ECC RESEARCH PROGRAM · P1.2

No Fixed-Length Subexponential Factor Base over Prime Fields

A support-size impossibility, a preprocessing loophole, and the coordinate-aware frontier of prime-field decomposition

math.st [1] **@math__street** [1]

with GPT 5.6 Sol · Claude Fable 5

[1]Independent research collective, math.st · July 2026

ABSTRACT

Index calculus breaks the discrete logarithm on elliptic curves over extension fields by decomposing random points over an algebraic factor base; over prime fields no such factor base is known, and the security of prime-field ECC rests in part on that gap. We work the problem posed as P1.2: given an ordinary curve $E/\mathbb{F}_p$, exhibit a factor base $\mathcal{F} \subseteq E(\mathbb{F}_p)$ of size at most $L_{p[1/2]}$ together with a polylogarithmic-time decoder that writes a uniform random point as a sum of a constant number m of factor-base points with inverse-polylogarithmic probability. Our central finding is negative and, under the standard reading of the notation, unconditional: a fixed-length sum over a base of genuinely subexponential size $L_{p[1/2,c]} = p^{o(1)}$ reaches only $p^{o(1)}$ targets, while Hasse's bound gives $\#E(\mathbb{F}_p) = p^{1+o(1)}$, so the success probability is at most $p^{-1+o(1)}$ — below every inverse polynomial in $\log p$. The literal specification is thus information-theoretically empty, and we quantify the minimum repairs: fixed m forces a base of size $p^{1/m-o(1)}$, and keeping a subexponential base forces m to grow like $\sqrt{\log p/\log\log p}$. We then show that the natural "square-root size" rescue is vacuous unless offline resources are charged: an elementary radix construction covers every target with a base of size $O(p^{1/3})$ but hides a complete discrete-logarithm table of $p^{1+o(1)}$ entries. For target-oblivious decoders we prove a genuine lower bound — a translate-probe searcher using only membership tests $R - a \in \mathcal{F}$ has success at most $T|\mathcal{F}|/r$ after T probes, so it needs $p^{1/2-o(1)}$ probes — and we verify it exhaustively on a toy group. Five explicit prime-field candidates (integer intervals, rational-height reconstruction, low-degree maps, integral lifts, smooth multiplicative subgroups) are measured and each falls to sparsity, near-universality, or a missing decoder; a validated extension-field control recovers planted logarithms end to end, isolating the elementary reason prime fields resist the Weil-restriction trick. The one honestly open door is a coordinate-aware decoder in the corrected square-root regime, which our lower bound does not reach.

Keywords. ECDLP · index calculus · factor base · prime fields · summation polynomials · lower bounds

1 Introduction

The elliptic-curve discrete logarithm problem (ECDLP) has a sharp asymmetry across the fields it is posed over. On a curve $E/\mathbb{F}_{q^n}$ of a suitable extension field, the index-calculus programme

of Semaev, Gaudry, and Diem [1], [2], [3] supplies an **algebraic factor base** — points whose x-coordinate lies in a proper subfield or in the fibre of a low-degree covering — over which a random point decomposes as a short sum, yielding subexponential-time algorithms. Over a **prime** field $\mathbb{F}_p$ no such factor base is known. The absence is not a minor gap: it is a large part of why a well-chosen prime-field curve is believed to offer square-root security, and why deployed curves are prime-field curves.

P1.2 asks whether the gap can be closed by fiat. Concretely, given p, an ordinary curve $E/\mathbb{F}_p$ with prime order $r = \#E(\mathbb{F}_p)$, and oracle access to uniform random points, one is to produce a pair $(\mathcal{F}, \mathcal{D})$ in which

1. $\mathcal{F} \subseteq E(\mathbb{F}_p)$ has size at most $L_{p[1/2]}$;
2. membership in $\mathcal{F}$ is decidable in $\mathrm{poly}(\log p)$ time;
3. the decoder $\mathcal{D}$ runs in $\mathrm{poly}(\log p)$ time and writes a uniform random R as $R = \sum_{i=1}^{m} P_i$ with $P_i \in \mathcal{F}$ and m a fixed constant, succeeding with probability at least $1/\mathrm{poly}(\log p)$.

The task splits into an **existence** question (do short decompositions exist for most targets?) and a strictly harder **findability** question (can one be located in polylogarithmic time?). Our work resolves the literal statement, audits the specification traps that make it ambiguous, and marks the single residual question that survives.

Main finding. Under the standard reading of $L_{p[1/2]}$ as a subexponential function and of "constant m" as independent of p, conditions (1) and (3) are **mutually inconsistent**: a fixed-length sum over a base of size $p^{o(1)}$ reaches at most $p^{o(1)}$ of the $p^{1+o(1)}$ group elements, so success is at most $p^{-1+o(1)}$. The specification is empty before any curve geometry is invoked. Replacing the size bound by $p^{1/2+o(1)}$ makes it **satisfiable but trivial** unless every offline resource is charged, and the honest remainder is a single coordinate-aware findability question in that corrected regime — for which we prove a lower bound against target-oblivious probes but no bound against coordinate-aware algorithms.

1.1 Contributions and honest scope

We contribute: (i) a support-size impossibility theorem for the literal specification, with a battery of adversarial checks and the exact necessary repairs (§3); (ii) an elementary radix construction exposing the nonuniform-preprocessing loophole in the square-root reading, and the resource discipline that closes it (§4); (iii) a measured taxonomy of five prime-field candidates against a random and an extension-field control, all grounded in real data (§5); (iv) a translate-probe lower bound of $p^{1/2-o(1)}$ queries for target-oblivious decoders, verified exhaustively at toy scale (§6); (v) an analysis of the coordinate-aware smooth-subgroup predicate, whose membership is succinct but whose decoder our solver did not realize (§7); and (vi) the elementary obstruction to Weil restriction over a prime field, with a validated extension-field positive control (§8).

We state the scope as plainly as the problem's ground rules demand. The central result is **negative and unconditional** under standard notation; it is a specification-level impossibility proof, not an ECDLP algorithm and not a break of prime-field ECC. It does not rule out the two resource-bounded corrected variants we isolate (§9), and it does not rule out a coordinate-aware decoder for a square-root-size base. Every candidate measurement below is toy-scale ($p \le 2^{20}$) and uses one curve per size, so it calibrates and falsifies but does not license asymptotic inference; we mark such statements EMPIRICAL: tested range throughout.

2 Setting and notation

Fix a prime $p > 3$ and a nonsingular short-Weierstrass curve

$$E : y^2 = x^3 + Ax + B, \quad A, B \in \mathbb{F}_p, \quad 4A^3 + 27B^2 \neq 0. \tag{1}$$

Write $r = \#E(\mathbb{F}_p)$ for the group order and assume E ordinary with r prime, as the target regime requires. For a point $P \neq O$ we write $x(P) \in \mathbb{F}_p$ for its abscissa; O is the point at infinity. A **factor base** is a subset $\mathcal{F} \subseteq E(\mathbb{F}_p)$; its size is $s = |\mathcal{F}|$. For a fixed summand count m the **ordered sum map** is

$$\sigma_m : \mathcal{F}^m \longrightarrow E(\mathbb{F}_p), \quad (P_1, \ldots, P_m) \mapsto \sum_{i=1}^{m} P_i, \tag{2}$$

and its image $S_m = \sigma_m(\mathcal{F}^m)$ is exactly the set of targets admitting an allowed length-m decomposition.

CITED We use the standard generalized L-notation [4]: for fixed parameters,

$$L_x[\rho, \psi] = \exp\big((\psi + o(1))(\log x)^{\rho}(\log\log x)^{1-\rho}\big), \tag{3}$$

with natural logarithms, and $0 < \rho < 1$ is called **subexponential** in $\log x$. In particular $L_{p[1/2,c]} = \exp\big((c + o(1))\sqrt{\log p \log\log p}\big)$. Two shorthands recur: a quantity is $p^{o(1)}$ when its logarithm is $o(\log p)$, and $p^{1+o(1)}$ when it is $p \cdot p^{o(1)}$. The Semaev polynomials [1] provide the algebraic decomposition test we use as a **verifier**; the third is

$$S_3(X_1, X_2, X_3) = (X_1 - X_2)^2 X_3^2 - 2((X_1 + X_2)(X_1X_2 + A) + 2B)X_3 + (X_1X_2 - A)^2 - 4B(X_1 + X_2),$$

with S_{m+1} built recursively by resultants; $S_{m+1}(x(P_1), \ldots, x(P_m), x(R)) = 0$ whenever $P_1 + \cdots + P_m = R$ with the correct sign choices.

3 The support-size impossibility

The problem's decisive property is visible before any elliptic-curve structure: a short sum over a small set cannot reach a large group.

Theorem 1 (support-size bound). PROVED Let $G_p = E(\mathbb{F}_p)$ have order r_p, let $\mathcal{F}_p \subseteq G_p$ have size $s_p \leq L_{p[1/2,c]}$ for a fixed $c > 0$, and fix an integer m independent of p. For a uniform $R \in G_p$, every algorithm outputting $R = P_1 + \cdots + P_m$ with $P_i \in \mathcal{F}_p$ has success probability at most $p^{-1+o(1)}$; in particular it is eventually below $1/q(\log p)$ for every fixed positive polynomial q. Thus conditions (1) and (3) cannot hold simultaneously under this interpretation.

Proof. The image of σ_m satisfies $|S_m| \leq |\mathcal{F}_p^m| = s_p^m$. A correct output on target R forces $R \in S_m$, independently of the algorithm's randomness, advice, preprocessing, or oracle queries; hence for any fixed factor base

$$\Pr[\text{success}] \leq \Pr[R \in S_m] = |S_m|/r_p \leq s_p^m/r_p. \tag{5}$$

If the base is sampled before the independent target is drawn, averaging this conditional inequality preserves it. Now CITED L-notation gives

$$\log(s_p^m) \leq m(c + o(1))\sqrt{\log p \log\log p} = o(\log p), \tag{6}$$

so $s_p^m = p^{o(1)}$. CITED Hasse's bound $\left|r_p - (p+1)\right| \le 2\sqrt{p}$ [5] gives $r_p = p^{1+o(1)}$, whence $\Pr[\text{success}] \le p^{-1+o(1)}$. For large p the right side is at most $2p^{-1/2}$; since $2q(\log p) < p^{1/2}$ eventually for any fixed polynomial q, the success probability is eventually below $1/q(\log p)$, contradicting condition (3). □

The argument never uses primality of r_p or ordinariness of E: it holds in **any** finite group of order $p^{1+o(1)}$. The elliptic curve is a red herring for the impossibility; only the cardinality matters.

> **Remark 2** (what the bound does and does not say). PROVED The problem statement's own hint — that a base of size roughly $p^{1/2}$ has abundant three-term decompositions — is correct but addresses a **different size regime**: for every fixed c, $L_{p[1/2,c]} = p^{o(1)} = o(p^{\varepsilon})$ for all fixed $\varepsilon > 0$, so a genuinely $p^{1/2}$-scale base already violates condition (1) under standard notation. The hint and the formal bound describe incompatible sizes; this is the notation ambiguity we return to in §9.

We attacked the theorem adversarially before relying on it. Each escape route in Table 1 was checked and closed; the toy row is an exhaustive enumeration on the order-19 curve $y^2 = x^3 + 2x + 2$ over $\mathbb{F}_{17}$.

Table 1: Adversarial checks on the support-size theorem (A007/`CLAIM.md`). Every route either falls inside the counting bound already used or fails to supply the fixed pair the specification demands.

Attempted escape	Disposition
At most m vs. exactly m terms	reaches $1 + s_p + \cdots + s_p^m \le (m+1)\max\left(1, s_p^m\right) = p^{o(1)}$
Repeated / distinct summands	repetition gives the larger domain $\mathcal{F}_p^m$, already used
Ordered vs. unordered tuples	ordered tuples are the larger space, already used
Signed summands $\pm P_i$	replace $\mathcal{F}_p$ by $\mathcal{F}_p \cup \left(-\mathcal{F}_p\right)$; size doubles, still $p^{o(1)}$
Randomized / nonuniform decoder	selects a representation only after $R \in S_m$; cannot enlarge S_m
Randomized base construction	condition on the realized base and average the pointwise bound
Uniform-point oracle	samples do not change whether the fixed target lies in S_m
$\mathcal{F}$ chosen after seeing R	no longer the fixed pair $(\mathcal{F}, \mathcal{D})$ required before the target
Toy counterexample search	EMPIRICAL: r=19 exhaustive over $1 \le s \le 5$, $0 \le m \le 4$: no violation; collisions often shrink S_m below s^m

The counting is also quantitatively informative: it dictates the **minimum** repair to a non-vacuous statement.

> **Corollary 3** (necessary repairs). PROVED If the desired success is $\delta_p \ge 1/q(\log p)$, then $s_p^m \ge \delta_p r_p$. For fixed m this forces $s_p \ge (p/2q(\log p))^{1/m} = p^{1/m - o(1)}$: a genuinely subexponential base is too small. Conversely, retaining $s_p \le L_{p[1/2,c]}$ forces
>
> $$m \ge (1/c + o(1))\sqrt{\log p / \log\log p}. \tag{7}$$
>
> Neither condition is an algorithm; both are pure counting necessities.

For fixed $m = 3$ the first branch demands a base of size $p^{1/3 - o(1)}$; the second branch is the entry point to a genuine subexponential index-calculus target, in which the summand count grows slowly with p. These two branches become the corrected Variants S and L of §9.

4 The square-root reading and the preprocessing loophole

The most charitable rescue reads $L_{p[1/2]}$ as $p^{1/2}$, so that s^3/r is large and short decompositions genuinely exist. Existence, however, is not the specification; findability is. We show that a **square-root reading that charges only online time is vacuous**, because it admits an input-specific lookup table that embeds a complete discrete-logarithm table.

Theorem 4 (radix factor base). PROVED Let $G = \langle G_0 \rangle$ be cyclic of order r, fix $m \geq 1$, and set $B = \lceil r^{1/m} \rceil$. Then

$$\mathcal{F} = \bigcup_{j=0}^{m-1} \{ [dB^j] G_0 : 0 \leq d < B \} \tag{8}$$

has $|\mathcal{F}| \leq mB$, and every element $[k]G_0$ with $0 \leq k < r$ is a sum of exactly m elements of $\mathcal{F}$, one per positional digit of the base-B expansion of k.

Proof. Since $B^m \geq r$, every $k \in \{0, \ldots, r-1\}$ has an m-digit base-B expansion $k = \sum_{j=0}^{m-1} d_j B^j$ with $0 \leq d_j < B$. Then $[k]G_0 = \sum_{j=0}^{m-1} [d_j B^j] G_0$ and each summand $[d_j B^j] G_0 \in \mathcal{F}$. The union of m digit sets of size B has at most mB points. □

For $m = 3$ and $r = p^{1+o(1)}$ this base has size $O(p^{1/3}) = o(p^{1/2})$ — comfortably under the square-root bar. Figure 1 shows the measured sizes: 120, 190, 304 points at the three test primes, each below the $\lfloor \sqrt{p} \rfloor$ diagnostics 255, 511, 1023. Coverage was exact: across all 1,373,865 targets on the three curves, every returned term passed membership and every returned sum was correct.

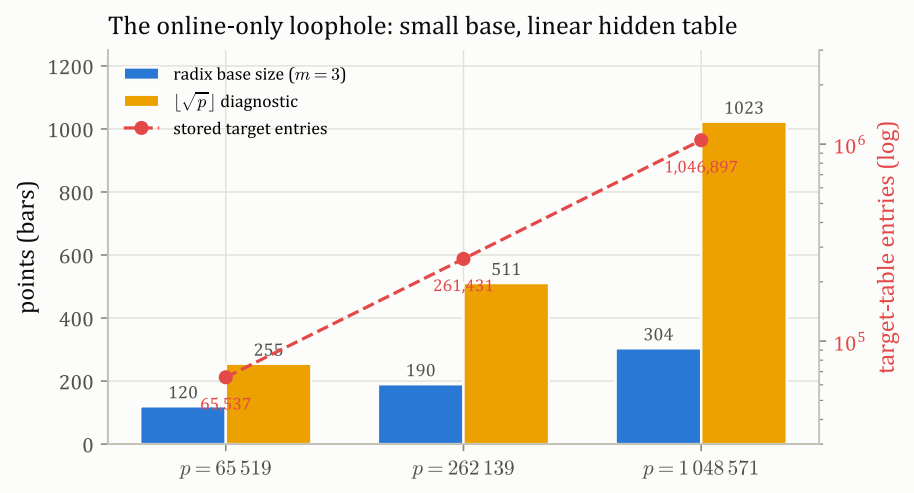

Figure 1: EMPIRICAL: p<=2^20, m=3 The radix construction of Theorem 4 (A008). Bars: the base size stays below the $\lfloor \sqrt{p} \rfloor$ diagnostic. Line (log axis): the decoder's target-indexed table holds exactly one entry per group element — 65,537, 261,431, 1,046,897 — so its description and preprocessing are $p^{1+o(1)}$. Data: `audit_preprocessing_loophole_p{65519,262139,1048571}_m3_20260630.csv`.

The decoder that makes online decomposition a lookup is where the trick hides: it stores, under the key $[k]G_0$, the m summands of k's digits, for **every** k. Membership and decomposition then run in polylogarithmic time with success one — but the table has exactly r entries.

Proposition 5 (the hidden table). CONDITIONAL: uncharged nonuniform advice With the radix base and a target-indexed decomposition table, online membership and decomposition take

poly$(\log p)$ time and succeed with probability one. PROVED This is purchased with r target entries, $\Theta(rm \log p)$ bits of input-specific description, and $\Omega(r)$ preprocessing steps — all $p^{1+o(1)}$, exponential in the input length $\Theta(\log p)$. The structure is a complete discrete-logarithm table in disguise and is not an efficient ECDLP attack.

The measured preprocessing times (0.24, 1.10, 6.04 s) and stored-reference counts (262,148, 1,045,724, 4,187,588 point references, four per target) confirm the $\Theta(r)$ scaling directly. The lesson is a **resource discipline**: a corrected statement must charge construction, description/advice, preprocessing, and storage, not only online lookup. We adopt the following convention throughout.

Definition 6 (charged resource model). A uniform constructor `Build(p, E)` outputs descriptions of the membership predicate and the decoder; its running time, output length, input-specific advice, preprocessing time, and persistent storage are all bounded explicitly. `Member` and `Decompose` are fixed uniform algorithms operating on those descriptions, and all preprocessing finishes before the independent target is drawn. PROVED Bounding only online operations does not enforce this convention — the radix table is the witness.

5 A measured taxonomy of prime-field candidates

With the literal statement resolved and the loophole named, we ask empirically whether any short-description prime-field predicate yields a square-root-size base with a better-than-generic decoder. Six constructions were measured at $p \in \{65519, 262139, 1048571\}$ (one ordinary prime-order curve per size, traces -17, 709, 1675, confirmed ordinary by the Waterhouse trace criterion [6]), using exact ordered three-term decomposition counts and a lexicographic pair scan as the generic finder.

The first calibration is that a **random** base of size $\lfloor\sqrt{p}\rfloor$ behaves exactly as the mean-value identity predicts.

Proposition 7 (expected decomposition count). PROVED For a fixed base of size s in a group of order r and a uniform target R, the expected number of ordered three-term decompositions is exactly s^3/r. Every ordered triple in $\mathcal{F}^3$ sums to one target, so summing target-wise counts over all r targets gives s^3; averaging gives s^3/r.

2 (left) shows the normalized mean count — the measured mean divided by s^3/r — hugging 1 for the random square-root base, the size-matched random base, and Candidate A (the integer-x interval $0 \le x(P) < \lfloor\sqrt{p}\rfloor$) at all three sizes; every 95% bootstrap interval contains 1. Candidate A therefore has the intended density but **no density advantage**: its normalized-count ratio to a size-matched random base is 1.030, 1.002, 1.004 with intervals $[0.963, 1.118]$, $[0.970, 1.035]$, $[0.986, 1.023]$, none excluding 1. Its generic pair-scan work is likewise indistinguishable (ratios 0.938, 1.050, 0.866, all intervals covering 1), and the descriptive log-log slope of pair checks versus p (2, right) is 0.408 for Candidate A against 0.437 and 0.529 for the random bases — a square-root-like growth, not a polylogarithmic finder.

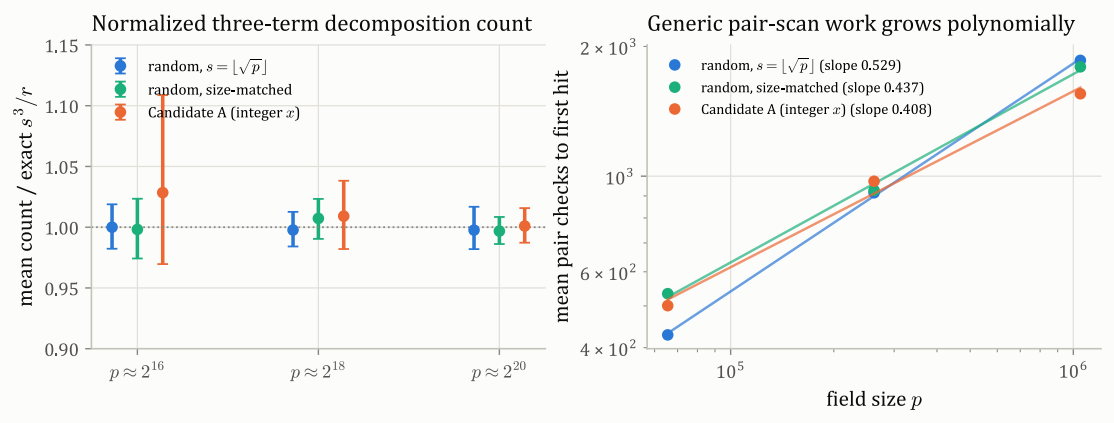

Figure 2: EMPIRICAL: one curve per size, seed 12022026 SG-01/SG-03. Left: normalized ordered-three-term counts match the exact s^3/r prediction for random and integer-x bases; error bars are 95% hierarchical-bootstrap intervals. Right: mean pair checks to the first hit grow with descriptive log-log slopes near $1/2$ for all three families, with no candidate-specific speedup. Data: `measure_factor_bases_b16-18-20_t96_r3_s12022026_20260622_{summary,scaling}.csv`.

The remaining candidates fall to the size or membership axis rather than density. EMPIRICAL: tested range Candidate B (bounded rational reconstruction, $x \equiv a/b$ with $|a|, |b| < \lfloor\sqrt{p}\rfloor$) selects 99.977%, 99.990%, 99.998% of the group — the symmetric square-root bounds are almost universal, so its trivial one-term success buys density by discarding the small-base requirement. EMPIRICAL: tested range Candidate D (the denominator-one integral-lift proxy, integral points with $|x| < \lfloor\sqrt{p}\rfloor$ on the centered lift) is the opposite extreme: base sizes 0, 0, 2, with three-term success 3.82×10^{-6} at 20 bits from the two signs above $x = -563$, $|y| = 7442$, canonical height 4.724. Candidate C is ruled out structurally, not empirically:

> **Proposition 8** (low-degree maps and auxiliary curves). PROVED Every rational map $\mathbb{P}^1 \dashrightarrow E$ is constant. It extends to a morphism because $\mathbb{P}^1$ is normal and E is proper [7]; factoring off the inseparable Frobenius part, a nonconstant separable morphism of genus-zero source and genus-one target would give, by Riemann–Hurwitz, $-2 = \deg(R) \geq 0$, a contradiction. PROVED A degree-d plane curve not containing E meets it in at most $3d$ points by Bézout [8], so a fixed-m base of this form reaches at most $(3d)^m$ targets; if the auxiliary curve contains E, the base becomes all of $E(\mathbb{F}_p)$.

CONDITIONAL: d, m = poly(log p) The first branch has success below $1/\operatorname{poly}(\log p)$ at prime-field scale, and the second fails the size condition; the bounded-degree plane subclass is thereby excluded, though not arbitrary higher-dimensional constructions. Table 2 consolidates the six families and their failure modes.

Table 2: The SG-07 candidate consolidation. Every prime-field predicate is too sparse, near-universal, or random-like at useful density; none supplies a small base **and** a better-than-generic decoder.

Candidate	Size behavior	Membership	Finder / outcome
Random baseline	$\approx \sqrt{p}$	explicit table	counts match s^3/r; calibration only
A: integer-x interval	$\approx \sqrt{p}$	PROVED polylog	density and scan indistinguishable from random
B: rational height	$> 99.97\%$ of group	not certified	trivial one-term; far too large
C: low-degree map / curve	singleton, $\leq 3d$, or whole curve	algebraic	PROVED too sparse or whole curve
D: integral-lift proxy	$0, 0, 2$ points	PROVED polylog	success $0, 0, 3.8 \times 10^{-6}$; too sparse
E: smooth subgroup	60 from $\lvert H\rvert = 64$	PROVED six quad. steps	density $\approx$ random; solver timed out (§7)

CONJECTURE The pattern across A001–A010 is a trichotomy: a short one-coordinate prime-field predicate tends to be too sparse for constant-length decomposition, random-like at useful density, or nearly universal. What the extension-field construction contributes and these do not is an **effective dimension drop** in the decomposition system, not merely a subset of the right cardinality. A square-root-size family with a proved polylogarithmic decoder on a growing prime-field range would refute the conjecture.

6 A lower bound for target-oblivious decoders

The measured $\sqrt{p}$-scale pair-scan work of §5 has, for a natural restricted class of decoders, a genuine proof rather than a three-point fit. Model the generic pair scan abstractly: preprocessing stores shifts $a = P + Q$ with $P, Q \in \mathcal{F}$, and an online **translate probe** tests whether $R - a \in \mathcal{F}$, returning $(P, Q, R - a)$ on a positive answer.

Theorem 9 (translate-probe bound). PROVED Fix $\mathcal{F} \subseteq G$ before drawing the uniform target R. For any schedule of at most T translate probes — fixed, failure-adaptive, or randomized — the uniform-target success probability is at most $T|\mathcal{F}|/r$. Consequently, achieving success δ requires $T \geq \delta r/|\mathcal{F}|$.

Proof. For a fixed shift a, the successful targets are exactly the translate $a + \mathcal{F}$, of size $s = |\mathcal{F}|$. For a fixed sequence of at most T probes the success set lies in $\bigcup_{j=1}^{T} (a_j + \mathcal{F})$, of size at most Ts, so success is at most Ts/r. For a failure-adaptive schedule, every observed bit before the first positive answer is zero, so — after fixing preprocessing — the shifts encountered follow the single all-failure path, and every successful target lies in one of the $\leq T$ translates on that path. For a randomized schedule, condition on the coins, apply the deterministic bound, and average. Target-independent preprocessing of $\mathcal{F}$ does not change the result. □

Corollary 10 (Candidate A needs square-root probes). PROVED Candidate A has $s \leq 2\lfloor\sqrt{p}\rfloor$ (a fixed abscissa admits at most two ordinates), while Hasse gives $r = p^{1+o(1)}$. Inverse-polylog-

arithmic success in the translate-probe model therefore requires $T \geq p^{1/2-o(1)}$ probes. The implemented lexicographic pair scan is inside this model.

We verified the union bound exhaustively on the order-19 fixture with a four-point base (Figure 3): across all 5,035 shift schedules of sizes one through four, **zero** exceeded the bound Ts, and 95, 950, 3,876 schedules for $T = 2, 3, 4$ fell strictly below it because translates overlapped — confirming that Ts is an upper bound, not an assumed independence equality.

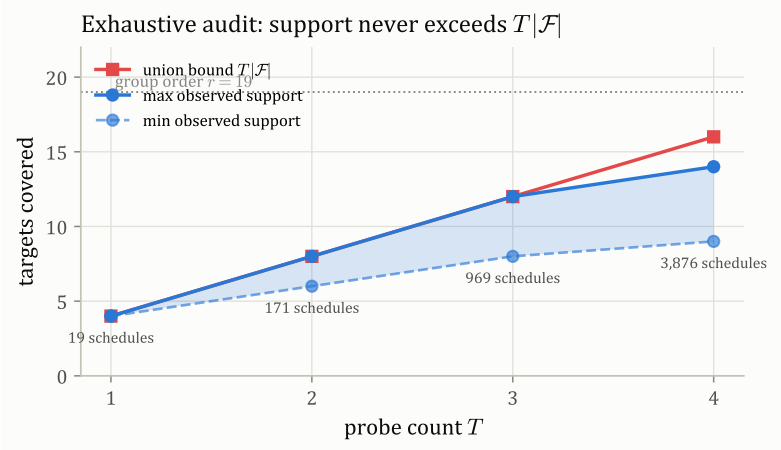

Figure 3: EMPIRICAL: r=19, exhaustive A009. Maximum and minimum observed support of every one-to-four-shift translate schedule against the union bound $T|\mathcal{F}|$, on the order-19 curve with a four-point base. No schedule exceeds the bound; collisions push many strictly below it. Data: `audit_translate_probe_r19_s4_t1-4_20260707.csv`.

Remark 11 (boundary of the result). PROVED The bound is for **target-oblivious** decoders. The radix table of §4 is outside it because it uses the full encoding of R to select a representation without walking the all-failure path. A decoder that reads $x(R)$, solves a coordinate equation, or otherwise chooses shifts from target coordinates is also outside it. This closes P1.2/Q001 for translate probes and leaves the coordinate-aware question P1.2/Q004 open.

7 The coordinate-aware frontier: smooth subgroups

The one construction that escapes both the near-universality of Candidate B and the translate-probe model is a **coordinate-aware** predicate: restrict abscissas to a smooth multiplicative subgroup, following Petit–Kosters–Messeng [9]. Its membership is genuinely succinct.

Lemma 12 (smooth-subgroup membership chain). PROVED If $H \subseteq \mathbb{F}_p^\times$ has smooth order $n = \prod_{j=1}^{t} \ell_j$, then $x \in H$ iff $x^n = 1$, and this single degree-n condition factors as the chain $z_0 = x$, $z_j = z_{j-1}^{\ell_j}$, $z_t = 1$, of t conditions of degrees ℓ_j. Since $\mathbb{F}_p^\times$ is cyclic, H is unique of its order and equals the root set of $x^n - 1$.

At $p = 65537$ the unique subgroup of order $64 = 2^6$ gives a chain of six quadratic squaring steps, evaluating membership in $O(\log 64)$ field operations. On the deterministic prime-order curve $E : y^2 = x^3 + 31771x + 14358$ (order 65809, trace -271), the predicate $x(P) \in H$ selects a 60-point factor base, and EMPIRICAL: all field and curve elements the direct predicate $x^{64} = 1$, explicit subgroup membership, and the six-step chain agreed with zero mismatches on all 65,537 field

elements and all 65,809 curve points. This predicate lies outside A009′s model, so the §6 lower bound says nothing about it.

Yet succinct membership is not a decoder. EMPIRICAL: 96 targets, seed 12022032 The subgroup base had normalized decomposition-count ratio 1.112 to matched random bases (95% interval $[0.819, 1.470]$) and pair-check ratio 0.969 ($[0.827, 1.130]$): no density or generic-search advantage. To probe an algebraic decoder we built the actual S_4-plus-subgroup polynomial systems and ran SymPy Gröbner bases in timeout-controlled subprocesses.

EMPIRICAL: SymPy 1.14.0, 5 s limit Both the direct and the repeated-squaring chain encodings completed only at $p = 17$ (direct: 7 basis polynomials in 0.28 s; chain: 18 in 1.13 s) and **timed out** at $p = 257$ and $p = 65537$ (Figure 4). At the largest fixture the chain traded degree for variables — maximum input degree 64 → 12, but variables 3 → 18 and equations 4 → 19. A timeout is neither a proof that a target lacks a decomposition nor a lower bound against another solver; the cited work leaves exactly this system-solving complexity open [9], [10].

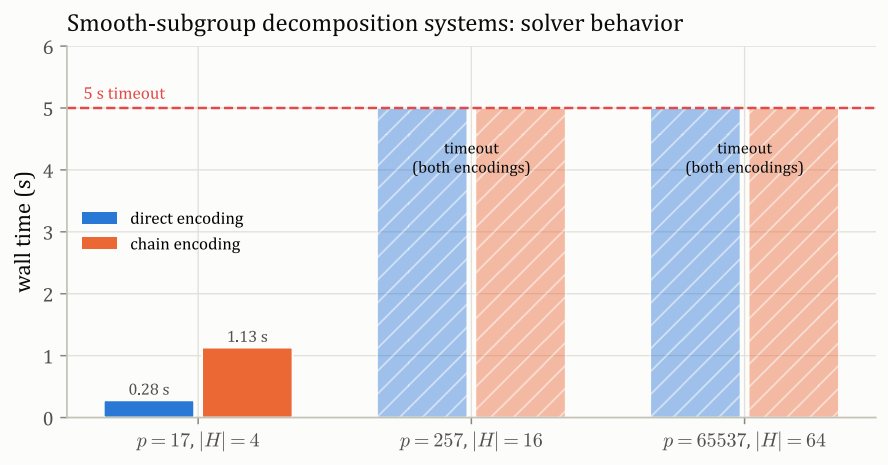

Figure 4: EMPIRICAL: SymPy 1.14.0, 5 s limit A010. Solver wall time for the smooth-subgroup decomposition systems under direct and chain encodings. Both complete only at $p = 17$; hatched bars mark five-second timeouts at $p = 257$ and $p = 65537$. The chain lowers degree but raises variable count, and neither becomes polynomial in $\log p$. Data: `benchmark_smooth_groebner_p17-65537_to5s_20260713.csv`.

PROVED The smooth-subgroup predicate is thus genuine coordinate structure outside the translate-probe model, with a low-degree membership chain but no realized efficient decoder. The recorded timeouts characterize only this solver and limit; they are not an asymptotic lower bound. P1.2/Q004 remains honestly open for the corrected square-root variant.

8 Why prime fields resist Weil restriction

The extension-field index calculus that motivates the whole problem works because a proper subfield supplies missing coordinates to eliminate. Over a prime field this mechanism is elementary but absolute.

Proposition 13 (no proper subfield of $\mathbb{F}_p$). PROVED In $\mathbb{F}_{q^n}$ the equation $x^q = x$ selects the embedded copy of $\mathbb{F}_q$; after choosing an $\mathbb{F}_q$-basis, a subfield element has one base-field coordinate against n in general, supplying $n - 1$ constraints. In the prime-field case $n = 1$ every $x \in \mathbb{F}_p$

satisfies $x^p = x$, and a subfield of $\mathbb{F}_p$ contains 1 and hence all of $\mathbb{F}_p$. Reusing the subfield equation therefore cuts out the whole field, not a factor base; the integer interval $0 \le x < \sqrt{p}$ is an external condition on a chosen representative that creates no smaller base field and hence no missing coordinate for Weil restriction to expose.

That the decomposition-and-linear-algebra machinery itself is sound — so that a prime-field failure cannot be blamed on a broken pipeline — is established by an extension-field positive control. EMPIRICAL: q=5,7,11 Over $\mathbb{F}_{q^3}$ with the Frobenius factor base $x^q = x$, the control decomposed sampled relations and recovered **every** planted logarithm end to end (Table 3), using nine relations in each case; scalar labels were constructed only after the linear solve, purely to validate it.

Table 3: SG-02 extension-field positive control (A002). The subfield predicate gives a small algebraic base and the relation/linear-algebra machinery recovers planted logarithms exactly. PROVED The control enumerates the whole extension curve and is intentionally exponential in the input length — a correctness control, not the asymptotic algorithm. Data: `run_extension_control_q{5,7,11}_*_summary.csv`.

q	$\lvert\mathbb{F}_{q^3}\rvert$	order r	base s	s^3/r	relations	secret recovered
5	125	139	8	3.68	9	37 (of 15 targets)
7	343	347	8	1.48	9	83 (of 34 targets)
11	1331	1367	8	0.38	9	123 (of 149 targets)

Diem's construction [3] makes this algebraic by requiring a degree-two covering value to lie in $\mathbb{P}^1(\mathbb{F}_q)$, and Gaudry [2] analyzes the related Weil-restriction index calculus; both rely on the proper base field that a prime field lacks. CONJECTURE What the extension construction supplies is an effective dimension drop; a prime-field predicate whose summation-polynomial system is provably polylogarithmic would refute this proposed common obstruction.

9 Obstructions, corrected variants, and open questions

The negative results above are precise enough to dictate the only coherent continuations. The specification has **two independent** defects — a size bound too small for constant length, and an online-only clock that permits lookup tables — and repairing one without the other yields either impossibility (§3) or triviality (§4). We isolate two resource-bounded replacement statements.

The two corrected variants. **Variant S (square-root size, fixed length).** Replace (1) by $|\mathcal{F}| \le p^{1/2+o(1)}$, fix $m = 3$, and require that `Build`, its output description, all input-specific advice, preprocessing, and storage be $p^{o(1)}$, with uniform polylogarithmic `Member` and `Decompose`. The support count does not exclude it ($|\mathcal{F}|^3/p$ can be large); the radix table **is** excluded because its preprocessing and storage are $p^{1+o(1)}$, and an explicit listing of a square-root-size base is excluded too — the predicate must be succinct.

Variant L (subexponential size, growing length). Retain $|\mathcal{F}| \le L_{p[1/2,c]}$ but permit $m(p) = \Theta\left(\sqrt{\log p / \log\log p}\right)$, with all offline resources bounded by a fixed-constant $L_{p[1/2,C]}$. The Corollary of §3 shows this growth of m is necessary up to its constant; this is the honest subexponential index-calculus target, for which no construction is supplied here.

Against this backdrop the residual questions are sharp. PROVED P1.2/Q001 (does the generic pair scan admit a polylogarithmic schedule?) is closed negatively for translate probes by §6.

CONJECTURE P1.2/Q004 — can a coordinate-aware decoder beat that bound in Variant S? — is open; the smooth-subgroup predicate of §7 is exactly the kind of construction that lives outside the lower bound, but our solver did not realize its decoder. Two further gaps are recorded rather than papered over: P1.2/Q002, the exact canonical-height membership predicate for the full Candidate D (only its denominator-one proxy was measured, base sizes $0, 0, 2$), and P1.2/Q003, the specification choice between Variants S and L that a future statement must make explicitly. Mixing the standard-L size, constant m, square-root baseline, and uncharged lookup — as the original prompt does — produces incompatible or vacuous requirements rather than a single research target.

We also flag the evidential limits of the empirical sections honestly. EMPIRICAL: recorded design Only one curve was selected per bit size, so no inference across the distribution of prime-order curves is licensed; counts are for ordered three-term decompositions with repetition; and the Candidate B and D measurements reuse one curve each without curve-to-curve variation. The descriptive log-log slopes of §5 are three-point fits, not asymptotic complexity estimates. None of these limitations affects the §3 impossibility, which is a counting theorem independent of any measurement.

10 Conclusion

We set out to construct a subexponential prime-field factor base with an efficient decoder, or to show the target is empty. The literal target is empty: under standard L-notation and constant m, a fixed-length sum over a $p^{o(1)}$-size base reaches only a $p^{-1+o(1)}$ fraction of the group, so conditions (1) and (3) are information-theoretically incompatible (Theorem 1), and the minimum repairs are the size floor $p^{1/m-o(1)}$ and the length growth $\sqrt{\log p/\log\log p}$ (Corollary 3). The natural square-root rescue is vacuous unless offline resources are charged: an elementary radix base covers every target while hiding a complete discrete-logarithm table (Theorem 4, Proposition 5). For target-oblivious decoders we proved a genuine $p^{1/2-o(1)}$-query lower bound (Theorem 9) and verified it exhaustively. Five explicit prime-field candidates were measured and each failed on size, density, or a missing decoder (§5); a validated extension-field control localizes the elementary reason prime fields resist Weil restriction (§8). The honest remainder is a single coordinate-aware findability question in the corrected square-root regime (§7, §9), for which the smooth-subgroup predicate is a live but unrealized candidate. We claim no positive construction and no break of prime-field ECC; the value here is a clean impossibility for the literal problem, a resource discipline that makes any successor problem well-posed, and a precise map of where a real prime-field factor base would have to live.

Reproducibility

All experiments run at the toy parameters stated inline ($p \le 2^{20}$; the extension control at $q \le 11$) using the repository's validated short-Weierstrass arithmetic and Semaev f_3, f_4, f_5 evaluation, checked against the known order-19 curve over $\mathbb{F}_{17}$ and independent symbolic resultants over $\mathbb{F}_{101}$. The random/Candidate-A baseline, extension control, structured candidates, preprocessing loophole, translate-probe audit, and smooth-subgroup density and Gröbner benchmarks are the scripts `measure_factor_bases.py`, `run_extension_control.py`, `measure_structured_candidates.py`, `audit_preprocessing_loophole.py`, `audit_translate_probe.py`, `measure_smooth_subgroup.py`, and `benchmark_smooth_groebner.py`; each writes the seeded CSV named in the corresponding caption. Confidence intervals are 95% hierarchical-

percentile bootstrap (2000 resamples) over base and target sampling, with Wilson intervals for success rates. Every mathematical claim above carries one of the epistemic tags PROVED, CITED, EMPIRICAL: range, CONDITIONAL: assumption, CONJECTURE exactly as the research log records it; untagged sentences are exposition, not claims.

References

[1] I. Semaev, "Summation polynomials and the discrete logarithm problem on elliptic curves." 2004.

[2] P. Gaudry, "Index calculus for abelian varieties of small dimension and the elliptic curve discrete logarithm problem," *Journal of Symbolic Computation*, vol. 44, no. 12, pp. 1690–1702, 2009.

[3] C. Diem, "On the discrete logarithm problem in elliptic curves," *Compositio Mathematica*, vol. 147, no. 1, pp. 75–104, 2011.

[4] A. K. Lenstra, "General purpose integer factoring," in *Topics in Computational Number Theory Inspired by Peter L. Montgomery*, Cambridge University Press, 2017, pp. 116–160.

[5] A. V. Sutherland, "Point counting." 2025.

[6] W. C. Waterhouse, "Abelian varieties over finite fields," *Annales scientifiques de l'École Normale Supérieure*, vol. 2, no. 4, pp. 521–560, 1969.

[7] The Stacks Project Authors, "The Stacks Project, Chapter 53: Algebraic Curves." 2026.

[8] Massachusetts Institute of Technology, "Notes for a course in algebraic geometry (Bézout's theorem)." 2021.

[9] C. Petit, M. Kosters, and A. Messeng, "Algebraic approaches for the elliptic curve discrete logarithm problem over prime fields," in *PKC 2016*, in LNCS, vol. 9615. 2016, pp. 3–18.

[10] A. Amadori, F. Pintore, and M. Sala, "On the discrete logarithm problem for prime-field elliptic curves," *Finite Fields and Their Applications*, vol. 51, pp. 168–182, 2018.