SEVEN PRIME DIVISORS IN ODD DISTINCT COVERING SYSTEMS

MICHAEL SCHROEDER
INDEPENDENT RESEARCHER

Abstract. More than seventy-five years after Erdős introduced distinct covering systems in 1950, the Erdős–Selfridge odd covering problem (Erdős Problem 7) remains open: does there exist a finite covering system with pairwise distinct odd moduli greater than one? We prove that the least common multiple N of the moduli in any such system has at least seven distinct prime divisors, without any squarefreeness or prime-power-height hypothesis. The proof passes to a residual product of prime trees on which all prime-power classes are absent, then charges every remaining cylinder at the last digit that it fixes. A normalized fibre distortion and a reverse causal comparison for increasing supermodular functionals reduce the estimate to five exact rational charges. Their sum is $0.914634\ldots < 1$, giving the contradiction uniformly in all prime-power heights. A complete Lean 4 formalization and an independent exact rational checker accompany the argument.

1. Introduction

A finite family

$$\mathcal{C} = \{a_\nu \pmod{n_\nu} : 1 \le \nu \le m\}$$

is a *covering system* if every integer belongs to at least one of its classes. It is *distinct* when the moduli n_ν are pairwise distinct, and it is *odd* when every n_ν is odd. We use $\omega(N)$ for the number of distinct prime divisors of N. Balister's recent survey [1] provides a broad guide to the classical questions and to the modern probabilistic methods for covering systems.

The study of distinct covering systems began with Erdős in 1950 [8]. The enduring Erdős–Selfridge question asks whether a distinct covering system can have every modulus odd and greater than one. Berger, Felzenbaum, and Fraenkel obtained the general necessary condition [4, 5]

$$\prod_{p \mid N} \frac{p-1}{p-2} - \sum_{p \mid N} \frac{1}{p-2} > 2. \tag{1.1}$$

Date: 1 September 2026.
2020 *Mathematics Subject Classification.* Primary 11B25; Secondary 11B75, 05D40.
Key words and phrases. Erdős–Selfridge odd covering problem, Erdős Problem 7, covering systems, distinct odd moduli, prime divisors, product cylinders, supermodular domination, Lean 4 formalization.

In particular, their condition implies $\omega(N) \geq 5$. Indeed, on writing $x_p = (p-2)^{-1}$, the left side of (1.1) becomes $\prod_p (1+x_p) - \sum_p x_p$, which is nondecreasing in every x_p and under the addition of a further positive variable. With at most four odd primes it is therefore at most its value at $3, 5, 7, 11$, namely

$$\frac{86}{45} < 2.$$

Much sharper conclusions were known under a squarefree hypothesis. Simpson and Zeilberger proved that a hypothetical system with squarefree moduli would require at least eighteen prime divisors [15]; Guo and Sun raised this to twenty-two [9]; and Balister, Bollobás, Morris, Sahasrabudhe, and Tiba ultimately ruled out the squarefree case altogether [2]. In the unrestricted setting, Hough's solution of the minimum-modulus problem introduced a powerful probabilistic sieve [11]. Hough and Nielsen used a refinement to show that every distinct covering contains a modulus divisible by 2 or 3 [12]. A further prime-by-prime distortion argument of Balister et al. shows that, if N is the least common multiple of a distinct covering, then $2 \mid N$, or $9 \mid N$, or $15 \mid N$ [3, Theorem 1.4].

Recent complementary work treats nearby questions. Odd systems with one repeated modulus are studied in [10, 6]; McNew and Setty investigate covering numbers and their density [13]; and Mian and Siddique give a kernel-checked exclusion of least common multiples up to 10^4 [14]. These works address variants, density, or least-common-multiple size rather than an exponent-uniform lower bound for $\omega(N)$.

Theorem 1.1 (Seven-prime support theorem)**.** *Let $\mathcal{C}$ be a finite distinct covering system with odd moduli $n_\nu > 1$, and put*

$$N = \operatorname{lcm}_{1 \leq \nu \leq m} n_\nu.$$

Then

$$\omega(N) \geq 7.$$

Theorem 1.1 strengthens the rank-five consequence of (1.1) while making no squarefree assumption. Its proof retains the prime-by-prime philosophy of [11, 3], but the finite tree structure is used much more rigidly. The Chinese remainder theorem turns a congruence class into a prefix cylinder in a product of prime trees, and distinct moduli give uniqueness of full depth vectors. We select a residual product avoiding every cylinder whose modulus is a prime power. A normalized distortion is then applied one digit at a time, while a reverse causal comparison replaces the surviving prefix events by independent gates. Finally, full-vector uniqueness bounds the number of cylinders *ending* at a digit. The resulting five block charges have exact sum

$$\frac{10123552563977963}{11068417920000000} < 1. \tag{1.2}$$

The word "ending" is essential. We never project cylinders that still fix future coordinates and then claim uniqueness for their projections. At the

stage where a cylinder is counted, all of its future depths are zero, so the depth data already exposed are its full depth vector.

The bottom ternary block also requires care: we use only unconditional prefix caps, never a false conditional product assertion. Unknown prime-power heights are removed by stochastic and increasing-convex domination, so the certificate is uniform in the exponents. Section 8 records the independent Lean formalization. Throughout, $x_+ = \max(x, 0)$; all probability spaces are finite until the infinite-height majorants of Section 6 are introduced.

2. CRT cylinders and a residual product

Write

$$N = \prod_{j=1}^{r} \pi_j^{a_j}, \qquad 3 \le \pi_1 < \pi_2 < \cdots < \pi_r. \tag{2.1}$$

By the Chinese remainder theorem,

$$\mathbb{Z}/N\mathbb{Z} \cong \prod_{j=1}^{r} \mathbb{Z}/\pi_j^{a_j}\mathbb{Z}. \tag{2.2}$$

Represent the j-th factor by its rooted π_j-ary digit tree of height a_j, reading base-π_j digits from least significant to most significant. Thus congruence modulo π_j^e fixes exactly the first e digits. If

$$n_\nu = \prod_{j=1}^{r} \pi_j^{e_{\nu j}}, \qquad 0 \le e_{\nu j} \le a_j,$$

then the lift of $a_\nu \pmod{n_\nu}$ to $\mathbb{Z}/N\mathbb{Z}$ is a product cylinder: in block j, it fixes a prefix of length $e_{\nu j}$. We call $e_\nu = (e_{\nu 1}, \ldots, e_{\nu r})$ its *depth vector*.

Lemma 2.1 (Full-vector uniqueness)**.** *For every nonzero depth vector e, at most one cylinder of $\mathcal{C}$ has depth vector e.*

Proof. The modulus belonging to e is $\prod_j \pi_j^{e_j}$. Two cylinders with the same depth vector therefore have the same modulus, contrary to the assumption that the moduli are distinct. □

Order digit coordinates block by block. A nonempty depth vector is assigned to the last coordinate that it fixes. Thus a cylinder assigned to depth $e \ge 1$ in block j has depth e in that block and depth zero in every block after j. A cylinder is *pure* when its depth vector is supported on only one block. Lemma 2.1 implies that, in each prime tree, there is at most one pure forbidden prefix at each positive depth.

We now remove all such prefixes.

Lemma 2.2 (Complete avoidance subtree)**.** *Let T be a full rooted p-ary tree of height a, and suppose that at most one vertex is forbidden at each positive depth. For every $q \le p - 1$, the tree contains a full rooted q-ary subtree of height a containing none of the forbidden vertices.*

Proof. Construct the subtree level by level. At a retained vertex, exclude the child that is the forbidden vertex at the next depth if that child exists; otherwise exclude an arbitrary child. At least $p - 1 \geq q$ children remain, from which we retain exactly q. Doing this at every retained vertex and every level gives the required subtree. A forbidden vertex whose parent was not retained is already absent. □

The first block admits a more efficient, non-product selection.

Lemma 2.3 (Bottom-packed ternary selection)**.** *Let T be a full ternary tree of height $a \geq 1$, with at most one forbidden vertex at each positive depth. There is a set Y_a of leaves, none at or below a forbidden vertex, such that*

$$|Y_a| = S_a := 3^{a-1} + 1, \tag{2.3}$$

and every prefix of depth d contains at most

$$2 \cdot 3^{a-1-d} \quad (1 \leq d < a), \qquad 1 \quad (d = a) \tag{2.4}$$

selected leaves.

Proof. Consider the 3^{a-1} vertices at depth $a-1$. A forbidden vertex at depth $d < a$ has exactly 3^{a-1-d} descendants at that level. Hence the forbidden vertices at depths $1, \ldots, a-1$ together exclude at most

$$\sum_{d=1}^{a-1} 3^{a-1-d} = \frac{3^{a-1} - 1}{2}$$

parents. At least $(3^{a-1} + 1)/2$ parents survive. Choose exactly that many, and below each chosen parent retain two leaves. If the single forbidden vertex at depth a lies below a chosen parent, choose the other two leaves there. Thus all selected leaves avoid every forbidden prefix and their number is twice the number of chosen parents, proving (2.3). A prefix at depth $d < a$ contains 3^{a-1-d} possible chosen parents and each contributes at most two leaves; a terminal prefix contains at most one. This proves (2.4). □

Suppose from now on, toward a contradiction, that $r \leq 6$. Since a covering system is nonempty and every modulus is greater than one, we also have $r \geq 1$ and $a_1 \geq 1$. Let

$$(p_1^*, \ldots, p_6^*) = (3, 5, 7, 11, 13, 17). \tag{2.5}$$

Because the primes in (2.1) are increasing, $\pi_j \geq p_j^*$. In the first block choose any full ternary subtree and apply Lemma 2.3 to the pure forbidden prefixes that remain inside it. In each block $j \geq 2$, apply Lemma 2.2 with

$$q_j = p_j^* - 1. \tag{2.6}$$

This is possible because $q_j \leq \pi_j - 1$.

Let T_j be the resulting complete q_j-ary tree in block j, and set

$$R = Y_{a_1} \times \prod_{j=2}^{r} T_j. \tag{2.7}$$

Every pure covering cylinder misses R. Put the uniform probability measure ν on R, using uniform measure on the leaf sets of the factors. A later-block digit is conditionally uniform on q_j retained children. In the first block, a prescribed prefix has ν-mass at most

$$s_{a,0} = 1, \qquad s_{a,d} = \frac{2 \cdot 3^{a-1-d}}{3^{a-1}+1} \quad (1 \le d < a), \qquad s_{a,a} = \frac{1}{3^{a-1}+1}, \tag{2.8}$$

where $a = a_1$. Notice that these bounds are unconditional; we shall not pretend that the uniform law on Y_a is a digitwise product law.

3. The normalized distortion

The whole first-block leaf is regarded as initial history. We then process all retained digit coordinates in blocks $2, \dots, r$, in lexicographic order by block and depth. For a stage s, let $B_s \subseteq R$ be the union of the nonpure cylinders assigned to that stage, intersected with R. Since an assigned cylinder fixes no future digit, B_s is measurable immediately after the current digit is exposed.

The following update is the engine of the proof.

Proposition 3.1 (Fibrewise distortion)**.** *Let B be measurable after a current finite-valued coordinate is exposed, let α be its conditional probability over each fibre of the prior history, and fix $0 \le \delta < 1$. Multiply the density on each fibre by*

	B	B^c
$\alpha \le \delta$	0	$(1-\alpha)^{-1}$
$\alpha > \delta$	$(\alpha-\delta)/(\alpha(1-\delta))$	$(1-\delta)^{-1}$.

(3.1)

Then the new fibre mass of B is

$$f_\delta(\alpha) = \frac{(\alpha-\delta)_+}{1-\delta}. \tag{3.2}$$

The average multiplier on every fibre is one, and every multiplier is at most $(1-\delta)^{-1}$.

Proof. If $\alpha \le \delta$, the average multiplier is $(1-\alpha)(1-\alpha)^{-1} = 1$, and the new mass of B is zero. If $\alpha > \delta$, the average is

$$\alpha \frac{\alpha-\delta}{\alpha(1-\delta)} + (1-\alpha)\frac{1}{1-\delta} = 1,$$

and the new mass of B is $(\alpha-\delta)/(1-\delta)$. In the first case, $(1-\alpha)^{-1} \le (1-\delta)^{-1}$; the remaining multiplier bounds are immediate. □

Start with $\mu_0 = \nu$ and apply Proposition 3.1 at every stage. Fibre normalization preserves the entire preceding marginal. In particular, every later update preserves every set measurable at an earlier stage. If μ_* is the final measure and μ_{s-1} is the measure just before stage s, then

$$\mu_*\left(\bigcup_s B_s\right) \le \sum_s \mu_*(B_s) = \sum_s \mathbb{E}_{\mu_{s-1}} f_{\delta_s}(\alpha_s). \tag{3.3}$$

Thus a total charge strictly smaller than one leaves a point of R outside all the nonpure cylinders.

There is a second consequence. Immediately before a digit is updated, the density with respect to ν depends only on the preceding coordinates; hence a retained digit in a q-ary block is conditionally uniform. Put $\delta = t/q$, where $0 \le t < q$. On a fixed fibre, B_s is a union of current symbols. Formula (3.1) shows, separately for symbols inside and outside that union, that after the update

$$\mathbb{P}(Z_s = y \mid Z_{<s}) \le \frac{1}{q-t}. \tag{3.4}$$

Indeed, an outside symbol receives mass at most $q^{-1}(1-t/q)^{-1} = 1/(q-t)$, and an inside symbol receives no more. Every later update preserves the joint marginal through stage s, so (3.4) remains true thereafter.

4. Supermodular domination

For a finite set L, a function $F : 2^L \to \mathbb{R}$ is *supermodular* if

$$F(A \cap B) + F(A \cup B) \ge F(A) + F(B) \qquad (A, B \subseteq L). \tag{4.1}$$

We shall use such functionals to count simultaneously surviving cylinder labels.

Lemma 4.1 (Bernoulli rearrangement)**.** *Let $(I_\lambda)_{\lambda \in L}$ be arbitrary Bernoulli variables satisfying $\mathbb{P}(I_\lambda = 1) \le \rho_\lambda$. If $F : 2^L \to \mathbb{R}$ is increasing and supermodular, then*

$$\mathbb{E}F(\{\lambda : I_\lambda = 1\}) \le \int_0^1 F(\{\lambda : u \le \rho_\lambda\})\, du. \tag{4.2}$$

Proof. First enlarge each marginal to its cap: whenever the current marginal of I_λ is $p < \rho_\lambda$, independently change a zero to a one with probability $(\rho_\lambda - p)/(1-p)$. This leaves all other coordinates unchanged and cannot decrease an increasing functional. We may therefore prescribe the marginals ρ_λ.

Among all probability distributions on 2^L with these marginals, choose one maximizing $\mathbb{E}F(S)$, and among those choose one maximizing $\mathbb{E}|S|^2$. If two incomparable sets A, B have positive mass, move the same sufficiently small mass from each of A, B to each of $A \cap B, A \cup B$. Every one-point marginal is unchanged. By (4.1), the first objective does not decrease, whereas

the second increases by

$$2|A \setminus B|\,|B \setminus A| > 0.$$

This contradicts the choice of the distribution. Its support is therefore a chain. A chain-supported Bernoulli law with the prescribed marginals is the common-uniform coupling $I_\lambda = \mathbf{1}_{\{U \le \rho_\lambda\}}$, $U \sim \mathrm{Unif}[0,1]$, which is precisely the right-hand side of (4.2). □

We next turn conditional atom bounds into gates. A label may impose either no condition at a coordinate or one required symbol there.

Theorem 4.2 (Reverse causal gate comparison)**.** *Let $Z_1, \dots, Z_m$ be a finite-valued adapted process such that, for every positive-probability history and every current symbol,*

$$\mathbb{P}(Z_i = y \mid Z_1, \dots, Z_{i-1}) \le r_i. \tag{4.3}$$

Let L be a finite labelled family of cylinder conditions on this process, and let $A(Z) \subseteq L$ be the set of satisfied labels. For every increasing supermodular $F : 2^L \to \mathbb{R}$,

$$\mathbb{E}F(A(Z)) \le \mathbb{E}F(A(G)), \tag{4.4}$$

where $G_1, \dots, G_m$ are independent Bernoulli variables with $\mathbb{P}(G_i = 1) = r_i$, and a label is placed in $A(G)$ when $G_i = 1$ at every coordinate at which that label fixes a symbol.

Proof. We eliminate coordinates in reverse order. Consider the final coordinate, and fix the preceding history. Let A be the set of labels still eligible at this point. Partition it as

$$A = A_0 \sqcup \bigsqcup_y A_y,$$

where A_0 contains labels not fixing the final coordinate and A_y contains labels requiring the symbol y. The set function

$$H(J) = F\left(A_0 \cup \bigcup_{y \in J} A_y\right)$$

on the symbol set is increasing and supermodular, because the bundles are disjoint and the bundle map preserves unions and intersections. Apply Lemma 4.1 to the mutually exclusive indicators $\mathbf{1}_{\{Z_m = y\}}$. Each has conditional mean at most r_m; the common-uniform upper coupling makes all symbol indicators equal to one Bernoulli gate G_m. Conditional on the preceding history, the resulting upper bound is

$$(1 - r_m)F(A_0) + r_m F(A). \tag{4.5}$$

As a function of the eligible label set A, expression (4.5) is again increasing and supermodular: it is the nonnegative linear combination of $F(A \cap L_0)$ and $F(A)$, where L_0 is the fixed set of labels not constraining the final coordinate. Consequently the same argument applies to coordinate $m - 1$,

then to $m-2$, and so on. The fresh common gate introduced at each step may be taken independent of all the others. This gives (4.4). □

The first residual block requires a different final step.

Corollary 4.3 (Hybrid prefix–causal comparison)**.** *Suppose each label $\lambda \in L$ additionally requires an event E_λ in an initial random variable, with*

$$\mathbb{P}(E_\lambda) \le s_{d(\lambda)}, \qquad 1 = s_0 \ge s_1 \ge \cdots \ge s_a \ge 0. \tag{4.6}$$

Conditional on that initial variable, let the remaining adapted coordinates satisfy (4.3). Then the expectation of every increasing supermodular functional of the active labels is at most its expectation in the following model: the later coordinates are replaced by the independent gates of Theorem 4.2, and independently one takes a run $D \in \{0, \dots, a\}$ with

$$\mathbb{P}(D \ge d) = s_d, \tag{4.7}$$

declaring the initial requirement of λ satisfied whenever $d(\lambda) \le D$.

Proof. Eliminate all later coordinates in reverse order exactly as in the proof of Theorem 4.2. The continuation functional of the labels satisfying their initial events remains increasing and supermodular. Apply Lemma 4.1 directly to the labelled event indicators $\mathbf{1}_{E_\lambda}$, using their unconditional caps in (4.6). In the common-uniform coupling, a label of depth d survives precisely when $U \le s_d$. Since the caps are nonincreasing, this is the run in (4.7). This last rearrangement uses only the marginal prefix bounds; it asserts no conditional digit bounds for the initial variable. □

For later use, note that for every integer $t \ge 0$,

$$F_t(A) = (|A| - t)_+ \tag{4.8}$$

is increasing and supermodular. Indeed, its discrete marginal increment on adjoining a new element is zero until the cardinality reaches t and one thereafter, so the increments are nondecreasing.

5. The block charge

Fix, for every later block j, an integer threshold

$$0 \le t_j \le q_j - 2, \qquad r_j = \frac{1}{q_j - t_j}, \tag{5.1}$$

and use $\delta_j = t_j/q_j$ at every digit of that block.

Consider depth e in block j. On a fixed preceding-history fibre, let M be the number of eligible cylinder labels ending at the current digit. Each such label requires one current symbol. Consequently the union $B_{j,e}$ occupies at most $\min(M, q_j)$ symbols, and its conditional mass α satisfies

$$\alpha \le \frac{\min(M, q_j)}{q_j}.$$

The distortion cost is therefore bounded by

$$f_{t_j/q_j}(\alpha) \le \frac{(\min(M, q_j) - t_j)_+}{q_j - t_j} \le \frac{(M - t_j)_+}{q_j - t_j}. \tag{5.2}$$

The second inequality uses only $\min(M, q_j) \le M$.

Apply Corollary 4.3 to the labels ending at this stage and to the hinge (4.8). Let D_{1,a_1} be the first run obtained from the caps (2.8). For $2 \le i < j$, let D_{i,a_i} be the initial run of open gates in block i; then

$$\mathbb{P}(D_{i,a_i} \ge d) = r_i^d \mathbf{1}_{\{d \le a_i\}}. \tag{5.3}$$

All these comparison runs are independent. Put

$$K_j = \prod_{i<j} (D_{i,a_i} + 1). \tag{5.4}$$

In the gate model, the first $e-1$ gates of the current block must all be open before any label ending at depth e can survive. This event has probability r_j^{e-1}. Once it occurs, a surviving label has an earlier depth vector

$$(d_1, \dots, d_{j-1}), \qquad 0 \le d_i \le D_{i,a_i}.$$

The all-zero vector would make the cylinder pure and has already been removed. For every other vector, Lemma 2.1 gives at most one cylinder: together with current depth e and zero future depths, this is the cylinder's full depth vector. Thus at most $K_j - 1$ labels survive. Combining this fact with (5.2) gives the depth charge

$$C_{j,e} \le \frac{r_j^{e-1}}{q_j - t_j} \mathbb{E}[(K_j - 1 - t_j)_+]. \tag{5.5}$$

Summing over $1 \le e \le a_j$ and enlarging the geometric sum to infinity yields the exponent-free block bound

$$\boxed{C_j \le \frac{\mathbb{E}[(K_j - (t_j + 1))_+]}{q_j - t_j - 1}.} \tag{5.6}$$

There is no first-block charge: after pure cylinders have been removed, no cylinder can end in block 1.

Remark 5.1 (The ending-only principle). The count $K_j - 1$ is not a bound on arbitrary cylinders visible after a projection. Such a bound would be false because different future depths can collapse to the same exposed prefix. Here every counted cylinder ends at the current digit, so every future depth is zero. This is why Lemma 2.1 applies without loss of information.

6. Uniform removal of the prime-power heights

It remains to replace the finite runs in (5.4) by fixed infinite majorants. From (2.8), for every integer $0 \le k < a$,

$$\mathbb{E}[(D_{1,a} - k)_+] = \sum_{d=k+1}^{a} s_{a,d} \tag{6.1}$$

$$= \frac{3^{a-k-1}}{3^{a-1}+1} < \frac{1}{3^k}. \tag{6.2}$$

For $k \ge a$, the left side is zero. Define an infinite run D_1 by

$$\mathbb{P}(D_1 \ge d) = \frac{2}{3^d} \qquad (d \ge 1). \tag{6.3}$$

Then

$$\mathbb{E}[(D_1 - k)_+] = \sum_{d=k+1}^{\infty} \frac{2}{3^d} = \frac{1}{3^k}.$$

The stop-loss inequalities at integer k imply $D_{1,a} \le_{\mathrm{icx}} D_1$. Indeed, for integer-valued variables the stop-loss transform is affine between consecutive integer thresholds. Conditional on all other runs, every hinge in (5.6) is an increasing convex function of $D_{1,a}$, so the replacement by D_1 is valid.

For $i \ge 2$, the finite survival function (5.3) is pointwise bounded by that of the infinite geometric run D_i satisfying

$$\mathbb{P}(D_i \ge d) = r_i^d \qquad (d \ge 1). \tag{6.4}$$

Hence D_{i,a_i} is stochastically dominated by D_i, and the coordinatewise increasing hinge permits this replacement as well.

Set $X_i = D_i + 1$. Equations (6.3) and (6.4) give the exact laws

$$\mathbb{P}(X_1 = 1) = \frac{1}{3}, \qquad \mathbb{P}(X_1 = n) = \frac{4}{3^n} \quad (n \ge 2), \qquad \mathbb{E}X_1 = 2, \tag{6.5}$$

and, with $h_i = q_i - t_i$,

$$\mathbb{P}(X_i = n) = \frac{h_i - 1}{h_i^n} \quad (n \ge 1), \qquad \mathbb{E}X_i = \frac{h_i}{h_i - 1}. \tag{6.6}$$

The right side of (5.6), with $K_j = \prod_{i<j} X_i$, is therefore a universal bound independent of all exponents a_i.

7. The exact six-block certificate

For blocks $2, \dots, 6$, choose the following thresholds:

(7.1)

comparison prime p_j^*	5	7	11	13	17
q_j	4	6	10	12	16
t_j	0	1	3	5	7
$h_j = q_j - t_j$	4	5	7	7	9

Substitution of (6.5)–(6.6) into (5.6) gives the five exact charges

(7.2)

block	C_j
5	$\frac{1}{3}$
7	$\frac{11}{48}$
11	$\frac{13609}{108000}$
13	$\frac{133996207321}{1089093600000}$
17	$\frac{30808237910202601}{298847283840000000}$

Appendix A derives every entry using finite rational arithmetic. The successive partial sums are

$$0,\quad \frac{1}{3},\quad \frac{9}{16},\quad \frac{74359}{108000},\quad \frac{883847235121}{1089093600000},\quad \frac{10123552563977963}{11068417920000000}. \tag{7.3}$$

In particular, the total charge through six prime blocks is

$$\begin{aligned} \Phi_6 &= \frac{10123552563977963}{11068417920000000} < 1, \\ \Phi_6 &\approx 0.914634109151705. \end{aligned} \tag{7.4}$$

with exact positive margin

$$\begin{aligned} 1-\Phi_6 &= \frac{944865356022037}{11068417920000000} > 0, \\ 1-\Phi_6 &\approx 0.0853658908482954. \end{aligned} \tag{7.5}$$

Remark 7.1 (Necessity and discrete optimality). The threshold schedule in (7.1) is the unique minimizer of the block-charge bound among all

$$\prod_{j=2}^{6}(q_j-1) = 3\cdot 5\cdot 9\cdot 11\cdot 15 = 22{,}275$$

integer schedules permitted by (5.1). Moreover, if the bottom-packed ternary selection is replaced by the ordinary binary product selection in the first block, exact minimization over the same schedules has the same unique minimizer but gives

$$\frac{22540693685733889}{22136835840000000} \approx 1.01824370242671 > 1.$$

Thus both the bottom-packed selection and the hybrid comparison are essential to this certificate. The ancillary rational checker performs both exhaustive computations.

Proof of Theorem 1.1. Assume that $r = \omega(N) \leq 6$. Construct the residual product R of (2.7); every pure cylinder misses it. Run the distortion of Section 3 in blocks $2, \dots, r$, using the corresponding prefix of the schedule (7.1). The block estimate (5.6), the height domination of Section 6, and the exact sums (7.3) imply

$$\mu_*\left(\bigcup_s B_s\right) < 1.$$

By (3.3), some point of R lies outside every nonpure cylinder. It also lies outside every pure cylinder by construction of R. Hence the lifted congruence classes fail to cover $\mathbb{Z}/N\mathbb{Z}$.

Every original class is periodic modulo N, so a cover of $\mathbb{Z}$ would induce a cover of $\mathbb{Z}/N\mathbb{Z}$. We have obtained a contradiction. Therefore $r \geq 7$. □

Remark 7.2 (Scope). The theorem establishes the lower bound $\omega(N) \geq 7$. It does not claim the stronger bound $\omega(N) \geq 8$, and it does not by itself rule out odd distinct covering systems of larger prime support.

8. Formal verification and reproducibility

The accompanying development in Lean 4 [7], using mathlib [16], formalizes the theorem from the arithmetic hypotheses onward. This complements the earlier kernel-checked least-common- multiple exclusion of Mian and Siddique [14]. Its input structure consists of a finite list of residues and moduli, a proof that every natural number lies in one of the corresponding congruence classes, proofs that all moduli are odd and greater than one, and injectivity of the modulus map. The final declaration is

`Erdos7.OddDistinctCoveringSystem.seven_prime_support`.

It concludes directly that the cardinality of the prime-factor set of the least common multiple is at least seven.

The formal arithmetic reduction proves all steps that are often suppressed in a cylinder argument. Prime valuations reconstruct each modulus and turn distinctness into injectivity of full depth vectors. Under the contrary assumption of at most six prime divisors, the ordered odd primes dominate $(3, 5, 7, 11, 13, 17)$. A reduced-alphabet word is embedded into each actual prime block and the blocks are combined by the finite Chinese remainder theorem. Congruence modulo p^d is proved to recover the first d base-p digits. Classes missing the comparison product are discarded; coverage is retained pointwise because the CRT integer of each comparison point belongs to an original class, which therefore survives. Zero-height blocks pad presentations with fewer than six primes.

The rational certificate is also proved semantically. In the largest five-factor calculation the identity

$$(x - t)_+ = x - t + (t - x)_+$$

separates the product mean from the complementary deficit. At threshold eight that deficit is nonzero on only 61 product states, which are propagated

through four small exact tables. Thus the final proof term does not use native evaluation.

With Lean version `v4.32.0-rc1` and the matching mathlib release, the development is reproduced from the unpacked arXiv source bundle by

```
cd anc/formal
lake build
lake env lean Audit.lean
```

The independent arithmetic checker is run from the project root with

```
python3 anc/scripts/verify_seven_prime_support.py
```

It uses exact rational arithmetic and reproduces all five charges, their sum, and the positive margin in (7.5). The Lean sources and the checker are supplied with this manuscript as ancillary files; neither is needed for the self-contained proof above.

The dependency audit for the final theorem reports only

`propext`, `Classical.choice`, `Quot.sound`,

the standard Lean/mathlib foundations used by the development. The project contains no additional axioms, admitted goals, unsafe declarations, or `native_decide` dependencies.

Appendix A. Exact evaluation of the certificate

This appendix expands the arithmetic behind (7.2). For a positive integer-valued random variable K and an integer $t \geq 0$,

$$\mathbb{E}[(K-(t+1))_+] = \mathbb{E}K - (t+1) + \sum_{k=1}^{t}(t+1-k)\mathbb{P}(K=k). \tag{A.1}$$

Thus only the mean and the probabilities below the hinge are needed.

For clarity, write X_5, X_7, X_{11}, X_{13} for the geometric variables created by the corresponding blocks of (7.1); their parameters are respectively $h = 4, 5, 7, 7$. Define

$$K_5 = X_1, \qquad K_7 = X_1X_5, \qquad K_{11} = X_1X_5X_7,$$
$$K_{13} = X_1X_5X_7X_{11}, \quad K_{17} = X_1X_5X_7X_{11}X_{13}.$$

Using the means in (6.5)–(6.6),

$$\mathbb{E}K_5 = 2, \quad \mathbb{E}K_7 = \frac{8}{3}, \quad \mathbb{E}K_{11} = \frac{10}{3}, \quad \mathbb{E}K_{13} = \frac{35}{9}, \quad \mathbb{E}K_{17} = \frac{245}{54}. \tag{A.2}$$

For any product $K = \prod_{i=1}^{m} X_i$ and integer $k \geq 1$, its small mass is the finite divisor sum

$$\mathbb{P}(K=k) = \sum_{\substack{n_1,\dots,n_m \geq 1 \\ n_1\cdots n_m = k}} \prod_{i=1}^{m} \mathbb{P}(X_i = n_i). \tag{A.3}$$

Applying (A.3) to the explicit laws gives

$$\mathbb{P}(K_7 = 1) = \frac{1}{4}, \tag{A.4}$$

$$\big(\mathbb{P}(K_{11} = k)\big)_{k=1}^{3} = \left(\frac{1}{5}, \frac{107}{300}, \frac{1969}{18000}\right), \tag{A.5}$$

$$\big(\mathbb{P}(K_{13} = k)\big)_{k=1}^{5} = \left(\frac{6}{35}, \frac{809}{2450}, \frac{100081}{1029000}, \frac{79974329}{432180000}, \frac{1720937761}{181515600000}\right), \tag{A.6}$$

$$\begin{aligned}\big(\mathbb{P}(K_{17} = k)\big)_{k=1}^{4} &= \left(\frac{36}{245}, \frac{2607}{8575}, \frac{103681}{1200500}, \frac{100577129}{504210000}\right),\\ \big(\mathbb{P}(K_{17} = k)\big)_{k=5}^{7} &= \left(\frac{1733897761}{211768200000}, \frac{7997427679049}{88942644000000},\right.\\ &\qquad\left.\frac{31902849672241}{37355910480000000}\right).\end{aligned} \tag{A.7, A.8}$$

We now insert these values into (A.1). The five hinge expectations, in schedule order, are

$$\mathbb{E}[(K_5 - 1)_+] = 2 - 1 = 1, \tag{A.9}$$

$$\mathbb{E}[(K_7 - 2)_+] = \frac{8}{3} - 2 + \frac{1}{4} = \frac{11}{12}, \tag{A.10}$$

$$\begin{aligned}\mathbb{E}[(K_{11} - 4)_+] &= \frac{10}{3} - 4 + 3\left(\frac{1}{5}\right) + 2\left(\frac{107}{300}\right) + \frac{1969}{18000} = \frac{13609}{18000},\\ \mathbb{E}[(K_{13} - 6)_+] &= \frac{35}{9} - 6 + 5\left(\frac{6}{35}\right) + 4\left(\frac{809}{2450}\right) + 3\left(\frac{100081}{1029000}\right)\\ &\quad + 2\left(\frac{79974329}{432180000}\right) + \frac{1720937761}{181515600000} = \frac{133996207321}{181515600000},\end{aligned} \tag{A.11, A.12}$$

$$\begin{aligned}\mathbb{E}[(K_{17} - 8)_+] &= \frac{245}{54} - 8 + 7\left(\frac{36}{245}\right) + 6\left(\frac{2607}{8575}\right) + 5\left(\frac{103681}{1200500}\right)\\ &\quad + 4\left(\frac{100577129}{504210000}\right) + 3\left(\frac{1733897761}{211768200000}\right)\\ &\quad + 2\left(\frac{7997427679049}{88942644000000}\right) + \frac{31902849672241}{37355910480000000}\\ &= \frac{30808237910202601}{37355910480000000}.\end{aligned} \tag{A.13}$$

Finally, formula (5.6) divides these five hinge expectations by

$$h_j - 1 = 3, 4, 6, 6, 8,$$

respectively. This gives exactly the five fractions in (7.2); their exact addition gives (7.4) and (7.5).

Appendix B. Logical dependency map

For reference, the proof depends on the following chain, with no asymptotic or computational premise hidden between the links:

(1) distinct moduli give full-vector uniqueness (Lemma 2.1);
(2) the residual selection removes every pure cylinder and supplies the exact ternary prefix caps (2.8);
(3) fibre normalization gives the union charge (3.3) and the causal symbol cap (3.4);
(4) uncrossing gives Bernoulli rearrangement, which in reverse order gives the causal gates and, only after those coordinates are gone, the ternary prefix run;
(5) ending-only uniqueness gives the block estimate (5.6);
(6) stop-loss and stochastic domination remove all unknown prime-power heights; and
(7) the exact certificate (7.2) has total strictly below one.

Acknowledgments

This work benefited from research assistance by AI systems developed by OpenAI and Anthropic, including support for proof exploration, proof development, and exact computational checks among others.

References

1. Paul Balister, *Erdős covering systems*, Surveys in Combinatorics 2024 (Felix Fischer and Robert Johnson, eds.), London Mathematical Society Lecture Note Series, vol. 493, Cambridge University Press, 2024, DOI, pp. 31–54.
2. Paul Balister, Béla Bollobás, Robert Morris, Julian Sahasrabudhe, and Marius Tiba, *The Erdős–Selfridge problem with square-free moduli*, Algebra & Number Theory **15** (2021), no. 3, 609–626, DOI; arXiv:1901.11465.
3. ———, *On the Erdős covering problem: The density of the uncovered set*, Inventiones Mathematicae **228** (2022), no. 1, 377–414, DOI; arXiv:1811.03547.
4. Marc Berger, Alexander Felzenbaum, and Aviezri S. Fraenkel, *Necessary condition for the existence of an incongruent covering system with odd moduli*, Acta Arithmetica **45** (1986), no. 4, 375–379, DOI.
5. ———, *Necessary condition for the existence of an incongruent covering system with odd moduli. II*, Acta Arithmetica **48** (1987), no. 1, 73–79, DOI.
6. Chris Bispels, Matthew Cohen, Joshua Harrington, Joshua Lowrance, Kaelyn Pontes, Leif Schaumann, and Tony W. H. Wong, *A further investigation on covering systems with odd moduli*, Discrete Mathematics **349** (2026), no. 7, 115013, DOI; arXiv:2507.16135.
7. Leonardo de Moura and Sebastian Ullrich, *The Lean 4 theorem prover and programming language*, Automated Deduction—CADE 28, Lecture Notes in Computer Science, vol. 12699, Springer, 2021, DOI, pp. 625–635.

8. Paul Erdős, *On integers of the form* 2^k+p *and some related problems*, Summa Brasiliensis Mathematicae **2** (1950), 113–123, Author's collected-paper PDF.
9. Song Guo and Zhi-Wei Sun, *On odd covering systems with distinct moduli*, Advances in Applied Mathematics **35** (2005), no. 2, 182–187, DOI; arXiv:math/0412217.
10. Joshua Harrington, Yewen Sun, and Wing Hong Tony Wong, *Covering systems with odd moduli*, Discrete Mathematics **345** (2022), no. 8, 112936, DOI; arXiv:2104.00602.
11. Robert D. Hough, *Solution of the minimum modulus problem for covering systems*, Annals of Mathematics **181** (2015), no. 1, 361–382, DOI.
12. Robert D. Hough and Pace P. Nielsen, *Covering systems with restricted divisibility*, Duke Mathematical Journal **168** (2019), no. 17, 3261–3295, DOI; arXiv:1703.02133.
13. Nathan McNew and Jai Setty, *On the densities of covering numbers and abundant numbers*, arXiv:2507.23041 [math.NT], 2025, Version 2, 10 February 2026.
14. Ibrahim Mian and Shayaan Siddique, *Kernel-checked exclusions for the Erdős–Selfridge odd covering problem: Any odd covering of* $\mathbb{Z}$ *has LCM exceeding 10000*, arXiv:2607.25628 [cs.LO], 2026.
15. R. Simpson and Doron Zeilberger, *Necessary conditions for distinct covering systems with square-free moduli*, Acta Arithmetica **59** (1991), no. 1, 59–70, DOI.
16. The mathlib Community, *The Lean mathematical library*, Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (New York), Association for Computing Machinery, 2020, DOI; arXiv:1910.09336, pp. 367–381.