Primes as the Source of New Distinctions

How repeating records gain information, and how much they can afford to forget

Keith Taylor · 24 November 2026

For the general reader

Picture a room full of clocks that all tick together. Each face has a different number of positions: two, three, four and so on. Looking at every face at once tells you where you are in a very long cycle. Sooner or later all the faces return to where they started, and from then on the room repeats itself.

Now install the clocks in order of length: two positions first, then three, and so on. Sometimes a new clock tells you something new, and moments that used to look identical now look different. Sometimes it tells you nothing, because the clocks already in the room fix what it must show. If we install the clocks in order of length, the room gains new information precisely when we reach a prime or one of its powers, such as 4, 8, 9 or 25. Every other clock is fully predictable from the smaller ones already there.

The amount of new information is fixed as well. Only one way of measuring it agrees with two modest requirements: clocks that share no common factor add their information together, and a bigger clock never carries less than a smaller one. On that measure the smallest possible gain is one bit, and it arrives at the powers of two.

So within this picture, primes are where a record can say something that no combination of earlier records already said. With them, the room's stock of distinguishable moments grows by about one unit for each clock added. How closely it keeps to that average is the subject of the Riemann hypothesis, whose proof was formally verified in November 2026 [11].

A real recorder will sometimes fail to keep a new distinction. The final result says how much it can afford to forget. It may lose distinctions at infinitely many primes and still follow the ideal count, provided its total losses grow no faster than the square root of the size of the clocks it has reached. Losing even one bit at every prime is too much. We also describe a simple heat experiment that would read one hidden distinction, and say what the VERSF programme would still have to supply. That includes a conjecture: short cycles, which come round often, may keep the present tied to what has already been recorded, while long cycles tell this moment apart from earlier ones that looked the same.

Abstract

A cycle of n phases has one autonomous coarse display for each divisor of n, and the capacity that its proper displays fail to supply is the von Mangoldt function $\Lambda(n)$. For any finite set of synchronised counters, the logarithm of the number of joint readings is the sum of Λ over the prime powers dividing at least one period. Two operational axioms, additivity over coprime cycles and monotonicity, force capacity to be a multiple of log n by a theorem of Erdős. These results come from the second supplement to [1] and are restated here with proofs. This paper adds four things. Cycles are classified as atomic (prime length), layered (higher prime power) or

composite. Known forms of the Riemann hypothesis are restated as bounds on the capacity of a room of counters. For any sufficiently smooth window, the windowed prime-power sum of [1] minus its benchmark equals the windowed surplus of capacity gains over their mean of one per counter, up to an error with an explicit bound. Finally, a record that loses part of each gain still tracks the ideal windowed reading to within a bounded error exactly when its total loss through X is $O(\sqrt{X})$. A heat comparison on a ring illustrates how an earlier record determines which new distinction must be read. The counters that return to their starting phase at tick T are those whose length divides T, and their gains sum to log T; a conjecture relates this to how a present moment is anchored.

1 Cycles, displays and the clock room

A cycle of length n is a system with n phases, labelled 0 to n − 1, that advances by one phase at each tick and returns to its starting phase after n ticks. A tick is one application of the update; no clock time is assigned to it. A display is any function of the phase. A display is *autonomous* when its present reading always determines its next reading, so that an observer who keeps only the display can advance it without knowing the phase. It is *proper* when it has fewer than n readings. Lemma 1 and Theorems 2 and 3 first appeared in the second supplement to [1] [3]. They are restated with proofs so that this paper can be read alone.

Lemma 1. Up to a relabelling of readings, the autonomous displays of an n-cycle are the maps $j \mapsto j \bmod d$, one for each divisor d of n.

Proof. Let f be an autonomous display, so $f(j + 1) = T(f(j))$ for a fixed update T. Then $f(j) = T^j(f(0))$, and T^n fixes f(0). The readings therefore form a single cycle under T, of some length d dividing n, and f(j) depends only on j mod d. Conversely j mod d is autonomous for each divisor d.

Theorem 2 (what shorter descriptions miss). Let $n \geq 2$ and read all proper autonomous displays of an n-cycle together. If $n = p^k$ is a power of a prime p, the joint reading takes n/p values and each value leaves exactly p phases possible. If n has two or more distinct prime factors, the joint reading takes n values and determines the phase. Hence

$$\log n - \log(\text{number of joint proper readings}) = \Lambda(n),$$

where $\Lambda(n) = \log p$ when n is a power of the prime p, and $\Lambda(n) = 0$ otherwise.

Proof. By Lemma 1 the joint reading records j modulo the least common multiple of the proper divisors of n. For $n = p^k$ that is p^{k-1}. Otherwise every maximal prime-power factor of n is a proper divisor, so the least common multiple is n.

We call $\Lambda(n)$ the *gain* of the n-cycle: the capacity it has that none of its shorter descriptions supply. Each prime factor p of n contributes log p once for each of its powers dividing n, so the gains of the divisors add up to the whole capacity [4]:

$$\sum_{d \mid n} \Lambda(d) = \log n.$$

Theorem 3 (any set of counters). Let S be a finite set of positive integers, and run one counter of each length in S from a shared tick. Write Q(S) for the set of prime powers that divide at least one member of S. The joint reading takes lcm(S) values, and

$$\log \operatorname{lcm}(S) = \sum_{q \in Q(S)} \Lambda(q).$$

Adding a counter of length n multiplies the number of joint readings by the exponential of the sum of Λ(q) over the prime powers q that divide n and are not in Q(S). Every old reading splits into the same number of new ones.

Proof. Two ticks give the same joint reading exactly when their difference is a multiple of every period. The readings therefore repeat with period lcm(S) and are distinct within one period. For each prime p let e be the largest exponent with p^e dividing a member of S. Then lcm(S) contains p to the power e, and the powers of p in Q(S) are p, p^2, …, p^e, each with Λ = log p. Summing over p gives the formula. The statement about a new counter follows by applying the formula to S and to S ∪ {n}. Finally, the old joint period D advances the new counter by D modulo n, which visits n/gcd(D, n) residues equally often.

Corollary 4 (the clock room). With S = {1, …, n − 1}, adding a counter of length n multiplies the number of joint readings by p if n is a power of the prime p, and by 1 otherwise. Hence L(n) = lcm(1, …, n) satisfies $\log L(n) = \sum_{m \le n} \Lambda(m) = \psi(n)$, Chebyshev's function [5].

The total in Theorem 3 does not depend on the order in which counters are added. Only the attribution of a gain to a particular counter does, as Table 1 shows. Theorem 2 is free of that choice, because it refers to one cycle alone.

Counters present	Counter added	New readings per old reading	Prime powers newly covered
All lengths 1 to 24	25	5	25
All lengths 1 to 25	26	1	none
50	25	1	none
2 and 3	12	2	4
6 and 10	15	1	none
4	6	3	3

Table 1. Gains under Theorem 3. A composite counter can carry a gain only when one of its prime-power factors is missing from the room.

These logarithms count alternatives. For any probability distribution of the tick, the conditional Shannon entropy of the new reading, given the old ones, is at most the gain. It equals the gain exactly when, within each old reading of positive probability, the new possibilities are equally likely. The old readings themselves need not be equally likely. The gain is therefore the most that the new counter can ever add.

2 Three kinds of cycle

Proposition 5 (classification). Let n ≥ 2.

(a) *Atomic.* n is prime exactly when the n-cycle has no proper autonomous display other than the constant one.

(b) *Layered.* $n = p^k$ with $k \geq 2$ exactly when the n-cycle has a proper autonomous display that is not constant and its autonomous displays are totally ordered, each being a function of the next. There are then $k + 1$ of them, with $1, p, \ldots, p^k$ readings, and each step up the ladder adds $\log p$.

(c) *Composite.* n has two or more distinct prime factors exactly when the proper autonomous displays together determine the phase. The n-cycle then carries the same information as its maximal prime-power cycles run from a shared tick.

Proof. Parts (a) and (c) restate Lemma 1 and Theorem 2. For (b), the display modulo d is a function of the display modulo d′ exactly when d divides d′, and the divisors of n are totally ordered by divisibility exactly when n is a prime power. By (a), a prime power has a proper display that is not constant exactly when its exponent is at least 2. The last statement of (c) is the Chinese remainder theorem. The three kinds are mutually exclusive and cover every $n \geq 2$.

Corollary 6 (the smallest positive gain is one bit). The gain of a cycle is either zero or $\log p$ for a prime p. The smallest positive gain is therefore $\log 2$, which is one bit, and it occurs exactly at the powers of two.

n	Kind	Gain in bits	n	Kind	Gain in bits
2	atomic	1	8	layered	1
3	atomic	1.585	9	layered	1.585
4	layered	1	10	composite	0
5	atomic	2.322	11	atomic	3.459
6	composite	0	12	composite	0
7	atomic	2.807	13	atomic	3.700

Table 2. The three kinds for n from 2 to 13.

In this model every gain is located at a prime or one of its powers. A room extended only by counters whose prime-power factors are already present gains nothing, however many are added (Theorem 3). That is the precise sense in which primes are the source of new distinctions here. It is a statement about systems that can be counted and run side by side.

3 Why the size of the gain is forced

Sections 1 and 2 measured capacity by the logarithm of a count. That choice can be derived, by an argument that also appears in [3]. Let $c(n)$ be any real number assigned to the n-cycle as its capacity, and require two things.

(A1) *Independent composition.* If m and n have no common factor, then $c(mn) = c(m) + c(n)$. By the Chinese remainder theorem, two cycles of coprime lengths run from a shared tick are the same system as one cycle of length mn, so this says that capacity adds when such cycles are run together.

(A2) *Monotonicity.* $c(n+1) \geq c(n)$ for every n: a cycle with more phases never has less capacity.

Theorem 7 (uniqueness). If c satisfies (A1) and (A2), there is a constant $\kappa \geq 0$ with $c(n) = \kappa \log n$ for all n. The capacity that the proper autonomous displays of the n-cycle fail to supply is then $\kappa\Lambda(n)$.

Proof. A real function with $c(mn) = c(m) + c(n)$ for coprime m and n is called additive. Erdős proved that a monotonic additive function is a constant multiple of the logarithm [6]. By Theorem 2 the joint proper display of the n-cycle is itself a cycle, of length n/p when n is a power of p and of length n otherwise. Its capacity is $\kappa \log(n/p)$ or $\kappa \log n$, and subtracting from $\kappa \log n$ gives $\kappa\Lambda(n)$.

Both axioms are needed. Without (A2), the values of an additive function on prime powers can be chosen freely. The first supplement to [1] considers altered weights 6 log 5 at 5 and −14 log 5 at 25, with every other weight unchanged [2]. These satisfy (A1). They give the 25-cycle the capacity 6 log 5 − 14 log 5 = −8 log 5, which is negative and smaller than the capacity of the 5-cycle, so (A2) fails. Without (A1), any increasing function of n qualifies.

The constant κ fixes the unit: $\kappa = 1$ gives natural units and $\kappa = 1/\log 2$ gives bits. From here on $\kappa = 1$.

4 How fast distinctions accumulate

By Corollary 4, the room with counters of every length up to n has capacity $\psi(n) = \log L(n)$. Four known facts say how this grows.

The average. The prime number theorem is the statement $\psi(n)/n \to 1$ [5]. On average each added counter contributes one natural unit.

A ceiling with no hypothesis. $\psi(n) < 1.03883\, n$ for every n [8]. For n up to 2×10^6 the largest ratio is 1.03882, at n = 113.

The Riemann hypothesis for the room. The hypothesis is equivalent to $\psi(n) = n + O(n^{1/2+\varepsilon})$ for every $\varepsilon > 0$ [13]. Under it Schoenfeld proved $|\psi(x) - x| < \sqrt{x}\log^2 x / (8\pi)$ for $x \geq 73.2$ [7]. Since the hypothesis has been formally verified [11], the capacity of the room differs from n by less than $\sqrt{n}\log^2 n / (8\pi)$ for every $n \geq 74$. For n up to 2×10^6 the difference uses at most 80.4% of that margin, at n = 96.

n	Capacity $\psi(n)$	$\psi(n) - n$	Margin $\sqrt{n}\log^2 n / (8\pi)$
100	94.045	−5.955	8.438
10,000	10,013.397	+13.397	337.529
1,000,000	999,586.597	−413.403	7,594.41

Table 3. Capacity of the clock room in natural units.

A fixed budget. Suppose instead that the total number of positions on all the faces is at most n. The largest number of distinguishable moments is then Landau's function g(n), the largest least common multiple of positive integers with sum at most n [10]. The best room uses only

counters of pairwise coprime prime-power lengths. With n = 30 it has counters of lengths 4, 3, 5, 7 and 11 and distinguishes 4,620 moments. Deléglise and Nicolas proved that the Riemann hypothesis is equivalent to $\log g(n) < \sqrt{(\mathrm{li}^{-1}(n))}$ for every $n \geq 1$, where li^{-1} is the inverse of the logarithmic integral [9].

5 The prime-count reading as a surplus of gains

Let G be a nonnegative even function of total mass one, supported on [−b, b]. Assume that for some $k \geq 2$ it has $k - 1$ continuous derivatives, the last of them absolutely continuous. For a log-centre t, the windowed prime-power sum of [1] is

$$P(t) = \sum_{n \geq 2} \Lambda(n)\, G(\log n - t) / \sqrt{n},$$

and its benchmark is $m\, e^{t/2}$, where $m = \int G(u)\, e^{u/2}\, du$. The factor $1/\sqrt{n}$ is discussed in §6.

Proposition 8 (the benchmark is the mean gain). Write D for differentiation and put $C_k(G) = \int |(D - 1/2)(D - 3/2) \cdots (D - k + 1/2)\, G(v)|\, dv$. Then for every real t

$$\left| \sum_{n \geq 1} G(\log n - t) / \sqrt{n} - m\, e^{t/2} \right| \leq (2\zeta(k) / (2\pi)^k)\, C_k(G)\, e^{-(k - 1/2)(t - b)}.$$

Proof. Let $f(x) = G(\log x - t)\, x^{-1/2}$ for $x > 0$ and $f(x) = 0$ otherwise. It is supported in $[e^{t-b}, e^{t+b}]$, and $\int f = m\, e^{t/2}$ by the substitution $x = e^{t+u}$. The Euler–Maclaurin formula for a compactly supported function gives $|\sum_n f(n) - \int f| \leq (2\zeta(k) / (2\pi)^k) \int |f^{(k)}|$, because the periodic Bernoulli function of order k, divided by k!, is bounded by $2\zeta(k) / (2\pi)^k$ [13]. Write $x = e^u$. The operator $x\, d/dx$ is D, so $x^k f^{(k)}$ is $D(D - 1) \cdots (D - k + 1)$ applied to $G(u - t)\, e^{-u/2}$, which equals $e^{-u/2}$ times $(D - 1/2) \cdots (D - k + 1/2)\, G$ evaluated at $u - t$. Therefore

$$\int |f^{(k)}(x)|\, dx = \int e^{-(k - 1/2)u}\, |(D - 1/2) \cdots (D - k + 1/2)\, G(u - t)|\, du \leq e^{-(k - 1/2)(t - b)}\, C_k(G),$$

since $u \geq t - b$ on the support.

Corollary 9 (the residual is a surplus). With $\Lambda(1) = 0$ and E(t) bounded as in Proposition 8,

$$P(t) - m\, e^{t/2} = \sum_{n \geq 1} (\Lambda(n) - 1)\, G(\log n - t) / \sqrt{n} + E(t).$$

So the benchmark is what the sum would be if every counter contributed the average gain of one, and the residual is the windowed surplus of gains over that average.

For the window of [1], the eightfold convolution of a box of half-width 1/10, we have b = 0.8 and m = 1.0033386, and one may take k = 7. The coefficient $(2\zeta(7) / (2\pi)^7)\, C_7(G)$ is 56.087, so the bound is $56.1\, e^{-6.5(t - 0.8)}$. The actual error is -2.3×10^{-5} at t = 2, -4.6×10^{-11} at t = 4 and 5×10^{-19} at t = 6.

What the surplus measures

The left side of Corollary 9 is the residual R(t) of [1]. That paper proves, for its class of windows, that the Riemann hypothesis holds exactly when R(t) is bounded above for large t. It also proves that under the hypothesis $R(t) = -K(t) - D(t)$ for $t > b$, where K is a sum of cosines at the heights of the zeta zeros with $|K(t)| \leq K(0)$, and D is positive and decays exponentially. For the worked window $K(0) = 0.1152942$.

In the present language: the Riemann hypothesis holds exactly when counters with lengths near e^t never deliver, in the windowed sense, more than a bounded amount above the average of one

unit each. By [11] the surplus stays within about 0.1153 of zero for large t. At t = 9 it is −0.0037322.

Sign and position

In [1], K(t) is the overlap of two placements of a packet separated by t, and 2K(0) − 2K(t) is the squared distance between them. Since K(t) = −R(t) − D(t), a surplus of gains near e^t lowers the overlap. With K(0) fixed, a decrease of δ in the overlap increases the squared distance by 2δ. More new distinctions mean less overlap.

The gain Λ(n) sits at log n because log n is the whole capacity of the n-cycle, of which Λ(n) is the newly available part (the divisor sum in §1).

6 Reading and keeping a distinction

The counting theorems say how many phases a display could distinguish. A physical comparison must actually do it. The second supplement to [1] gives a heat model on a ring, the two results below and a protocol for testing whether a mark is retained [3]. We restate the results because §7 depends on what a comparison retains.

Take N ≥ 3 identical sites on an insulated ring, each exchanging heat with its two neighbours at rate κ. Place one unit of excess heat at site 0, and let $h_j(\tau)$ be the excess at site j at age τ. With $\theta_m = 2\pi m/N$ and $\lambda_m = 2\kappa(1 - \cos\theta_m)$,

$$h_j(\tau) = (1/N) \sum_{m=0}^{N-1} \exp(-\lambda_m \tau) \cos(j\theta_m).$$

Proposition 10 (a margin that lasts). For every j not divisible by N and every τ ≥ 0,

$$h_0(\tau) - h_j(\tau) \geq \max\{ \exp(-4\kappa\tau), (2/N)(1 - \cos(2\pi j/N)) \exp(-\lambda_1 \tau) \}, \quad \lambda_1 = 2\kappa(1 - \cos(2\pi/N)).$$

Proof. The difference equals $(1/N) \sum_m \exp(-\lambda_m \tau)\,(1 - \cos(j\theta_m))$, a sum of nonnegative terms. Since $\lambda_m \leq 4\kappa$, it is at least $\exp(-4\kappa\tau)\,(1/N) \sum_m (1 - \cos(j\theta_m)) = \exp(-4\kappa\tau)$. Keeping only the terms m = 1 and m = N − 1 gives the second bound.

The first bound is exact at age zero. The second governs later ages, because it decays at the slowest diffusion rate.

Take N = 25, and suppose a five-position display has already fixed the phase modulo 5. The candidate source sites are then 0, 5, 10, 15 and 20, and the comparison only has to tell the source from the other four. The nearest of them is five sites away. For j = 5, (2/25)(1 − cos 72°) = 0.0553 and $\lambda_1 = 0.0628\,\kappa$. At age 20/κ the true gap $h_0 - h_5$ is 0.0168. The second bound gives 0.0157 and the first gives 10^{-35}. A thermometer whose absolute error is below half that gap answers the question “is this candidate the source?” with one bit.

At the same age the gap between sites 0 and 1 is only 0.000786. The thermometer does not need to resolve it, because the shorter display has already ruled site 1 out. The old record determines which new distinction has to be read.

Proposition 11 (what persists). As $\tau \to \infty$, the heat vector tends to $(1/N, \ldots, 1/N) = (1/\sqrt{N})\,u$, where u is the unit vector with equal entries. Its distance from that limit is at most $\exp(-\lambda_1 \tau)$.

Proof. The component of the unit pulse along u is $(1/\sqrt{N})$ u and is conserved. Every other Fourier mode decays at a rate of at least λ_1, and the pulse minus its component along u has length less than one.

So the lasting part of one unit event on an N-site ring has size $1/\sqrt{N}$. This is the weight attached to the n-cycle in P(t). The first supplement obtains the same factor from the norm-preserving rescaling of a count [2]. The agreement is suggestive. It is not a derivation of the weight.

Keeping the answer

Fix the readings of the proper displays of a p^k-cycle, which leaves p candidate phases. Suppose a comparison writes a mark that is a function of the phase. The full gain log p is retained exactly when the p candidates receive p different marks. If all receive the same mark, nothing is retained, even if the write is irreversible [3]. The proposed test and its controls are described in [3]. It has not been performed.

7 How much a record can afford to forget

Section 6 ended with a recorder that may keep less than the full gain. Suppose that at each n the record retains r(n) of the gain and loses $\ell(n) = \Lambda(n) - r(n)$, with $0 \le \ell(n) \le \Lambda(n)$. If the marks written at $n = p^k$ sort the p candidate phases into classes, and the candidates are equally likely, then $\ell(n)$ is the average over candidates of the logarithm of the size of their class. It is zero when all marks differ and log p when all coincide. Write

$$\mathrm{Lost}(X) = \sum_{n \le X} \ell(n), \qquad W(t) = \sum_{n \ge 2} \ell(n)\, G(\log n - t) / \sqrt{n}.$$

W(t) is the amount by which the retained reading $P_{\mathrm{ret}}(t) = \sum r(n)\, G(\log n - t) / \sqrt{n}$ falls short of the ideal reading P(t) of §5.

Theorem 12 (retained capacity). Let G be nonnegative with support $[-b, b]$ and maximum $G_{\max}$, and suppose $G \ge g_0 > 0$ on $[-c, c]$ for some $0 < c \le b$. Let $\ell(n) \ge 0$ for $n \ge 2$.

(i) If $\mathrm{Lost}(X) \le A\sqrt{X}$ for every $X \ge 1$, then $W(t) \le G_{\max}\, e^{b}\, A$ for every t.

(ii) If $W(t) \le B$ for every t, then $\mathrm{Lost}(X) \le B\sqrt{X} / (g_0(1 - e^{-c}))$ for every $X \ge 1$.

In particular W is bounded exactly when $\mathrm{Lost}(X) = O(\sqrt{X})$.

Proof. (i) Only n between e^{t-b} and e^{t+b} contribute, and for these $1/\sqrt{n} \le e^{-(t-b)/2}$. Hence $W(t) \le G_{\max}\, e^{-(t-b)/2}\, \mathrm{Lost}(e^{t+b}) \le G_{\max}\, e^{-(t-b)/2}\, A\, e^{(t+b)/2}$. (ii) For $X > 0$ take $t = \log X - c$. The n with $X e^{-2c} < n \le X$ have $|\log n - t| \le c$ and $1/\sqrt{n} \ge X^{-1/2}$, so $B \ge W(t) \ge g_0\, X^{-1/2}\, [\mathrm{Lost}(X) - \mathrm{Lost}(X e^{-2c})]$. Apply this at X, $X e^{-2c}$, $X e^{-4c}$, … and add. The right sides form a geometric series with ratio e^{-c}, and Lost vanishes below 2.

Corollary 13 (tracking the ideal reading). Let $0 \le \ell(n) \le \Lambda(n)$, and let G be one of the windows of [1]. By [1] and [11] the ideal residual $P(t) - m\, e^{t/2}$ is bounded. The retained residual $P_{\mathrm{ret}}(t) - m\, e^{t/2}$ is therefore bounded exactly when $\mathrm{Lost}(X) = O(\sqrt{X})$. For the worked window $G_{\max} = 151/63$ and $b = 0.8$, so a loss of at most $A\sqrt{X}$ lowers the reading by at most 5.34 A. With $c = 0.2$, a reading that is never more than B low has lost at most $4.67\, B\sqrt{X}$.

Three recorders

Forgetting sparsely. Lose the whole gain at the first prime after each fourth power. Infinitely many primes are forgotten, yet Lost(X) ≤ $X^{1/4}$ log X, and in fact Lost(X) ≤ 0.80 √X for every X. By Theorem 12 the reading is never more than 4.27 low. For t up to 15 the largest shortfall is 1.65.

Forgetting a shrinking share. Lose the share $1/\sqrt{n}$ of every gain. Then Lost(X) is close to $2\sqrt{X}$, and W(t) tends to one by the prime number theorem: 0.985 at t = 4, 0.998 at t = 8 and 1.000 at t = 12. The reading settles one unit low.

Forgetting one bit at every prime. Now Lost(X) = π(X) log 2, where π(X) counts the primes up to X. This outgrows √X, and the shortfall is unbounded: 1.14 at t = 4, 4.68 at t = 8 and 23.3 at t = 12.

The budget √X is of the same order, up to logarithms, as the fluctuation of the room's capacity about n in §4. A record can afford to forget an amount of the same order as those fluctuations. It cannot afford a fixed loss at every prime.

8 What is classical and what is added

Most of the individual facts used here are classical, and several of the constructions first appeared in the second supplement [3]. Table 4 separates them from what this paper contributes.

Result	Source	What this paper adds
log lcm(1, …, n) = ψ(n); $\sum_{d \mid n} \Lambda(d) = \log n$	Classical [4, 5]	Nothing
Lemma 1, Theorem 2: the gain of a cycle is Λ(n)	[3]	Restated with proof
Theorem 3: any set of counters	[3]	Restated with proof
Proposition 5, Corollary 6	Elementary	The three kinds; the smallest positive gain
Theorem 7: uniqueness	Erdős [6], applied in [3]	Restated; the control at 5 as a failure of monotonicity
Section 4	Known [7, 8, 9, 13]	Translation into capacities
Proposition 8, Corollary 9	Euler–Maclaurin summation	The bound for a general smooth window, with an explicit coefficient
Propositions 10 and 11	[3]	Restated; the reading in terms of candidate sites
Theorem 12, Corollary 13	New here	The retained-capacity criterion with explicit constants
Proposition 14	The divisor sum [4]	The returning counters name the tick; Conjecture 15

Table 4. Status of the results.

Nothing here proves or reproves the Riemann hypothesis. Sections 1 to 3 and 6, Theorem 12 and Proposition 14 use no hypothesis at all.

9 What VERSF would have to supply

VERSF distinguishes reversible changes of state from an irreversible commitment that leaves a persistent outcome [12]. In that language a cycle carries its phase reversibly, a comparison such as that of §6 commits one bit, and Theorem 2 says which cycles have a distinction left to commit.

The results above fix the arithmetic: where gains occur, how large they must be, and how their surplus over the mean relates to the overlap in [1]. Three things are not supplied. The first is the periods: a process within the framework that generates cycles of the lengths that matter. The second is the shared tick and the comparison: which physical handshake distinguishes, for example, phases 0 and 5 of a 25-cycle before the mark is read. The third is the window, which here is the observer's choice. Theorem 12 adds a quantitative requirement on whatever process commits the distinctions: its total loss through X must be $O(\sqrt{X})$.

Recurrence and the present

One more piece of arithmetic bears on how a present moment could be tied to earlier facts. Start one counter of every length together at phase 0 and call that tick 0.

Proposition 14 (the returning counters name the tick). At tick $T \geq 1$ the counters at phase 0 are exactly those whose length divides T. Their gains add up to $\log T$, and the longest of them has length T. So the set of counters that are returning determines T, and their total gain is the capacity needed to name one tick among the first T.

Proof. The n-counter is at phase 0 exactly when n divides T. The sum of the gains is the divisor sum of §1.

At tick 30 the returning counters have lengths 1, 2, 3, 5, 6, 10, 15 and 30. Those with a positive gain are 2, 3 and 5, and $\log 2 + \log 3 + \log 5 = \log 30$.

Two rates follow. An n-counter returns once in every n ticks, so short counters return often and long ones rarely. Among counters of length at most n, the average number returning at a tick is $1 + 1/2 + \ldots + 1/n$, which is about $\log n + 0.58$. In the other direction, Corollary 4 shows that the counters of length at most n give different joint readings at every tick from 0 to $L(n) - 1$, and $\log L(n) = \psi(n)$ is close to n by §4. Telling tick T apart from every earlier tick therefore needs counters only up to a length of about $\log T$. For example $L(16) = 720{,}720$ and $L(17) = 12{,}252{,}240$, so at tick one million the counters of length at most 17 already single out the present.

A return, by itself, is only the same state again. It becomes an opportunity to recognise a match if an earlier comparison left a mark that is still held, and if that mark takes part in forming the next commitment. The first condition is the subject of Theorem 12, which says how much of those marks may be lost. The second is a claim about the dynamics, and it is where the VERSF notion of fact momentum would enter [14]. We state it as a conjecture.

Conjecture 15 (anchoring and discrimination). The present may be anchored by short cycles that repeatedly reconnect it to retained facts, while longer cycles allow this moment to be distinguished from earlier moments that looked the same.

Proposition 14 and the two rates show that the opportunities exist and how often they arise. Whether the dynamics use them is open.

Status

Until the three items above are derived and the conjecture is tested, the statements of this paper are theorems about a model and predictions for a tabletop test. Their identification with VERSF records is a proposal.

Verification

The listing below uses only the Python standard library and runs in about a second. It checks Theorems 2 and 3 and Proposition 5 on small cases, the negative capacity in §3, the bounds of §4 up to n = 20,000, the budget example, the coefficient of Proposition 8 and its bound at t = 2, 3 and 4, Propositions 10 and 11 with the candidate-site gaps on 25 sites, the constants and first recorder of Theorem 12, and Proposition 14 for ticks up to 2,000. The high-precision values quoted in §5, the values to 2×10^6 in §4 and the shortfalls for t up to 15 in §7 were computed separately.

```
from math import gcd, lcm, log, exp, sqrt, pi, cos, sinh, comb, factorial
from itertools import combinations
from fractions import Fraction

def mangoldt(n):
    p = next(q for q in range(2, n + 1) if n % q == 0)
    while n % p == 0:
        n //= p
    return log(p) if n == 1 else 0.0

LAM = [0.0, 0.0] + [mangoldt(n) for n in range(2, 20001)]

# Theorems 2 and 3, Proposition 5
for n in range(2, 257):
    proper = [d for d in range(1, n) if n % d == 0]
    readings = {tuple(j % d for d in proper) for j in range(n)}
    assert abs(log(n) - log(len(readings)) - LAM[n]) < 1e-12
    chain = all(c % d == 0 for d, c in zip(proper, proper[1:]))
    assert chain == (LAM[n] > 0)          # totally ordered: prime powers only
    assert (len(proper) == 1) == (LAM[n] == log(n))    # atomic: primes only
for S in combinations(range(1, 31), 3):
    Q = [q for q in range(2, 31) if LAM[q] and any(d % q == 0 for d in S)]
    assert abs(log(lcm(*S)) - sum(LAM[q] for q in Q)) < 1e-9

# Section 3: the altered weights give a 25-cycle negative capacity (Theorem 7)
assert 6 * log(5) > 0 > 6 * log(5) - 14 * log(5)

# Section 4: capacity of the room, and the best room on a budget of 30
psi, top = 0.0, 0.0
for n in range(1, 20001):
    psi += LAM[n]
    top = max(top, psi / n)
    if n >= 74:
        assert abs(psi - n) < sqrt(n) * log(n) ** 2 / (8 * pi)
assert 1.0388 < top < 1.03883
best = [1] * 31
for p in (2, 3, 5, 7, 11, 13, 17, 19, 23, 29):
    new, q = best[:], p
```

```
    while q <= 30:
        for s in range(q, 31):
            new[s] = max(new[s], best[s - q] * q)
        q *= p
    best = new
assert max(best) == 4620 == 4 * 3 * 5 * 7 * 11

# Proposition 8: the benchmark is the mean gain
a, b = 0.1, 0.8
def G(x):
    u = (x + b) / (2 * a)
    if not 0 < u < 8:
        return 0.0
    return sum((-1) ** k * comb(8, k) * (u - k) ** 7
               for k in range(int(u) + 1)) / (factorial(7) * 2 * a)
m = (sinh(a / 2) / (a / 2)) ** 8
C7 = 0.0                # C_7(G), piece by piece, with u = 5x + 4 and D = 5 d/du
for i in range(8):
    q = [Fraction(0)] * 8
    for k in range(i + 1):
        for r in range(8):
            term = comb(8, k) * comb(7, r) * Fraction(-k) ** (7 - r)
            q[r] += (-1) ** k * term
    q = [5 * v / factorial(7) for v in q]
    for j in range(7):
        dq = [5 * (r + 1) * q[r + 1] for r in range(7)] + [Fraction(0)]
        q = [d - (j + Fraction(1, 2)) * v for d, v in zip(dq, q)]
    def size(s, steps=1000):
        u = i + Fraction(s, steps)
        return abs(float(sum(v * u ** r for r, v in enumerate(q))))
    inner = sum((4 if s % 2 else 2) * size(s) for s in range(1, 1000))
    simpson = size(0) + size(1000) + inner
    C7 += simpson / 3000 / 5
zeta7 = sum(k ** -7.0 for k in range(1, 2000))
coefficient = 2 * zeta7 / (2 * pi) ** 7 * C7
assert 56.08 < coefficient < 56.1
for t in (2, 3, 4):
    lo, hi = int(exp(t - b)), int(exp(t + b)) + 1
    mean = sum(G(log(n) - t) / sqrt(n) for n in range(max(lo, 1), hi + 1))
    assert abs(mean - m * exp(t / 2)) < coefficient * exp(-6.5 * (t - b))
    assert abs(mean - m * exp(t / 2)) < 3e-5

# Propositions 10 and 11: the heat comparison on 25 sites
N = 25
lam = [2 * (1 - cos(2 * pi * k / N)) for k in range(N)]
def h(j, age):
    modes = (exp(-lam[k] * age) * cos(2 * pi * j * k / N) for k in range(N))
    return sum(modes) / N
for age in (0, 1, 5, 20, 50, 100):
    for j in range(1, N):
        gap = h(0, age) - h(j, age)
        slow = (2 / N) * (1 - cos(2 * pi * j / N)) * exp(-lam[1] * age)
        assert gap >= max(exp(-4 * age), slow) - 1e-12
assert abs(sqrt(sum(h(j, 400) ** 2 for j in range(N))) - 1 / sqrt(N)) < 1e-9
gaps = {j: h(0, 20) - h(j, 20) for j in range(1, N)}
assert 0.0167 < min(gaps[j] for j in (5, 10, 15, 20)) < 0.0168 < gaps[10]
assert 0.00078 < gaps[1] < 0.00079

# Theorem 12: forgetting the first prime after each fourth power
gmax, c = G(0.0), 0.2
assert abs(gmax - 151 / 63) < 1e-12 and gmax * exp(b) < 5.335
assert 1 / (G(c) * (1 - exp(-c))) < 4.67
forgotten, k = set(), 2
while k ** 4 < 20000:
    forgotten.add(next(n for n in range(k ** 4 + 1, 20001) if LAM[n] == log(n)))
    k += 1
lost, A = 0.0, 0.0
for n in range(2, 20001):
    if n in forgotten:
        lost += LAM[n]
        A = max(A, lost / sqrt(n))
assert A < 0.8
```

```
for t in (3, 5, 7, 9):
    short = sum(LAM[n] * G(log(n) - t) / sqrt(n) for n in forgotten)
    assert short <= gmax * exp(b) * A

# Proposition 14: the counters returning at tick T name the tick
for T in range(1, 2001):
    back = [n for n in range(1, T + 1) if T % n == 0]
    assert max(back) == T
    assert abs(sum(LAM[n] for n in back) - log(T)) < 1e-9
assert lcm(*range(1, 17)) == 720720 and lcm(*range(1, 18)) == 12252240
assert abs(sum(1 / d for d in range(1, 18)) - (log(17) + 0.58)) < 0.03
print("all checks pass")
```

References

[1] K. Taylor. Prime Counts at Fourth Powers and the Riemann Zeros (2026).

[2] K. Taylor. What the VERSF Ring Explains About Its Silent Frequencies. First supplement to [1] (2026), §25 and Proposition 24.

[3] K. Taylor. Prime Powers and the Formation of Distinguishable Records. Second supplement to [1] (2026).

[4] T. M. Apostol. Introduction to Analytic Number Theory. Springer (1976), Theorem 2.10.

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

[6] P. Erdős. On the distribution function of additive functions. Annals of Mathematics 47 (1946), 1–20.

[7] L. Schoenfeld. Sharper bounds for the Chebyshev functions θ(x) and ψ(x). II. Mathematics of Computation 30 (1976), 337–360.

[8] J. B. Rosser and L. Schoenfeld. Approximate formulas for some functions of prime numbers. Illinois Journal of Mathematics 6 (1962), 64–94, Theorem 12.

[9] M. Deléglise and J.-L. Nicolas. The Landau function and the Riemann hypothesis. arXiv:1907.07664 (2019).

[10] E. Landau. Über die Maximalordnung der Permutationen gegebenen Grades. Archiv der Mathematik und Physik (3) 5 (1903), 92–103.

[11] Lean FRO and Mathlib Initiative. Lean verification of the Riemann Hypothesis. Joint statement, 21 November 2026. leanprover-community.github.io/blog/posts/riemann-hypothesis/

[12] K. Taylor. From Necessary Facts to Physical Structure (2026), §§4.6 and 12.5.

[13] H. L. Montgomery and R. C. Vaughan. Multiplicative Number Theory I: Classical Theory. Cambridge University Press (2007), Chapters 13 and 15 and Appendix B.

[14] K. Taylor. Fact Momentum, Record Overlap and the Emergence of Stable Aggregates in VERSF (2026).