hexagon:2611.00264v1

WIEFERICH CO-OCCURRENCE ALONG EXPONENTIAL SEQUENCES: A WEIGHTED TAIL ESTIMATE AND ITS APPLICATION TO $(\star')$

CHUNPING ZENG

ABSTRACT. Fix an integer $r \geq 2$ and a bound $A \geq 1$, and write $b_e = r^e - a'$ for $a' \leq A$. A recent preprint (hexagon:2610.00630, Zeng) reduces the *abc* conjecture, within a lattice-theoretic framework, to a single remaining hypothesis $(\star')$. We show that on the $k = 2$ slice ($\omega(ab) = 2$), the density-zero requirement on the exceptional set attached to $(\star')$ follows from one easily quoted sufficient condition $K_{r,a'} < \infty$ — the exact threshold of the first-moment route being the strictly weaker $\Phi_r(E) = o(E)$ (§9) — where

$$K_{r,a'} = \sum_{q \in \mathcal{C}_{r,a'}} s_q \log q \Big(\tfrac{1}{D_3(q)} + \tfrac{1}{e_0(q)} \Big),$$

$\mathcal{C}_{r,a'}$ is the set of primes passing a three-filter congruence sieve, $D_3(q) = \mathrm{ord}_{q^3}(r)$, and $e_0(q)$ is the least solution of $q^3 \mid r^e - a'$. The condition $K_{r,a'} < \infty$ holds under $P_4 + P_{6a}$, where P_4 is convergence of the classical Wieferich reciprocal sum and P_{6a} is the finiteness of super-Wieferich primes; the remaining gap is isolated in a single weighted hypothesis (H-j) on the "bad set" $\{q : e_0(q) < q\}$. We determine the exact unconditional threshold: an elementary pointwise bound disposes of the $s_q = 2$ layer in Case A unconditionally, the deep layer reduces to the single harmonic-weighted sum $\Phi_r(E)$, and we prove the unconditional linear cap $\Phi_r(E) \leq \frac{6}{\pi^2}(1+\gamma)\log(r+1)E + o(E)$; the missing input is an $\Omega(\log E)$ saving — P1/P2-level counting of $\mathcal{W}_r^{(3)}$ — not a constant and not a methodological abyss, and a congruence-expansion (information-theoretic) barrier shows the deep layer cannot be reached by any finite-order Fermat-quotient expansion. We prove that (H-j) is *not* a GRH-conditional statement (GRH controls only the $\mathrm{ord}_q(r)$ side, which the Hensel wall already absorbs), locate it precisely as a pointwise-discretized refinement of Katz's 2015 equidistribution conjecture lying strictly between Katz and Murata's 1991 conjectures, and support it by computation up to 10^6. Neither P_4 nor P_{6a} is necessary for the density conclusion; the exact sufficient-and-necessary threshold of the first-moment route is $\Phi_r(E) = o(E)$ (Theorems 9.4, 9.6).

1. INTRODUCTION

1.1. **The problem.** For a prime $q \nmid r$ write $d_q = \mathrm{ord}_q(r)$ for the multiplicative order and $s_q = v_q(r^{q-1} - 1) = v_q(r^{d_q} - 1) \geq 1$ for the *Wieferich depth*; the Fermat quotient is $q_q(u) = (u^{q-1} - 1)/q$, and $q \in \mathcal{W}_r := \{q : s_q \geq 2\}$ are the Wieferich primes to base r. The set $\mathcal{W}_r$ is the basic obstruction in several classical problems: the first case of Fermat's Last Theorem (Wieferich), the infinitude of non-Wieferich primes (Silverman), and—most recently—the remaining hypothesis $(\star')$ in Zeng's lattice-theoretic approach [1] to the *abc* conjecture, which is the motivating application of this note.

On the "fixed-modulus" side (fixed p, varying base) the Fermat quotient is well understood: Heath–Brown's Burgess-based estimates, the GRH coverage results of Alexandru–Cobeli–Vâjâitu–Zaharescu, and Shparlinski's unconditional improvements of the Ihara sum. On the "varying-modulus" side (fixed r, q varying) not even the zero-density of $\mathcal{W}_r$ is known unconditionally. The present note shows that the *application layer* of $(\star')$ can be reduced to a single summability condition on this varying-modulus side, and isolates the remaining analytic input as a precisely located, numerically supported hypothesis.

1.2. **Main results.** Throughout, q denotes an odd prime; the prime $q = 2$ is handled by the LTE branch of Theorem 1.1 (so $D_3(2) \in \{1, 2\}$) and contributes $O_{r,a'}(E)$ to $S(E)$, absorbed into the constant C of a first-moment bound $S(E) \le CE$. Also, $r \ge 2$ is a fixed integer (not a perfect power, without loss of generality), $A \ge 1$ is fixed, $a' \in \{1, \dots, A\}$, $\gcd(a', r) = 1$, and $b_e = r^e - a'$. Set $V_q(e) = v_q(b_e)$, $L_e = e \log r$, and define the exceptional mass

$$T_s(e) = \sum_{\substack{q \mid b_e \\ V_q(e) \ge 3 \\ s_q \ge 2}} s_q \log q, \qquad S(E) = \sum_{e \le E} T_s(e).$$

Call e *$\mathcal{W}$-exceptional* if $T_s(e) \ge \frac{1}{20} L_e$.

Theorem 1.1 (Order-lifting formula)**.** *For odd $q \nmid r$ and every $V \ge 1$,*

$$D_V(q) := \operatorname{ord}_{q^V}(r) = d_q \, q^{\max(0,\, V - s_q)}.$$

For $q = 2$: if $s_2 = v_2(r-1) \ge 2$ the same formula holds; if $s_2 = 1$ (i.e. $r \equiv 3 \pmod 4$) then $D_V(2) = 2^{\max(1,\, V - v_2(r+1))}$.

The *classical layer* $s_q = 2$ carries $D_3 = d_q q$ (a "Hensel wall" of one extra factor q), while the super-Wieferich layer $s_q \ge 3$ has $D_3 = d_q$ (wall-free). The equivalent form $s_q \ge 2 \iff \operatorname{ord}_{q^2}(r) = \operatorname{ord}_q(r)$ identifies $\mathcal{W}_r$ with the order-lift locus of Dillon [12].

Theorem 1.2 (Two-base necessary condition)**.** *Let q be odd, $s_q \ge 2$. If $q^3 \mid r^e - a'$ for some $e \ge 1$, then $q \nmid a'$ and $q^2 \mid a'^{q-1} - 1$ (i.e. $q \in \mathcal{W}_{a'}$); if moreover $s_q \ge 3$ then $q^3 \mid a'^{q-1} - 1$. Writing $w_q = (r^{q-1} - 1)/q^2 \bmod q$, $\beta_q = (a'^{q-1} - 1)/q^2 \bmod q$ (defined when $s_q = 2$), the least solution $e_0(q)$ satisfies $e_0(q) \equiv \beta_q w_q^{-1} \pmod q$, and $\beta_q = 0$ forces $e_0(q) \ge q$. Finally $e_0(q) \ge \frac{3 \log q}{\log r} - O_{r,a'}(1)$ unconditionally.*

Theorem 1.2 compresses the candidate enumeration from a scan over e to a three-filter sieve per prime: $q \in \mathcal{W}_r$, $q \in \mathcal{W}_{a'}$, and $a' \in \langle r \rangle \subseteq (\mathbb{Z}/q^3\mathbb{Z})^\times$; the last filter is necessary but not sufficient [17, numerics].

Theorem 1.3 (One-condition closure)**.** *For q in the compatible set $\mathcal{C}_{r,a'}$ (the three-filter set above), let $e_0(q)$ be as in Theorem 1.2 and define*

$$K_{r,a'} = \sum_{q \in \mathcal{C}_{r,a'}} s_q \log q \Big(\frac{1}{D_3(q)} + \frac{1}{e_0(q)} \Big).$$

If $K_{r,a'} < \infty$, then $S(E) \le E \, K_{r,a'}$, and consequently

$$\frac{1}{E} \#\{e \le E : T_s(e) \ge \tfrac{e \log r}{20}\} \;\le\; \frac{20 \, K_{r,a'}}{\log r} \cdot \frac{1 + \ln E}{E} \;\longrightarrow\; 0.$$

Theorem 1.3 uses no GRH, no sieve, and no information on the distribution of $\operatorname{ord}_q(r)$: the hypothesis $V_q(e) \ge 3$ has already converted the weight $1/\operatorname{ord}_q(r)$ into $1/\operatorname{ord}_{q^3}(r)$.

Theorem 1.4 (Bad-set structure)**.** *Write $e_0(q) = b_q + j_q q$ with $b_q = (\beta_q w_q^{-1} \bmod q) \in [0, q)$ and $j_q \in [0, d_q)$ (layer $s_q = 2$). If $j_q \ge 1$ then $e_0(q) \ge q$ and $s_q \log q / e_0(q) \le s_q \log q / q$, so the $s_q = 2$ part of this layer is controlled directly by $P_4 : \sum_{q \in \mathcal{W}_r} \log q / q < \infty$, and the $s_q \ge 3$ part by P_{6a} (finiteness of $\mathcal{W}_r^{(3)}$). The unique remaining layer is the* bad set $\mathcal{B}_{r,a'} = \{q \in \mathcal{C}_{r,a'} : s_q = 2,\ j_q = 0\} = \{q \in \mathcal{C}_{r,a'} : s_q = 2,\ e_0(q) < q\}$ *(the super-Wieferich layer $s_q \ge 3$, for which $e_0 \le d_q$*

may be $< q$, is controlled by P_{6a} and is not part of $\mathcal{B}_{r,a'}$; see §12), and the remaining hypothesis is

$$\text{(H-j)}: \qquad \sum_{q\in\mathcal{B}_{r,a'}} \frac{s_q \log q}{e_0(q)} < \infty.$$

The set $\mathcal{B}_{r,a'}$ is nonempty already at $(r, a', q) = (17, 19, 3)$. Layerwise, $\{q : e_0(q) = b\} \subseteq \{q : q^3 \mid r^b - a'\}$ is finite for each fixed b [17, Layer census]*, so divergence of (H-j) can only come from a bad sequence with $e_0(q_n) \to \infty$.*

The subtraction chain organizing this note:

$$\text{(AP)} \longrightarrow K_{r,a'} < \infty \longrightarrow \Xi_r < \infty \longrightarrow \underbrace{\Xi_r^{(2)} < \infty}_{\text{Theorem 7.5, unconditional}} + \underbrace{\Xi_r^{(3)} < \infty}_{\text{sufficient, not necessary (Thm. 9.4)}}.$$

The exact threshold of the Markov route is $\Phi_r(E) = o(E)$, equivalently $A_r(t) = o(t)$ (Theorem 9.6); $\Xi_r^{(3)} < \infty$ and its weakenings $\Xi_r^{(3)}(E) = o(E/\log E)$, $K_{r,a'} < \infty$ are *sufficient rungs* of the ladder of §9, none of them necessary.

Contributions. (i) Theorem 7.5 (main, unconditional): $\Xi_r^{(2)} < \infty$ for every base r — block decomposition by d_q plus Stieltjes partial summation, with explicit bound $O(1 + \log\log r)$. (ii) Theorems 8.1–8.4 and Proposition 8.5: the remaining gap is squeezed strictly between P_{6a} and $P_4^{(3)}$, explained structurally (the Hensel wall is the only unconditional decay source), and shown unreachable by block/partial-summation methods along three independent barriers, with a fourth (radical) withdrawn as a bookkeeping artifact and replaced by an identity (Theorem 8.6). (iii) Theorems 8.10 and 10.3: under the Katz double-zero hypothesis (D4), $\Xi_r^{(3)} < \infty$ and P_{6a} hold almost surely (Borel–Cantelli), and (K^+) unifies the main line and the bad-set line into one conjecture family; the counting-density reading of (D4) is equivalent to P_{6a} (Theorem 14.4). (iv) Section 11: the unconditional $\sum_p \log p/(d_p(p-1)) < \infty$ (Erdős–Murty), the exact identity $C_r = S_W + S_{\mathrm{ord}} + LTE_2$, and the strict one-way $S_W < \infty \Rightarrow \Xi_r < \infty$; GRH is redundant throughout the chain. (v) Section 9: the exact threshold of the unconditional question — an elementary pointwise treatment of Case A, the unconditional linear cap $\Phi_r(E) \le \frac{6}{\pi^2}(1+\gamma)\log(r+1)E + o(E)$, the *exact* threshold $\Phi_r(E) = o(E) \iff A_r(t) = o(t)$ (Theorem 9.6), the weakest sufficient rung of the ladder $\Xi_r^{(3)}(E) = o(E/\log E)$ (strictly above the threshold, Theorem 9.4), and the information-theoretic barrier showing the deep layer is unreachable by any finite-order congruence expansion.

1.3. **Locating (H-j).** Katz's Wieferich quotient $W_u(q) = (u^{q-1} - 1)/q^2$ takes values in $\frac{1}{q}\mathbb{Z}/\mathbb{Z} \subset \mathbb{R}/\mathbb{Z}$, and Katz conjectured Haar equidistribution for Zariski-dense points on group schemes [2]. Writing $q_q(u) = k_u + q\, m_u$ ($k_u, m_u \in [0, q)$), one has $q \in \mathcal{W}_u \iff k_u = 0$, and then $w_q = m_r$, $\beta_q = m_{a'}$: the pair (w_q, β_q) is the *mod-q higher layer* of the Katz quotient at its Haar-null zeros. The hypothesis (H-j) is controlled by the three-layer discretized joint refinement (K)=(D1)+(D2)+(D3) of §10, and we show (Proposition 10.1) that (K) implies both (H-j) and $K_{r,a'} < \infty$. Moreover:

Theorem 1.5 (not a GRH statement)**.** *No known theorem gives, under GRH, a nontrivial bound on $N_r(x)$, the density of $\mathcal{W}_r$, or P_4; the natural Chebotarev fields $L_q = \mathbb{Q}(\zeta_{q^2}, r^{1/q^2})$ have discriminants of size $q^{O(q^2)}$, so quantitative GRH–Chebotarev error terms are uncontrolled; and GRH controls only the* $\mathrm{ord}_q(r)$ *(Artin) side, which the Hensel wall $1/D_3$ has already absorbed. Hence "GRH $\Rightarrow$ (H-j)" would mispackage an open problem as a conditional theorem.*

Instead (H-j) sits *strictly between* Katz's conjecture and Murata's 1991 conjectures [3]: it is stronger than Katz (pointwise discretization plus mod-q^2 two-base joint equidistribution plus discrete-logarithm uniformity) and weaker than Murata (ii) (finiteness of $\mathcal{W}_r \cap \mathcal{W}_{a'}$); Murata (i) is exactly P_{6a}.

1.4. **Relation to** $(\star')$**.** In Zeng's framework [1] the remaining gap is stated in Section 7 ("The remaining gap, exactly") as Working hypothesis 6.4: $|\log x| = o(\log c)$, equivalently $O(\log\log c)$, with $x = \det(\Lambda)^{(k-1)/k}/\Delta$ and $\Delta = N|a'|_2$; Theorem 6.5 there gives $\eta \le \frac{1}{k} + O(\frac{\log\log c}{\log c})$ and Theorem 7.1 the equivalent form $N|a'|_2 = (c\theta_c/g)^{(k-1)/k} c^{o(1)}$ (cf. §13). The expression $\frac{k-1}{k}\log c - \log\mathrm{rad}(ab)$ is a heuristic transcription, not a formula of [1] (see Remark 13.1); and "degeneration to *abc* strength" should be read as the structural ceiling of Theorem 10.1 there ($\Pi \le 1/k + o(1)$, attained at $c = r^e$; full *abc* is structurally unreachable inside the framework). On the line $a' = O(1)$, $b_e = r^e - a'$, one has $k_e \to \infty$, so the requirement degenerates to *abc*-strength already for bounded a', and Theorem 1.3 governs the $k = 2$ slice. Theorem 1.2 pins the Layer-III Pillai candidates to $\mathcal{W}_r \cap \mathcal{W}_{a'}$ with an explicit class e mod q; Theorem 1.3 pins the Layer-III mass to $J'_{r,a'}$, the Pillai-single-solution layer; and (H-j) shows the application-layer gap of Layer III does not require Baker-type bounds—Baker belongs to the deeper $k \ge 3$, $V_q = 2$ layer, untouched here. The defect/order-core framework of [15] independently corroborates this picture on the Mersenne line.

2. Preliminaries and notation

For an odd prime $q \nmid r$, LTE in the form $v_q(r^{d_q m} - 1) = s_q + v_q(m)$ (valid since $v_q(r^{d_q} - 1) = s_q \ge 1$) is used repeatedly. The unconditional bounds

$$d_q \ge \frac{\log q}{\log r}, \qquad s_q \log q < d_q \log r \tag{2.1}$$

follow from $q \le r^{d_q} - 1$ and $q^{s_q} \le r^{d_q} - 1$. We write $\Sigma_r = \sum_{q\in\mathcal{W}_r} \frac{\log q}{q}$ (P_4 is $\Sigma_r < \infty$), $\Theta_r^{[s]}(\infty) = \sum_{q\in\mathcal{W}_r} \frac{s_q \log q}{q}$ (P_{6b}), and $N_r(x) = \#\{q \le x : q \in \mathcal{W}_r\}$. The chain (H-$\Theta$) $\Rightarrow P_{6b} \Rightarrow \sum_{q\in\mathcal{W}_r^{(3)}} \log q/q < \infty$ holds by the Fubini identity $s_q = \sum_{k\ge 2} \mathbf{1}_{\mathcal{W}_r^{(k)}}(q)$ (nonnegativity of the summands); P_{6a} (finiteness of $\mathcal{W}_r^{(3)}$) is strictly stronger in general — an infinite $\mathcal{W}_r^{(3)}$ can be sparse enough for the weighted sum to converge. Both reverse implications fail [17]. We use $\sim$, $\ll$, $O(\cdot)$ with their standard meanings; all implied constants may depend on r and a' unless stated.

3. Proof of Theorem 1.1

Write $r^{d_q} = 1 + q^{s_q} u$ with $q \nmid u$. For odd q and $m \ge 1$,

$$r^{d_q m} - 1 = (r^{d_q})^m - 1 = q^{s_q} u \sum_{i=0}^{m-1} r^{d_q i}, \qquad v_q(r^{d_q m} - 1) = s_q + v_q(m),$$

because the sum $\sum_{i=0}^{m-1} r^{d_q i} \equiv m \pmod q$ and induction on $v_q(m)$ gives $v_q\big(\sum_{i=0}^{m-1} r^{d_q i}\big) = v_q(m)$ (the standard LTE step). Moreover $\mathrm{ord}_{q^V}(r)$ is a multiple of $d_q = \mathrm{ord}_q(r)$. Hence the least m with $q^V \mid r^{d_q m} - 1$ is $m = q^{\max(0, V - s_q)}$, giving $D_V(q) = d_q q^{\max(0,V-s_q)}$. For $q = 2$ with $s_2 \ge 2$ the same LTE applies; if $s_2 = 1$ then $w := v_2(r^2 - 1) = 1 + v_2(r+1) \ge 3$, so $r^2 \equiv 1 \pmod 8$, $v_2(r^{2m} - 1) = w + v_2(m)$, and the least e with $2^V \mid r^e - 1$ is $2^{\max(1, V-w+1)}$, i.e. $D_V(2) = 2^{\max(1, V - v_2(r+1))}$. $\square$

Remark 3.1. A wall-illustration: $(a', r, q) = (2, 23, 3)$ gives $v_3(23^e - 2) \geq 10 \iff e \equiv 5 \pmod{39366}$; the unique hit below $r^e \leq 10^{14}$ is $e = 5$ (the Reyssat triple), and the next candidate has $c \approx 10^{53600}$.

4. Proof of Theorem 1.2

From $r^e \equiv a' \pmod{q^3}$ raise to the $(q-1)$-st power: $a'^{q-1} \equiv r^{e(q-1)} \pmod{q^3}$. Write $t = \min(s_q, 3)$ and $r^{q-1} = 1 + q^t w$, $q \nmid w$ (for $s_q = 2$). The binomial expansion gives, since $2t \geq 4 > 3$,

$$r^{e(q-1)} = (1 + q^t w)^e \equiv 1 + e\, q^t w \pmod{q^3}.$$

If $s_q \geq 3$ then $t = 3$ and $r^{q-1} \equiv 1 \pmod{q^3}$, so $a'^{q-1} \equiv 1 \pmod{q^3}$. If $s_q = 2$ then $a'^{q-1} \equiv 1 + eq^2 w \pmod{q^3}$, giving $q^2 \mid a'^{q-1} - 1$ and $(a'^{q-1} - 1)/q^2 \equiv ew \pmod{q}$; solving for the least representative gives $e_0 \equiv \beta_q w_q^{-1} \pmod{q}$, and $\beta_q = 0$ forces $e_0 \equiv 0 \pmod{q}$, i.e. $e_0 \geq q$. Finally $q^3 \leq |r^{e_0} - a'| \leq r^{e_0} + |a'|$ (using $b_{e_0} \neq 0$) yields $e_0 \geq \frac{3 \log q}{\log r} - O_{r,a'}(1)$. □

5. Proof of Theorem 1.3

We need two lemmas.

Proof. Partition $[1, E]$ into dyadic blocks $(E/2^{j+1}, E/2^j]$, $0 \leq j < \lceil \log_2 E \rceil$. Inside block j one has $e > E_j := E/2^{j+1}$, hence the exceptional condition $T_s(e) \geq ce > cE_j$; Markov with the fixed threshold cE_j gives at most $S(E/2^j)/(cE_j) \leq (CE/2^j)/(cE/2^{j+1}) = 2C/c$ exceptions per block. □

Lemma 5.1 (Correct counting)**.** *For $q \in \mathcal{C}_{r,a'}$ and $E \geq 1$,*

$$N_q(E) := \#\{e \leq E : q^3 \mid b_e\} \;\leq\; \frac{E}{D_3(q)} + \mathbf{1}_{\{e_0(q) \leq E\}} \;\leq\; E\Big(\frac{1}{D_3(q)} + \frac{1}{e_0(q)}\Big).$$

Proof. The solution set of $q^3 \mid r^e - a'$ is the single class $e \equiv e_0 \pmod{D_3}$ with $e_0 \leq D_3$, so $N_q(E) = \lfloor (E - e_0)/D_3 \rfloor + 1$ if $e_0 \leq E$ and 0 otherwise; the second inequality uses $\mathbf{1}_{\{e_0 \leq E\}} \leq E/e_0$. □

Proof of Theorem 1.3. By Theorem 1.2, $N_q(E) > 0$ forces $q \in \mathcal{W}_r \cap \mathcal{W}_{a'}$ with $a' \in \langle r \rangle \bmod q^3$, i.e. $q \in \mathcal{C}_{r,a'}$; hence

$$S(E) = \sum_{q \in \mathcal{C}_{r,a'}} s_q \log q\, N_q(E) \leq E \sum_{q \in \mathcal{C}_{r,a'}} s_q \log q \Big(\frac{1}{D_3(q)} + \frac{1}{e_0(q)}\Big) = E\, K_{r,a'},$$

For the density bound, the moving threshold is handled directly: with $c = \log r/20$,

$$\#\{e \leq E : T_s(e) \geq ce\} \;\leq\; \frac{1}{c} \sum_{e \leq E} \frac{T_s(e)}{e}, \qquad \sum_{e \leq E} \frac{T_s(e)}{e} = \sum_{q \in \mathcal{C}_{r,a'}} s_q \log q \sum_{\substack{e \leq E \\ q^3 \mid b_e}} \frac{1}{e} \;\leq\; (1 + \ln E)\, K_{r,a'},$$

the inner sum over the class $e \equiv e_0 \pmod{D_3}$ being at most $(1 + \ln E)\big(1/D_3(q) + 1/e_0(q)\big)$; dividing by cE gives the claim. The argument never uses $D_3 \leq E$: the "+1" term is absorbed by $1/e_0$. □

Remark 5.2 ($K_{r,a'} < \infty$: sufficient, not necessary)**.** The condition $K_{r,a'} < \infty$ is the most quotable member of a family of sufficient conditions, not the exact requirement: the threshold of the first-moment route is the strictly weaker (M) $\sum_{e \leq E} T_s(e)/e = o(E)$ (§9), which allows $\Xi_r^{(3)} = \infty$.

Proposition 5.3 (Splitting and sufficient conditions)**.** *Write* $\Xi_r = \sum_{q\in\mathcal{W}_r} \frac{s_q \log q}{D_3(q)} = \Xi_r^{(2)} + \Xi_r^{(3)}$. *Then* $\Xi_r^{(2)}$ *is finite* unconditionally *(Theorem 7.5), and* $\Xi_r^{(3)} = \sum_{q\in\mathcal{W}_r^{(3)}} \frac{s_q \log q}{d_q} \le (\log r)\,\#\mathcal{W}_r^{(3)}$ *by* (2.1)*. Hence* $P_{6a} \Rightarrow \Xi_r < \infty$*; the precise squeeze around* $\Xi_r^{(3)}$ *and the barriers to its unconditional proof are given in §8. If* $a' = r^k$ *then* $b_e = r^k(r^{e-k}-1)$ *and* $T_s^{(a')}(e) = T_s^{(1)}(e-k)$*; compatibility* $a' \in \langle r\rangle \bmod q^3$ *is automatic, and* $q \in \mathcal{W}_r \Rightarrow q \in \mathcal{W}_{r^k}$ *by* $q_q(r^k) \equiv k\, q_q(r) \pmod q$*, so* $\mathcal{C}_{r,r^k} = \mathcal{W}_r$*. Moreover* $e_0(q) = k + D_3(q) \ge D_3(q)$*, so* $K_{r,r^k} \le 2\Xi_r$*:* Case A is closed unconditionally on (H-j)*. If* $a' \notin r^{\mathbb{Z}}$ *then* $e_0 \le D_3$ *and* $K_{r,a'} \le 2J'_{r,a'}$ *where* $J'_{r,a'} = \sum_{q\in\mathcal{C}_{r,a'}} \frac{s_q \log q}{e_0(q)}$*; in particular* $\Xi_{r,a'} := \sum_{\mathcal{C}_{r,a'}} \frac{s_q \log q}{D_3(q)} \le J'_{r,a'}$ *is a consequence, not an independent hypothesis.*

For $a' = 1$ the count is exact with no boundary term: $V_q(e) = s_q + v_q(e/d_q)$ for $d_q \mid e$, so $\{e : V_q(e) \ge 3\} = D_3(q)\mathbb{Z}$ and $S(E) = \sum_{q\in\mathcal{W}_r} s_q \log q\, \lfloor E/D_3(q)\rfloor \le E\Xi_r$.

6. The bad set: proof of Theorem 1.4

For the layer $s_q = 2$ we have $D_3 = d_q q$. By Theorem 1.2, $e_0 \equiv b_q \pmod q$ with $b_q = \beta_q w_q^{-1} \bmod q \in [0,q)$; writing $e_0 = b_q + j_q q$ gives $j_q \in [0, d_q)$ (since $e_0 \le D_3 = d_q q$). If $j_q \ge 1$ then $e_0 \ge q$ and $\frac{s_q \log q}{e_0} \le \frac{s_q \log q}{q}$; summing over the layer $s_q = 2$ gives at most $2\Sigma_r$; the layer $s_q \ge 3$ is controlled separately by P_{6a}, via $s_q \log q/e_0 \le s_q \log q \le d_q \log r$ and the finiteness of $\mathcal{W}_r^{(3)}$. If $j_q = 0$ then $e_0 = b_q < q$ with $b_q \ge \frac{3\log q}{\log r} - O(1)$ (Theorem 1.2). The witness $(r, a', q) = (17, 19, 3)$: $17^2 \equiv 19 \pmod{27}$ gives $e_0 = 2 < 3$, $j_q = 0$, contributing $2\log 3 > \frac{2\log 3}{3}$. For the layerwise finiteness, $e_0(q) = b$ forces $q^3 \mid r^b - a'$ with $q > b$, and q is a cube-divisor prime of the fixed integer $r^b - a'$; each layer is finite and enumerable. □

Proposition 6.1 (Closing conditions)**.** *(H-j) follows from the pair*

(SC1) $\#\mathcal{B}_{r,a'}(x) \ll (\log x)^A$ *for some* $A < 1+\varepsilon$*, where* $\varepsilon > 0$ *is as in* (SC2)*;*

(SC2) $e_0(q) \ge (\log q)^{1+\varepsilon}$ *for all but finitely many* $q \in \mathcal{B}_{r,a'}$*;*

with A and ε linked*: Abel summation gives* $\sum_{q\in\mathcal{B}(x)} (\log q)^{-1-\varepsilon} \ll \int^{\log x} u^{A-2-\varepsilon}\, du$*, which converges if and only if* $A < 1+\varepsilon$*. In particular the cleanest sufficient form is* $\#\mathcal{B}_{r,a'}(x) \ll \log x$ *with any* $\varepsilon > 0$*. Note (SC1) implies a P2-type bound for* $N_r(x)$ *(the bad set is a subset of* $\mathcal{W}_r$*); it is an* intersection-strengthened *version of* P_2*, not a weakening.*

Remark 6.2 (Gallagher skeleton)**.** One has $\#\mathcal{B}_{r,a'}(x) = |\mathcal{M}(x)|$ in the layer $s_q = 2$ (each q contributes at most one $b \in [1,q)$ since $D_3 = d_q q > q$). Expanding the Wieferich indicator as $\mathbf{1}_{\mathcal{W}_r}(q) = \frac{1}{q}\sum_{t \bmod q} e\big(\frac{t\, q_r(q)}{q}\big)$ splits $\sum_b \#\{\mathcal{Q}_b \cap \mathcal{W}_r\}$ into a main term $\sum_b \frac{1}{q}|\mathcal{Q}_b|$ (whose random expectation is $\asymp \sum_q q^{-4}$) and a remainder of varying-modulus Fermat-quotient character sums. A mean-value hypothesis

$$\text{(FQM-avg)} \qquad \sum_{q\le x} \Big| \sum_{q'\in\mathcal{Q}_b} e\Big(\frac{t\, q_r(q')}{q'}\Big)\Big| \ll x^\theta \quad (\theta < 1),$$

uniform in the truncated range $b \le B_0 = 3\log x/\log r$. The *main term* is unconditionally $O(\log x \cdot \log\log x)$: $\sum_{b\le B_0}\sum_{\substack{q\in\mathcal{Q}_b\\ q>b}} q^{-1} \le \sum_{b\le B_0} |\mathcal{Q}_b|/b \le \frac{B_0 \log r}{\log 8}\sum_{b\le B_0}\frac{1}{b} \ll \log x \cdot \log\log x$. Hence (FQM-avg) yields the *weighted* bound $\sum_{q\in\mathcal{B}(x)} q^{-1} = O(\log x)$ (not the unweighted (SC1)); and the weighted bound combined with (SC2) closes (H-j), since $\sum_{\mathcal{B}(x)} (\log q)^{-1-\varepsilon} \ll \mathcal{B}_1(x)(\log x)^{-1-\varepsilon} + \int^x \mathcal{B}_1(t)\,\frac{dt}{t(\log t)^{2+\varepsilon}}$ with $\mathcal{B}_1(t) = \sum_{q\in\mathcal{B}(t)} q^{-1} = O(\log t)$ converges for every

$\varepsilon > 0$. The small-b range $b \le B_0$ is covered by the trivial $O((\log x)^2)$ bound. Fixed-modulus Fermat-quotient sums are understood (Heath–Brown, Shparlinski); (FQM-avg) is the varying-modulus input, at the same depth as Katz's conjecture.

7. The unconditional layer: $\Xi_r^{(2)} < \infty$

Split the governing weight by depth:

$$\Xi_r = \Xi_r^{(2)} + \Xi_r^{(3)}, \qquad \Xi_r^{(2)} = \sum_{\substack{q \in \mathcal{W}_r \\ s_q = 2}} \frac{2\log q}{d_q q}, \qquad \Xi_r^{(3)} = \sum_{\substack{q \in \mathcal{W}_r \\ s_q \ge 3}} \frac{s_q \log q}{d_q}.$$

7.1. **Block decomposition.** For each integer $d \ge 1$ set $S_d = \{q \in \mathcal{W}_r : d_q = d,\ s_q = 2,\ q \text{ odd}\}$; the prime $q = 2$ is listed separately and contributes a finite $O(\log r)$ term.

Lemma 7.1 (LTE identification)**.** *For $q \in S_d$ we have $d \mid q-1$ and*

$$v_q(r^{q-1} - 1) = v_q(r^d - 1) + v_q\Big(\frac{q-1}{d}\Big) = v_q(r^d - 1),$$

since $0 < (q-1)/d < q$. Hence $s_q = 2 \iff q^2 \parallel r^d - 1$.

Lemma 7.2 (Unconditional mass constraint)**.** *The primes $q \in S_d$ are pairwise coprime divisors with $q^2 \mid r^d - 1$, so*

$$\Big(\prod_{q \in S_d} q\Big)^2 \,\Big|\, r^d - 1 \quad \Longrightarrow \quad \sum_{q \in S_d} \log q < \frac{d \log r}{2}.$$

Also $d \mid q-1$ gives $q_i \ge id + 1$ for the i-th element, hence

$$|S_d| \log(d+1) \le \sum_{q \in S_d} \log q < \frac{d \log r}{2} \quad \Longrightarrow \quad |S_d| < K_d := \frac{d \log r}{2 \log(d+1)}.$$

With $A(t) = \#\{q \in S_d : q \le t\}$: $A(t) \le \min(K_d, t/d)$.

Lemma 7.3 (Stieltjes partial summation)**.** *For $f(t) = \log t/t$ ($t \ge 3$, positive, strictly decreasing, $-f'(t) = \frac{\log t - 1}{t^2} \le \frac{\log t}{t^2}$), the boundary terms vanish ($A(t) \le K_d$, $f(t) \to 0$) and*

$$\sum_{q \in S_d} \frac{\log q}{q} = \int_{d+1^-}^{\infty} f \, dA = -\int_{d+1}^{\infty} A(t) f'(t)\, dt \le \int_d^{\infty} \min\Big(K_d, \frac{t}{d}\Big) \frac{\log t}{t^2}\, dt =: I_d.$$

Lemma 7.4 (Explicit I_d)**.** *With $t_* = K_d d = \frac{d^2 \log r}{2\log(d+1)}$: if $t_* \le d$ then $I_d \le K_d \frac{\log d + 1}{d}$; otherwise*

$$I_d \le \frac{1}{d} \int_d^{t_*} \frac{\log t}{t}\, dt + K_d \int_{t_*}^{\infty} \frac{\log t}{t^2}\, dt = \frac{(\log t_*)^2 - (\log d)^2}{2d} + \frac{\log t_* + 1}{d}.$$

Since $\log t_* = 2 \log d + O_r(1)$, in all cases

$$I_d \le \frac{3(\log d)^2}{2d} + O_r\Big(\frac{\log d}{d}\Big),$$

and $d = 1$ contributes a finite $O_r(1)$ constant.

Theorem 7.5 (Main theorem, unconditional)**.** *For every fixed $r \ge 2$,*

$$\boxed{\Xi_r^{(2)} = \sum_{d \ge 1} \frac{2}{d} \sum_{q \in S_d} \frac{\log q}{q} \le \sum_{d \ge 1} \frac{2 I_d}{d} \le \sum_{d \ge 1} \Big(\frac{3(\log d)^2}{d^2} + O_r\Big(\frac{\log d}{d^2}\Big)\Big) < \infty}$$

since $\sum_{d\ge 2}(\log d)^2/d^2 = |\zeta''(2)| + O(1) < \infty$. *No abc, no GRH, no Artin, no Katz, and no count, density, or sparsity of* $\mathcal{W}_r$*: the set* $\mathcal{W}_r$ *may even be infinite or of positive density.*

Remark 7.6 (Ancestry and priority)**.** The mass constraint of Lemma 7.2 is a Wieferich-depth refinement of the classical engine of Erdős–Murty [18] (reproduced in Moree's survey [27]): primes with $\mathrm{ord}_q(r) = d$ divide $r^d - 1$, so their product divides $r^d - 1$; this is how Erdős–Murty bound the density of small-order primes. Here the depth condition $s_q = 2$ upgrades the divisibility to $q^2 \mid r^d - 1$ (Lemma 7.1), and that single extra factor of q is what converts an open density question into a convergent sum (cf. Theorem 8.3). We are not aware of an unconditional convergence result for any ord-weighted reciprocal sum over the Wieferich primes in the literature; the closest object is $\sum_p 1/\mathrm{ord}_{p^2}(2)$ of Granville–Soundararajan [26], whose convergence is open even under their weaker conjectures — the weight $1/\mathrm{ord}$ there lacks the saving factor $\log p/p$ that the Hensel wall supplies here.

Corollary 7.7 (Explicit bound and numerics)**.** $\Xi_r^{(2)} \le \sum_{d\ge 1} \frac{2I_d}{d} = O(1 + \log\log r)$*; summing the block bounds to* $d \le 2\times 10^5$ *gives* $5.02, 7.13, 9.37, 10.65$ *for* $r = 2, 3, 5, 7$*, while the true values (*$q < 10^6$*) are* $0.00003782, 0.08719619, 0.00000011, 0.16094379$*. A softer Cauchy proof:* $\min(K_d, t/d) \le \sqrt{K_d t/d}$ *gives* $\Xi_d^{(2)} \le 2\sqrt{2\log r}\,\sqrt{\log d}/d^{3/2}$ *and* $\sum_d d^{-3/2}\sqrt{\log d} < \infty$.

7.2. **Self-consistency: why this does not contradict the openness of** P_4**.** The same block controls three weights with three fates:

weight	per-block bound	$\sum_d$	conclusion
$\Theta_d = \sum_{q\in S_d} \frac{\log q}{q}$	$\le I_d \asymp \frac{3(\log d)^2}{2d}$	diverges ($\sim (\log D)^3$)	the method gives no P_4
$\Xi_d^{(2)} = \frac{2}{d}\Theta_d$	$\le \frac{2I_d}{d} \asymp \frac{3(\log d)^2}{d^2}$	converges	Theorem 7.5
$S_{\mathcal{W},d} = \frac{1}{d}\sum_{q\in S_d} \log q$	$\le \frac{\log r}{2}$ per d	diverges	consistent with Theorem 11.3

Numerically, at $D = 10^2/10^3/10^4/10^5$ the $\sum \Theta_d$ upper bounds are $37.4/120.5/288.1/573.3$, while the $\sum_d \Xi_d^{(2)}$ bounds plateau at $4.85/5.40/5.52/5.54$. Since $\Xi_q^{(2)} = \frac{2\log q}{d_q q} \le \frac{\log q}{q}$, the theorem proves a *smaller* quantity than P_4 — but precisely the one the application layer needs. A second consistency point: $|S_d| \le K_d \asymp d/\log d$ may grow linearly in d, so $\sum_d |S_d|$ may diverge ($\mathcal{W}_r$ infinite or even of positive density) while $\sum_d \Xi_d^{(2)}$ converges. Open counting and convergent weighting do not contradict each other.

8. The remaining layer: $\Xi_r^{(3)}$

Write $P_4^{(3)}$ for the assertion $\sum_{q\in\mathcal{W}_r^{(3)}} \frac{\log q}{q} < \infty$ (P_4 restricted to the deep layer; a weakening of P_4).

Theorem 8.1 (Three-layer squeeze, both ends strict)**.** *Pointwise* $\dfrac{3\log q}{q-1} \le \Xi_q^{(3)} := \dfrac{s_q \log q}{d_q} \le \log r$ (U2)*, and*

$$P_{6a} \implies \Xi_r^{(3)} < \infty \implies P_4^{(3)},$$

with both implications strict.

Proof. The left implication is $\Xi_r^{(3)} \le (\log r)\,\#\mathcal{W}_r^{(3)}$ by (2.1); the right one is $d_q \le q - 1$ with $s_q \ge 3$. Left strict: take $\mathcal{W}^{(3)}$ infinite, $d_q = q - 1$, $s = 3$, and q_j so sparse that $\sum \log q_j/q_j < \infty$; then $\Xi_r^{(3)} < \infty$ while P_{6a} fails. Right strict: take $\log q_j = 2^j \log 2$ (the $P_4^{(3)}$ partial sums

converge to $0.54169\ldots$), $d_j = \frac{3\log q_j}{\log r}$ (the minimal order allowed by $q^3 \le r^d$, cf. (U1)), $s = 3$; then $\Xi_{q_j}^{(3)} = \log r$ constantly and $\Xi_r^{(3)} \sim j \log r \to \infty$ ($j = 24$ already gives 26.37). □

Thus $\Xi_r^{(3)} < \infty$ sits strictly between P_{6a} and $P_4^{(3)}$: it requires *both* sparsity of $\mathcal{W}_r^{(3)}$ and a lower bound on d_q.

Remark 8.2 (The counter-models sit on the equality case)**.** The structural lower bound $q^{s_q} \mid \Phi_{d_q}(r) \le (r+1)^{d_q}$ gives $d_q \ge \frac{s_q \log q}{\log(r+1)} \ge \frac{3\log q}{\log(r+1)}$ on $\mathcal{W}_r^{(3)}$; the right-strictness model of Theorem 8.1 takes $d_j = \frac{3\log q_j}{\log r}$, i.e. the minimal order allowed by $q^3 \le r^d$ (cf. (U1)). The counter-models are therefore the pointwise equality cases of the mass constraint, not arbitrary configurations.

Proposition 8.3 (Structural reading: the Hensel dividend is spent once)**.** *The exact and unique dividing line between Theorem 7.5 and the unconditionality barrier for $\Xi_r^{(3)}$ is the power of q in the order-lifting formula (Theorem 1.1):*

layer	$D_3(q)$	*weight*	*moment order*	*partial summation*	$\sum_d$
$s = 2$	$d\,q$	$\frac{2\log q}{d\,q}$ *carries* $1/q$	*first*	$\frac{2I_d}{d} \asymp \frac{3(\log d)^2}{d^2}$	*converges*
$s \ge 3$	d	$\frac{s\log q}{d}$, *no* $1/q$	*zeroth*	$\le \log r$ *(mass) or* $\frac{3\Delta_d}{2d}$ *(rad)*	*diverges*

The extra factor q from the Hensel wall is the only *object supplying unconditional decay in the whole chain; the $s \ge 3$ layer has spent it ($q^{\max(0,3-s)} = 1$). Any further progress must import counting/distribution information about $\mathcal{W}_r^{(3)}$ itself, not a cleverer block decomposition.*

Theorem 8.4 (A weaker sufficient condition (S))**.** *If there exist a slowly growing f and a constant S with*

$$d_q \ge \frac{q}{f(q)}, \qquad s_q \le S, \qquad \sum_{q \in \mathcal{W}_r^{(3)}} \frac{f(q)\log q}{q} < \infty \quad \text{on } \mathcal{W}_r^{(3)},$$

then $\Xi_q^{(3)} = \frac{s_q q}{d_q} \cdot \frac{\log q}{q} \le S f(q) \frac{\log q}{q}$, so $\Xi_r^{(3)} < \infty$. Numerically $f = 1, \log q, (\log q)^2$ give the $P_4^{(3)}$-type sums $0.542, 1.083, 2.685$ (all convergent). But $d_q \ge q/f(q)$ is not automatic for a sparse $\mathcal{W}_r^{(3)}$: the exceptional set in Erdős–Murty [18] *is $o(\pi(x))$ and may contain $\mathcal{W}_r^{(3)}$ entirely. Hence (S) hangs on an "order × Wieferich" correlation, and is not unconditional progress.*

Proposition 8.5 (Three barriers and an identity)**.** *Three independent routes to an unconditional $\Xi_r^{(3)} < \infty$ are closed; a fourth (the radical channel) is not a barrier at all — it collapses into the threshold itself (Theorem 8.6).*

(1) Per-d constraints are insufficient. *The block constraints $\sum_{q:\, d_q = d} s_q \log q \le d \log r$ (from $q^{s_q} \mid r^d - 1$, Lemma 7.1), $q \equiv 1 \pmod d$, and partial summation give only $\Xi_d^{(3)} \le \log r$ per block. There exist pseudo-configurations meeting all constraints with $\Xi^{(3)}(d \le 2 \times 10^5) = 219549 \sim \log r \log D$ divergent (each block still $\le \log 3 = 1.0986$). This is a method-level impossibility, not a claim about the true factor sets of $r^d - 1$.*

(2) The radical channel is not an independent barrier (erratum). *With $\Delta_d = \log(r^d - 1) - \log \mathrm{rad}(r^d - 1)$: since* log *is additive but* log rad *is not,*

$$\Delta_d = \underbrace{\sum_{c \mid d} \Delta_c^{\Phi}}_{=:\mathrm{PRIM}_d} + \underbrace{\sum_{p:\, d_p \mid d} v_p(d/d_p)\log p}_{=:\mathrm{BOOK}_d}, \qquad \Delta_c^{\Phi} = \sum_{p:\, d_p = c} (s_p - 1)\log p \tag{8.1}$$

for odd p (the prime 2 contributes a bounded LTE correction). The primes contributing to BOOK_d *divide d, whereas $q \in S_d^{(3)}$ forces $q \equiv 1 \pmod d$, hence $q \geq d+1$: the two sets are disjoint, and the "unconditional divergence" $\sum_{d\leq D} \Delta_d/d \geq c \log D$ claimed in an earlier version is produced* entirely *by* BOOK_d *— in fact $\sum_{d\leq D} \mathrm{BOOK}_d/d = (S_{\mathrm{ord}} + o(1)) \log D$ (Theorem 8.6) — carrying no information about $\mathcal{W}_r^{(3)}$. The correct layer bound uses only the $s_p \geq 3$ part of* PRIM_d*:*

$$\Xi_d^{(3)} \;\leq\; \frac{3}{2d} \sum_{c\mid d} \Delta_c^{\Phi,(3)}, \qquad \Delta_c^{\Phi,(3)} = \sum_{\substack{p:\, d_p = c \\ s_p \geq 3}} (s_p - 1) \log p,$$

since $(s_p - 1) \geq \frac{2}{3} s_p$ for $s_p \geq 3$; charging the full PRIM_d *would re-introduce the $s_q = 2$ layer, whose harmonic weight $R_2(D) = \sum_{q\in\mathcal{W}_r^{(2)}} \frac{\log q}{d_q} H_{\lfloor D/d_q \rfloor}$ is $O(D)$, not $O(\log D)$ (numerically* 3.345 *against* $\mathrm{BOOK} = 3.091$ *at $r = 3$, $D = 3000$). Summed, the radical route is the identity of Theorem 8.6: the radical channel* is *the threshold $\Phi_r(D) = o(D)$ of §9, up to the $s_q = 2$ remainder $R_2(D)$ — neither closed nor independently open.*

(3) External radical inputs are all insufficient. *The abc-type input* [22] *gives $\Delta_d/d \leq (1 - \frac{\varphi(d)}{d} + \varepsilon) \log r = O(1)$; Stewart–Yu* [23] *gives unconditionally $\log c \leq \kappa R^{1/3} (\log R)^3$, i.e. only $\log \mathrm{rad} = O(\log d)$; Stewart* [24] *gives $P(a^n - b^n) > n \exp(\frac{\log n}{104 \log\log n})$ — the same; Balaji–Luca* [25]*, Theorem 1.2, give the closest unconditional sparsity, $\#\mathcal{A}_{K,c,u} \cap [1, N] \ll N \exp(-\frac{\log N}{156 \log\log N})$, All four inputs control Δ_d (including* BOOK_d*) and yield only $\Delta_d/d = O(1)$, i.e. $\Phi_r(D) = O(D)$; the Chen–Ding/Graves–Murty bound $C_n' \gg a^{\varphi(n) - \varepsilon n}$ controls the squarefree part of Φ_n and gives $\mathrm{PRIM}_d \leq \varepsilon d \log r$, still merely $\Phi_r(D) = O(\varepsilon D)$. None reaches $o(D)$.*

(4) The DeKoninck–Doyon dichotomy points the other way [21]*: its other branch is "small radical $\Rightarrow$ $\mathcal{W}$ infinite".*

Conclusion: a family of sufficient conditions remains, of which the most easily quoted is

$$\underbrace{\Xi_r^{(2)} < \infty}_{\text{Theorem 7.5, unconditional}} \qquad + \qquad \underbrace{\Xi_r^{(3)} < \infty}_{P_{6a}\ \text{sufficient; not necessary}} \qquad ,$$

and the exact threshold of the unconditional question, together with strictly weaker rungs ($\Xi_r^{(3)}(E) = o(E/\log E)$, the block-counting condition (C)*), is determined in §9.*

Theorem 8.6 (Radical identity)**.** *For every $r \geq 2$ and $D \geq 2$, with $\Delta_d = \log(r^d - 1) - \log \mathrm{rad}(r^d - 1)$,*

$$\sum_{d\leq D} \frac{\Delta_d}{d} = \sum_{q\in\mathcal{W}_r^{(3)}} \frac{(s_q - 1) \log q}{d_q} H_{\lfloor D/d_q \rfloor} + R_2(D) + \big(S_{\mathrm{ord}} + o(1)\big) \log D, \tag{8.2}$$

where $R_2(D) := \sum_{q\in\mathcal{W}_r^{(2)}} \frac{\log q}{d_q} H_{\lfloor D/d_q \rfloor} \leq \frac{\log r}{2} \sum_{c\leq D} H_{\lfloor D/c \rfloor} = O(D)$ by Lemma 7.2, and $R_2(D) = O(\log D)$ whenever $\mathcal{W}_r^{(2)}$ is finite (in particular for every base examined). Hence

$$\frac{2}{3} \Phi_r(D) \;\leq\; \sum_{d\leq D} \frac{\Delta_d}{d} \;\leq\; \Phi_r(D) + R_2(D) + O(\log D),$$

so $\sum_{d\leq D} \Delta_d/d = o(D) \Rightarrow \Phi_r(D) = o(D)$ unconditionally, and if $R_2(D) = o(D)$ — in particular if $\mathcal{W}_r^{(2)}$ is finite — then $\Phi_r(D) = o(D) \iff \sum_{d\leq D} \Delta_d/d = o(D)$.

Proof. By (8.1), swapping the order of summation,

$$\sum_{d\le D}\frac{\mathrm{PRIM}_d}{d}=\sum_{c\le D}\frac{\Delta_c^{\Phi}}{c}H_{\lfloor D/c\rfloor}=\sum_{p\in\mathcal{W}_r}\frac{(s_p-1)\log p}{d_p}H_{\lfloor D/d_p\rfloor}$$

(primes with $s_p=1$ contribute 0), and, since only primes $p\le D$ can divide a $d\le D$ with $d_p\mid d$,

$$\sum_{d\le D}\frac{\mathrm{BOOK}_d}{d}=\sum_{p\le D}\frac{\log p}{d_p}\sum_{m\le D/d_p}\frac{v_p(m)}{m}=\big(S_{\mathrm{ord}}+o(1)\big)\log D,$$

using $\sum_{m\le M}v_p(m)/m=\log M/(p-1)+O_p(1)$ and the convergence of $S_{\mathrm{ord}}=\sum_p\frac{\log p}{d_p(p-1)}$ (Theorem 11.1). Splitting the primitive sum by depth: the coefficients $(s_q-1)/s_q\in[2/3,1)$ on $\mathcal{W}_r^{(3)}$ give the $[\frac23,1)\cdot\Phi_r(D)$ part, and the remaining $s_q=2$ part is $R_2(D)$, bounded by $O(D)$ via Lemma 7.2 exactly as in Theorem 9.2. □

Remark 8.7 (Numerics)**.** For $r=3$, $d=20$: $\Delta_{20}=6.0868=\mathrm{PRIM}\ 3.0910+\mathrm{BOOK}\ 2.9957$ ($3^{20}-1=2^4\cdot5^2\cdot11^2\cdot61\cdot1181$; the only Wieferich contribution is $\Delta_5^{\Phi}=\log 11$). The bookkeeping part grows like $S_{\mathrm{ord}}\log D$ (BOOK $=3.058/4.113/5.340/6.491$ at $D=10^3/3\cdot10^3/10^4/3\cdot10^4$), while the layer-specific bound $\Xi_d^{(3)}\le\frac{3}{2d}\sum_{c\mid d}\Delta_c^{\Phi,(3)}$ keeps the $s_q=2$ layer out: charging PRIM_d in full would add $\frac{\log 11}{5}H_{\lfloor D/5\rfloor}=3.345$ at $D=3000$ — an $O(\log D)$, not $O(D)$, term, but one that must not be charged to $\mathcal{W}_r^{(3)}$.

8.1. **An almost-sure route: the double-zero hypothesis (D4).**

Definition 8.8 ((D4))**.** There is a constant $C_r>0$ such that for every odd prime $q\nmid r$,

$$\mathbb{P}\big(s_q\ge3\big)=\mathbb{P}\big(q^3\mid r^{q-1}-1\big)=\mathbb{P}\big(Q_q^{(1)}=0,\ Q_q^{(2)}=0\big)\le\frac{C_r}{q^2}.$$

A pointwise bound for the Katz quotient at the double-zero scale. (D4) is *not* a consequence of Katz's 2015 conjecture (continuous $\mathbb{R}/\mathbb{Z}$ equidistribution): attempting to place $(Q^{(1)},Q^{(2)})$ in the group-scheme framework $G=\mathbb{G}_m^2$, the cyclic subgroup generated by a single base lies in a proper subtorus and is likely excluded by the Zariski-density hypothesis of Katz's Lemma 4.2 (isomorphic to the $a'=r^k$ degeneration of $(\overline{W}_r,\overline{W}_{a'})$). Hence (D4) must be listed as an independent hypothesis, and it must be distinguished from (D2): (D2) is "two bases × first quotient" modulo q^2 joint, (D4) is "one base × first and second quotients"; both give $1/q^2$, but (D2) does not imply (D4).

Ensemble separation (why (D4) is not free). Let $\mathcal{P}=\prod_{q\ge3}(\mathbb{Z}/q^3\mathbb{Z})^\times$ carry its product Haar measure $\nu=\bigotimes_q\nu_q$, and let $T_q=\mu_{q-1}(\mathbb{Z}/q^3\mathbb{Z})$. Then $\nu_q(T_q)=|\mu_{q-1}|/|(\mathbb{Z}/q^3)^\times|=(q-1)/(q^2(q-1))=1/q^2$ *exactly*, and with the independent indicators $\mathbf{1}_{T_q}$

$$\lambda=\sum_{q\ge3}q^{-2}=P(2)-\tfrac14=0.2022475\ldots,\qquad\nu(K=0)=\prod_{q\ge3}\big(1-q^{-2}\big)=\frac{8}{\pi^2}=0.8105695\ldots,$$

where $K=\#\{q:x_q\in T_q\}$; the exact law of K is Poisson-binomial (the Poisson value $e^{-\lambda}=0.8168927$ is only an approximation, off by 0.78%), and Borel–Cantelli gives $K<\infty$ ν-a.s. Thus in the *Haar ensemble* the pointwise bound of Definition 8.8 holds with $C_r=1$: (D4) is a free theorem of product measure, and Theorem 8.10 becomes an a.s. statement about a ν-random element of $\mathcal{P}$.

The primal question, however, concerns the single point

$$\iota(r)=(r\bmod q^3)_{q\ge3}\in\mathcal{P},\qquad\iota:\mathbb{Z}\hookrightarrow\mathcal{P},$$

the diagonal embedding of the integers. Its image is countable in an uncountable compact group, hence $\nu(\iota(\mathbb{Z})) = 0$. Every ν-a.s. statement is therefore void on $\iota(r)$, and no product-measure argument can yield a statement about a fixed integer r. This resolves the apparent tension between the counting-density reading of (D4) (equivalent to P_{6a}, Theorem 14.4) and the probabilistic reading (Definition 8.8, finite mean, strictly stronger): the two readings live in *different ensembles*, and

$$(\mathrm{D4})_{\mathrm{Haar}} \not\Longrightarrow (\mathrm{D4})_{\mathrm{diag}}.$$

What is missing is precisely a *pointwise rigidity of $\iota(r)$ relative to $\bigotimes \nu_q$*: the assertion that a fixed integer has no systematic preference among the index-q^2 Teichmüller subgroups T_q as q varies. This is the exact content of specification S3 (and of the dual→primal transfer of Proposition 14.5); it is neither a consequence of Katz's continuous $\mathbb{R}/\mathbb{Z}$ equidistribution nor of the free Haar statement above.

Proposition 8.9 (Teichmüller anchor)**.** *For odd $q \nmid r$,*

$$q \in \mathcal{W}_r^{(3)} \iff q^3 \mid r^{q^2} - r \iff r \bmod q^3 \in \mu_{q-1}\big((\mathbb{Z}/q^3\mathbb{Z})^\times\big),$$

where $(\mathbb{Z}/q^3\mathbb{Z})^\times \cong C_{q-1} \times C_{q^2}$ and $|\mu_{q-1}| = q-1$: the subgroup has relative size exactly $1/q^2$.

Proof. Write $r = \omega u$ with $\omega^{q-1} = 1$ (the Teichmüller part) and $u \in 1 + q\mathbb{Z}_q$ (the principal unit). Then $r^{q^2} = \omega^{q^2} u^{q^2} \equiv \omega \pmod{q^3}$, since $q^2 \equiv 1 \pmod{q-1}$ and $(1+qx)^{q^2} \equiv 1 \pmod{q^3}$. Hence $r^{q^2} \equiv r = \omega u \pmod{q^3}$ iff $u \equiv 1 \pmod{q^3}$, i.e. $r \equiv \omega \pmod{q^3}$, i.e. $\mathrm{ord}_{q^3}(r) \mid q-1$; by Theorem 1.1, $\mathrm{ord}_{q^3}(r) = d_q\, q^{\max(0,3-s_q)}$ divides $q-1$ iff $s_q \geq 3$. □

Layer distinction (mod q^2 vs. mod q^3). The subgroup governing (D4) and Proposition 8.9 is $\mu_{q-1}(\mathbb{Z}/q^3\mathbb{Z})$, corresponding to $s_q \geq 3$. The cluster $\{19, 22, 23\}$ at $q = 13$ of §12.3 lies in $\mu_{12}(\mathbb{Z}/169\mathbb{Z})$ (mod q^2, the $s_q \geq 2$ Katz/(D2) layer): $19 \in \mu(\mathbb{Z}/169)$ but $19 \notin \mu(\mathbb{Z}/2197)$. These are different subgroups of different groups; there is no contradiction, and the census of §12.3 tests (D2), not (D4). In the *dual* ensemble (fixed q, r uniform in $(\mathbb{Z}/q^3)^\times$) one has the exact identity $\mathbb{P}(r^{q-1} \equiv 1 \bmod q^3) = \frac{|\mu_{q-1}|}{|(\mathbb{Z}/q^3)^\times|} = \frac{q-1}{q^2(q-1)} = \frac{1}{q^2}$; in the primal ensemble (fixed r, q varying) there is no probability space, so the missing tool is precisely the transfer dual→primal (Proposition 14.5).

Consequently (D4) is the purely algebraic assertion that a fixed integer r has no systematic preference among the index-q^2 Teichmüller subgroups as q varies; the constant $\sum_q q^{-2} = P(2) = 0.452247514\ldots$ is the reciprocal sum of the subgroup indices, not a heuristic input. The Chebotarev obstruction is unchanged: the group varies with q. The running example $\langle 19\rangle \bmod 343 = \{\pm 1, \pm 18, \pm 19\}$ of §12 is the image of μ_6. Numerically: zero counterexamples among $q < 10^5 \times 12$ bases, and $|\mu_{q-1} \bmod q| = q-1$ for every $q \leq 97$ enumerated (see the ensemble separation above).

Set $X_q = \frac{s_q \log q}{d_q} \mathbf{1}_{s_q \geq 3} \geq 0$, so $\Xi_r^{(3)} = \sum_q X_q$.

Theorem 8.10 ((D4) ⇒ $\Xi_r^{(3)} < \infty$ and P_{6a}, a.s.)**.** *Under* (D4)*,*

$$\mathbb{E}\big[\Xi_r^{(3)}\big] \leq C_r(\log r)\, P(2) = 0.452247514\ldots \cdot C_r \log r < \infty,$$

where $P(2) = \sum_q q^{-2}$ (over all primes). Hence $\Xi_r^{(3)} < \infty$ a.s.; and by the first Borel–Cantelli lemma (no independence needed), $\#\mathcal{W}_r^{(3)} < \infty$ a.s., i.e. P_{6a}/Murata's conjecture (i) holds almost surely under (D4)*.*

Proof. By (U2) ((2.1)) the pointwise deterministic bound $X_q \le (\log r)\mathbf{1}_{s_q\ge 3}$ holds, independent of d_q; hence $\mathbb{E}[X_q] \le (\log r)\,\mathbb{P}(s_q \ge 3) \le C_r \log r/q^2$. Since $X_q \ge 0$, Tonelli permits term-by-term expectation, and $\sum_q q^{-2} < \infty$. Finite expectation implies a.s. finiteness; Borel–Cantelli I needs only $\sum_q \mathbb{P}(s_q \ge 3) \le C_r P(2) < \infty$ and yields $\#\mathcal{W}_r^{(3)} < \infty$ a.s. □

Remark 8.11. The proof needs no conditioning on d_q, no $k \ge 4$ layer hypothesis, no lower bound on orders, and no GRH: (U2) absorbs the ord side entirely, bypassing the "order × Wieferich" correlation of the (D2)/(D3) side. If (D4) is stated in conditional form one gets $\mathbb{E}[\Xi_r^{(3)}] \le C_r \sum_q \frac{\log q}{d_q q^2}$, whose true values are 0.0346–0.1470 for $r = 2, 3, 5, 7$; the proof itself does not use this refinement.

Scale. (D4) yields $\mathbb{E}[\#\mathcal{W}_r^{(3)}(x)] \le C_r P(2) = O(1)$, not $\ll \log\log x$ (the latter is the $\sum_q 1/q$ scale of (D1)). Read as a *counting density*, $\#\mathcal{W}_r^{(3)}(x) \le C_r \sum_{q\le x} q^{-2} = O(1)$, (D4) is *equivalent* to P_{6a} (Theorem 14.4); it is the *probabilistic* form of Definition 8.8, with the finite mean $\sum_q C_r/q^2$, that is strictly stronger.

(K) alone is insufficient. (D1) gives only $O(1/q)$, and $\sum_{q<Q} 1/q = 2.198/2.483/2.705/2.887$ at $Q = 10^3, 10^4, 10^5, 10^6$ diverges like $\log\log Q$; hence "(K) $\Rightarrow \Xi_r^{(3)} < \infty$" fails, and one must pass to $(\mathrm{K}^+) = (\mathrm{K}) \cup \{(\mathrm{D4})\}$.

9. The exact threshold: what an unconditional proof requires

Write $\delta(E) = \frac{1}{E}\#\{e \le E : e \text{ exceptional}\}$. Markov with the moving threshold $c = \log r/20$ gives, unconditionally,

$$\delta(E) \;\le\; \frac{20}{E\log r}\sum_{e\le E}\frac{T_s(e)}{e}, \tag{9.1}$$

so the exact threshold of the first-moment route is

$$(\mathrm{M}): \qquad \sum_{e\le E}\frac{T_s(e)}{e} = o(E). \tag{9.2}$$

Condition (M) is sufficient but *not* necessary for $\delta(E) \to 0$ (the Markov upper bound may be non-tight: the mass of T_s can concentrate on few e). We claim only that (M) is the weakest operational sufficient condition of the first-moment route, and it is strictly weaker than every condition used above (Theorem 9.4).

9.1. **Case A: the $s_q = 2$ layer is elementary.** Let $a' = r^k$ (including $a' = 1$), so $b_e = r^k(r^{e-k} - 1)$. For $q \nmid r$, LTE gives, when $d_q \mid e - k$,

$$V_q(e) = v_q(r^{e-k} - 1) = s_q + v_q\Big(\frac{e-k}{d_q}\Big) \;\ge\; s_q,$$

and $V_q(e) = 0$ otherwise; $q \mid e - k \iff d_q \mid e - k$ since $d_q \mid q - 1$ gives $\gcd(d_q, q) = 1$.

Proposition 9.1 (Pointwise bounds, Case A, unconditional)**.** *For $e > k$,*

$$T_s^{(2)}(e) = \sum_{\substack{q\mid e-k\\ q\in\mathcal{W}_r^{(2)}}} 2\log q \;\le\; 2\sum_{q\mid e-k}\log q \;\le\; 2\log(e-k) \;\le\; 2\log e,$$

hence $\sum_{e\le E} T_s^{(2)}(e)/e \le 2\sum_{e\le E}\frac{\log e}{e} = O\big((\log E)^2\big) = o(E)$, unconditionally. Moreover

$$T_s(e) \;\le\; \sum_{q^3\mid b_e} V_q(e)\log q \;\le\; \log(r^{e-k} - 1) \;<\; e\log r,$$

so the exception threshold $e\log r/20$ *sits a factor* 20 *below the unconditional pointwise bound: density zero needs the* average *to be far below the bound. Numerically (*$r=2,3,5,7$*;* $\mathcal{W}_r^{(2)}$ *below* 10^6 *is* $\{5,11,1093,3511,20771,40487\}$*):* $\max_{e\le E} T_s^{(2)}(e) = 8.0147/8.0147/17.2122/22.0080$ *at* $E=10^2/10^3/10^4/10^5$*, against* $2\log E = 9.2103/13.8155/18.4207/23.0259$ *— zero violations.*

Thus in Case A the density-zero conclusion for the $s_q=2$ layer needs neither Theorem 7.5 nor any hypothesis: Proposition 9.1 suffices. Theorem 7.5 remains the stronger result and the required input for Case B (where $V_q(e)\ge s_q$ fails and no pointwise bound holds). Note $\mathcal{C}_{r,r^k}=\mathcal{W}_r$ (Proposition 5.3), so Case A never meets the bad set at all: (H-j) enters only through Case B.

9.2. **The harmonic threshold** $\Phi_r(E)$**.** For $q\in\mathcal{W}_r^{(3)}$ the solution class of $q^3\mid b_e$ is $e\equiv e_0 \pmod{d_q}$ (wall-free, $D_3=d_q$), and for any a' the class sum satisfies

$$\sum_{\substack{e\le E\\ q^3\mid b_e}}\frac{1}{e} \;\le\; \frac{1}{e_0(q)}+\frac{H_{\lfloor E/d_q\rfloor}}{d_q}.$$

Define

$$\Phi_r(E) := \sum_{\substack{q\in\mathcal{W}_r^{(3)}\\ d_q\le E}}\frac{s_q\log q}{d_q}\,H_{\lfloor E/d_q\rfloor}. \tag{9.3}$$

Then $\sum_{e\le E}\frac{T_s^{(3)}(e)}{e}\le \Phi_r(E)+O(1)$, and together with Proposition 9.1 (Case A; the same bound holds for general a'):

$$\sum_{e\le E}\frac{T_s(e)}{e} \;\le\; (\log E)^2+\Phi_r(E)+O(\log E).$$

Theorem 9.2 (Unconditional linear cap)**.** *Let* $S_d^{(3)}=\{q\in\mathcal{W}_r^{(3)}: d_q=d\}$*. Since* $d_q=d$ *forces* $d\mid q-1$ *(hence* $q\nmid d$*) and* $v_q(r^d-1)=s_q$*, we have* $\prod_{q\in S_d^{(3)}}q^{s_q}\mid \Phi_d(r)$ *(the* d*-th cyclotomic value), and* $\Phi_d(r)\le (r+1)^{\varphi(d)}$*; therefore*

$$\mathrm{mass}_d := \sum_{q\in S_d^{(3)}} s_q\log q \le \varphi(d)\log(r+1),$$

and, with $H_m\le \log m+\gamma+o(1)$ *uniformly,*

$$\Phi_r(E) \;\le\; \log(r+1)\sum_{d\le E}\frac{\varphi(d)}{d}\Big(\log\frac{E}{d}+\gamma+o(1)\Big) \;=\; \Big(\frac{6}{\pi^2}+o(1)\Big)(1+\gamma)\log(r+1)\,E.$$

The constant is $C=\frac{6}{\pi^2}(1+\gamma)=0.958832\ldots$*; the resulting cap* $\frac{20C\log(r+1)}{\log r}$ *equals* $30.39, 24.20, 21.35, 20.49, 19.87$ *for* $r=2,3,5,7,11$ *(limit* 19.18 *as* $r\to\infty$*) — all trivial, since* $\delta(E)\le 1$*. The same mass bound gives the linear estimate* $\Xi_r^{(3)}(E):=\sum_{d_q\le E}\frac{s_q\log q}{d_q}\le \frac{6}{\pi^2}\log(r+1)E$ *(*$0.668/0.843/1.089/1.264$ *for* $r=2,3,5,7$*).*

So the unconditional bound is *linear*, while the target is $o(E)$: the required saving is a factor $\Omega(\log E)$ — not a constant factor and not an infinite methodological gulf. It lands precisely on "P1/P2-level counting of $\mathcal{W}_r^{(3)}$", which is the varying-modulus obstruction itself.

Tighter budget and its saturation. For each d, $\prod_{q:\, d_q=d} q^{s_q} \mid \Phi_d(r) \le (r+1)^{\varphi(d)}$, so $\mathrm{mass}_d \le \varphi(d)\log(r+1)$, a factor 1.25–2.95 tighter than $\log(r^d-1) < d\log r$ (e.g. $r=19$, $d=6$: $\varphi(6)\log 20 = 5.9915$ versus $6\log 19 = 17.67$). The bound is *attained* at the observed deep hits: $\Phi_2(26) = 27 = 3^3$, $\Phi_2(124) = 125 = 5^3$ and $\Phi_2(80) = 81 = 3^4$ give $\rho_d^{\Phi} := \mathrm{mass}_d/(\varphi(d)\log(r+1)) = 1$ exactly, and $\Phi_6(19) = 343 = 7^3$ gives 0.974, $\Phi_4(239) = 2\cdot 13^4$ gives 0.936. Hence the volume bound carries no cross-q slack.
A stratification trap. Collecting primes from $q^3 \mid r^d - 1$ under the weaker condition $d_q \mid d$ is unsafe: LTE gives $v_q(r^d-1) = s_q + v_q(d/d_q)$, so a *classical* prime with $q \mid d/d_q$ reaches exponent 3 for free. Concrete false positives: $(r,q,d_q,d) = (3,11,5,55)$ has $v_{11}(3^{55}-1) = 3$ but $s_{11} = 2$; likewise $(7,5,4,20)$ and $(2,1093,364,397852)$. One must stratify by $d_q = d$ exactly and exclude $q \mid d/d_q$. Equivalently, since $d_q \mid q-1$ forces $q > d_q$ and $q \nmid d_q$, the bookkeeping contamination requires $d \ge q\, d_q > q^2$; hence for $d \le E$ and $q > \sqrt{E}$ one has $q^3 \mid r^d - 1 \iff s_q \ge 3$ and $d_q \mid d$, with no contamination.

Remark 9.3 (Structural empty layers and the role of smoothness)**.** Since $q \in S_d^{(3)}$ forces $d \mid q-1$ with q odd, $2d \mid q-1$ for odd d; writing $q_{\min}(d)$ for the least prime $\equiv 1 \pmod d$, the divisibility $q^3 \mid \Phi_d(r) \le (r+1)^{\varphi(d)}$ gives the necessary condition

$$3\log q_{\min}(d) \;\le\; \varphi(d)\log(r+1),$$

stronger than $q \ge d+1$. For $r=2$ the structurally empty $d \le 299$ are $\{2,\dots,10,12,14,15,18,20,24,30\}$, the first admissible being $d=11$; the witness $(r,q) = (19,7)$ has $d_q = 6$ and sits on the boundary ($3\log 7 = 5.8377$ against $\varphi(6)\log 20 = 5.9915$). For $d \ge 50$ the condition holds automatically, so it excludes only finitely many small d and has no asymptotic strength. Moreover $\varphi(d)/d = \prod_{p\mid d}(1-1/p)$, so large $\varphi(d)/\log d$ occurs for *rough* d, not smooth d (pass rates 58.8% versus 71.9% for $r=2$, $d \in [2,50]$); summing the mass over d gives $\sum_{d\le x}\varphi(d) \asymp \frac{3}{\pi^2}x^2$, a quadratic bound, as recorded before Theorem 9.2.

Theorem 9.4 (The weakest sufficient condition; none of the quoted conditions is necessary)**.** *Since $H_{\lfloor E/d\rfloor} \le \log E + \gamma$,*

$$\Xi_r^{(3)}(E) = o\Big(\frac{E}{\log E}\Big) \implies \Phi_r(E) = o(E) \implies \delta(E) \to 0.$$

Neither $K_{r,a'} < \infty$ nor $\Xi_r^{(3)} < \infty$ is necessary for density zero. Indeed there are configurations meeting every per-block structural constraint with

$$\Xi_d = \frac{1}{(\log d)^2} \quad (d \ge 5), \qquad \Xi_r^{(3)}(E) \sim \frac{E}{(\log E)^2} \to \infty, \qquad \Phi(E) = \sum_{d\le E} \Xi_d\, H_{\lfloor E/d\rfloor} = o(E):$$

E	$\Xi_r^{(3)}(E)$	$\Phi_r(E)$	$\Phi_r(E)/E$	*bound (9.1) on* $\delta(E)$
10	32.53	69.11	0.0691	2.127
10^4	160.1	323.9	0.0324	0.744
10^5	943.6	1782	0.0178	0.349
10^6	6245	11260	0.0113	0.208
10^7	44500	77980	0.0078	0.142

divergent $\Xi_r^{(3)}$ with $\delta(E) \to 0$. These are logical models (satisfying all per-block constraints), not claims about the true factor sets of $r^d - 1$.

The ladder of sufficient conditions, weakest first, all unproved unconditionally:

#	statement	status
(M)	$\Phi_r(E) = o(E)$	the exact threshold of the Markov route; equivalent to $\sum_e T_s(e)/e = o(E)$
(M')	$\Xi_r^{(3)}(E) = o(E/\log E)$	sufficient for (M); strictly weaker than $\Xi_r^{(3)} < \infty$
(C)	$N_3(E) := \#\{d \le E : S_d^{(3)} \neq \emptyset\} = o(E/\log E)$	sufficient for (M') (per-block $\Xi_d^{(3)} \le \log(r+1)$); this is P1/P2-level counting of $\mathcal{W}_r^{(3)}$
(P)	$\Xi_r^{(3)} < \infty$	sufficient for (M'); the previously quoted condition — not necessary
P_{6a}	$\mathcal{W}_r^{(3)}$ finite (Murata 1991 (i))	the classical anchor; sufficient for (P)
(K^+)	(K) + (D4)	gives (P) almost surely (Theorem 8.10), not unconditionally

For orientation: $N_3(E) \le \sum_{d\le E} \frac{\varphi(d)\log(r+1)}{3\log(d+1)} \asymp \frac{(\log r)E^2}{6\log E}$ is quadratic — useless; a genuine sub-quadratic bound on $N_3(E)$ is itself open. In one sentence: an unconditional proof requires a sublinear bound on the deep-Wieferich mass weighted by $1/d_q$ — a P1/P2-level statement restricted to $\mathcal{W}_r^{(3)}$ and weighted by ord; it is neither manifestly easier nor manifestly harder than P1, but it lives in the same no-fixed-modulus no-man's-land. Silverman's *abc* route gives a *non-Wieferich lower* bound — the opposite direction; the other branch of [21] gives $\mathcal{W}$ *infinite*. No unconditional tool currently produces upper bounds of this type.

Theorem 9.5 (Information-theoretic barrier)**.** *Let $q \in \mathcal{W}_r$, $r^{q-1} = 1 + q^{s_q}u$, $q \nmid u$, and suppose $q^3 \mid r^e - a'$. Raising to the $(q-1)$-st power,*

$$a'^{q-1} \equiv (1+q^{s_q}u)^e \equiv 1 + e\,q^{s_q}u \pmod{q^3},$$

the $k \ge 2$ binomial terms carrying $q^{2s_q} \ge q^4$. For $s_q = 2$ the q^2 term survives and $e \equiv \beta_q w_q^{-1} \pmod q$ is solvable — this is the second-Fermat-quotient formula of Theorem 1.2, the entire information content of (H-j)*/(K). For $s_q \ge 3$ the term $e\,q^{s_q}u \equiv 0 \pmod{q^3}$: $a'^{q-1} \equiv 1 \pmod{q^3}$ independent of e, and no congruence expansion of any finite order recovers e; the least solution $e_0(q)$ is determined only by the discrete-logarithm structure of $\langle r\rangle \subseteq (\mathbb{Z}/q^3\mathbb{Z})^\times$ — a varying-modulus discrete logarithm.*

Numerically: four artificial bases $r \equiv 1, 57, 68, 124 \pmod{125}$ with $r^4 \equiv 1 \pmod{125}$ (forced $s \ge 3$ at $q = 5$) give $a'^4 \equiv 1 \pmod{125}$ identically in $e = 1, \dots, 29$ — no e-information; while for the genuine pair $(r, q) = (3, 11)$, $w_q = 4$ and $a'^{10} \equiv 1 + 100e \cdot 4 \pmod{1331}$ predicts e mod 11 exactly for $e = 1, 2, 3, 5, 7$. Conclusion: all information about the $s \ge 3$ layer lives in the ord and counting structure of $\mathcal{W}_r^{(3)}$, not in any congruence expansion — the same obstruction, which now has seven guises: **(T1)** GRH insufficiency (Theorem 1.5); **(T2)** the radical channel — withdrawn as a barrier and replaced by the identity of Theorem 8.6; **(T3)** the expansion barrier above (Theorem 9.5); **(T4)** no cross-d repulsion from divisibility identities — a fixed q has a unique d_q with $q \equiv 1 \pmod{d_q}$, so distinct d's carry disjoint q-sets, all multiples $e = md_q$ are consistent by LTE, and $\prod_{d\mid n} \Phi_d(r) = r^n - 1$ with $n = \mathrm{lcm}(1..E) \sim e^E$ is too loose by ~ 33 orders of magnitude; **(T5)** no cited theorem controls the cube-full part of $\Phi_n(a)$ (Theorem 14.2); **(T6)** all structural constraints together give *zero* asymptotic improvement on $A_r(t)$ (Theorem 14.2); and **(T7)** the missing dual→primal transfer of (D4) (Proposition 14.5).

Theorem 9.6 (Exact threshold, dyadic form)**.** *Put* $M_d = \sum_{q:\, d_q = d,\, s_q \ge 3} s_q \log q$ *and* $A_r(t) = \sum_{d \le t} \frac{M_d}{d}$*. Then* $\Phi_r(E) = o(E)$ if and only if $A_r(t) = o(t)$*; precisely*

$$\tfrac{3}{2}\, A_r(E/2) \;\le\; \Phi_r(E) \;\le\; (1+\log 2)\, A_r(E) + \log 2 \sum_{j \ge 1} A_r(E/2^j),$$

so $A_r(t) \le \varepsilon t$ *gives* $\Phi_r(E) \le (1 + 2\log 2)\,\varepsilon E = 2.386294\,\varepsilon E$*. A saving* $M_d = d/\log d$ *is* not *enough (*$\Phi(E)/E \approx 0.17$ *at* $E = 10^5$*); one needs* $M_d/d \to 0$ *in Cesàro mean. In particular* $\Xi_r^{(3)}(E) = o(E/\log E)$ *of Theorem 9.4 is a sufficient rung, not the threshold itself.*

Proof. Group $\Phi_r(E) = \sum_{d \le E} \frac{M_d}{d} H_{\lfloor E/d \rfloor}$ (Theorem 9.2). For $d \le E/2$ the floor is ≥ 2, so $H_{\lfloor E/d \rfloor} \ge H_2 = \frac{3}{2}$, giving the lower bound. For the upper bound split $[1, E]$ dyadically: on $(E/2^{j+1}, E/2^j]$ one has $H_{\lfloor E/d \rfloor} \le (j+1)\log 2 + 1$, and Abel reordering gives the stated bound. If $A_r(t) \le \varepsilon t$ for all $t \le E$, then $\sum_{j \ge 1} A_r(E/2^j) \le \varepsilon E \sum_{j \ge 1} 2^{-j} = \varepsilon E$. □

Theorem 9.7 (Container; the gap is exactly one $o(1)$ factor)**.** *Let* $C_E = \prod_{d \le E} \Phi_d(r)$*. Then* $\prod_{q \in \mathcal{W}_r^{(3)},\, d_q \le E} q^{s_q} \mid C_E$ *and*

$$\log C_E = \sum_{d \le E} \Big(\varphi(d) \log r + O\big(2^{\omega(d)} \log r\big) \Big) = \frac{3}{\pi^2} E^2 \log r \,\big(1 + o(1)\big),$$

since $\sum_{d \le E} \varphi(d) = \frac{3}{\pi^2} E^2 + O(E \log E)$ *and* $\sum_{d \le E} 2^{\omega(d)} = O(E \log E)$*. The volume bound gives*

$$A_r(t) \;\le\; \log(r+1) \sum_{d \le t} \frac{\varphi(d)}{d} = \frac{6}{\pi^2} \log(r+1)\, t \,\big(1 + o(1)\big),$$

with $\sum_{d \le t} \frac{\varphi(d)}{d} \big/ \big(\frac{6}{\pi^2} t\big) = 1.000658, 1.000761, 1.000375, 1.000095$ *at* $t = 10^2, 3 \cdot 10^2, 10^3, 3 \cdot 10^3$*. Thus the container budget is quadratic, the first moment is linear, and what is known differs from what is needed by exactly one factor* $o(1)$ *per* d *— the* $\Omega(\log E)$ *saving of §9 in closed form. In particular the earlier* $\operatorname{lcm}(1..E) \sim e^E$ *argument is unnecessary:* C_E *itself is a fixed integer of size* $\exp(\Theta(E^2))$*.*

10. Locating (H-j): Katz, Murata, and the GRH ruling

Write $q_q(u) = k_u + q\, m_u$ with $k_u, m_u \in [0, q)$. Then $\overline{W}_u(q) := W_u(q) \bmod \mathbb{Z} = \frac{q_q(u)}{q} \bmod \mathbb{Z}$ has fractional part k_u/q, so $q \in \mathcal{W}_u \iff k_u = 0 \iff \overline{W}_u(q) = 0$; and when $q \in \mathcal{W}_u$ one has $m_u(q) \equiv (u^{q-1} - 1)/q^2 \pmod q$, i.e. $m_r = w_q$, $m_{a'} = \beta_q$. Thus (w_q, β_q) is the mod-q higher layer of the Katz quotient at its Haar-null zeros: one order higher than $\overline{W}$ (the mod-q^2 information, equivalently $u^{q-1} - 1 \bmod q^3$). If $a' = r^k$ then $q_q(r^k) \equiv k\, q_q(r) \pmod q$ and the pair $(\overline{W}_r, \overline{W}_{a'})$ lies on the one-dimensional subtorus $\{(x, kx)\}$: exactly the degenerate case excluded by the Zariski-density hypothesis of the group-scheme version, matching the Case A/B dichotomy. (Numerically, the joint χ^2 for $(3, 9)$ and $(3, 27)$ explodes to 7736 and 3454 against a critical value 36.4, while independent bases pass; see §12.)

The continuous equidistribution of Katz does not control point masses at scale $1/q$. (H-j) requires the three-layer discretized joint refinement:

- (D1) *Pointwise discretization*: $\#\{q \le x : \overline{W}_u(q) = 0\} \sim \sum_{q \le x} 1/q \sim \log\log x$ (Lang's fixed-residue statement [4] and the Crandall–Dilcher–Pomerance model [5]);
- (D2) *Mod-q^2 two-base joint*: for multiplicatively independent $\langle r, a' \rangle$, the pair (w_q, β_q) is jointly uniform on $\mathbb{F}_q^2$ and $P(q \in \mathcal{W}_r \cap \mathcal{W}_{a'}) \asymp 1/q^2$;

(D3) *Discrete-logarithm layer*: conditional on compatibility, j_q is uniform on $[0, d_q)$, so $P(j_q = 0) = 1/d_q$.

Set (K)=(D1)+(D2)+(D3). None of (D1)–(D3) is a GRH theorem or a direct consequence of Katz's conjecture.

Proposition 10.1 (conditional on the random model (K))**.** *Assume* (K) *and that* r, a' *are multiplicatively independent (Case B). Then*

$$\mathbb{E}\Big[\sum_{q\in\mathcal{B}_{r,a'}} \frac{s_q \log q}{b_q}\Big] \ \ll\ \log r \sum_q \frac{\log q}{q^2} < \infty,$$

so (H-j) *holds almost surely; together with* P_4 *and* $\Xi_r^{(3)} < \infty$ *one obtains* $K_{r,a'} < \infty$ *and the density bound of Theorem 1.3.*

Proof. Here (K) is used with its built-in independence: (D2) and (D3) are joint uniformities, so (w_q, β_q) and j_q are independent, and $s_q \log q$ (which depends only on r) is independent of b_q. By (D2)–(D3) $P(q \in \mathcal{B}) \le P(q \in \mathcal{W}_r \cap \mathcal{W}_{a'}) \cdot P(j_q = 0) \le q^{-2} \cdot d_q^{-1} \le q^{-2}$; by (D2) conditional uniformity b_q is uniform on $\mathbb{F}_q^\times$ (note $w_q \neq 0$ on the layer $s_q = 2$), so $\mathbb{E}[1/b_q] = H_{q-1}/q \asymp (\log q)/q$; and $s_q \log q < d_q \log r \le q \log r$ by (2.1). Multiplying, $\mathbb{E}[\text{contribution of } q] \ll q^{-2} \cdot q \log r \cdot (\log q)/q = \log r \cdot (\log q)/q^2$, summable. Nonnegativity plus Markov give a.s. convergence. The same estimate controls $\Xi_{r,a'} \ll \sum_q (\log q)/q^2$ directly. □

Sketch of the GRH ruling (Theorem 1.5). (i) *Literature*: no theorem of the form GRH $\Rightarrow$ $N_r(x) \ll x^{1/2} \log^A x$, GRH $\Rightarrow$ $\mathcal{W}_r$ has density 0, or GRH $\Rightarrow$ $\Sigma_r < \infty$ is known; the varying-modulus side is precisely where the classical literature is silent. (ii) *Algebra*: the congruence $q^2 \mid r^{q-1} - 1$ has modulus q itself; the natural Chebotarev fields $L_q = \mathbb{Q}(\zeta_{q^2}, r^{1/q^2})$ vary with q, are wildly ramified, and satisfy $\log d_{L_q} \asymp q^2 \log q$, so the termwise error in quantitative GRH–Chebotarev, $\sum_{q\le x} x^{1/2} \log(d_{L_q} x^{n_{L_q}})/q^{1/2}$, is uncontrolled; the averaged versions require a family with bounded discriminant growth, which this family is not; Hooley's GRH Artin method handles only fixed d in $\mathbb{Q}(\zeta_d, r^{1/d})$. (iii) *Direction*: GRH controls only the $\mathrm{ord}_q(r)$ (Artin-type) distribution, and the Hensel wall has already converted the weight $1/\mathrm{ord}_q(r)$ into $1/D_3 = 1/(d_q q^{\max(0,3-s_q)})$; hence in any composite "GRH + W-side assumption" package, GRH is redundant. □

Remark 10.2 (Classical coordinates)**.** Murata's 1991 conjectures [3]: (i) $p^3 \mid a^p - a$ for only finitely many p — exactly P_{6a}; (ii) for coprime $b, c > 1$, $p^2 \mid b^p - b$ and $p^2 \mid c^p - c$ simultaneously for only finitely many p — exactly $\#(\mathcal{W}_r \cap \mathcal{W}_{a'}) < \infty$, the strong form of (HW). Our (H-j) is *weaker* than Murata (ii) (the intersection is allowed to be infinite; only the bad subset must be $1/b_q$-summable) and *stronger* than Katz's conjecture (which is continuous equidistribution and silent on all three discretized layers). The correct description of (H-j) is: a weighted refinement interpolating between Katz 2015 and Murata 1991.

10.1. **The unified family** (K^+) **and the main line.** Recall $(\mathrm{K}^+) = (\mathrm{K}) \cup \{(\mathrm{D4})\}$ with (D4) from §8.1.

Theorem 10.3 (Unified conclusion)**.** *Under* (K^+)*, for every fixed* $r \ge 2$ *and every* a' *with* $\gcd(a', r) = 1$*:* $K_{r,a'} < \infty$ *almost surely; consequently the density bound of Theorem 1.3 holds almost surely for every* a'*.*

Proof. Case A: by Proposition 5.3, $K_{r,r^k} \le 2\Xi_r = 2(\Xi_r^{(2)} + \Xi_r^{(3)})$; $\Xi_r^{(2)} < \infty$ is the unconditional Theorem 7.5, and $\Xi_r^{(3)} < \infty$ a.s. is Theorem 8.10. Case B: $K_{r,a'} = \Xi_{r,a'} + J' \le \Xi_r + J'$; the Ξ_r

part is controlled as before (Theorems 7.5 and 8.10); J' is controlled by (D1)–(D3): on the bad set $\mathbb{P}(q \in \mathcal{B}) \ll q^{-2} d_q^{-1} \le q^{-2}$, conditionally $\mathbb{E}[1/b_q] \asymp (\log q)/q$, and with (U2),

$$\mathbb{E}\Big[\sum_{q \in \mathcal{B}_{r,a'}} \frac{s_q \log q}{b_q} \Big] \ll \log r \sum_q \frac{\log q}{q^2} < \infty,$$

i.e. (H-j) holds a.s., so $J' < \infty$ a.s. Finally apply Theorem 1.3. □

Erratum note. An earlier draft claimed that "(D4) does not participate in Case B"; this is inaccurate. In Case B, $\Xi_{r,a'} \le \Xi_r$ still contains the $s \ge 3$ layer, which must be controlled by (D4); (K) alone controls only J'.

Honest strength ordering.

$$(\mathrm{K}^+) \;\Rightarrow\; P_{6a} \text{ a.s. } \Rightarrow\; \Xi_r^{(3)} < \infty \;\Rightarrow\; P_4^{(3)},$$

the last three arrows by Theorem 8.1. Read as a counting density, (D4) is *equivalent* to P_{6a} (Theorem 14.4); its probabilistic form (Definition 8.8, finite mean) is strictly stronger, yielding $\Xi_r^{(3)} < \infty$ a.s. Its value lies in being same-origin with (K) and in covering the main line and the bad-set line uniformly. The final condition set of the chain may be stated either as the classical anchor P_{6a} or as the unified conjecture family (K^+).

11. Ancillary results and the redundancy of GRH

Three results decide which hypotheses are redundant in the chain; none of them fills the application-layer gap, but together they show GRH is nowhere needed.

Theorem 11.1 ($S_{\mathrm{ord}} < \infty$, unconditional)**.** $S_{\mathrm{ord}} = \sum_{p \nmid r} \frac{\log p}{d_p(p-1)} < \infty$.

Proof. Erdős–Murty [18] prove unconditionally $\sum_{p \le x} 1/d_p \ll \sqrt{x}/(\log x)^{1+\delta}$ (Murty–Srinivasan [19] already gave $\ll \sqrt{x}$). With the decreasing weight $w(t) = \frac{\log t}{t-1}$ $(t \ge 3)$, Stieltjes partial summation gives

$$\sum_{3 \le p \le x} \frac{w(p)}{d_p} = A(x)w(x) - \int_3^x A(t)w'(t)\,dt,$$

with $A(x) = \sum_{p \le x} 1/d_p$; the boundary term is $\ll x^{-1/2}(\log x)^{-\delta} \to 0$, and after $t = e^u$ the integral is $\ll \int^\infty e^{-u/2} u^{-\delta}\,du < \infty$. □

The GRH version of the order bound (Kurlberg–Pomerance [16], $d_p \ge p/\varepsilon(p)$ on average) is not needed anywhere: GRH is redundant throughout the chain.

Theorem 11.2 (C_r exact identity)**.** *With* $X_n = \log \frac{n}{\operatorname{rad} n}$*, for* $a' = 1$*,*

$$\sum_{e \le E} X_{r^e - 1} = \sum_{p \nmid r} \log p \sum_{k \ge 2} \Big\lfloor \frac{E}{D_k(p)} \Big\rfloor + \mathrm{LTE}_2(E),$$

so $\frac{1}{E}\sum_{e \le E} X_{r^e-1} \nearrow C_r = S_W + S_{\mathrm{ord}} + LTE_2(r)$*, where* $S_W = \sum_{q \in \mathcal{W}_r} \frac{(s_q - 1)\log q}{d_q}$*. Hence* (H-X) $\iff$ $S_W < \infty$*. Numerically* $C_2 = 0.6938$*,* $C_3 = 1.9112$*,* $C_5 = 1.9745$*; for* $r = 3$ *the decomposition is* $S_W = 0.479579$*,* $S_{\mathrm{ord}} = 0.391941$*,* $LTE_2 = 1.039721$*, summing to* $C_3 = 1.911241$*.*

Indeed $X_n = \sum_p \log p \sum_{k \ge 2} \mathbf{1}_{v_p(n) \ge k}$, and $v_p(r^e - 1) \ge k \iff D_k(p) \mid e$ (the LTE tower of Theorem 1.1); the prime $p = 2$ contributes the finite correction $\mathrm{LTE}_2(E)$. For $p \in \mathcal{W}_r$ $(s_p \ge 2)$ the limiting contribution is $\log p \sum_{k=2}^{s_p} 1/D_k(p) = \frac{(s_p - 1)\log p}{d_p}$; for $p \notin \mathcal{W}_r$ it is $\frac{\log p}{d_p(p-1)}$.

Theorem 11.3 (Strict one-way)**.** *Termwise* $\Xi_r \le \frac{3}{2}S_W + O(1)$ *(ratios* $\Xi_q/S_{\mathcal{W},q}$*:* $s=2$ *gives* $2/q$*,* $s=3$ *gives* $3/2$*,* $s\ge 4$ *gives* $s/(s-1)$*), so* $S_W<\infty \Rightarrow \Xi_r<\infty$*. The converse fails: the exact identity* $S_W^{(2)} = \sum_q \frac{q}{2}\Xi_q^{(2)}$ *shows* S_W *is the* q*-weighted first moment of* Ξ_r*; the universal bound* $\Xi_q^{(2)} \le 2\log^2 q/q^2$ *is summable over all primes (*1.2429*), while the corresponding* S_W *terms* $\sim \log^2 q/q$ *diverge (*92.65*). A configuration with* $d_q = q-1$ *and* $\mathcal{W}_r$ *of positive density makes* $\Xi_r \approx 0.818$ *convergent while* $S_W \sim \frac{1}{m}\log X$ *diverges; since P1 is open, no theorem excludes it.*

Theorem 11.4 (Evolution of the condition set)**.**

$$P_4 + \Xi_r^{(3)} \;\longrightarrow\; P_4' + \Xi_r^{(3)} \;\longrightarrow\; \Xi_r^{(3)} \text{ alone},$$

where $P_4' = \sum_{q\in\mathcal{W}_r} 1/q < \infty$ *was once supported by* $d_q \ge \log q/\log r$ *((*2.1*), giving* $\Xi_q^{(2)} \ll 1/q$*. Theorem 7.5 removes* P_4*,* P_4' *and* P_1 *from the condition set entirely (they remain open as statements about* $\mathcal{W}_r$*). The (H-X)/*S_W *route is a stronger-than-necessary fallback: strictly one-way (Theorem 11.3), not equivalent to the main line.*

route	Wieferich side	ord side	GRH
A (main line; the Ξ part of Cases A and B)	$\Xi_r^{(2)}$ unconditional + $\Xi_r^{(3)}$ single condition	Hensel wall $1/D_3$ and (U2) absorb it	redundant
B (side line; the J' part of Case B)	(H-j) / (K)	not needed	insufficient (needs Katz pointwise strengthening)

(K^+) collects both routes into one family: (D4) governs the deep layer of A, (D1)–(D3) govern the bad set of B.

12. Computational evidence

Variable direction. Every statistic in this section is a fixed-r, varying-q scan, and the hypotheses they test — (D1), (D2), (D4), (K) — are likewise fixed-r, varying-q pointwise statements. A fixed-q census over bases r is a different, easier object: by Proposition 8.9 the solutions of $u^{q-1}\equiv 1 \pmod{q^2}$ are exactly the Teichmüller subgroup μ_{q-1}, so a scan over r hits a residue class with rate $(q-1)/q^2$ — even the trivial family $r\equiv 1 \pmod{q^3}$ ($d_q=1$) is included. Such a census cannot test (D1)/(D2)/(D4); see §12.3 for the correction this distinction forces.

All computations were performed with exact integer arithmetic (`pow` with modulus q^3); scripts and data are available with this note.

12.1. **The bad-set scan.** For $q<10^6$, twelve bases $r\le 37$ and $a'<30$: the set $\mathcal{W}_r$ has 23 elements; the bad set has *five* elements, all with $b_q\le 4$:

$$(7,26,5,4),\ (17,19,3,2),\ (19,10,3,2),\ (37,10,3,1),\ (37,19,3,2),$$

and the partial sum in (H-j) stabilizes at 6.30. The bad-layer primes are only 3 and 5. A layer census for $b\le 15$ finds *no* cube-divisor candidates at all for (r,a') equal to $(2,1)$, $(2,3)$, $(3,1)$, $(3,2)$ or $(5,2)$, matching the random expectation $\sum_q q^{-3}\approx 0.17$ per integer.

12.2. **First super-Wieferich sample.** The scan exhibits the first observed super-Wieferich prime: $s_7(19) = 3$ exactly, since $19^6 - 1 = 2^3 \cdot 3^3 \cdot 5 \cdot 7^3 \cdot 127$, and 19 is a primitive root modulo 7. Note $\langle 19 \rangle \bmod 343 = \{\pm 1, \pm 18, \pm 19\}$ with $D_3 = d_q = 6$: on the super-Wieferich layer $e_0 \in [1, d_q]$ is automatically $< q$ whenever compatible, but this layer is *not* the (H-j) bad layer—it is controlled separately by P_{6a}, via the weight $s_q \log q/d_q = 3 \log 7/6 \approx 0.973$.

Systematic super-Wieferich sample. A full scan over $2 \le r \le 500$, $q < 2 \times 10^5$ yields 71 genuine pairs with $s_q \ge 3$ (15 with $s_q \ge 4$), on 12 distinct primes $q \le 113$ and 16 distinct orders $d_q \le 112$. All satisfy $q^3 \mid \Phi_{d_q}(r)$, the structural bound $3 \log q_{\min}(d_q) \le \varphi(d_q) \log(r+1)$ (Remark 9.3) and $d_q \ge \frac{3 \log q}{\log(r+1)}$ with zero violations; the accumulated $\Xi^{(3)}$ mass is 159.947 (a lower bound), and $N_3(112) \ge 16$. The pair $(19, 7)$ above is the equality case: $\Phi_6(19) = 343 = 7^3$ exactly.

12.3. **Effective twin-Wieferich.** *Erratum (null hypothesis).* The 27 effective pairs are reproduced exactly by counting *distinct ordered pairs* (r, a') with $a' > 1$, $\gcd(r, a') = 1$, $q < 10^6$; a scan over all triples gives 78 (23 trivial, 55 effective), reducing to 27 distinct ordered pairs. (The earlier "$49 = 22 + 27$" under-counted both layers; the trivial-pair count 22 differs by one from the number of $\mathcal{W}_r$ hits and awaits the author's reconciliation.) The expectation $27 \sum_q q^{-2} \approx 12.2$ of an earlier version uses the observed number of pairs as the number of candidates and is circular. More importantly, at fixed q the congruence $u^{q-1} \equiv 1 \pmod{q^2}$ has exactly $q - 1$ solutions in $(\mathbb{Z}/q^2\mathbb{Z})^\times$, namely the Teichmüller subgroup μ_{q-1} (Proposition 8.9); the correct null density for a scan *over bases* is $(q-1)/q^2$, not $1/q^2$. Measured densities for $q = 3, 5, 7, 11, 13, 17, 19$ are $0.222, 0.160, 0.122, 0.082, 0.0695, 0.055, 0.049$, agreeing with $(q-1)/q^2$ to three decimals; the "factor 4–6 excess" reported in an earlier version is this missing factor $q - 1 \in [2, 18]$, not positive correlation. Likewise the four pairs at $q = 13$ are four of the six ordered pairs inside $\mu_{12}(\mathbb{Z}/169\mathbb{Z}) \cap [2, 30] = \{19, 22, 23\}$, and the $q = 3$ layer is the structural condition $u \equiv \pm 1 \pmod 9 = \mu_2(\mathbb{Z}/9\mathbb{Z})$. All of these are consequences of the Teichmüller anchor, not statistical anomalies.

What remains open. This does not restore numerical support for (D2): its conditional-on-$k = 0$ content remains untested, since below $q < 3 \times 10^5$ only 10 primes carry a second hit; a meaningful test needs $q \sim 10^9$–10^{10}, conditioned on $q \in \mathcal{W}_r$, and is beyond present reach. The bad-set conclusions of this section (five hits, partial sum 6.30) are unaffected.

12.4. **Katz layer tests.** On the full odd-prime population $q < 5 \cdot 10^4$ the four-dimensional statistic (k_u, m_u) for independent bases passes marginal uniformity and joint independence (KS ≤ 0.0145 against critical 0.0190; joint $\chi^2 \le 29.9$ against 36.4), while Case A pairs $(3, 9)$, $(3, 27)$ explode ($\chi^2 = 7736, 3454$), exactly reproducing the Zariski-density restriction. The twin-Wieferich numerics of §12 cut against (D2) over the full observable range, so (K) is a heuristic whose second layer is currently unsupported and whose use in Proposition 10.1 must be read conditionally.

Limitation: below $5 \cdot 10^4$ the expected number of Wieferich primes is ≈ 2.4 and of twin-Wieferich ≈ 0.45, so the conditional-on-$k = 0$ content of (D1), (D2) has no numerical evidence; the $s_q \ge 3$ branch has the single sample $(19, 7)$.

Observable summary. All deep hits ever observed have $q \le 13$ and $d_q \le 6$; $A_{\text{obs}}(t)$ is a truncated lower bound (saturating for $t \ge 6$ on the scanned range), not a proof of $A(t) = o(t)$. The statistics of this section and of §14.4–14.5 are the first fixed-census data for $\mathcal{W}_r^{(3)}$: the $m(q)$ table tests P_{6a} for *all* $r \le X$ simultaneously at fixed X, while the pair scans test (D1)/(D2) pointwise at fixed r — a strictly stronger requirement (§12.3, Proposition 14.5).

13. Relation to ($\star'$)

In the framework of [1] the remaining gap is stated in Section 7 ("The remaining gap, exactly") as Working hypothesis 6.4: $|\log x| = o(\log c)$, equivalently $O(\log\log c)$, with $x = \det(\Lambda)^{(k-1)/k}/\Delta$ and $\Delta = N|a'|_2$; Theorem 6.5 there gives $\eta \le \frac{1}{k} + O(\frac{\log\log c}{\log c})$ and Theorem 7.1 the equivalent form $N|a'|_2 = (c\theta_c/g)^{(k-1)/k} c^{o(1)}$. The expression $\frac{k-1}{k}\log c - \log\mathrm{rad}(ab)$ used in an earlier version of this note is a heuristic transcription, not a formula of [1]; and the "degeneration to *abc* strength" should be read as the structural ceiling of Theorem 10.1 there ($\Pi \le 1/k + o(1)$, attained at $c = r^e$; full *abc* is structurally unreachable inside the framework). Three regimes: 0.3 (the present results, $\frac{1}{2} - \frac{1}{5}$), 0.5 ($k = 2$), and $(k-1)/k \to 1$ ($k \to \infty$, i.e. *abc*). Along the line $a' = O(1)$, $b_e = r^e - a'$, one has $k_e = \omega(a'b_e) \to \infty$, so the hypothesis degenerates to *abc*-strength already for bounded a'; the present note therefore governs the $k = 2$ slice. Theorem 1.2 pins the Layer-III Pillai candidates $r^e - \prod q_i^{V_i} = a'$ to $\mathcal{W}_r \cap \mathcal{W}_{a'}$ with an explicit class e mod q; Theorem 1.3 pins the Layer-III mass to $J'_{r,a'}$, the Pillai-single-solution layer ($D_3 > E \ge e_0$ means "q hits exactly once in $[1, E]$"); and (H-j) shows the application-layer residual gap does not require Baker-type bounds—the Baker layer belongs to $k \ge 3$ and the $V_q = 2$ squarefull layer, where the first moment of the union bound diverges ($t_q = V_q - 1 - s_q \le -1$, no Hensel wall), and which is untouched here. The defect/order-core analysis of [15] on the Mersenne line independently corroborates this trichotomy.

Remark 13.1 (Notation and quantifiers in the ($\star'$) interface)**.** The symbol a' denotes three different objects in the sources combined here, and no identification between them is asserted: (i) in [1], Definition 3.4, $a' = h/g_h$ is a *primitive normal vector* in $\mathbb{Z}^k$, with $N = [\ker H : \Lambda_0]$ and $\Delta = \det(\Lambda_0) = N|a'|_2$ — this is the $N|a'|_2$ occurring in Working hypothesis 6.4 and Theorem 7.1; (ii) in §9 of [1] (Theorem 9.8), a' is a *scalar Pillai remainder* with $|a'| \le c^{o(1)}$; (iii) throughout the present note $a' \le A$ is a *fixed integer* and $b_e = r^e - a'$. In particular the line $a' = O(1)$ of the present note is a proper subfamily of Layer III in (ii), which is why the first-moment Markov argument can produce an e-density-zero conclusion here without resolving that layer in general.

Two uses of k must likewise be separated: $k = \omega(ab)$ is a global quantity attached to a triple (a, b, c) (so $k_e = \omega(a'b_e) \to \infty$ along $b_e = r^e - a'$), whereas the "$k = 2$ slice" governed by Theorem 1.3 is the *local* per-prime valuation layer $V_q \ge 3$. Working hypothesis 6.4 is a pointwise hypothesis $|\log x| = o(\log c)$ with no "almost all" quantifier; Theorem 6.5 is a conditional statement along a sequence of triples with $c \to \infty$; and the only density statement in [1], Theorem 11.7(i), is on the c-axis and asserts that *every* coprime decomposition of a density-one set of c has $q \le 1 + o(1)$, proved by one-variable counting. There is no quantifier position in [1] for averaging over $a' \in [1, A]$ or discarding a zero-density set of a'; nor would such an average transfer to the pointwise statement of Theorem 6.5. Accordingly no averaging step is used or claimed below.

Acknowledgments

The reductions of §§4–10 were found and machine-verified in the course of an extended collaboration on the Wieferich tail-estimation problem; the numerical protocols follow the standards of [17].

The condition set has now evolved to its minimal form. The $s_q = 2$ layer is unconditionally closed (Theorem 7.5). The *exact* threshold of the first-moment route is

$$\Phi_r(E) = o(E) \quad \Longleftrightarrow \quad A_r(t) := \sum_{d \le t} \frac{M_d}{d} = o(t), \qquad M_d = \sum_{\substack{q:\, d_q = d \\ s_q \ge 3}} s_q \log q$$

(Theorem 9.6); this is equivalent to the statement that the cube-full part of $\Phi_d(r)$ has logarithmic mass $o(d)$ in Cesàro mean. $\Xi_r^{(3)} < \infty$ is a convenient sufficient rung but is *not* necessary (Theorem 9.4): a configuration with $M_d = d/(\log d)^2$ has $\Xi_r^{(3)}(E) \sim E/(\log E)^2 \to \infty$ while $\Phi_r(E) = o(E)$. Classically the remaining input is anchored at P_{6a} (Murata 1991 (i)), probabilistically at (D4); GRH is redundant throughout (Theorem 11.1 and the Hensel-wall absorption of the ord side, Theorem 8.10).

14. Postscript

This section records the outcome of the fifteenth–eighteenth steps of the working chain: what a new tool would have to look like, and the single-variable reformulation of the deep-layer finiteness.

14.1. **Literature verdict: no existing tool applies.** A systematic check of the six natural candidates finds *no* unconditional bound $\log \operatorname{rad} \Phi_n(a) \ge \varphi(n) \log a - o(n)$ — i.e. no theorem giving $M_d = o(d)$:

- Stewart [24, 23] controls the *greatest prime factor*, $P(\Phi_n(\alpha, \beta)) > n \exp(\frac{\log n}{104 \log\log n})$; a large single-power prime factor is compatible with the remaining mass being concentrated in high powers of small primes.
- Bugeaud–Corvaja–Zannier [28] gives $\log \gcd(a^n - 1, b^n - 1) = o(n)$: the *common* factor of two sequences, not the repeated-prime mass of one.
- Balaji–Luca [25] gives $\#\mathcal{A}_{K,c,u} \cap [1, N] \ll N \exp(-\frac{\log N}{156 \log\log N})$ for the Kn-smooth part, for *most* n; smooth parts include first powers and do not exclude high powers.
- Zsigmondy / Bilu–Hanrot–Voutier [29] guarantees *one* primitive divisor, not that all primitive divisors occur once. The paraphrase "$p \nmid n$ implies p occurs at most once in $\Phi_n(r)$" is *false*: $v_{1093}(\Phi_{364}(2)) = 2$ with $1093 \nmid 364$, and $\Phi_5(3) = 121 = 11^2$. Indeed *all* deep Wieferich primes satisfy $q \nmid d_q$ (since $d_q \mid q - 1$), so high exponents occur precisely at primitive divisors.
- Murty–Séguin [8] gives conditional greatest-prime-factor bounds and "super-Wieferich finite ⇒ non-Wieferich infinite"; no upper bound on $W_r^{(3)}(x)$ better than $O(x^2)$.
- Fellini–Murty [20], Silverman [6], Graves–Murty [22] and Ding [22] give *lower* bounds for non-Wieferich primes or *abc*-conditional results; Ding's $C_n' \gg a^{\varphi(n) - \varepsilon n}$ yields $\log \operatorname{rad} \Phi_n(a) \ge (1 - \varepsilon)\varphi(n) \log a$, which controls M_d only with $\varepsilon \to 0$ *and abc*.
- W. Xie, *Primary components and large prime factors of cyclotomic values*, arXiv: 2611.00075v1 (math.RA, 36 pp., preprint, not peer-reviewed). This is the closest same-object reference: Theorem 3.2 gives an unconditional lower bound for the primary part, but Theorem 3.6 takes $\log \operatorname{rad} \Phi_r(a) \ge q^{1-\gamma}$ as a *hypothesis* and its counting input is $\vartheta(Y; r, 1)$ with a *fixed* conductor and an average over bases. It therefore supplies no unconditional $\log \operatorname{rad} \ge \varphi(n) \log a - o(n)$ and does not transfer to the fixed-base, varying-modulus side; the verdict above is unchanged. *Citation caveat*: the GenPDF header prints "arXiv:2611.00075v1" followed by "[math.RA] 16 Oct 2026"; the "2611" identifier and the October date are incompatible at the metadata level. We cite the v1 PDF as deposited and do not assert a submission date.

The unconditional side has no lower bound of strength $\log\operatorname{rad}\Phi_n(a) \geq \varphi(n)\log a - o(n)$; the strongest unconditional statements are $\log\operatorname{rad} = O(\log n)$ type. (D4) has no standard name: Katz [2] is a continuous $\mathbb{R}/\mathbb{Z}$ equidistribution conjecture, Gras [30] is a $1/p$ heuristic, and Berger [31]'s formal-group rigidity treats fixed formal groups. If (D4) is read as a counting density it is already P_{6a}; a probabilistic reading must specify its probability space (Proposition 14.5).

Remark 14.1 (Transcription check)**.** Three paraphrases in the commissioned verification report were wrong and are corrected here; the body of this note is unaffected: (i) $v_q(r^{q-1}-1) = s_q$, not $s_q + v_q(q-1)$, since $q \nmid (q-1)/d_q$ (Lemma 7.1); (ii) the translation $r \bmod q^3 \in \mu_{q-1}$ of Proposition 8.9 stands ($q \leq 97$ enumeration, zero counterexamples); (iii) the "primitive divisors occur at most once" paraphrase is false (counterexamples above).

14.2. **The power of structural constraints: zero.**

Theorem 14.2 (Zero saving)**.** *The small-q layer $q \leq \sqrt{t}$ is unconditionally $o(t)$, and for $q > \sqrt{t}$ one has $d_q > c_r \log t$ with $c_r = 3/(2\log r)$ and $q \nmid d_q$ (no BOOK contamination, cf. the stratification trap above). Nevertheless* every *volume-type upper bound obtained from this split is strictly worse than the trivial volume bound:*

$$C(t) - V(t) = \sqrt{t}\left[\log r - \frac{6}{\pi^2}\log(r+1)\right] + O(\log t) > 0 \qquad \text{for every } r \geq 2$$

(coefficient 0.0253 *at* $r = 2$*,* 0.2558 *at* $r = 3$*,* 1.7995 *at* $r = 100$*, since* $\log r > \frac{6}{\pi^2}\log(r+1)$ *for all* $r \geq 2$*). Hence all known structural constraints together save* zero *linear mass: the small-d structural emptiness of Remark 9.3 excludes only finitely many d and has no asymptotic force.*

Moreover, a saturated pseudo-configuration satisfying all *per-d constraints* $3\log q \leq \min(\varphi(d)\log(r+1), d\log r)$*, $q > \sqrt{t}$, $q \equiv 1 \pmod d$ still gives* $A(t)/t = 0.552/0.775/1.051/1.234$ *for* $r = 2,3,5,7$ *at* $t = 10^3$*, stabilising at* $0.561/0.783/1.058/1.243$ *at* $t = 10^4$*. Thus block/volume/partial-summation methods are exhausted (logical pseudo-configurations, not claims about real factor sets).*

14.3. **The ladder has two independent weakenings.**

Theorem 14.3 ($A_r(t) = o(t)$ and $P_4^{(3)}$ are incomparable)**.** *Model M1 ($d_q = q-1$, $s_q = 3$, all primes) gives, by Mertens' first theorem* $\sum_{p\leq x}\frac{\log p}{p} = \log x + O(1)$*, $A(t) \approx 3\log t = o(t)$ while $P_4^{(3)} \approx \log t \to \infty$; model M2 (one saturated deep prime per d with $\varphi(d)/d \geq \log r/\log(r+1)$) gives $P_4^{(3)} < \infty$ while $A(t) = \delta_r t$ with $\delta_r = 0.417/0.289/0.214/0.184/0.129/0.095$ for $r = 2,3,5,7,19,100$. Both satisfy the tight structural necessary condition. Hence the ladder*

$$P_{6a} \Rightarrow (\mathrm{D4})_{\mathrm{count}} \Rightarrow \Xi_r^{(3)} < \infty \Rightarrow \{A = o(t)\} \wedge \{P_4^{(3)}\}$$

has two independent weakenings, and $\Xi_r^{(3)} < \infty$ is strictly stronger than either.

14.4. **The single-variable translation of P_{6a}.**

Theorem 14.4 ($P_{6a} \iff m(q) \to \infty$)**.** *Let $m(q) = \min\big(\mu_{q-1}(\mathbb{Z}/q^3\mathbb{Z}) \setminus \{1\}\big)$ be the least nontrivial Teichmüller representative modulo q^3. Then*

$$P_{6a} \iff m(q) \to \infty.$$

Proof. If $q > r^{1/3}$ then $r < q^3$, so by Proposition 8.9 $q \in \mathcal{W}_r^{(3)} \iff r \in \mu_{q-1}$, which forces $r \geq m(q)$. Hence $\{q \in \mathcal{W}_r^{(3)} : q > r^{1/3}\} \subseteq \{q : m(q) \leq r\}$; if $m(q) \to \infty$ the right side is finite, and the finitely many primes $q \leq r^{1/3}$ add nothing: $\mathcal{W}_r^{(3)}$ is finite. Conversely, if $m(q) \not\to \infty$, there is X with infinitely many q having $m(q) \leq X$; choosing for each such q an element of

$\mu_{q-1} \setminus \{1\}$ in $[2, X]$, the pigeonhole principle gives a fixed $r \in \{2, \dots, X\}$ lying in $\mu_{q-1} \bmod q^3$ for infinitely many q, all but finitely many with $q > r^{1/3}$; by Proposition 8.9 again, $\mathcal{W}_r^{(3)}$ is infinite. □

Numerically, among the 1228 odd primes $q < 10^4$ the median of $m(q)/q^2$ is 0.7222 (so $m(q) \asymp q^2$, the exact source of the $1/q^2$ heuristic, since $q-1$ elements are spread over $[0, q^3)$). The ten smallest $(q, m(q))$ are $(7,18)$,$(3,26)$,$(23,42)$,$(5,57)$,$(113,68)$,$(11,124)$,$(17,158)$,$(13,239)$, $(47,295)$,$(19,333)$. For each fixed $X \le 10^3$ the set $\{q : m(q) \le X\}$ saturates already by $Q = 300$ (3/5/12/15 primes for $X = 50/100/500/1000$, largest such $q = 113$), and for $X = 10^4$ by $Q = 3000$ (40 primes, largest $q = 1483$). This is the first *fixed-X, all-$r \le X$* numerical evidence for P_{6a}; all previous scans were fixed-r, varying-q. Consequently the counting-density reading of (D4) is *equivalent* to P_{6a}, while the probabilistic form with finite mean C_r/q^2 is strictly stronger (yielding $\Xi_r^{(3)} < \infty$ a.s.).

14.5. **Dual exactness and the trivial family.**

Proposition 14.5 (Dual ensemble)**.** *In the dual ensemble (fixed q, r uniform in $(\mathbb{Z}/q^3)^\times$) one has the exact identity $\mathbb{P}(s_q \ge 3) = 1/q^2$, verified for $q = 3, 5, 7, 11, 13, 17, 19$ ($|\mu_{q-1}| = q-1$ throughout). In short intervals $r \ll q^3$ the expectation is correct but the variance is large (at $R = 100$ the counts for $q = 11, 13, 17, 19$ are all 0 against expectations $0.82, 0.59, 0.34, 0.27$). In the primal ensemble (fixed r, q varying) there is* no *probability space; the missing tool is exactly the transfer dual→primal (cf. the ensemble separation of §8.1). (D4) is therefore not a Katz/Gras theorem and has no standard name.*

Theorem 14.6 (The trivial family is thin)**.** *$d_q = d$ implies $q^3 \mid r^d - 1 \le r^d$, hence $q \le r^{d/3}$; in particular $d_q = 1$ forces $q \le (r-1)^{1/3}$, so the trivial family $r \equiv 1 \pmod{q^3}$ contributes only finitely many q for each fixed r. A sweep over $q < 5 \cdot 10^6$ (*348512 *odd primes) and* 24 *fixed bases (including the $r = 1 + q^3$ families* $28, 126, 344, 1332, 2198$ *and large bases* $10^4, 10^5$*),* $8.36 \cdot 10^6$ *congruence tests, yields* 15 *deep hits,* all *with $q \in \{3, 5, 7, 11, 13\}$, $d_q \le 6$, $s_q \in \{3, 4, 5\}$; the bound $q \le r^{d/3}$ has zero violations and the $r = 1 + q^3$ families give $d_q = 1$ exactly. The observed $A_{\mathrm{obs}}(t)$ saturates for $t \ge 6$ and is a lower bound truncated at $q < Q$: it is not a proof of $A(t) = o(t)$, since primes with $d_q \le t$ may reach $q = r^{t/3}$.*

Remark 14.7 (The trivial family is thin *per base*, but not uniformly)**.** Theorem 14.6 is a statement for each *fixed* r. There is no bound uniform in r: for any finite set S of odd primes put $N = \prod_{q \in S} q^3$ and $r_S = 1 + N$. Then $r_S \equiv 1 \pmod{q^3}$ for every $q \in S$, and LTE gives

$$s_q(r_S) = v_q\big((1+N)^{q-1} - 1\big) = v_q(N) + v_q(q-1) = 3 + 0 = 3$$

exactly, so $|W_{r_S}^{(3)}| \ge |S|$. Hence $\sup_r |W_r^{(3)}| = +\infty$ unconditionally; the family $r = 1 + q^3$ of Theorem 14.6 is the case $|S| = 1$. Numerically, for S the first $k \le 10$ odd primes, $\log_{10} r_S$ runs from 1.45 to 33.00, every $q \in S$ has $s_q = 3$, and no further deep hit occurs below $2 \cdot 10^4$. The two quantifier orders

$$\forall r\ \exists \text{ finite } W_r^{(3)} \quad (P_{6a},\ \text{open}), \qquad \forall \text{ finite } S\ \exists r \text{ with } W_r^{(3)} \supseteq S \quad (\text{this remark})$$

are compatible because r_S grows with S. Consequently any proof plan that would bound $|W_r^{(3)}|$ or the constant C_r of (D4) *uniformly in r* is ruled out; P_{6a} can only be read as finiteness *for each fixed r*.

14.6. **What the new tool must look like.** The missing object must *cash out a codimension-two double zero on a varying-modulus family*: (D4) is not more codimension — the $s_q \geq 3$ condition $(Q^{(1)}, Q^{(2)}) = (0, 0)$ already has codimension 2, and the $s_q = 2$ layer (codimension 1) is closed only because the Hensel wall $D_3 = d_q q$ donates the extra $1/q$. The Hensel dividend can be spent *once*. Seven specifications:

- **S1** uniform in the varying modulus q, with pointwise error $< q^{-2-\varepsilon}$ independent of $\operatorname{disc} L_q$ (GRH–Chebotarev is unavailable: q ramifies in $L_q = \mathbb{Q}(\zeta_{q^2}, r^{1/q^2})$, $\operatorname{disc} \sim q^{O(q^2)}$);
- **S2** produces $A_r(t) = o(t)$, equivalently $\Phi_r(E) = o(E)$ (Theorem 9.6), not merely $\Xi_r^{(3)} < \infty$;
- **S3** discrete, pointwise mod q^3; continuous $\mathbb{R}/\mathbb{Z}$ equidistribution does not control $1/q$ point masses — equivalently, pointwise rigidity of $\iota(\mathbb{Z})$ in $\mathcal{P}$;
- **S4** penetrates the index-q^2 Teichmüller subgroup μ_{q-1} (Theorem 14.4);
- **S5** layer-specific: isolate both the $s_q = 2$ PRIM layer and the BOOK term $q \mid d/d_q$ (equivalently cut at $q > \sqrt{E}$);
- **S6** bypass finite-order congruence expansions (Theorem 9.5) or handle the varying-modulus discrete logarithm.
- **S7** *No r-uniform output*. By the CRT family of Remark 14.7, any bound uniform in r is false; the constant C_r of (D4) must depend on r, and the tool must produce a statement for each fixed base.

Self-checks: the tool must not imply P_4; must give $\Phi = o(E)$ rather than $\Xi^{(3)} < \infty$; must control the $1/d_q$-weighted mass, not $N_3(E)$ alone. **T1**–**T6**: GRH-only, gcd/divisibility-identity-only, greatest-prime-factor/single-power/smooth-part/gcd-only, and structure-constraint/block/volume-only mechanisms all fail (see §9 and §14.1–14.2). The *minimal* tool is a single deterministic statement, $m(q) \to \infty$ (Theorem 14.4), or its quantitative form $\#\{q \leq x : m(q) \leq X\} = O_X(1)$.

References

[1] C. Zeng, *A lattice-theoretic approach to the abc conjecture*: hexagon:2610.00630v1; Zenodo 22995922 (v1, 2026-09-27); Zenodo 23745797 (v2.0, 2026-10-28, revised). Not peer-reviewed; the v1 abstract ("min–max barrier") was withdrawn by errata E12/E30 after Theorem 4.5. Hexagon is a deposition platform, not arXiv; no arXiv record 2610.00630 exists. The results of this note are proved independently and use only the problem setting.

[2] N. M. Katz, *Wieferich past and future*, Contemp. Math. 632 (2015), 253–270, Conjecture 3.1. DOI 10.1090/conm/632/12632.

[3] L. Murata, *The distribution of primes satisfying the condition* $a^{p-1} \equiv 1 \pmod{p^2}$, Acta Arith. **58**(2) (1991), 103–111, Conjectures (i), (ii).

[4] S. Lang, *Old and new conjectured Diophantine inequalities*, Bull. AMS **23** (1990), 37–75, p. 42.

[5] R. Crandall, K. Dilcher, C. Pomerance, *A search for Wieferich and Wilson primes*, Math. Comp. **66** (1997), 433–449.

[6] J. H. Silverman, *Wieferich's criterion and the abc-conjecture*, J. Number Theory **30** (1988), 226–237.

[7] D. Dorais, D. Klyve, *A Wieferich prime search up to* 6.7×10^{15}, J. Integer Seq. **14** (2011).

[8] M. R. Murty, F. Séguin, *Prime divisors of sparse values of cyclotomic polynomials and Wieferich primes*, J. Number Theory **201** (2019), 1–22.

[9] C. Hooley, *On Artin's conjecture*, J. Reine Angew. Math. **225** (1967), 209–220.

[10] D. R. Heath–Brown, *The least square-free number in an arithmetic progression*, and Fermat-quotient estimates via Burgess bounds (see Shparlinski's survey, arXiv:1104.3909, for the Ihara-sum improvement).

[11] A. Alexandru, C. Cobeli, M. Vâjâitu, A. Zaharescu, *Fast covering of* $\mathbb{Z}/p\mathbb{Z}$ *by Fermat quotients*, J. Number Theory **161** (2022).

[12] R. Dillon, *The Wieferich locus of an algebraic unit*, Zenodo 21456694 (order-lift dichotomy).

[13] K. Chua, *A note on p-adic valuations*, arXiv:2112.04173.

[14] A. Caruso, Q. Gazda, G. Lucas, *On ϕ-Wieferich places of Drinfeld modules*, arXiv:2412.11588 (Thm. C gives the family average and explicitly not the fixed-module infinitude).

[15] W. Laniewski, *Defect and order-core analysis of abc data*, arXiv:2609.00039 (Mersenne-line density 47/210).
[16] P. Kurlberg, C. Pomerance, *On the periods of the linear congruential and power generators*, Acta Arith. **119**(2) (2005), 149–169 (unconditional $\mathrm{ord}_q(r) > q^{1/2+\varepsilon(q)}$ on a positive density set; GRH $> q^{1-\varepsilon(q)}$).
[17] Companion verification report and scripts for this note (machine-assisted proofs of Theorems 1.1–1.4 and Proposition 10.1 on all tested samples; available from the authors).
[18] P. Erdős, M. R. Murty, *On the order of a (mod p)*, CRM Proc. Lecture Notes **19** (1999), 87–97.
[19] M. R. Murty, S. Srinivasan, *Some remarks on Artin's conjecture*, Canad. Math. Bull. **30**(1) (1987), 78–88.
[20] N. Fellini, M. R. Murty, *Wieferich primes in number fields and the conjectures of Ankeny–Artin–Chowla and Mordell*, J. Number Theory **278** (2026), 109883 (published version of arXiv:2508.08472v2; the authors note explicitly that no unconditional result in this direction is known).
[21] J.-M. DeKoninck, N. Doyon, *On the set of Wieferich primes and of its complement*, Ann. Univ. Sci. Budapest. Sect. Comput. **27** (2007), 3–13 (notation: their $\mathcal{W}_r$ denotes the depth-r layer for the fixed base 2, i.e. $W_2^{(r)}$; not the base-r set $\mathcal{W}_r$ used here).
[22] Y. Ding, *Non-Wieferich primes under the abc conjecture*, C. R. Math. Acad. Sci. Paris **357** (2019), 483–486 (the intermediate bound $C_n' \gg a^{\varphi(n)-\varepsilon n}$ on the squarefree part of $\Phi_n(a)$). See also H. K. Graves, M. R. Murty, *The abc conjecture and non-Wieferich primes in arithmetic progressions*, J. Number Theory **133** (2013), 1809–1813.
[23] C. L. Stewart, K. Yu, *On the abc conjecture, II*, Duke Math. J. **108**(1) (2001), 169–181 (the $\kappa R^{1/3}(\log R)^3$ bound).
[24] C. L. Stewart, *On divisors of Lucas and Lehmer numbers*, Acta Math. **211** (2013), 291–314. Cf. also Acta Arith. **26** (1975), 427–433; Proc. London Math. Soc. (3) **35** (1977), 425–447.
[25] N. Balaji, F. Luca, *Terms of Lucas sequences having a large smooth divisor*, Canad. Math. Bull. **66** (2023), 225–231 (arXiv:2203.12364, Theorem 1.2).
[26] A. Granville, K. Soundararajan, *A binary additive problem of Erdős and the order of* 2 *(mod* p^2*)*, Ramanujan J. **2** (1998), 283–298.
[27] P. Moree, *Artin's primitive root conjecture — a survey*, Integers **12A** (2012), #A13.
[28] Y. Bugeaud, P. Corvaja, U. Zannier, *An upper bound for the greatest common divisor of* $a^n - 1$ *and* $b^n - 1$, Math. Z. **243** (2003), 79–84.
[29] Y. Bilu, G. Hanrot, P. M. Voutier, *Existence of primitive divisors of Lucas and Lehmer numbers*, J. Reine Angew. Math. **539** (2001), 75–122.
[30] G. Gras, *Probabilistic study of Fermat quotients*, Comment. Math. **54** (2016), 115–140 (heuristic, not a pointwise theorem).
[31] L. Berger, *Rigidity and unlikely intersections for formal groups*, arXiv:1811.05824v2 (framework only; not a pointwise theorem for fixed r/varying q).
[32] W. Xie, *Primary components and large prime factors of cyclotomic values*, arXiv:2611.00075v1 (math.RA; v1 PDF header dated 16 Oct 2026, identifier/date incompatible; not peer-reviewed).