2}.$$

Proof. Since $G_{\mathcal{P}}(a)$ is 1×1, its logarithm and inverse are scalars:

$$A_{\mathcal{P}}(a)^{-1} = \frac{1}{\log G_{\mathcal{P}}(a)} = \frac{1}{2\log\|\Phi(p)\|}.$$

Likewise $A_{\mathcal{P}}(b) = 2\log\|\Phi(q)\|$ and $A_{\mathcal{P}}(c) = 2\log\|\Phi(r)\|$. Therefore

$$\begin{aligned}
g_a(b, c) &= \operatorname{tr}\big(A_{\mathcal{P}}(a)^{-1}A_{\mathcal{P}}(b)\,A_{\mathcal{P}}(a)^{-1}A_{\mathcal{P}}(c)\big) \\
&= \big(A_{\mathcal{P}}(a)^{-1}A_{\mathcal{P}}(b)\,A_{\mathcal{P}}(a)^{-1}A_{\mathcal{P}}(c)\big) \\
&= \frac{1}{2\log\|\Phi(p)\|} \cdot 2\log\|\Phi(q)\| \cdot \frac{1}{2\log\|\Phi(p)\|} \cdot 2\log\|\Phi(r)\| \\
&= \frac{\log\|\Phi(q)\|\ \log\|\Phi(r)\|}{\big(\log\|\Phi(p)\|\big)^2}.
\end{aligned}$$

□

Proposition 9 (Affine–invariant distance between prime powers)**.** *In the same sector, the affine–invariant geodesic distance between $b = q^y$ and $c = r^z$ is*

$$\mathrm{dist}_{\mathrm{geo}}(b, c) = 2\left|\log\|\Phi(q)\| - \log\|\Phi(r)\|\right|.$$

Proof. In the 1×1 case we have

$$G_{\mathcal{P}}(b)^{-1/2} = \left[\|\Phi(q)\|^{-1}\right], \qquad G_{\mathcal{P}}(c) = \left[\|\Phi(r)\|^{2}\right],$$

so

$$G_{\mathcal{P}}(b)^{-1/2} G_{\mathcal{P}}(c) G_{\mathcal{P}}(b)^{-1/2} = \left[\frac{\|\Phi(r)\|^2}{\|\Phi(q)\|^2}\right].$$

Taking the logarithm gives the scalar

$$\log\bigl(G_{\mathcal{P}}(b)^{-1/2} G_{\mathcal{P}}(c) G_{\mathcal{P}}(b)^{-1/2}\bigr) = \left[\,\log(\|\Phi(r)\|^2/\|\Phi(q)\|^2)\,\right] = \left[\,2\log\|\Phi(r)\| - 2\log\|\Phi(q)\|\,\right].$$

Its Frobenius norm squared is just the square of this scalar, hence

$$\mathrm{dist}_{\mathrm{geo}}(b, c) = \sqrt{\mathrm{tr}\bigl(\log(\cdots)^2\bigr)} = \left|2\log\|\Phi(r)\| - 2\log\|\Phi(q)\|\right| = 2\left|\log\|\Phi(q)\| - \log\|\Phi(r)\|\right|.$$

□

27 A physical reading of the Φ–geometry

27.1 Overall geometric picture: integers as extended atoms

In the Φ–framework each natural number $n \in \mathbb{N}$ comes with a feature vector

$$\Phi(n) \in \mathcal{H} := \mathbb{R} \oplus \ell^2,$$

constructed from the prime-step recursion of the polynomials $f_n(x)$. We think of $\mathcal{H}$ as a *background Hilbert space* that carries the full arithmetic information of n in a linear, coordinate-free way. In particular, the map $n \mapsto \Phi(n)$ is injective and

$$\log n = L(\Phi(n))$$

for a fixed continuous linear functional $L : \mathcal{H} \to \mathbb{R}$.

For each integer

$$n \geq 2, \qquad \mathcal{P}(n) := \{p : \ p \mid n, \ p \text{ prime}\}$$

we form the *prime-restricted Gram matrix*

$$G_{\mathcal{P}}(n) := \big(\langle \Phi(p_i), \Phi(p_j) \rangle \big)_{p_i, p_j \in \mathcal{P}(n)},$$

a positive definite matrix of size $\omega(n) \times \omega(n)$, where $\omega(n) = |\mathcal{P}(n)|$ is the number of distinct prime factors. The collection

$$\mathcal{M} := \{G_{\mathcal{P}}(n) : n \in \mathbb{N},\ \omega(n) \geq 1\}$$

lives inside the manifold $\mathrm{SPD}(k)$ of $k \times k$ symmetric positive definite matrices (for varying k), which we interpret as the *spatial / extension space* of integer-atoms. In this picture:

- the background Hilbert space $\mathcal{H}$ carries the "internal" arithmetic degrees of freedom;
- the point $G_{\mathcal{P}}(n) \in \mathrm{SPD}(\omega(n))$ is the *extended spatial configuration* of the atom n, obtained from its prime constituents;
- the eigenvalues of $G_{\mathcal{P}}(n)$ form an effective internal spectrum of n, and their logarithms are encoded in the observable

  $$A_{\mathcal{P}}(n) := \log G_{\mathcal{P}}(n),$$

  a Hermitian matrix acting on the prime-subspace of n.

Energetically, we assign to each integer the "mass"

$$E_n := \log n,$$

as in the primon gas: E_n plays the role of an energy or rest mass (in units with $c = 1$). For a typical large n one has, by the Erdős–Kac heuristic,

$$\omega(n) \approx \log \log n = \log E_n,$$

so heavy atoms (large E_n) are expected to have many distinct prime factors and therefore large intrinsic spatial dimension $\dim G_{\mathcal{P}}(n) = \omega(n)$.

The basic geometric quantities are:

- the *Gram-space point* of n, namely $G_{\mathcal{P}}(n)$;
- the *observable* of n, $A_{\mathcal{P}}(n) = \log G_{\mathcal{P}}(n)$;

- the *affine-invariant geodesic distance* between two atoms m, n, given by
$$d_{\mathrm{AI}}(m,n) := \left\| \log\bigl(G_{\mathcal{P}}(m)^{-1/2} G_{\mathcal{P}}(n) G_{\mathcal{P}}(m)^{-1/2}\bigr) \right\|_{\mathrm{F}},$$
where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm.

This d_{AI} is the standard Riemannian distance on the manifold of positive definite matrices; it measures how far the extended prime-geometry of m must be moved along geodesics in SPD to reach that of n.

A second, purely linear measure is the direct Frobenius distance

$$\|G_{\mathcal{P}}(m) - G_{\mathcal{P}}(n)\|_{\mathrm{F}},$$

which compares the Gram matrices in a flat (Euclidean) way. For matrices that are already very close, the two notions agree to first order; for large atoms with structured prime content, the affine-invariant distance $d_{\mathrm{AI}}(m,n)$ retains information about multiplicative deformations, whereas the raw Frobenius distance only sees additive differences of entries.

27.2 General-relativistic viewpoint: curvature in Gram space

For each integer $n \geq 1$ we associate

$$E_n := \operatorname{span}\{\Phi(p) : p \mid n,\ p \text{ prime}\} \subset \mathcal{H},$$

of dimension
$$\dim E_n = \omega(n),$$
and the prime Gram matrix

$$G_{\mathcal{P}}(n) := \bigl(\langle \Phi(p_i), \Phi(p_j) \rangle \bigr)_{p_i, p_j \in \mathcal{P}(n)} \ \in\ \mathrm{SPD}\bigl(\omega(n)\bigr),$$

where $\mathrm{SPD}(k)$ denotes the manifold of real $k \times k$ symmetric positive definite matrices.

The *expansion space* of n in the spatial sense is then the Riemannian manifold $\mathrm{SPD}\bigl(\omega(n)\bigr)$ equipped with the standard affine-invariant metric

$$\langle U, V \rangle_X := \operatorname{tr}\bigl(X^{-1} U X^{-1} V\bigr), \qquad X \in \mathrm{SPD}(k),\ U, V \in T_X \mathrm{SPD}(k).$$

This metric turns $\mathrm{SPD}(k)$ into a homogeneous Riemannian symmetric space of $\mathbb{R}$-dimension
$$\dim \mathrm{SPD}(k) = \frac{k(k+1)}{2}.$$

Thus:

- For $\omega(n) = 0$ there is no SPD block at all (the "vacuum" case $n = 1$).
- For $\omega(n) = 1$ the expansion space is $\mathrm{SPD}(1) \cong (0, \infty)$, a one-dimensional half-line.
- For $\omega(n) = 2$ the expansion space is $\mathrm{SPD}(2)$, a *three*-dimensional Einstein manifold; this is the smallest nontrivial spatial expansion in this sense.
- In general, as $\omega(n)$ grows, the spatial expansion space of n has dimension $1, 3, 6, 10, \ldots, \frac{\omega(n)(\omega(n)+1)}{2}$, with rapidly increasing geometric complexity.

We regard the disjoint union

$$\mathcal{M}_{\mathrm{Gram}} := \bigcup_{k \geq 0} \mathrm{SPD}(k),$$

equipped with the affine-invariant metric on each component $\mathrm{SPD}(k)$, as the *spatial background* of the theory. Each integer n picks out a distinguished point $G_{\mathcal{P}}(n) \in \mathrm{SPD}\big(\omega(n)\big)$; its *mass* is $E_n = \log n$, and its *internal dimension* is $\omega(n)$.

Geodesic distance. On a fixed component $\mathrm{SPD}(k)$ we measure the distance between two points $X, Y \in \mathrm{SPD}(k)$ by the affine-invariant Riemannian distance

$$d_{\mathrm{AI}}(X, Y) := \left\| \log\big(X^{-1/2} Y X^{-1/2}\big) \right\|_{\mathrm{F}},$$

where $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm and log is the matrix logarithm. For two integers m, n with $\omega(m) = \omega(n) = k$ we set

$$d_{\mathrm{AI}}(m, n) := d_{\mathrm{AI}}\big(G_{\mathcal{P}}(m), G_{\mathcal{P}}(n)\big).$$

This distance depends on the eigenvalues of

$$S_{m,n} := G_{\mathcal{P}}(m)^{-1/2} G_{\mathcal{P}}(n) G_{\mathcal{P}}(m)^{-1/2} \in \mathrm{SPD}(k):$$

if $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of $S_{m,n}$, then

$$d_{\mathrm{AI}}(m, n)^2 = \sum_{j=1}^{k} \big(\log \lambda_j\big)^2.$$

27.3 Triangle experiments in Gram space: light vs. heavy atoms

We now illustrate the negative curvature of the affine–invariant metric on Gram space by an explicit "triangle experiment". In both cases we work in the 5–dimensional expansion sector $\omega(n) = 5$, i.e. each integer has five distinct prime factors and hence an internal Gram matrix in SPD(5). The three vertices of each triangle are integer–atoms m_1, m_2, m_3 with pairwise coprime prime sets; the edges are the affine–invariant distances

$$d_{\mathrm{AI}}(m_i, m_j) = \left\| \log\bigl(G_{\mathcal{P}}(m_i)^{-1/2} G_{\mathcal{P}}(m_j) G_{\mathcal{P}}(m_i)^{-1/2} \bigr) \right\|_F,$$

and the interior angles are computed from the usual law of cosines in the Riemannian metric:

$$\cos\angle(m_i) = \frac{d_{\mathrm{AI}}(m_j, m_k)^2 + d_{\mathrm{AI}}(m_i, m_k)^2 - d_{\mathrm{AI}}(m_i, m_j)^2}{2\, d_{\mathrm{AI}}(m_j, m_k)\, d_{\mathrm{AI}}(m_i, m_k)}.$$

Because the affine–invariant metric has negative sectional curvature, one expects the sum of angles of such triangles to be *less* than π. We compare a "light" triangle built from small primes with a "heavy" triangle built from much larger primes.

Light atoms: small primes and almost flat geometry

Consider

$$\begin{aligned} m_1 &= 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 = 2310, \\ m_2 &= 13 \cdot 17 \cdot 19 \cdot 23 \cdot 29 = 2\,800\,733, \\ m_3 &= 31 \cdot 37 \cdot 41 \cdot 43 \cdot 47 = 95\,041\,567, \end{aligned}$$

with

$$\gcd(m_i, m_j) = 1, \qquad \omega(m_1) = \omega(m_2) = \omega(m_3) = 5.$$

The associated prime sets and Gram matrices are

$$\mathcal{P}(m_1) = \{2, 3, 5, 7, 11\}, \quad G_{\mathcal{P}}(m_1) \approx \begin{pmatrix} 1 & 1 & 2 & 2 & 3 \\ 1 & 2 & 2 & 3 & 3 \\ 2 & 2 & 5 & 4 & 7 \\ 2 & 3 & 4 & 6 & 6 \\ 3 & 3 & 7 & 6 & 11 \end{pmatrix},$$

$$\mathcal{P}(m_2) = \{13, 17, 19, 23, 29\}, \quad G_{\mathcal{P}}(m_2) \approx \begin{pmatrix} 11 & 12 & 11 & 12 & 13 \\ 12 & 17 & 12 & 16 & 16 \\ 11 & 12 & 14 & 12 & 14 \\ 12 & 16 & 12 & 19 & 16 \\ 13 & 16 & 14 & 16 & 19 \end{pmatrix},$$

$$\mathcal{P}(m_3) = \{31, 37, 41, 43, 47\}, \quad G_{\mathcal{P}}(m_3) \approx \begin{pmatrix} 19 & 18 & 21 & 18 & 21 \\ 18 & 21 & 20 & 20 & 20 \\ 21 & 20 & 27 & 20 & 26 \\ 18 & 20 & 20 & 22 & 20 \\ 21 & 20 & 26 & 20 & 29 \end{pmatrix}.$$

For orientation we also record the naive Frobenius distances

$$\begin{aligned} \|G_{\mathcal{P}}(m_1) - G_{\mathcal{P}}(m_2)\|_F &\approx 52.39, \\ \|G_{\mathcal{P}}(m_2) - G_{\mathcal{P}}(m_3)\|_F &\approx 38.34, \\ \|G_{\mathcal{P}}(m_1) - G_{\mathcal{P}}(m_3)\|_F &\approx 87.40. \end{aligned}$$

The affine–invariant geodesic distances in SPD(5) are

$$\begin{aligned} d_{\mathrm{AI}}(m_1, m_2) &\approx 4.58, \\ d_{\mathrm{AI}}(m_2, m_3) &\approx 1.37, \\ d_{\mathrm{AI}}(m_1, m_3) &\approx 4.93. \end{aligned}$$

From these we compute the interior angles at the vertices m_1, m_2, m_3:

$$\begin{aligned} \angle(m_1) &\approx 15.09^\circ, \\ \angle(m_2) &\approx 95.45^\circ, \\ \angle(m_3) &\approx 66.53^\circ. \end{aligned}$$

Hence the sum of angles is

$$\angle(m_1) + \angle(m_2) + \angle(m_3) \approx 177.06^\circ,$$

which is only slightly smaller than the Euclidean value 180°. In this sense the triangle determined by these three "light" integer atoms is almost flat: the negative curvature in this region of Gram space is quite mild.

Heavy atoms: large primes and visible negative curvature

We now repeat the construction with much larger primes:

$$\begin{aligned} m_1 &= 101 \cdot 103 \cdot 107 \cdot 109 \cdot 113 = 13\,710\,311\,357, \\ m_2 &= 127 \cdot 131 \cdot 137 \cdot 139 \cdot 149 = 47\,205\,940\,259, \\ m_3 &= 151 \cdot 157 \cdot 163 \cdot 167 \cdot 173 = 111\,641\,786\,731, \end{aligned}$$

again with

$$\gcd(m_i, m_j) = 1, \qquad \omega(m_1) = \omega(m_2) = \omega(m_3) = 5.$$

The prime sets and Gram matrices are

$$\mathcal{P}(m_1) = \{101, 103, 107, 109, 113\}, \quad G_{\mathcal{P}}(m_1) \approx \begin{pmatrix} 41 & 36 & 36 & 30 & 36 \\ 36 & 39 & 37 & 33 & 37 \\ 36 & 37 & 40 & 33 & 37 \\ 30 & 33 & 33 & 35 & 33 \\ 36 & 37 & 37 & 33 & 39 \end{pmatrix},$$

$$\mathcal{P}(m_2) = \{127, 131, 137, 139, 149\}, \quad G_{\mathcal{P}}(m_2) \approx \begin{pmatrix} 36 & 33 & 35 & 33 & 36 \\ 33 & 40 & 42 & 38 & 38 \\ 35 & 42 & 51 & 42 & 42 \\ 33 & 38 & 42 & 41 & 38 \\ 36 & 38 & 42 & 38 & 42 \end{pmatrix},$$

$$\mathcal{P}(m_3) = \{151, 157, 163, 167, 173\}, \quad G_{\mathcal{P}}(m_3) \approx \begin{pmatrix} 42 & 38 & 34 & 44 & 38 \\ 38 & 42 & 38 & 42 & 40 \\ 34 & 38 & 42 & 35 & 38 \\ 44 & 42 & 35 & 53 & 42 \\ 38 & 40 & 38 & 42 & 43 \end{pmatrix}.$$

The Frobenius distances now read

$$\begin{aligned} \|G_{\mathcal{P}}(m_1) - G_{\mathcal{P}}(m_2)\|_F &\approx 24.37, \\ \|G_{\mathcal{P}}(m_2) - G_{\mathcal{P}}(m_3)\|_F &\approx 28.32, \\ \|G_{\mathcal{P}}(m_1) - G_{\mathcal{P}}(m_3)\|_F &\approx 33.53, \end{aligned}$$

so in the naive Euclidean sense the heavy atoms are *closer together* than the light ones.

However, the affine–invariant distances in Gram space behave differently:

$$\begin{aligned} d_{\mathrm{AI}}(m_1, m_2) &\approx 1.33, \\ d_{\mathrm{AI}}(m_2, m_3) &\approx 1.88, \\ d_{\mathrm{AI}}(m_1, m_3) &\approx 1.91. \end{aligned}$$

These are shorter than in the light case (because the Gram matrices are larger), but the relative shape of the triangle is noticeably distorted by curvature. The interior angles are

$$\begin{aligned} \angle(m_1) &\approx 64.97^\circ, \\ \angle(m_2) &\approx 66.71^\circ, \\ \angle(m_3) &\approx 37.50^\circ, \end{aligned}$$

with

$$\angle(m_1) + \angle(m_2) + \angle(m_3) \approx 169.18^\circ.$$

The deficit from 180°,

$$\delta_{\text{light}} \approx 180^\circ - 177.06^\circ \approx 2.94^\circ, \qquad \delta_{\text{heavy}} \approx 180^\circ - 169.18^\circ \approx 10.82^\circ,$$

is significantly larger for the heavy triple. Since $E_n = \log n$ and $\omega(n)$ are both larger for the second triangle, this numerical experiment is consistent with the idea that "heavy" atoms with many prime factors live deeper in the negatively curved region of Gram space: their internal Gram matrices produce triangles whose angle sums depart more strongly from the Euclidean value.

In summary, even though the Frobenius distances between the heavy Gram matrices are smaller, the affine–invariant geometry on SPD(5) reveals a stronger negative curvature through the larger angle deficit. This concretely visualizes how the mass scale $E_n = \log n$ (and, via Erdős–Kac, the typical size of $\omega(n)$) correlates with the way integer–atoms sit in the curved Einstein manifold of prime Gram matrices.

27.4 Quantum-mechanical viewpoint: states, observables, spectra

The Φ-geometry also admits a static quantum-mechanical interpretation, focused on states and spectra rather than on time evolution.

Sectors by intrinsic dimension. Each integer n determines an intrinsic internal dimension $\omega(n)$ and hence a finite-dimensional space E_n together with an internal Gram matrix $G_{\mathcal{P}}(n)$. This stratifies $\mathbb{N}$ into sectors

$$\mathcal{N}_k := \{n \in \mathbb{N} : \omega(n) = k\},$$

with associated internal spaces

$$E_n \cong \mathbb{R}^k, \qquad G_{\mathcal{P}}(n) \in \mathrm{SPD}(k), \qquad A_{\mathcal{P}}(n) = \log G_{\mathcal{P}}(n) \text{ Hermitian.}$$

From this viewpoint:

- $\omega(n) = 0$ (only $n = 1$) is a vacuum-like sector with no internal degrees of freedom.
- $\omega(n) = 1$ (prime powers) gives one-level internal systems, with $A_{\mathcal{P}}(n)$ a scalar and no nontrivial internal mixing.

- $\omega(n) = 2$ (bi-prime atoms) yields true two-level systems; the off-diagonal entry of $G_{\mathcal{P}}(n)$ encodes the hybridization between the two prime modes.
- Higher $\omega(n)$ produce k-level systems with $k \times k$ Hermitian observables $A_{\mathcal{P}}(n)$ and increasingly rich internal spectra.

Global state space and integer basis. It is natural to assemble all integer-atoms into a single Hilbert space

$$\mathcal{H}_{\text{arith}} := \overline{\text{span}}\{\Phi(n) : n \in \mathbb{N}\}$$

or, abstractly, to work with an orthonormal basis $\{|n\rangle : n \in \mathbb{N}\}$ indexed by the integers. Pure integer states correspond to these basis vectors, while general states are superpositions

$$|\psi\rangle = \sum_n c_n |n\rangle, \qquad \sum_n |c_n|^2 = 1.$$

In this language, each n labels a distinguished internal expansion geometry through $G_{\mathcal{P}}(n)$; the abstract integer basis is static, while the nontrivial structure appears in the family of observables attached to it.

Observables and spectra. The construction provides two closely related classes of observables:

- A diagonal "mass operator"

  $$H_{\text{arith}} := \sum_{n \geq 1} E_n \, |n\rangle\langle n|, \qquad E_n = \log n,$$

  whose eigenvalues are the arithmetic energies E_n of the integer-atoms.
- For each n, an internal Hermitian matrix

  $$A_{\mathcal{P}}(n) := \log G_{\mathcal{P}}(n)$$

  acting on the k-dimensional expansion space E_n with $k = \omega(n)$, whose eigenvalues describe the internal spectrum of the prime configuration of n.

One may think of H_{arith} as fixing the "rest energy" of each integer-atom, while $A_{\mathcal{P}}(n)$ refines this into an internal k-level structure within the prime sector of n.

Static spectral statistics and universality. Because $A_{\mathcal{P}}(n)$ is Hermitian, its eigenvalues can be studied using the tools of random-matrix theory, but in a purely static way: for a given n (or for an ensemble of n), one examines level spacings and spacing-ratio statistics. Empirically:

- For small $\omega(n)$, the spectra of $A_{\mathcal{P}}(n)$ (equivalently of $G_{\mathcal{P}}(n)$) are low-dimensional and dominated by finite-size, almost-integrable structure; there is little room for universal statistics.

- For larger $\omega(n)$ and for random ensembles of prime sets, the spacing-ratio statistics of the Gram eigenvalues move toward the $\beta = 1$ universality class, reminiscent of GOE behavior for real symmetric matrices.

This mirrors the usual quantum intuition: simple systems with few degrees of freedom tend to have highly structured, non-universal spectra, while large, complicated systems with many interacting modes exhibit random-matrix-type statistics.

In summary, from the quantum-mechanical standpoint the Φ-geometry supplies:

- a natural Hilbert space of integer-atoms and their superpositions,

- a stratification into internal sectors by $\omega(n)$, each with its own finite-dimensional expansion space E_n,

- and, for every n, a Hermitian observable $A_{\mathcal{P}}(n)$ whose eigenvalues encode an internal spectrum arising from the prime configuration of n and whose statistics can be compared to standard random-matrix universality classes.

No time evolution is needed to formulate this analogy: it is already meaningful at the level of static states and spectra.