THE SPLIT LOCUS OF SUPERELLIPTIC CURVES: DETERMINATION FOR SMALL FIELDS, AND MEMBERS THAT ARE NEITHER MAXIMAL NOR HERMITIAN QUOTIENTS

KOREA SUPERINTELLIGENCE LABS

Abstract. Let $X\colon y^n = f(x)$ be a superelliptic curve over $\mathbb{F}_q$ with f separable of degree $d \geq 3$ splitting into distinct linear factors and with $n \mid q-1$. Say that the pair (n, f) lies in the *split locus* when every value of f outside its roots is an n-th power in $\mathbb{F}_q^*$, so that every unramified fibre of x splits into n rational points. A recent preprint isolated this locus, exhibited one family inside it — a family of quotients of the Hermitian curve, all of them maximal — wrote that it did not know how far the list extends, and asked whether every member of the locus is a maximal Hermitian quotient. We answer both halves in the negative and determine the locus over small fields. Over $\mathbb{F}_{16}$ the pair $n = 3$, $f = \mathrm{Tr}_{\mathbb{F}_{16}/\mathbb{F}_2}$ lies in the locus, and its curve has 33 rational points against a maximality bound of 73 and genus 7 against the Hermitian genus 6; this is proved with no appeal to any axiom. It is not an accident of one field: for every even $m = 2e \geq 4$, $q = 2^m$, and every $n \geq 3$ dividing $q-1$, the curve $y^n = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_2}(x)$ lies in the locus and is neither maximal nor a quotient of the Hermitian curve, and at $e = 1$ both inequalities degenerate into equalities — which is exactly the Hermitian case. We further determine the locus completely, by exhaustive search over all 2^q root sets: for nine fields $q \leq 19$ inside a formal development, and, outside it, for six further fields, up to $q = 32$. Over $\mathbb{F}_{16}$ the locus consists of 2101 pairs (n, R), of which the previously known Hermitian family is a single cell of 48. Finally, Weil's character-sum bound yields $d \geq \sqrt{q}$, sharper than the pigeonhole bound of the source for small n, and over the square fields $q = 9, 16, 25$ the minimum $d = \sqrt{q}$ is attained exactly by the Hermitian quotients. The answer to the source's question is therefore yes at minimal degree and no above it. Every numbered statement of Sections 3–5 is machine-checked in Lean 4, with one documented gap: the finite-field input to the infinite family is verified in three concrete field models rather than uniformly in m. The computations that lie outside the formal development are labelled as such where they appear.

1. Introduction

Let $q = p^m$ be a prime power, let $n \geq 2$ with $p \nmid n$, and let $f \in \mathbb{F}_q[x]$ be separable of degree d. The *superelliptic curve* $X\colon y^n = f(x)$ is a cyclic Kummer cover of the projective line of degree n, and when $n \mid q-1$ its geometry over $\mathbb{F}_q$ is governed by one arithmetic question: for each $a \in \mathbb{F}_q$, is the value $f(a)$ an n-th power? If it is and $f(a) \neq 0$, the fibre over a consists of n rational points; if it is not, the fibre carries no rational point at all; if $f(a) = 0$ the fibre is a single ramified rational point. Curves for which the second case never occurs are the extreme case of this dichotomy, and they are the subject of this paper.

The source and its problem. A recent preprint [1] studies hulls and linear equivalence for the algebraic-geometry codes attached to such curves, under two standing hypotheses, which we quote verbatim, in its own notation and under its own labels:

- **(H1′)** [1, §7.1]: $p \nmid n$, the polynomial $f \in \mathbb{F}_q[x]$ is separable of degree $d \geq 3$, and f splits into distinct linear factors over $\mathbb{F}_q$;
- **(H4)** [1, §4]: $n \mid q-1$, so that $\zeta_n \in \mathbb{F}_q$.

Writing $\Delta = \{a \in \mathbb{F}_q : f(a) \neq 0,\ f(a) \notin (\mathbb{F}_q^*)^n\}$ and $\delta = |\Delta|$, the source isolates the locus $\delta = 0$ and exhibits one family in it [1, Thm. 7.10]: for $q = q_0^2$ and $n \mid q_0 + 1$, the curves $X_n\colon y^n = -(x^{q_0+1} + 1)$, all of which are maximal and are quotients of the Hermitian curve. Immediately after the sentence

"We do not know how far the list extends" it poses the following problem, which we quote verbatim from the e-print, transcribed into the notation above.

Problem 1.1 ([1, Problem 7.11])**.** Determine the pairs (n, f) over $\mathbb{F}_q$ satisfying (H1′) and (H4) for which $\delta = 0$, that is, for which f splits into d distinct linear factors over $\mathbb{F}_q$ and $f(a) \in (\mathbb{F}_q^*)^n$ for every $a \in \mathbb{F}_q$ outside the roots of f. Are they all quotients of Hermitian curves, and are they all maximal?

We call the set of such pairs the *split locus* over $\mathbb{F}_q$. The problem has three parts: an enumeration (*determine the pairs*), and two yes/no questions about the members found. This paper settles the two yes/no questions and carries out the enumeration as far as exhaustive search reaches.

Results. Both yes/no halves are *no*, and the counterexample is elementary once one looks at the right polynomials. Over $\mathbb{F}_{16}$ take $n = 3$ and let f be the product of $x - r$ over the eight elements of an $\mathbb{F}_2$-hyperplane; then f is the absolute trace, its value set is $\{0, 1\}$, and 1 is a cube, so $(3, f)$ lies in the split locus (Theorem 3.2). Its curve has 33 rational points where a maximal curve of its genus would have 73, and genus 7 where the Hermitian curve over $\mathbb{F}_{16}$ has genus 6; a quotient cannot have larger genus than its cover, so the curve is neither maximal nor a Hermitian quotient (Theorem 3.3). The refutation happens over a *square* field, so the charitable reading $q = q_0^2$ — under which maximality is defined at all — does not rescue the question.

The counterexample is not an accident of one small field. Theorem 4.1 gives an infinite family: for every even $m = 2e \geq 4$, $q = 2^m$, and every $n \geq 3$ dividing $q - 1$, the curve $y^n = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_2}(x)$ lies in the split locus, misses the Hasse–Weil bound and has genus strictly above the Hermitian genus. The hypothesis $e \geq 2$ is exactly what is needed: at $e = 1$, where the absolute trace coincides with the trace to $\mathbb{F}_{\sqrt{q}}$, both inequalities collapse to equalities and the curve *is* a Hermitian quotient (Proposition 4.3). So the family degenerates precisely into the source's own example.

For the enumeration, a reduction (Lemma 2.2) turns $\delta = 0$ into a condition on the root set of f alone together with a free choice of leading coefficient inside one coset of $(\mathbb{F}_q^*)^n$, which makes the locus a subset of the 2^q subsets of $\mathbb{F}_q$ and the determination a finite search. Theorems 5.1 and 5.2 carry that search out for the nine fields $q \in \{4, 5, 7, 8, 9, 11, 13, 16, 19\}$, and §6 reports it, outside the formal development, for six further fields, up to $q = 32$. Over $\mathbb{F}_{16}$ the locus has 2101 pairs (n, R) and 7395 pairs (n, f), and the source's Hermitian family is a single cell of 48: the locus is more than forty times as large as the part of it that was known.

Two structural facts emerge from the data (§7). First, Weil's character-sum bound gives a degree bound $d \geq \sqrt{q}$ (Proposition 7.4) which is sharper than the pigeonhole bound $d \geq qn/(q + n - 1)$ of [1] for small n; it is attained only over square fields. Second, over the square fields our search reaches ($q = 9, 16, 25$) the minimum degree $\sqrt{q}$ is attained exactly for those n dividing $q_0 + 1$, by the translates of the $\mathbb{F}_{q_0}$-lines — which are precisely the Hermitian quotients, and are maximal. The source's guess is therefore right at minimal degree and wrong above it, and that is the shape of the answer to Problem 1.1.

What is classical here, and what is not. The construction behind every counterexample in this paper — a Kummer cover $y^N = f(x)$ in which f is the additive polynomial of an $\mathbb{F}_p$-subspace of $\mathbb{F}_q$ — goes back to van der Geer and van der Vlugt [2], who build Kummer covers of the projective line out of such polynomials. Their §3 takes $L = \ker \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_p}$ and the additive polynomial $R(x) = \sum_i x^{p^i}$ exactly as we do, and their Proposition (3.5) exhibits the *maximal* member of the construction, identified as a quotient of the Hermitian curve in the sentence that follows it. The mechanism, and the "yes" side of Problem 1.1, are therefore more than twenty-five years old. What is new, as far as our searches reach, is that the *other* members of the same construction populate the $\delta = 0$ locus of [1], that they answer Problem 1.1 in the negative for infinitely many q, and the determination of the locus itself.

The negative bibliographic evidence behind that claim is worth stating precisely, because a reader should be able to judge it rather than re-run it. The source is live at v1, published on 3 August 2026

and never updated, with no recorded citations; a full-text search of a local arXiv mirror (730 shards, 85 GB) for the identifier, for the phrase *weighted superelliptic*, for *split locus* and for the source's own phrasing of (H1′) returns only the source itself and unrelated uses of the words; a listing query returns exactly one arXiv abstract containing *superelliptic* in the window since publication, again the source. One bibliographic service (Semantic Scholar) returned HTTP 429 on every attempt, including on its own positive control, and therefore contributed nothing; we record that as a gap in the search, not as a negative result. Searches of the OEIS for the totals of §5 and for the minimal-degree column of §7 returned no match, with a positive control confirming the queries were well formed.

Organisation. §2 fixes notation, records the invariants we use and proves the reduction. §3 gives the $\mathbb{F}_{16}$ counterexample and the positive control that the counting machinery reproduces the source's own maximal curve. §4 gives the infinite family. §5 gives the determination for nine fields $q \le 19$ and §6 its unformalised extension to six more, up to $q = 32$. §7 describes the three structures visible in the answer and the two degree bounds. §8 describes the formal verification.

2. Preliminaries

Throughout, $q = p^m$, $\mathbb{F}_q$ is the field with q elements, $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$ and $(\mathbb{F}_q^*)^n$ is the subgroup of n-th powers, of index $\gcd(n, q-1)$. We work under (H1′) and (H4) of §1; in particular $n \mid q-1$, so $(\mathbb{F}_q^*)^n$ has index n and order $(q-1)/n$, and $\mu_n \subseteq \mathbb{F}_q^*$.

Let $f \in \mathbb{F}_q[x]$ be separable of degree $d \ge 3$, split over $\mathbb{F}_q$, with root set $R \subseteq \mathbb{F}_q$, $|R| = d$, and leading coefficient $c \in \mathbb{F}_q^*$, so that $f = c\prod_{r\in R}(x-r)$. Let X be the smooth projective model of $y^n = f(x)$. Since f is separable and $p \nmid n$, the affine plane model is nonsingular, and the standard theory of Kummer extensions [4, Prop. 3.7.3, Cor. 3.7.4] gives the invariants we use, which are also the ones the source records [1, §2 and Lem. 7.1]:

$$\#\{\text{places of } X \text{ over } x=\infty\} = c_\infty := \gcd(n,d), \tag{1}$$

$$\#X(\mathbb{F}_q) = d + n\big(q - d - \delta\big) + \#\{\text{rational places over } \infty\}, \tag{2}$$

$$2g_X - 2 = nd - n - d - c_\infty, \tag{3}$$

and when $c_\infty = 1$ the unique place over $x = \infty$ is rational. Here (2) counts the d ramified affine points $(r, 0)$, $r \in R$, and the n points over each of the $q - d - \delta$ unramified a with $f(a) \in (\mathbb{F}_q^*)^n$; the middle term is the quantity called N in [1, Lem. 7.1]. A curve X over $\mathbb{F}_q$ is *maximal* when q is a square and $\#X(\mathbb{F}_q) = q + 1 + 2g_X\sqrt{q}$, the upper end of the Hasse–Weil bound. The Hermitian curve over $\mathbb{F}_{q_0^2}$ is $H\colon y^{q_0+1} = x^{q_0} + x$, of genus $q_0(q_0-1)/2$, and it is maximal. We use one further standard fact: if G is a group of automorphisms of a curve Y, then the quotient map $Y \to Y/G$ is separable, so by Riemann–Hurwitz $g_{Y/G} \le g_Y$. In particular a quotient of the Hermitian curve over $\mathbb{F}_{16}$ has genus at most 6. We never need Serre's theorem that quotients of maximal curves are maximal; the genus comparison is enough and is cheaper.

Definition 2.1. The *split locus* over $\mathbb{F}_q$ is the set of pairs (n, f) satisfying (H1′) and (H4) with $\delta = 0$. We also speak of the set of admissible pairs (n, R), R being the root set of f; the passage between the two is Lemma 2.2.

Lemma 2.2. *Fix a primitive element of $\mathbb{F}_q^*$ and write* ind *for the resulting discrete logarithm. Let $R \subseteq \mathbb{F}_q$ with $|R| = d \ge 3$ and let $f = c\prod_{r\in R}(x-r)$. For $a \notin R$,*

$$f(a) \in (\mathbb{F}_q^*)^n \iff \operatorname{ind}(c) + \sum_{r\in R} \operatorname{ind}(a-r) \equiv 0 \pmod{n}.$$

Consequently $\delta = 0$ holds for some choice of c if and only if $S_R(a) := \sum_{r\in R}\operatorname{ind}(a-r) \bmod n$ is constant on $\mathbb{F}_q \setminus R$, and in that case the admissible c form a single coset of $(\mathbb{F}_q^*)^n$, of size $(q-1)/n$ — all $q-1$ elements when $d = q$, where the condition is vacuous.

Proof. An element of $\mathbb{F}_q^*$ is an n-th power exactly when its discrete logarithm is divisible by n, since $\mathbb{F}_q^*$ is cyclic of order $q-1$ and $n \mid q-1$. Apply this to $f(a) = c\prod_{r\in R}(a-r)$ and use additivity of ind on products. The last sentence is immediate: the condition constrains $\mathrm{ind}(c)$ modulo n to a single residue. □

Lemma 2.2 is what makes Problem 1.1 finite: the split locus over $\mathbb{F}_q$ is determined by the family of subsets $R \subseteq \mathbb{F}_q$ with $|R| \geq 3$ on which S_R is constant off R, of which there are at most 2^q to inspect for each of the divisors $n \geq 2$ of $q-1$. Note also that (H4) already implies $p \nmid n$, since $p \nmid q-1$.

We record the arithmetic used repeatedly below. If $\gcd(n,d) = 1$ then by (1) and (3) there is a single rational place at infinity and $2g_X = (n-1)(d-1)$, so with $\delta = 0$,

$$\#X(\mathbb{F}_q) = d + n(q-d) + 1, \qquad 2g_X = (n-1)(d-1). \tag{4}$$

3. A pair over $\mathbb{F}_{16}$ that is neither maximal nor a Hermitian quotient

We use the model $\mathbb{F}_{16} = \mathbb{F}_2[X]/(X^4+X+1)$ and encode its elements as the integers $0,\dots,15$ whose binary digits are the coordinates in the basis $1, X, X^2, X^3$. Under this encoding the set $\{0,1,\dots,7\}$ is the $\mathbb{F}_2$-span of $\{1, X, X^2\}$, an $\mathbb{F}_2$-hyperplane.

Proposition 3.1. *The encoded model is a commutative ring with* 1 *— addition and multiplication are commutative and associative, multiplication distributes over addition,* 0 *and* 1 *are the identities, and every element is its own additive inverse — and the element X has multiplicative order* 15*. Hence every nonzero element is a power of X and the model is a field with* 16 *elements.*

Proof. The ring axioms are checked on all 16^3 triples and the order of X on its fifteen powers, by direct evaluation of the tables; nothing else is used. □

Theorem 3.2. *Let $q = 16$, let $H = \{0,1,\dots,7\}$ be the $\mathbb{F}_2$-hyperplane above, let $n = 3$ and*

$$f(x) = \prod_{r\in H}(x-r) = x^8 + x^4 + x^2 + x = \mathrm{Tr}_{\mathbb{F}_{16}/\mathbb{F}_2}(x).$$

Then (n,f) satisfies (H1′) *and* (H4) *and has $\delta = 0$: f is separable of degree $d = 8 \geq 3$ and splits into eight distinct linear factors over $\mathbb{F}_{16}$, $3 \mid 15$ and $2 \nmid 3$, the value set of f on $\mathbb{F}_{16}$ is $\{0,1\}$ with exactly eight zeros, and $1 = 1^3$ is a cube. The smooth model $X\colon y^3 = f(x)$ has* 32 *affine rational points, one rational place over $x = \infty$, hence $\#X(\mathbb{F}_{16}) = 33$, and genus $g_X = 7$.*

Proof. The identification $f = \mathrm{Tr}_{\mathbb{F}_{16}/\mathbb{F}_2}$ is checked at all sixteen points of $\mathbb{F}_{16}$; since both sides are monic of degree $8 < 16$, agreement at sixteen points is equality of polynomials. The value set and the count of zeros are read off the same sixteen evaluations. For $\delta = 0$ one checks, for each of the eight $a \notin H$, that $f(a) = 1 = b^3$ for some $b \in \mathbb{F}_q^*$; the search runs over all sixteen candidates b for each a, so the verification is exhaustive over 16×16 pairs. The affine point count 32 is obtained by evaluating $b^3 = f(a)$ on all 256 pairs $(a,b) \in \mathbb{F}_{16}^2$, and the split $32 = 8 + 24$ into ramified and unramified points reproduces $N = n(q-d-\delta) = 24$ of (2). Since $\gcd(3,8) = 1$, (1) gives one place over $x = \infty$, necessarily rational, so $\#X(\mathbb{F}_{16}) = 33$, and (4) gives $2g_X = (3-1)(8-1) = 14$. Every step other than the invariants (1)–(3) quoted from §2 is a finite evaluation in the field tables of Proposition 3.1. □

Theorem 3.3. *The curve X of Theorem 3.2 is not maximal and is not a quotient of the Hermitian curve over $\mathbb{F}_{16}$. Both halves of Problem 1.1 are therefore answered in the negative, over a square field.*

Proof. A maximal curve of genus 7 over $\mathbb{F}_{16}$ has $q + 1 + 2g\sqrt{q} = 16 + 1 + 2\cdot 7\cdot 4 = 73$ rational points, and $\#X(\mathbb{F}_{16}) = 33 < 73$. The Hermitian curve over $\mathbb{F}_{16}$ is $H\colon y^5 = x^4 + x$, of genus $q_0(q_0-1)/2 = 6$; a quotient of a curve by a group of automorphisms has genus at most that of the curve, while $g_X = 7 > 6$. So X is not such a quotient. The arithmetic ($33 < 73$ and $6 < 7$) is the only computation; everything else is Theorem 3.2 and the standard facts of §2. □

Remark 3.4. The genus comparison is what makes the second half cheap: it needs no property of maximal curves, only that a quotient map is separable. In particular the refutation does not depend on Serre's theorem.

The counting machinery of Theorem 3.2 is calibrated against the source's own example over the same field.

Proposition 3.5. *Over the same model of* $\mathbb{F}_{16}$*, take* $n = 5$ *and let* $R = \mu_5 = \{1, X^3, X^6, X^9, X^{12}\}$ *be the set of fifth roots of unity, so that* $\prod_{r\in R}(x-r) = x^5+1$*. This is the source's family at* $q_0 = 4$*,* $n = 5$*. Then* R *has five distinct elements,* $\delta = 0$*, the unramified affine count is* $N = n(q_0^2 - q_0 - 1) = 55$*, the genus is* $(n-1)(q_0-1)/2 = 6$*, and with the* $c_\infty = \gcd(5,5) = 5$ *places at infinity all rational* [1, Thm. 7.10]*,*

$$\#X(\mathbb{F}_{16}) = 55 + 5 + 5 = 65 = 16 + 1 + 2 \cdot 6 \cdot 4,$$

the maximality bound exactly: the curve is maximal, as [1, Thm. 7.10] *states.*

Proof. As in Theorem 3.2: all statements are finite evaluations in the same field tables, over the sixteen elements and the 256 pairs, together with (1)–(3) and the rationality of the places at infinity quoted above. □

Proposition 3.6. *The three computations above are not vacuous.*

(1) $\delta = 0$ *is not automatic for eight-element root sets: replacing one element of* H *by an element outside it, in either of two ways, gives* $\delta \neq 0$ *at* $n = 3$*.*
(2) *The point counter separates cases: it returns* 65 *on the curve of Proposition 3.5 and* 33 *on the curve of Theorem 3.2.*
(3) *Both hypotheses are live at the witness: (H4) fails for* $n = 4$ *over* $\mathbb{F}_{16}$ *(since* $4 \nmid 15$*) and the degree condition* $d \geq 3$ *fails for a two-element root set, while both hold for* $(3, H)$*.*

Proof. Each item is a finite evaluation in the tables of Proposition 3.1: two further root sets for (1), the two point counts for (2), two divisibility tests and two cardinality tests for (3). □

4. An infinite family

Let $q = 2^m$ and let $f = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_2}(x) = \sum_{i<m} x^{2^i}$, of degree $d = 2^{m-1}$. In characteristic two,

$$f(x)^2 + f(x) = \sum_{i<m} x^{2^{i+1}} + \sum_{i<m} x^{2^i} = x^{2^m} + x = \prod_{a\in\mathbb{F}_q} (x-a), \tag{5}$$

an identity of polynomials. Three consequences are immediate. First, f divides $\prod_a (x-a)$, which is squarefree and split over $\mathbb{F}_q$; hence f is separable and splits into $d = 2^{m-1}$ distinct linear factors over $\mathbb{F}_q$, and $d \geq 8 \geq 3$ as soon as $m \geq 4$. Second, evaluating (5) at $a \in \mathbb{F}_q$ gives $f(a)^2 = f(a)$, so the value set of f on $\mathbb{F}_q$ is $\{0,1\}$, with d zeros by the first point. Third, $1 = 1^n$ for every n, so $\delta = 0$ for every $n \mid q-1$; and $n \mid 2^m - 1$ forces n odd, so $p = 2$ does not divide n. Thus $(n, \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_2})$ lies in the split locus for every $n \geq 2$ dividing $2^m - 1$.

Theorem 4.1. *Let* $m = 2e$ *with* $e \geq 2$*, let* $q = 2^m$ *and let* $n \geq 3$ *divide* $q-1$*. Then the pair* $(n, \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_2})$ *lies in the split locus over* $\mathbb{F}_q$*, and the curve* $X\colon y^n = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_2}(x)$ *satisfies, with* $d = q/2$ *and* $Q = 2^e = \sqrt{q}$*,*

$$\#X(\mathbb{F}_q) = d(n+1) + 1 \;<\; q + 1 + 2g_X\, Q, \qquad 2g_X = (n-1)(d-1) \;>\; Q(Q-1) = 2g_H,$$

H being the Hermitian curve over $\mathbb{F}_q$*. Hence no member of this family is maximal, and no member is a quotient of the Hermitian curve over its field of definition. The family is nonempty for every even* $m \geq 4$*, since* $3 \mid 2^{2e} - 1$*.*

Proof. Membership in the split locus is the discussion around (5). As n is odd we have $c_\infty = \gcd(n, 2^{m-1}) = 1$, so (4) applies and gives, using $q - d = d$, both $\#X(\mathbb{F}_q) = d + nd + 1$ and $2g_X = (n-1)(d-1)$. With $q = Q^2 = 2d$ the two claimed inequalities read

$$d(n+1) + 1 < 2d + 1 + (n-1)(d-1)Q \qquad \text{and} \qquad Q(Q-1) < (n-1)(d-1).$$

The first cancels to $d < (d-1)Q$, which holds because $Q \geq 4$ and $d \geq 8$. The second follows from $Q(Q-1) = 2d - Q \leq 2d - 4 < 2(d-1) \leq (n-1)(d-1)$. Both are inequalities between natural numbers in the two parameters $e \geq 2$ and $n \geq 3$, proved uniformly, with no finite search anywhere. Finally $3 \mid 2^{2e} - 1$ by induction on e, so an admissible $n \geq 3$ exists for every even $m \geq 4$. □

Remark 4.2. The division of labour in the proof is worth stating exactly, and it is reflected in the formal development (§8). The two inequalities are proved uniformly in e and n. The finite-field input — that $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_2}$ splits into 2^{m-1} distinct linear factors with value set $\{0, 1\}$, i.e. identity (5) and its three consequences — is elementary, but in the formal development it is verified in three concrete field models ($q = 16, 64, 256$; Theorem 3.2 and Proposition 4.4) rather than proved uniformly in m.

Proposition 4.3. *At $e = 1$, that is over $\mathbb{F}_4$, both inequalities of Theorem 4.1 become equalities at $n = 3$: the point count equals the maximality bound, and $2g_X = 2 = 2g_H$. Moreover at $n = 1$ the construction degenerates, $2g_X = 0$ and the point count again equals the bound. Hence the hypotheses $e \geq 2$ and $n \geq 3$ are exactly what separate the Hermitian case from the rest.*

Proof. Direct evaluation of the four quantities at $(e, n) = (1, 3)$, and the identity $2g_X = (n-1)(d-1) = 0$ at $n = 1$ for every e. □

Proposition 4.3 is the reason the family is best read as a deformation of the source's own example. Over $\mathbb{F}_4$ the absolute trace *is* the trace to $\mathbb{F}_{\sqrt{q}} = \mathbb{F}_2$, the curve is a Hermitian quotient, and everything is tight. Raising e keeps the pair in the split locus and breaks both tightness properties at once.

Proposition 4.4. *In the models $\mathbb{F}_{64} = \mathbb{F}_2[X]/(X^6+X+1)$ and $\mathbb{F}_{256} = \mathbb{F}_2[X]/(X^8+X^4+X^3+X^2+1)$ the absolute trace has exactly $q/2$ roots and takes the value 1 on the remaining $q/2$ elements, and each admissible exponent ($n \in \{3, 7, 9, 21, 63\}$ for $q = 64$, $n \in \{3, 5, 15, 17, 51, 85, 255\}$ for $q = 256$) has exactly n n-th roots of unity. Hence $\#X(\mathbb{F}_q) = q/2 + n\,q/2 + 1$ for each of them, as in Theorem 4.1. In particular*

q	n	$\#X(\mathbb{F}_q)$	$2g_X$	$q + 1 + 2g_X\sqrt{q}$	$2g_H$
16	3	33	14	73	12
64	3	129	62	561	56
256	3	513	254	4321	240

and every one of the fifteen admissible pairs (q, n) with $q \in \{16, 64, 256\}$ falls under Theorem 4.1, hence satisfies $\#X(\mathbb{F}_q) < q + 1 + 2g_X\sqrt{q}$ and $2g_X > 2g_H$.

Proof. For each field, the trace is evaluated at all q elements and the n-th roots of unity are counted by evaluating b^n at all q elements, for each admissible n; the tabulated invariants are then integer arithmetic. Each model is certified before use: at $q = 64$ the commutative-ring axioms are checked on all 64^3 triples, and at both $q = 64$ and $q = 256$ a cheaper certificate is used — X has multiplicative order $q - 1$, multiplication is commutative, and $0, 1$ behave — which forces every nonzero element to be a power of X and hence invertible. □

Proposition 4.5. *The field certificate of Proposition 4.4 has content: it rejects the reducible modulus $X^6 + X^2 + 1 = (X^3 + X + 1)^2$, whose quotient ring has seven non-invertible nonzero elements, and it rejects the irreducible AES modulus $X^8 + X^4 + X^3 + X + 1$, in which X has multiplicative order 51 rather than 255.*

Proof. Two evaluations of the certificate, each returning false. □

5. The determination for nine fields, $q \le 19$

By Lemma 2.2 the split locus over $\mathbb{F}_q$ is determined by the pairs (n, R) with $n \ge 2$, $n \mid q-1$, $R \subseteq \mathbb{F}_q$, $|R| = d \ge 3$ and S_R constant off R; each such pair carries $(q-1)/n$ admissible leading coefficients, or $q-1$ when $d = q$. The following two theorems record the result of enumerating *all* 2^q subsets R, for every admissible n, for each field named. We write the answer as the array of counts indexed by d, which is the finest invariant the enumeration produces and the one the structural discussion of §7 needs.

Theorem 5.1. *For each $q \in \{4, 5, 7, 8, 9, 11, 13, 19\}$ and each $n \ge 2$ dividing $q-1$, the numbers of admissible root sets R of each size d are exactly as follows; unlisted sizes have count zero.*

q	n	*nonzero cells* (d : *count*)
4	3	3:4, 4:1
5	2	3:10, 4:5, 5:1
5	4	4:5, 5:1
7	2	3:14, 4:14, 6:7, 7:1
7	3	5:21, 6:7, 7:1
7	6	6:7, 7:1
8	7	4:14, 6:28, 7:8, 8:1
9	2	3:12, 4:18, 5:18, 6:12, 7:36, 8:9, 9:1
9	4	3:12, 4:18, 6:12, 7:36, 8:9, 9:1
9	8	6:12, 8:9, 9:1
11	2	6:22, 7:110, 8:55, 10:11, 11:1
11	5	9:55, 10:11, 11:1
11	10	10:11, 11:1
13	2	7:156, 8:117, 9:130, 10:52, 11:78, 12:13, 13:1
13	3	5:39, 6:52, 9:39, 11:78, 12:13, 13:1
13	4	7:52, 8:39, 10:52, 12:13, 13:1
13	6	11:78, 12:13, 13:1
13	12	12:13, 13:1
19	2	7:114, 8:171, 10:380, 11:1026, 12:684, 14:684, 15:1140, 16:285, 18:19, 19:1
19	3	11:684, 12:456, 13:228, 14:513, 15:171, 16:114, 17:171, 18:19, 19:1
19	6	16:114, 18:19, 19:1
19	9	17:171, 18:19, 19:1
19	18	18:19, 19:1

Summing over n and d, the numbers of admissible pairs (n, R) and, after Lemma 2.2, of pairs (n, f) in the split locus are

q	4	5	7	8	9	11	13	19
pairs (n, R)	5	22	73	51	216	278	1032	7206
pairs (n, f)	7	43	186	57	639	1163	4883	55551

Each field model used is certified to be a field of the stated size. Of the sixteen totals, the fourteen with $q \le 13$ are themselves machine-checked; the two for $q = 19$ are obtained by summing the array above.

Theorem 5.2. *Over $\mathbb{F}_{16}$ the split locus is, in the same format,*

q	n	*nonzero cells* (d : *count*)
16	3	8:270, 9:240, 10:48, 11:48, 12:140, 14:120, 15:16, 16:1
16	5	4:20, 5:48, 8:30, 9:240, 10:168, 12:140, 13:80, 14:120, 15:16, 16:1
16	15	8:30, 10:48, 12:140, 14:120, 15:16, 16:1

In total there are 2101 *admissible pairs (n, R) and* 7395 *pairs (n, f). The source's Hermitian family is the single cell $(n, d) = (5, 5)$, of size* 48*; the counterexample of Theorem 3.2 lies in the cell $(3, 8)$, of size* 270*.*

Proof of Theorems 5.1 and 5.2. Both are exhaustive computations, and we describe exactly what was run. Prime fields are represented as $\{0, \dots, q-1\}$ with arithmetic modulo q; $\mathbb{F}_4, \mathbb{F}_8, \mathbb{F}_{16}$ as

$\mathbb{F}_2[X]/(X^m + X + 1)$ with X primitive; $\mathbb{F}_9$ as $\mathbb{F}_3[X]/(X^2 + 1)$ with $1 + X$ primitive. Each model is certified before any counting: the commutative-ring axioms, the behaviour of 0 and 1, and the invertibility of subtraction are checked on all triples — q^3 of them — and the designated generator is checked to have multiplicative order exactly $q - 1$, which makes every nonzero element a power of it and hence invertible. The discrete-logarithm and subtraction tables are then built. For each q and each $n \geq 2$ dividing $q - 1$, every one of the 2^q subsets $R \subseteq \mathbb{F}_q$ is enumerated as a bit mask; subsets with $|R| < 3$ are discarded, and for the rest the value $S_R(a) = \sum_{r \in R} \operatorname{ind}(a - r) \bmod n$ is computed for every $a \notin R$ and tested for constancy. This is $O(q^2)$ table lookups per subset. The largest instance is $q = 19$, with five admissible exponents and $2^{19} = 524{,}288$ subsets each. By Lemma 2.2 the count of pairs (n, f) is obtained from the count of pairs (n, R) by weighting each cell of size $d < q$ by $(q - 1)/n$ and the cell $d = q$ by $q - 1$. No step other than the certified field arithmetic is used, and the enumeration is complete, so the arrays are the full answer for these fields and not a sample. The identification of the cell $(5, 5)$ over $\mathbb{F}_{16}$ with the source's family is Observation 7.3 below. □

Proposition 5.3. *The enumeration is selective, and its top cells are degenerate.*

(1) *Over $\mathbb{F}_{16}$ at $n = 3$ the count at $d = 13$ is zero, and at $d = 8$ only 270 of the $\binom{16}{8} = 12{,}870$ eight-element subsets are admissible.*
(2) *For every field and every admissible n the counts at $d = q$ and $d = q - 1$ are 1 and q: these subsets are admissible because S_R has at most one value to be constant on.*

Consequently the informative content of the arrays lies strictly below $d = q - 1$.

Proof. Item (1) is read off the arrays of Theorem 5.2 together with the binomial coefficient; item (2) is read off the arrays for $(q, n) \in \{(16, 3), (13, 4), (9, 8)\}$ and holds for the stated reason in general. □

6. Beyond the formal development: six further fields

The computations in this section were produced by a separate implementation in C and are *not* part of the formal development; we state them as computations, not as theorems, and no result elsewhere in this paper depends on them. The same enumeration as in §5 — all 2^q subsets, all $n \geq 2$ dividing $q - 1$, degree at least 3 — was run for the six further fields $q = 23, 25, 27, 29, 31, 32$, at a measured cost of 24 minutes of a single low-priority thread and 2.1 MB of memory. The resulting totals are:

q	23	25	27	29	31	32
pairs (n, R)	26,476	124,662	107,988	388,924	752,780	3939
pairs (n, f)	288,555	1,275,528	1,370,337	5,364,415	10,760,062	3969

The same C program reproduces every cell of Theorems 5.1 and 5.2, and an independent implementation in Python reproduces every cell for $q \leq 16$; the three implementations agree wherever they overlap. Together with §5 this covers fifteen fields. It is not an interval: of the prime powers at most 32, the fields $q = 2, 3$ and $q = 17$ were not run. The field $q = 32$ is the ceiling of subset enumeration: $q = 37$ would require 2^{37} subsets for the exponent $n = 2$ alone, which we priced at one and a half to three CPU-hours for that single field, and $q \geq 41$ is out of reach of this method entirely. The next fields worth having, $q = 49, 64, 81$, need a structural enumeration — over $\mathbb{F}_p$-subspaces and multiplicative cosets, with exhaustiveness proved — rather than a larger machine.

7. The structure of the answer

Three structures account for the arrays of §5. We describe them, then turn to the degree bounds.

A degenerate tail. By Proposition 5.3(2), every R with $d = q$ or $d = q - 1$ is admissible for every n. A third degenerate layer sits just below: $R = \mathbb{F}_q \setminus \{a_1, a_2\}$ is admissible exactly when $-1 \in (\mathbb{F}_q^*)^n$, because the derivative of $x^q - x$ is -1 and hence $f(a_1)/f(a_2) = -1$. In characteristic two this is automatic, which is why the cell $(3, 14)$ over $\mathbb{F}_{16}$ has the full count $\binom{16}{2} = 120$. These are genuine

solutions of Problem 1.1 as literally posed, but the source's $N = n(q-d-\delta)$ is then 0, n or $2n$, so nothing of coding-theoretic interest survives. The content of the problem is at small d.

The linearized family. This is the classical construction of van der Geer and van der Vlugt [2, §§1, 3]. Let $V \subseteq \mathbb{F}_q$ be an $\mathbb{F}_p$-subspace and let $L_V = \prod_{v\in V}(x-v)$, the additive polynomial of V; take $R = V + t$ and $f = c\,L_V(x-t)$. Since L_V is $\mathbb{F}_p$-linear, the values of f off R form the nonzero part of the image $L_V(\mathbb{F}_q)$, up to the constant c. For the natural choices this is decisive.

Proposition 7.1. *Let $j \mid m$ with $j < m$, let $V = \ker \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{p^j}}$ and $t \in \mathbb{F}_q$. Then $L_V = \mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{p^j}}$, $d = |V| = q/p^j$, and for $n \geq 2$ dividing $q-1$ with $d \geq 3$ the pair $(n, L_V(x-t))$ satisfies* (H1′), (H4) *and $\delta = 0$ if and only if*

$$n \;\Big|\; \frac{q-1}{p^j-1}.$$

Proof. The monic additive polynomial with kernel exactly V is unique of degree $|V|$, and $\mathrm{Tr}_{\mathbb{F}_q/\mathbb{F}_{p^j}}$ is monic of degree q/p^j with kernel V; so L_V is that trace, and it is separable and split with root set $V+t$ after the translation. Its image is $\mathbb{F}_{p^j}$, so the values off the roots are exactly $\mathbb{F}_{p^j}^*$. As 1 is among them, they lie in a single coset of $(\mathbb{F}_q^*)^n$ if and only if $\mathbb{F}_{p^j}^* \subseteq (\mathbb{F}_q^*)^n$; in the cyclic group $\mathbb{F}_q^*$ this says that the subgroup of order p^j-1 lies in the subgroup of order $(q-1)/n$, i.e. $(p^j-1) \mid (q-1)/n$. □

Proposition 7.1 is elementary and is proved here on paper only; it is not part of the formal development. Its predictions match the arrays cell by cell. Over $\mathbb{F}_{16}$ with $j=1$ it predicts admissibility at $d=8$ for every $n \mid 15$, and the number of affine $\mathbb{F}_2$-hyperplanes is $15\cdot 2 = 30$: exactly the counts in the cells $(5,8)$ and $(15,8)$ of Theorem 5.2. With $j=2$ it needs $n \mid 15/3 = 5$, and indeed $n=5$ is the only exponent with a nonzero cell at $d=4$, of size $20 = 5\cdot 4$ (5 lines over $\mathbb{F}_4$, 4 translates each). Over $\mathbb{F}_{27}$ with $j=1$ it needs $n \mid 26/2 = 13$, and the C computation of §6 finds the cell $(13,9)$ nonempty with $39 = 13\cdot 3$ entries and no other exponent nonempty at $d=9$; over $\mathbb{F}_{25}$ with $j=1$ it needs $n \mid 24/4 = 6$, and the cells $(2,5),(3,5),(6,5)$ each have 30 entries while $(4,5),(8,5),(12,5),(24,5)$ are empty.

Note that $j \mid m$ and $j < m$ force $j \leq m/2$, so $d = q/p^j \geq \sqrt{q}$, with equality exactly at $j = m/2$. The extreme case $j = m/2$ is where the construction gives the Hermitian quotients — the case in which [2, (3.4)–(3.5)] finds a maximal curve and writes down the Hermitian equation $y^{\sqrt{q}+1} = a(z^{\sqrt{q}} - z)$ — and $j=1$ in characteristic two is Theorem 4.1. Cases with $1 < j < m/2$ occur first at $q=64$, $j=2$; they are covered by neither our theorems nor our enumeration, and we do not claim anything about them.

Remark 7.2. The cell $(3,8)$ over $\mathbb{F}_{16}$ has 270 entries, of which only 30 are affine $\mathbb{F}_2$-hyperplanes. So the split locus is strictly larger than the linearized construction even at the degree where the linearized construction lives, and the description of the whole locus by explicit families remains open.

The Hermitian family. Finally the family the source already knew, which the enumeration locates exactly.

Observation 7.3. Over $q = q_0^2$ the source's family [1, Thm. 7.10] consists, up to the substitutions $x \mapsto \lambda x + t$ which preserve the split locus, of the root sets $R = \lambda\mu_{q_0+1} + t$ with $\lambda \in \mathbb{F}_q^*$ and $t \in \mathbb{F}_q$, for $n \mid q_0+1$; these are $q(q-1)/(q_0+1)$ distinct sets of size $d = q_0+1$. Over $\mathbb{F}_{16}$ that number is 48, which is exactly the count in the cell $(5,5)$ of Theorem 5.2; hence that cell consists precisely of the source's curves. Over $\mathbb{F}_9$ the number is 18, exactly the count in the cells $(2,4)$ and $(4,4)$ of Theorem 5.1.

The sets $\lambda\mu_{q_0+1} + t$ are distinct: the elements of μ_{q_0+1} sum to 0, so $\sum_{r\in R} r = (q_0+1)t$, and $q_0+1 \equiv 1 \pmod p$, so t is recovered from R and then the coset $\lambda\mu_{q_0+1}$ is too. Over $\mathbb{F}_{16}$ there are 16 translates and $15/5 = 3$ multiplicative cosets, giving 48. Since each such R is admissible and the cell has exactly 48 elements, the cell is exhausted by them. So, over $\mathbb{F}_{16}$, the previously known part of the split locus is one cell out of the 2101 pairs found — less than a fortieth of the total.

Two degree bounds, and where the minimum sits. The source records the pigeonhole bound $d \geq qn/(q+n-1)$ [1, after Prob. 7.11]. Weil's character-sum bound gives a second one, sharper for small n.

Proposition 7.4. *If (n, f) lies in the split locus over $\mathbb{F}_q$ then $d \geq \sqrt{q}$.*

Proof. Let χ be a multiplicative character of $\mathbb{F}_q$ of order n, extended by $\chi(0) = 0$, and write $f = c\,g$ with $g = \prod_{r \in R}(x - r)$ monic. Since g has $d \geq 3$ distinct roots it is not the n-th power of a polynomial, so [3, Thm. 5.41], applied to g with the scalar c, gives $\left|\sum_{a \in \mathbb{F}_q} \chi(c\,g(a))\right| \leq (d-1)\sqrt{q}$, the number of distinct roots of g being d. On the other hand $\delta = 0$ means $\chi(f(a)) = 1$ for each of the $q - d$ elements $a \notin R$, while $\chi(f(a)) = 0$ for the d roots, so the sum equals $q - d$. Hence $q - d \leq (d-1)\sqrt{q}$, i.e. $d \geq (q + \sqrt{q})/(1 + \sqrt{q}) = \sqrt{q}$. □

Proposition 7.4 is a routine consequence of a classical bound and is proved here on paper only; it is not part of the formal development. Both bounds were checked against the minimal admissible degree in each of the 52 cells (q, n) that §5 and §6 determine, with no violation. The single cell over $\mathbb{F}_{32}$, namely $n = 31$, has minimal admissible degree 16, meeting the source's bound $qn/(q+n-1) = 16$ exactly.

Over a square field $q = q_0^2$ the bound $d \geq \sqrt{q} = q_0$ can be attained, and by Proposition 7.1 the linearized member with $j = m/2$ attains it exactly when $n \mid (q-1)/(q_0 - 1) = q_0 + 1$. The enumeration says this is the whole story for the square fields it reaches:

q	exponents n with $\min d = \sqrt{q}$	other exponents, with their $\min d$
9	$2, 4$ (the divisors ≥ 2 of $q_0 + 1 = 4$)	8: $\min d = 6$
16	5 (the divisors ≥ 2 of $q_0 + 1 = 5$)	$3, 15$: $\min d = 8$
25	$2, 3, 6$ (the divisors ≥ 2 of $q_0 + 1 = 6$)	$4, 8, 12, 24$: $\min d = 15, 19, 15, 20$

(the $q = 25$ row rests on the computation of §6, the other two on Theorems 5.1 and 5.2). The field $\mathbb{F}_4$ is the exception that proves the rule: there $q_0 = 2$, so $d = \sqrt{q} = 2$ is excluded by the hypothesis $d \geq 3$ of (H1′), and the minimum is 3.

At $d = \sqrt{q} = q_0$ the curve is maximal, by a two-line computation. Here $\gcd(n, d) = 1$ automatically, because $n \mid q - 1$ while $d = q_0$ is a power of p, so (4) applies and gives

$$\#X(\mathbb{F}_q) = q_0 + n(q - q_0) + 1 = q + (n-1)(q - q_0) + 1, \qquad q + 1 + 2g_X\sqrt{q} = q + 1 + (n-1)(q_0 - 1)q_0,$$

and $(q_0 - 1)q_0 = q - q_0$, so the two agree: the point count meets the Hasse–Weil bound. Combining this with the table above and with Theorems 3.3 and 4.1: within the range determined here, the answer to Problem 1.1 is *yes at the minimal degree $d = \sqrt{q}$, and no above it.* We regard the positive half as the part of the source's question that deserves a proof; the statement to aim at is that a member of the split locus with $d = \sqrt{q}$ is necessarily linearized over $\mathbb{F}_{\sqrt{q}}$, and hence a maximal Hermitian quotient.

8. Verification

Every statement labelled Theorem or Proposition in §§3–5 has been formally verified in Lean 4 against *Mathlib* [5, 6]; it contains no `sorry` and no hand-written `axiom`, and it is organised in four files: the $\mathbb{F}_{16}$ witness with its controls, the uniform family, the enumeration, and the two further field models. The axiom audit is the sharpest thing we can say about §3: all fourteen theorems of the $\mathbb{F}_{16}$ file — the field model of Proposition 3.1, the hypotheses, the identification $f = \mathrm{Tr}$, the value set, $\delta = 0$, the point count 33, the genus 7, the two refutations of Theorem 3.3, the maximal control of Proposition 3.5 and the three controls of Proposition 3.6 — report `does not depend on any axioms` under `#print axioms`, the empty axiom set. (This required re-implementing bitwise exclusive-or by structural recursion on binary digits, since Mathlib's `Nat.xor` is defined by well-founded recursion and does not reduce in the kernel.) The uniform inequalities of Theorem 4.1 and the non-vacuity $3 \mid 2^{2e} - 1$ depend only on `propext`, `Classical.choice` and `Quot.sound`: they use no finite data and no compiled evaluation, so the family is established for all $e \geq 2$ and $n \geq 3$

at once, subject to the finite-field input discussed in Remark 4.2. What is delegated to compiled evaluation, through Lean's `native_decide`, is exactly the two exhaustive searches that are too large for the kernel: the enumeration of §5 (all 2^q subsets for each of the nine fields named there, together with the certification of each field model on all q^3 triples; 214 seconds of elaboration) and the two extra field models of Propositions 4.4 and 4.5. Every cell of Theorems 5.1 and 5.2 was produced independently twice, by the Lean development and by a C program, and every cell with $q \leq 16$ a third time by a Python program; the implementations agree wherever their ranges overlap. The cells of §6 come from the C program alone. The computations of §6, and Propositions 7.1 and 7.4, are outside the formal development, as stated where they appear.

References

[1] J. Mezinaj and T. Shaska, *Hulls, linear equivalence, and weighted superelliptic codes*, arXiv:2608.02850v1 (3 August 2026), §7 *Split ramification and maximal curves*. Live at v1, with no journal reference, when checked on 28 August 2026.

[2] G. van der Geer and M. van der Vlugt, *Kummer covers with many points*, Finite Fields Appl. **6** (2000), no. 4, 327–341; arXiv:math/9909037.

[3] R. Lidl and H. Niederreiter, *Finite Fields*, 2nd edition, Encyclopedia of Mathematics and its Applications 20, Cambridge University Press, 1997. The character-sum bound used in Proposition 7.4 is Theorem 5.41 (p. 225).

[4] H. Stichtenoth, *Algebraic Function Fields and Codes*, 2nd edition, Graduate Texts in Mathematics 254, Springer, 2009. The Kummer-extension data quoted in our Section 2 is Proposition 3.7.3(b), for the ramification indices, and Corollary 3.7.4, for the genus.

[5] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28 (A. Platzer and G. Sutcliffe, eds.), Lecture Notes in Computer Science 12699, Springer, 2021, pp. 625–635.

[6] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, pp. 367–381.

[7] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, `https://oeis.org` (searches of 28 August 2026, no match).