ABELIAN VARIETIES OVER $\mathbb{F}_2$ OF PRESCRIBED ORDER

KIRAN S. KEDLAYA

Abstract. We prove that for every positive integer m, there exist infinitely many simple abelian varieties over $\mathbb{F}_2$ of order m. The method is constructive, building on the work of Madan–Pal in the case $m = 1$ to produce an explicit sequence of Weil polynomials giving rise to abelian varieties over $\mathbb{F}_2$ of order m. This sequence itself depends on the choice of a suitable generalized binary representation of m; by making careful choices of this representation, we can ensure that the the resulting sequence of polynomials have 2-adic Newton polygons which guarantee the existence of suitable irreducible factors.

1. Introduction

How can a given positive integer m occur as the order of the group of rational points of an abelian variety A over $\mathbb{F}_q$ (or for short, the *order of* A)? While this is nominally a question in arithmetic algebraic geometry, it immediately translates into a pure matter of algebraic number theory. To wit, Weil's theorems on the zeta function of A (e.g., see [13]) imply that $\#A(\mathbb{F}_q) = P(1)$ where $P(x)$ is the characteristic polynomial of Frobenius on A. The polynomial $P(x)$ is monic of degree $2g$ where $g = \dim(A)$ and its complex roots can be labeled $\alpha_1, \dots, \alpha_{2g}$ so that

$$|\alpha_i| = \sqrt{q}, \qquad \alpha_{g+i} = \overline{\alpha}_i \quad (i = 1, \dots, g).$$

Moreover, the Honda–Tate theorem asserts that any polynomial $P(x)$ satisfying these conditions (plus a mild additional hypothesis, which is automatic if $q = p$ is prime) occurs for some abelian variety. Understanding our original question thus becomes a matter of studying the space of *Weil polynomials*; some important foundational work on this issue was done by DiPippo–Howe [2].

While it may seem at this point that our original question is mostly resolved by prior work, it should be emphasized that we did *not* ask about a specific value of g, and this has a profound effect on the nature of the question. For example, we may read off from Weil's results the bounds

$$(\sqrt{q} - 1)^{2g} \leq \#A(\mathbb{F}_q) \leq (\sqrt{q} + 1)^{2g};$$

if we distinguish these intervals based on g, then as q increases they become more and more separated, to the extent that $\#A(\mathbb{F}_q)$ eventually determines g uniquely. By contrast, if we fix q, then as g increases the intervals eventually start to overlap, so specifying $\#A(\mathbb{F}_q)$ does not fix g at all.

Keeping this in mind, let us now narrow our original question and ask: for a given prime power q, which integers occur *at least once* as the order of an abelian variety A over $\mathbb{F}_q$? Howe–Kedlaya [6] showed that *every* positive integer occurs as the order of an abelian variety over $\mathbb{F}_2$, which can further be taken to be ordinary. Building on this, van Bommel–Costa–Li–Poonen–Smith [1] showed that for any fixed q, every *sufficiently large* positive integer occurs as the order of an abelian variety over $\mathbb{F}_q$. This abelian variety can be further taken to be ordinary, geometrically simple, and/or principally polarizable, and for each combination of conditions one can in principle establish an effective "sufficiently large" cutoff; for example, for $q > 4$, every integer $m \geq q^{3\sqrt{q}\log q}$ occurs as the order of an ordinary abelian variety over $\mathbb{F}_q$ [1, Theorem 1.13(b)], and the lower bound on m is best possible up to replacing 3 with a smaller constant [1, Remark 1.15].

Another natural question to ask is, for a fixed $\mathbb{F}_q$, *how often* a given order can occur. For $q > 2$, a result of Kadets [7] implies that for all but finitely many simple abelian varieties A over $\mathbb{F}_q$,

$$\#A(\mathbb{F}_q) \geq 1.359^{\dim(A)};$$

Date: February 7, 2024.

Thanks to Francesc Fité for discussions that led to this work as well as that of [6], and to the anonymous referee for valuable feedback on the readability of the SageMath code. Financial support was provided by NSF (grants DMS-1802161, DMS-2053473) and UC San Diego (Warschawski Professorship).

in particular, there are only finitely many simple abelian varieties over $\mathbb{F}_q$ of any given order (regardless of dimension). It is natural to try to count these, but we do not address this here.

Instead, we focus on the case $q = 2$ and prove the following theorem.

Theorem 1.1. *For every positive integer m, there exist* infinitely many *simple abelian varieties over $\mathbb{F}_2$ (of various dimensions) of order m.*

One key motivation for Theorem 1.1 is the fact that it holds for $m = 1$ by an old theorem of Madan–Pal [9]. That result gives a complete classification (up to isogeny) of simple abelian varieties over $\mathbb{F}_2$ of order 1 using work of Robinson on algebraic integers with all conjugates in a short real interval [15, 16]. The case $m = 2$ resolves a question of Kadets [7, §1]; the general question was raised in [6, §1].

The first step towards Theorem 1.1 is to produce some sequences of Weil polynomials giving rise to abelian varieties over $\mathbb{F}_2$ of order m (without apparent common simple factors). This builds upon the work of Madan–Pal, and also uses some careful choices of generalized binary representations of m as in [6] and [1], including the *nonadjacent binary representations* of Reitwiesner [14]. One convenient feature of the construction is that each sequence we produce satisfies a second-order linear recurrence (closely linked to the recurrence relation satisfied by Chebyshev polynomials); this implies that every irreducible factor shared by more than one term of the sequence corresponds to an abelian variety of order 1 (Lemma 5.3). Consequently, if infinitely many terms of our sequence have irreducible factors of bounded codegree, then all but finitely many of the corresponding simple abelian varieties have order m (Lemma 5.4). This observation by itself is enough to establish Theorem 1.1 for m prime (Lemma 6.6), and thus to answer the question of Kadets.

To finish the proof of Theorem 1.1, we establish this irreducibility using 2-adic calculations (mostly Newton polygons). For m even, we get by with a slight variant of the nonadjacent binary representation (Lemma 7.9); for m odd, we need a representation of a more restricted form, which we construct using a short computer calculation (Lemma 8.7) in SageMath [17]. We include the relevant SageMath code as an appendix; it is also available as a Jupyter notebook from the author's web site.

We conclude this introduction with some discussion of related questions that we do not treat.

- It is not clear whether our approach can be upgraded to ensure that the simple abelian varieties we obtain are ordinary, geometrically simple, or principally polarizable; the ordinary condition in particular is incompatible with our use of 2-adic methods. For order greater than 1, it is possible that a suitable adaptation of [1, Construction 9.1] can be used for this purpose. Such a construction may also shed some light on the *number* of isogeny classes of simple abelian varieties over $\mathbb{F}_2$ of fixed dimension and order, which is known for dimensions up to 6 by the exhaustive tables in LMFDB; see [5] for more on this data and its tabulation.
- For simple abelian varieties of order 1, one cannot hope to enforce the ordinary, geometrically simple, and principally polarizable conditions simultaneously because the Madan–Pal classification demonstrates a "rigidity" of these abelian varieties. In fact, it can be shown that *no* simple abelian variety of order 1 is both ordinary and geometrically simple; see [4].
- Marseglia–Springer [11] consider the question of finding abelian varieties realizing specific *groups* of rational points; this is a problem of a somewhat different nature because the group of rational points of an abelian variety over a finite field is *not* an isogeny invariant (whereas its order is). Using the result of [6] and [1], Marseglia–Springer show that every finite abelian group occurs as the group of rational points of some ordinary abelian variety over $\mathbb{F}_2$, $\mathbb{F}_3$, and $\mathbb{F}_5$ (and a slightly weaker analogue over $\mathbb{F}_4$). Using Theorem 1.1, Marseglia–Springer show that any fixed finite abelian group occurs as the group of rational points of infinitely many pairwise coprime abelian varieties.
- In contrast with Theorem 1.1, for any positive integer m there are are only finitely many isomorphism classes of curves whose Jacobians have order m; this is even true if we vary over all finite fields (modulo the trivial exception of curves of genus 0 in the case $m = 1$). It would be interesting to identify these curves for some small values of m. For example, it is known from work of Madan–Queen [10], Stirpe [19], Mercuri–Stirpe [12], and Shen–Shi [18] that there are eight isomorphism classes of curves of positive genus whose Jacobians have order 1: one curve of genus 1 over each of $\mathbb{F}_3$ and $\mathbb{F}_4$, plus six more curves of genera at most 4 over $\mathbb{F}_2$.

2. Setup

We first introduce the setup used by Madan–Pal to study abelian varieties over $\mathbb{F}_2$ with small order, building on work of Robinson [15]. Throughout this paper, we consider the interval

$$[a, b] := [3 - 2\sqrt{2}, 3 + 2\sqrt{2}].$$

Lemma 2.1. *Let $P(x) \in \mathbb{Z}[x]$ be an irreducible monic polynomial with all roots in $[a, b]$ and set*

$$Q(x) := (-1)^{\deg P(x)} P(3 - x), \qquad R(x) := x^{\deg P(x)} Q(x + 2x^{-1}).$$

Then $R(x)$ occurs as the characteristic polynomial of Frobenius of some simple abelian variety A over $\mathbb{F}_2$ with $\#A(\mathbb{F}_2) = (-1)^{\deg P(x)} P(0)$.

Proof. Put $m = (-1)^{\deg P(x)} P(0)$. The conditions on P imply that $Q(x) \in \mathbb{Z}[x]$ is a monic irreducible polynomial with all roots in the interval $[-2\sqrt{2}, 2\sqrt{2}]$ with $Q(3) = m$, and then that $R(x) \in \mathbb{Z}[x]$ is a monic irreducible polynomial with all roots on the circle $|x| = \sqrt{2}$. By the Honda–Tate theorem [20], [21], $R(x)$ occurs as the characteristic polynomial of Frobenius of some simple abelian variety A over $\mathbb{F}_2$ (there being no Brauer obstruction because we are working over a prime field and we avoid the exceptional case $R(x) = x^2 - 2$); for any such A, we have $\#A(\mathbb{F}_2) = R(1) = Q(3) = m$. □

For n a positive integer, let $T_n(x) \in \mathbb{Z}[x]$ be the n-th Chebyshev polynomial of the first kind for the "arithmetic" normalization (i.e., the Dickson polynomials of the first kind with parameter 1):

$$T_n(2\cos\theta) = 2\cos n\theta.$$

For $n \geq 0$, define the polynomial $f_n(x)$ of degree $2n$ by the formula

$$f_n(x) := x^n T_n(x + x^{-1} - 4).$$

Since $x \mapsto x + x^{-1} - 4$ maps $[a, b]$ two-to-one onto $[-2, 2]$, $f_n(x)$ has all roots in $[a, b]$. In the ring

$$R := \frac{\mathbb{Z}[x^{\pm 1}, y^{\pm 1}, (x-1)^{-1}]}{(x + x^{-1} - 4 - y - y^{-1})} \cong \mathbb{Z}[x^{\pm 1}, (x-1)^{-1}, \sqrt{x^2 - 6x + 1}], \tag{2.2}$$

we have

$$f_n(x) = x^n T_n(y + y^{-1}) = x^n (y^n + y^{-n}). \tag{2.3}$$

We finally introduce a key modification that will give rise to abelian varieties of prescribed orders greater than 1. For $n, k \geq 0$, define the rational function

$$g_{n,k}(x) := (x-1)^{-k} \sum_{j=0}^{k} \binom{k}{j} f_{n+j}(x),$$

so that $g_{n,0}(x) = f_n(x)$. In the ring R, we have

$$g_{n,k}(x) = (xy)^n \left(\frac{xy+1}{x-1}\right)^k + (xy^{-1})^n \left(\frac{xy^{-1}+1}{x-1}\right)^k. \tag{2.4}$$

We will see later that $g_{n,k}(x)$ is a polynomial of degree $2n+k$ (Lemma 3.11) with constant term $(-2)^k$ (3.12) having all roots in $[a, b]$ (Lemma 4.1).

3. Recurrence relations and algebraic corollaries

We next introduce some recurrence relations satisfied by $f_n(x)$ and $g_{n,k}(x)$, and use these to derive some additional algebraic properties, notably that $g_{n,k}(x)$ is indeed a polynomial (Lemma 3.11). Many spot verifications of these properties can also be found in the associated Jupyter notebook.

To begin with, recall that the Chebyshev polynomials are characterized by the recurrence relation and initial conditions:

$$T_n(x) - xT_{n-1}(x) + T_{n-2}(x) = 0, \quad T_0(x) = 2, \quad T_1(x) = x. \tag{3.1}$$

This translates into a corresponding recurrence relation and initial conditions for $f_n(x)$:

$$f_n(x) - (x^2 - 4x + 1) f_{n-1}(x) + x^2 f_{n-2}(x) = 0, \quad f_0(x) = 2, \quad f_1(x) = x^2 - 4x + 1. \tag{3.2}$$

From (3.2), it is easy to deduce by induction that

$$f_n(x) \equiv x^{2n} + 1 \pmod{x} \tag{3.3}$$

$$f_n(x) \equiv (-1)^n 2 \pmod{x-1} \tag{3.4}$$

$$f_n(x) \equiv x^{2n} + 4n(x^{2n-1} + x^{2n-3} + \cdots + x) + 1 \pmod{8}. \tag{3.5}$$

(In (3.3), the term x^{2n} is only relevant when $n = 0$.) The recurrence relation (3.2) for f_n translates into the recurrence relation

$$g_{n,k}(x) - (x^2 - 4x + 1)g_{n-1,k}(x) + x^2 g_{n-2,k}(x) = 0. \tag{3.6}$$

We can also formulate recurrence relations for $g_{n,k}$ in which k varies. To begin with, for $k \geq 1$,

$$(x-1)g_{n,k}(x) = g_{n,k-1}(x) + g_{n+1,k-1}(x). \tag{3.7}$$

We can also avoid division by $x - 1$ at the expense of lengthening the recurrence in the k-aspect.

Lemma 3.8. *For $k \geq 2$,*

$$g_{n,k}(x) - (x-3)g_{n,k-1}(x) + 2g_{n,k-2}(x) = 0, \tag{3.9}$$

$$g_{n,k}(x) + 4g_{n,k-1}(x) + 4g_{n,k-2}(x) - x^2 g_{n-1,k}(x) = 0. \tag{3.10}$$

Proof. The equalities can be seen to hold for $k = 2$ by expanding $g_{n,k}(x)$ in terms of $f_{n+j}(x)$ and applying (3.2), and then for $k > 2$ by induction using (3.7). □

We are now ready to establish that $g_{n,k}(x)$ is in fact a polynomial.

Lemma 3.11. *For $n, k \geq 0$, $g_{n,k}(x)$ is a polynomial of degree $2n + k$.*

Proof. We have $g_{n,k}(x) \in \mathbb{Z}[x]$ for $k = 0$ because $g_{n,0}(x) = f_n(x)$, and for $k = 1$ by (3.4) and (3.7). By (3.9) we deduce that $g_{n,k}(x) \in \mathbb{Z}[x]$ for $n, k \geq 0$. The degree assertion then follows from the fact that $\deg f_{n+j}(x) = 2n + 2j$. □

Using (3.7), we may formally promote (3.3): for $n > 0$,

$$g_{n,k}(x) \equiv (-2)^k \pmod{x}. \tag{3.12}$$

We may also promote (3.4) as follows.

Lemma 3.13. *For $n, k \geq 0$,*

$$g_{n,k}(x) \equiv (-1)^{n-k}((1+i)^k + (1-i)^k) \pmod{x-1}. \tag{3.14}$$

Proof. For $k = 0$ this is a restatement of (3.4). For $k = 1$, we may check the claim for $n = 0, 1$ from the values

$$f_0(x) = 2, \quad f_1(x) = x^2 - 4x + 1, \quad f_2(x) = x^4 - 8x^3 + 16x^2 - 8x + 1$$

and then for general n by (3.6). We may then extend to general k using (3.9). □

We next consider analogues of (3.5) for $g_{n,k}$ for $k > 0$. We start with a mod 2 congruence: from (3.5) and (3.7),

$$g_{n,k}(x) \equiv x^{2n}(x+1)^k + 2^k \pmod{2}. \tag{3.15}$$

We can also establish congruences modulo a higher power of 2 provided that we ignore some leading coefficients.

Lemma 3.16. *For $n \geq 0$,*

$$g_{n,k}(x) \equiv 0 \pmod{(x^{2n}, 2^k)}. \tag{3.17}$$

Proof. The claim holds for $n = 0$ and $k = 0$ vacuously, and for $k = 1$ by (3.15). We may then deduce the general case by (3.10). □

Finally, from (3.5) and (3.7) we obtain some congruences modulo higher powers of 2 relative to k:

$$g_{n,1}(x) \equiv \sum_{i=0}^{2n-1} (-1)^{\lfloor (i-1)/2 \rfloor} 2x^i \pmod{(x^{2n}, 8)} \tag{3.18}$$

$$g_{n,2}(x) \equiv \sum_{i=0}^{n-1} 4x^{2i} \pmod{(x^{2n}, 8)}. \tag{3.19}$$

4. Counting roots

To count zeros of polynomials in the interval $[a, b]$, we use an approach based on *winding numbers*.

Lemma 4.1. *Let $a_0, \dots, a_k$ be a sequence of real numbers with $a_k = 1$, such that the polynomial $Q(z) := \sum_{i=0}^k a_i z^i$ has all of its complex roots inside the closed disc $|z| \leq \sqrt{2}$ (e.g., by condition (4.4) below). Then for each $n \geq 0$, the roots of the polynomial*

$$P_n(x) := \sum_{i=0}^k a_i g_{n,i}(x)$$

are all real and contained in $[a, b]$. If in fact $Q(z)$ has all of its complex roots inside the open disc $|z| < \sqrt{2}$, then the roots of $P_n(x)$ are pairwise distinct.

Proof. By continuity (of the roots of a polynomial as a function of the coefficients), we may reduce to the case where $Q(z)$ has all of its complex roots in the open disc $|z| < \sqrt{2}$. For $\theta \in [-2\pi, 2\pi]$, we define a parametric complex solution of the equation

$$x + x^{-1} - 4 = y + y^{-1}$$

by setting $y(\theta) := e^{2\pi i \theta}$ and

$$x(\theta) := \cos\theta + 2 + \sqrt{\cos^2\theta + 4\cos\theta + 3},$$

choosing the branch of the square root so that $x(\theta)$ varies continuously and

$$x(-2\pi) = b, x(0) = a, x(2\pi) = b.$$

Define the function

$$s(\theta) := \begin{cases} \frac{x(\theta)y(\theta)+1}{x(\theta)-1} & (\theta \neq \pm\pi) \\ -1 - i & (\theta = -\pi) \\ -1 + i & (\theta = \pi); \end{cases}$$

one may check using L'Hôpital's rule that this function is continuous. By writing

$$|s(\theta)|^2 = \frac{(x(\theta)y(\theta)+1)(x(\theta)y(-\theta)+1)}{(x(\theta)-1)^2} = \frac{x(\theta)^2 + x(\theta)(x(\theta) + x(\theta)^{-1} - 4) + 1}{x(\theta)^2 - 2x(\theta) + 1} = 2, \tag{4.2}$$

we deduce that s carries $[-2\pi, 2\pi]$ into the circle $|z| = \sqrt{2}$ (making one full counterclockwise circuit).

By (2.4),

$$P_n(x(\theta)) = 2x(\theta)^n \mathrm{Real}\left(y(\theta)^n \sum_{i=0}^k a_i s(\theta)^i\right). \tag{4.3}$$

Since $x(\theta)$ is monotone, the zeros of $P_n(x)$ in the interval $[a, b]$ (counted without multiplicity) are in bijection with zeros of $P_n(x(\theta))$ in either of the intervals $[-2\pi, 0]$ or $[0, 2\pi]$. We will estimate the number of zeros of $P_n(x(\theta))$ by computing the displacement of

$$\arg\left(y(\theta)^n \sum_{i=0}^k a_i s(\theta)^i\right) = n \arg y(\theta) + k \arg s(\theta) + \arg\left(\sum_{i=0}^k a_i s(\theta)^{i-k}\right)$$

over the interval $[-2\pi, 2\pi]$ (choosing all of the arguments to vary continuously in θ).

As θ varies from -2π to 2π, the displacement of $n \arg y(\theta) + k \arg s(\theta)$ equals $(4n+2k)\pi$. Meanwhile, we may see that $\arg\left(\sum_{i=0}^{k} a_i s(\theta)^{i-k}\right)$ has displacement 0 by combining (4.2), our condition on the roots of $Q(z)$, and the argument principle.

Since $\arg\left(y(\theta)^n \sum_{i=0}^{k} a_i s(\theta)^i\right)$ varies continuously from 0 to $(4n+2k)\pi$ as θ runs from -2π to 2π, by the intermediate value theorem it evaluates to an odd multiple of π at no fewer than $4n+2k$ distinct values in this range. By (4.3), these values are zeros of $P_n(x(\theta))$ in $[a,b]$, each counted at most twice. Since $P_n(x)$ is a polynomial of degree $2n+k$, we deduce that all of its zeros are pairwise distinct real numbers in $[a,b]$. □

Note that in Lemma 4.1, one way to enforce the condition on $Q(z)$ is to assume

$$\sum_{i=0}^{k-1} |a_i| 2^{(i-k)/2} \leq 1, \tag{4.4}$$

as then the triangle inequality implies that $|z^{-k}Q(z)| > 0$ for $|z| > \sqrt{2}$ (compare [6, Lemma 2]). This restricted setting will be enough to prove Theorem 1.1 for m even (Lemma 7.9), but we will need to exercise more flexibility for m odd (Lemma 8.7).

5. Repeated zeros in a recurrent sequence

Note that for any fixed sequence $\{a_i\}$, the sequence of polynomials $P_n(x)$ considered in Lemma 4.1 satisfies the same second-order recurrence as the ones satisfied by $f_n(x)$ (3.2) and $g_{n,k}(x)$ (3.6). Using this, we can show that the polynomials $P_n(x)$ have very few common zeros.

Lemma 5.1. *Let $\{P_n(x)\}_{n\geq 0}$ be a sequence of monic integer polynomials satisfying the recurrence relation*

$$P_n(x) - (x^2 - 4x + 1)P_{n-1}(x) + x^2 P_{n-2}(x) = 0. \tag{5.2}$$

Suppose that $\alpha \in \mathbb{C}^\times$ is a root of both $P_n(x)$ and $P_{n'}(x)$ for some $n < n'$. Then α is a unit in the ring of algebraic integers.

Proof. In the ring R from (2.2), we can solve the recurrence (5.2) to obtain an analogue of (2.3): for some $P_+, P_- \in R$ (independent of n),

$$P_n = P_+(xy)^n + P_-(xy^{-1})^n.$$

Define a specialization homomorphism $\pi : R \to \mathbb{C}$ taking x to α by picking a square root of $\alpha + \alpha^{-1} - 4$; then solving the system of equations

$$\pi(P_+)\pi(xy)^n + \pi(P_-)\pi(xy^{-1})^n = \pi(P_+)\pi(xy)^{n'} + \pi(P_-)\pi(xy^{-1})^{n'} = 0$$

yields

$$\pi(xy)^{n'-n} = \pi(xy^{-1})^{n'-n}$$

and so $\pi(y)^{2(n'-n)} = 1$. By (2.3) this yields $f_{2(n'-n)}(\alpha) = 2\alpha^{2(n'-n)}$; since $\deg f_{2(n'-n)}(x) = 4(n'-n) > 2(n'-n)$ and $f_{2(n'-n)}(0) = 1$ by (3.3), α is a root of the monic polynomial $f_{2(n'-n)}(x) - 2x^{2(n'-n)}$ with constant coefficient 1. □

This has the following implication for Theorem 1.1.

Lemma 5.3. *Let $m > 1$ be an integer and fix a sequence $a_0, \dots, a_k$ of integers satisfying the hypotheses of Lemma 4.1. Suppose that for infinitely many n, the polynomial $P_n(x) = \sum_i a_i g_{n,i}(x)$ over $\mathbb{Q}$ has an irreducible factor $Q(x)$ with $P(0) = \pm m$. Then there exist infinitely many simple abelian varieties A over $\mathbb{F}_2$ with $\#A(\mathbb{F}_2) = m$.*

Proof. By Lemma 4.1, the polynomial $P_n(x)$ has all of its roots in $[a,b]$, as then does $Q(x)$. By Lemma 5.1, the factors $Q(x)$ are pairwise distinct. We may thus apply Lemma 2.1 to conclude. □

We also need a slightly modified version of Lemma 5.3.

Lemma 5.4. *Let $m > 1$ be an integer and fix a sequence $a_0, \dots, a_k$ of integers satisfying the hypotheses of Lemma 4.1, and additionally satisfying $\sum_{i=0}^k a_i 2^i \in \{\pm m\}$. Suppose that for infinitely many n, the polynomial $P_n(x) = \sum_i a_i g_{n,i}(x)$ over $\mathbb{Q}$ has a monic irreducible factor $Q(x)$ whose codegree (i.e., $\deg P_n(x) - \deg Q(x)$) is bounded by a function of m alone. Then there exist infinitely many simple abelian varieties A over $\mathbb{F}_2$ with $\#A(\mathbb{F}_2) = m$.*

Proof. By hypothesis, we can write $P_n(x) = Q(x)R(x)$ where $\deg R(x)$ is bounded by a function of m alone. By Lemma 4.1, $P_n(x)$ has all roots in $[a, b]$, as then do $Q(x)$ and $R(x)$. Since $R(x)$ has integer coefficients and roots in a fixed interval, $R(x)$ itself is contained in a finite set determined by m. By Lemma 5.1, there are only finitely many values of n for which $R(x)$ has constant term not in $\{\pm 1\}$. For the remaining values, $Q(x)$ is a monic irreducible polynomial with $Q(0) = \pm m$. We may thus apply Lemma 5.3 to conclude. □

6. Nonadjacent binary representations

In order to apply Lemma 4.1, we need to find ways to represent a given positive integer m as the evaluation at $z = 2$ of a monic integer polynomial $Q(z)$ having all complex roots in the disc $|z| \leq \sqrt{2}$. That is, we need a *binary representation* of m which is "efficient" in a suitable sense.

One good candidate is the *nonadjacent binary representation* of m in the sense of Reitwiesner [14]:

$$m = \sum_{i=0}^{\infty} a_i 2^i \qquad \text{where} \qquad a_i \in \{-1, 0, 1\},\ a_k = 1,\ a_i a_{i+1} = 0 \quad (i \geq 0). \tag{6.1}$$

The sequence $a_0, \dots, a_k$ can be generated efficiently from m using the rule

$$a_0 = \begin{cases} \pm 1 & m \equiv \pm 1 \pmod 4 \\ 0 & m \equiv 0 \pmod 2. \end{cases}$$

Moreover, the largest index k with $a_k \neq 0$ (and hence $a_k = 1$) is $k(m) = \lfloor \log_2(3m) \rfloor - 1$.

Define the polynomial

$$h_{n,m}(x) := \sum_{i=0}^{k} (-1)^{i+k} a_i g_{n,i}(x);$$

by Lemma 3.11, $h_{n,m}(x)$ is a monic polynomial of degree $2n + k$. By (3.12),

$$h_{n,m}(0) = (-1)^k m. \tag{6.2}$$

Since we chose $a_0, \dots, a_k$ without reference to n, we deduce from (3.6) that

$$h_{n,k}(x) - (x^2 - 4x + 1) h_{n-1,k}(x) + x^2 h_{n-2,k}(x) = 0. \tag{6.3}$$

Lemma 6.4. *The roots of the polynomial $h_{n,m}(x)$ are all real, pairwise distinct, and contained in the interval $[a, b] = [3 - 2\sqrt{2}, 3 + 2\sqrt{2}]$.*

Proof. From the definition of nonadjacent binary representations, we see that

$$\sum_{i=0}^{k-1} |a_i| 2^{(i-k)/2} \leq 1 + 2^{-1} + \cdots + 2^{-\lfloor k/2 \rfloor} \leq 1 - 2^{-k/2} < 1. \tag{6.5}$$

We may thus apply Lemma 4.1. □

In passing, we can already derive some cases of Theorem 1.1, including the case $m = 2$ considered in [7]. While this case is logically necessary for the rest of the proof, it does illustrate the key ideas with limited technical complications compared to the general case.

Lemma 6.6. *Theorem 1.1 holds when m is prime.*

Proof. This is immediate from Lemma 5.3 and Lemma 6.4: if m is prime, then $h_{n,m}(x)$ admits a unique irreducible factor with constant coefficient $\pm m$. □

7. 2-adic congruences: even order case

In this section, we prove Theorem 1.1 for m even, using factorizations over the 2-adic field $\mathbb{Q}_2$. Let $v_2(m)$ denote the 2-adic valuation of m. By convention, our Newton polygons are convex with left endpoint $(0,0)$.

As a warmup, we treat the case where $v_2(m) = 1$.

Lemma 7.1. *Theorem 1.1 holds when $m \equiv 2 \pmod 4$.*

Proof. From (3.15),

$$h_{n,m}(x) \equiv x^{2n} \sum_{i=0}^{k(m)} a_i (x+1)^i + m \pmod 2. \tag{7.2}$$

By (6.2) and (7.2), the 2-adic Newton polygon of $h_{n,m}(x)$ has vertices

$$(0,0), (k(m) - d, 0), (2n + k(m), 1)$$

for some $d \in \{0, \dots, k(m)\}$. The last segment corresponds to an irreducible factor of $h_{n,m}(x)$ over $\mathbb{Q}_2$; hence over $\mathbb{Q}$, $h_{n,m}(x)$ has an irreducible factor of codegree bounded by a function of m. We may thus combine Lemma 5.4 and Lemma 6.4 to conclude. □

We next generalize the Newton polygon calculation from the previous argument.

Lemma 7.3. *For m even and $n \gg 0$, the 2-adic Newton polygon of $h_{n,m}(x)$ has vertices*

$$(0,0), (k(m) - d, 0), (2n + k(m), v_2(m))$$

where d is the order of vanishing of $\sum_{i=0}^{k} a_i (x+1)^k$ at $x = 0$ over $\mathbb{F}_2$.

Proof. From (3.17), we have

$$h_{n,m}(x) \equiv 0 \pmod{(x^{2n}, 2^{v_2(m)})}. \tag{7.4}$$

By combining (7.2) with (7.4), we deduce the claim. □

This gives us a direct adaptation of Lemma 7.1 when $v_2(m)$ is odd.

Lemma 7.5. *Theorem 1.1 holds when $v_2(m)$ is odd.*

Proof. Define d as in Lemma 7.3. Since $v_2(m)$ is odd, by restricting n to a suitable arithmetic progression we can ensure that $\gcd(v_2(m), 2n + d) = 1$; then the final segment of the 2-adic Newton polygon of $h_{n,m}(x)$ corresponds to an irreducible factor of $h_{n,m}(x)$ over $\mathbb{Q}_2$. For such n, $h_{n,m}(x)$ has an irreducible factor over $\mathbb{Q}$ of codegree bounded by a function of m; we may thus combine Lemma 5.4 and Lemma 6.4 to conclude. □

To handle the case where $v_2(m)$ is even, it is convenient to separate off the case $v_2(m) = 2$, which we can handle in a similar manner.

Lemma 7.6. *Theorem 1.1 holds when $v_2(m) = 2$.*

Proof. Define d as in Lemma 7.3. The final segment of the 2-adic Newton polygon of $h_{n,m}(x)$ corresponds to either an irreducible factor of $h_{n,m}(x)$ over $\mathbb{Q}_2$ or a pair of irreducible factors, each of degree $1/(n + d/2)$. In the latter case (which only occurs if d is even), the coefficient of $x^{n+d/2}$ must be congruent to 0 or 4 modulo 8 according to whether $(-1)^k m$ is congruent to -4 or 4 modulo 16; since (3.17) and (3.19) together imply

$$h_{n,m}(x) \equiv 4(x^{2n} + \cdots + x^2 + 1) \pmod{(x^{2n}, 8)}, \tag{7.7}$$

this case can be ruled out by fixing the parity of n appropriately. For such n, $h_{n,m}(x)$ has an irreducible factor over $\mathbb{Q}$ of codegree bounded by a function of m. We may thus combine Lemma 5.4 and Lemma 6.4 to conclude. □

To handle higher values of $v_2(m)$, we modify the polynomial $h_{n,m}(x)$ so that we can better emulate the case $v_2(m) = 1$.

Lemma 7.8. *Theorem 1.1 holds when $v_2(m) \geq 4$.*

Proof. Since $a_0 = a_1 = a_2 = a_3 = 0$, the sequence

$$(a'_0, \dots, a'_k) = (2, 1, 0, 0, a_4, \dots, a_k)$$

satisfies

$$\sum_{i=0}^{k} |a'_i| 2^{(i-k)/2} \leq 2 \cdot 2^{-k/2} + 2^{(1-k)/2} + 1 - 2^{(4-k)/2} < 1;$$

hence the polynomial

$$h'_{n,m}(x) = h_{n,m}(x) + (-1)^k (2g_{n,0}(x) + g_{n,1}(x))$$

satisfies the hypothesis of Lemma 4.1. We may again compute its 2-adic Newton polygon using (3.15), (3.18), (7.2), and (7.4): its vertices are

$$(0,0), (k-d,0), (2n + k(m) - 1, 1), (2n + k(m), v_2(m))$$

where d is the order of vanishing of $x + 1 + \sum_{i=0}^{k} a_i (x+1)^k$ at $x = 0$ over $\mathbb{F}_2$. Over $\mathbb{Q}_2$, the middle segment corresponds to a single irreducible factor of $h_{n,m}(x)$; we may thus argue as in the proof of Lemma 7.5 to conclude. □

To summarize, by combining Lemma 7.5, Lemma 7.6, and Lemma 7.8, we deduce the following.

Lemma 7.9. *Theorem 1.1 holds when m is even.*

8. 2-adic congruences: odd order case

In this section, we prove Theorem 1.1 for m odd. For this, we cannot use the 2-adic Newton polygon of $h_{n,m}(x)$ because it has all slopes equal to 0; instead, we use the 2-adic Newton polygon of $h_{n,m}(x+1)$. To begin with, note that for m odd, by (3.14) we have

$$h_{n,m}(1) \equiv 2 \pmod 4; \tag{8.1}$$

more precisely, we are using here the fact that a_0 is odd, a_1 is even, and $g_{n,k}(1) \equiv 0 \pmod 4$ for $k \geq 2$.

To illustrate the method, we first prove some isolated cases of Theorem 1.1.

Lemma 8.2. *Theorem 1.1 holds for any odd m such that $k(m)$ is even and*

$$\sum_{i=0}^{k(m)} a_i x^i \equiv (x+1)^{k(m)} \pmod 2.$$

For example, this holds for $m = 15, 45, 51, 75, 77, 85$.

Proof. For $n = 2^j - k(m)/2$, we have from (7.2) that

$$h_{n,m}(x) \equiv (x+1)^{2^{j+1}} \pmod 2.$$

By Lemma 3.13 and (8.1), the Newton polygon of $h_{n,m}(x+1)$ has vertices

$$(0,0), (2^{j+1}, 1);$$

that is, $h_{n,m}(x+1)$ satisfies the Schönemann–Eisenstein irreducibility criterion at 2. We may thus combine Lemma 2.1 and Lemma 6.4 to conclude. □

To cover the remaining values of m, we use a variant construction that preserves the key features of this method. We say that a monic integer polynomial $Q(z)$ is a *compliant representation* of the odd positive integer m if $Q(2) = m$, $Q(z) \equiv (z-1)^{\deg Q(z)} \pmod 2$, and $Q(z)$ has all complex roots in the disc $|z| < \sqrt{2}$.

Lemma 8.3. *Theorem 1.1 holds for m admitting a compliant representation.*

Proof. Let $Q(z)$ be a compliant representation of m; by multiplying by $z-1$ as needed, we may ensure that $k := \deg Q(z)$ is even. Write $Q(z) = \sum_{i=0}^{k} c_i z^i$. For each n, the polynomial $P_n(x) = \sum_i (-1)^{i+k} c_i g_{n,i}(x)$ satisfies $P_n(0) = (-1)^k m$. By Lemma 4.1, $P_n(x)$ has all roots in $[a, b]$. By (3.14) (as in the proof of (8.1)), $P_n(1) \equiv 2 \pmod 4$. For $n = 2^j - k/2$, we see from the proof of Lemma 8.2 that $P_n(x)$ is Eisenstein at 2 and hence irreducible. We may thus directly apply Lemma 2.1 to conclude. □

In order to produce compliant representations, it will be convenient to further quantify the condition on the roots. To this end, for $Q(z)$ a compliant representation of some integer m, define the *quality* of $Q(z)$ as

$$\operatorname{qual}(Q(z)) := \min\{|Q(z)| : |z| = \sqrt{2}\} = \min\{|Q(z)| : |z| \geq \sqrt{2}\}. \tag{8.4}$$

Keep in mind that the last equality in (8.4) is a consequence of the maximum modulus principle, and is only valid under the assumption that $Q(z)$ is compliant.

Lemma 8.5. *Let m be a positive odd integer.*

(a) *If $m \leq 3094$, then m admits a compliant representation.*

(b) *If $3094 \leq m \leq 50000$, then m admits a compliant representation of quality at least 7.*

Proof. We describe a computer-assisted proof; the associated computations run in SageMath (version 9.6) in under 5 minutes on a standard laptop (we used one core on an Intel iCore i5-6200U @2.30GHz). As the SageMath code is quite short, we have included it in its entirety in the appendix.

We first observe that SageMath provides an exact representation of the subfield $\overline{\mathbb{Q}}$ of $\mathbb{C}$ based on interval arithmetic. Using this, given a monic integer polynomial $Q(z)$, we may compute the roots of $Q(z)$ in $\overline{\mathbb{Q}}$ and then test rigorously whether they all lie in the disc $|z| < \sqrt{2}$. If so, we may then rigorously compute $\operatorname{qual}(Q(z)) \in \overline{\mathbb{Q}}$ as follows. Since $Q(z)Q(\overline{z})$ is a symmetric integer polynomial in z and $\overline{z}$, we may rewrite it as an integer polynomial in $z + \overline{z}$ and $z\overline{z}$; specializing these to t and 2, respectively, yields an integer polynomial $R(t)$ such that

$$\operatorname{qual}(Q(z))^2 = \min\{R(t) : t \in [-2\sqrt{2}, 2\sqrt{2}]\}. \tag{8.6}$$

The minimum is achieved either at $\pm 2\sqrt{2}$ or at some zero of $R'(t)$ in $[-2\sqrt{2}, 2\sqrt{2}]$.

We now describe the main computation. We first run an exhaust over monic polynomials of degree at most 7 with all coefficients in $\{-3, \dots, 3\}$, checking whether each polynomial is compliant. (While it would be feasible to perform an exhaustive search for compliant polynomials of degree up to 7 by adapting the search strategy for Weil polynomials described in [8], we did not need to implement this here.) In this way we find compliant representations of 167 distinct integers in the range $\{1, \dots, 459\}$; for each of these we record the maximum observed quality, rounded down to the nearest multiple of $1/7$.

We then compute, for each odd integer $m \in \{1, \dots, 50000\}$, a lower bound on the maximum quality of a compliant representation of m using the following logic. Given odd integers $m_1 < m$, let m_2 be one of the nearest odd integers to m/m_1 and set $c = m - m_1 m_2$. Let $Q_1(z), Q_2(z)$ be compliant representations of m_1, m_2 of respective qualities q_1, q_2. If $q_1 q_2 > |c|$, then $R(z) = Q_1(z)Q_2(z) + c$ is a compliant representation of m of quality at least $q_1 q_2 - |c|$.

From the results of this computation, we read off (a) and (b). □

Lemma 8.7. *Every positive odd integer admits a compliant representation. Hence by Lemma 8.5 (or the theorem of Madan–Pal in the case $m = 1$), Theorem 1.1 holds for m odd.*

Proof. By Lemma 8.5(a), it will suffices to check that every odd integer $m \geq 3095$ admits a compliant representation of quality at least 7. We check by induction on m that each of $\{m, m+2, \dots, 15m - 16\}$ admits such a representation, this being true for $m = 3095$ by Lemma 8.5(b) because $15 \cdot 3095 - 16 < 50000$.

Given the claim for some m, let $Q(z)$ be a compliant representation of m of quality at least 7. For c even with $|c| \leq 14$,

$$R(z) = (z^4 - 1)Q(z) + c$$

is a compliant representation of $15m + c$ of quality at least

$$\operatorname{qual}(z^4 - 1)\operatorname{qual}(Q(z)) - |c| \geq 3 \cdot 7 - 14 \geq 7.$$

It follows that each of $\{m + 2, m + 4, \dots, 15m + 14 = 15(m + 2) - 16\}$ admits a compliant representation of quality at least 7; that is, the induction hypothesis holds with m replaced by $m + 2$, as desired. □

Appendix A. SageMath code for Lemma 8.5

This code uses the following features of Python and SageMath:

- `all` returns `True` iff all of its inputs evaluate to `True` in a boolean context.
- Python indexing starts from 0 rather than 1, so `range(n)` returns $0, \ldots, n-1$ and `range(1,n)` returns $1, \ldots, n-1$.
- `AA` and `QQbar` are predefined in SageMath as the fields of algebraic real and complex numbers, respectively. Computations in these fields is rigorous, not subject to roundoff errors.
- For `f` a polynomial over a field, `f.roots(K)` computes its roots in the field `K` (defaulting to the base field of `f` if `K` is omitted). The output consists of pairs (α, m) where α is a root and m is the multiplicity of the root.

```
import itertools
R.<z> = QQ[] # Univariate polynomial ring

# Check that the polynomial f has all complex roots
# in the disc |z| < sqrt(2).
def all_roots_in_disc(f):
    return all(abs(i)^2 < 2 for (i,_) in f.roots(QQbar))

# Compute the quality of f (see (8.4)), multiplied by 7
# and rounded down to the nearest integer.
def quality_lower_bound(f):
    P.<x,y,t> = QQ[]
    I = P.ideal(x+y-t, x*y-2)
    # Compute a representative of f(x)*f(y) modulo I.
    # This corresponds to the polynomial R(t) appearing
    # in the proof of Lemma 8.5.
    g1 = I.reduce(f(x) * f(y))
    # The polynomial g1 is currently a univariate polynomial in t,
    # but in the ring P.
    # We next create g by substituting t -> z to land in the ring R.
    g = R(g1(0, 0, z))
    # Make the list of roots of this polynomial,
    # together with +/- 2*sqrt(2).
    rootlist = (g.derivative()*(z^2 - 8)).roots(AA)
    # Implement (8.7).
    ans = min((g(i) * 49).floor() for (i,_) in rootlist if i^2 <= 8)
    return floor(sqrt(ans))

# Create a table of compliant representations of small integers.
compliant_reps = {}
rep_quality = {}
for n in range(1, 8): # step through n = 1, ..., 7
    # Iterate over n-tuples in which the i-th term
    # (starting wih i=0) runs over -3,-1,1,3 if (n choose i) is odd
    # and -2,0,2 otherwise.
    for t in itertools.product(*((range(-3,4,2) if binomial(n,i)%2
                          else range(-2,3,2)) for i in range(n))):
        # Convert t to a polynomial, after appending 1
        # for the leading coefficient.
        u = R(list(t) + [1])
        if all_roots_in_disc(u):
            m = u(2)
```

```
            q = quality_lower_bound(u)
            if m not in rep_quality or rep_quality[m] < q:
                compliant_reps[m] = u
                rep_quality[m] = q

# Compute lower bounds of qualities of compliant representations of
# larger integers. To save time, rather than optimizing fully,
# we quit as soon as we find a representation of quality at least 8.
n = 50000
for m in range(1, n, 2): # step by 2
    for m1 in range(3, ceil(sqrt(m)), 2): # step by 2
        if m1 in rep_quality:
            # Let m2 be one of the nearest odd integers to m/m1.
            tmp = (QQ(m)/m1+1) / 2
            for m2 in [tmp.floor()*2-1, tmp.ceil()*2-1]:
                if m2 < m and m2 in rep_quality:
                    c = m - m1*m2
                    q = (rep_quality[m1]*rep_quality[m2])//7 - abs(c)
                    if m not in rep_quality or rep_quality[m] < q:
                        compliant_reps[m] = compliant_reps[m1] * \
                            compliant_reps[m2] + c
                        if q < 56:
                            q = quality_lower_bound(compliant_reps[m])
                        rep_quality[m] = q
            if m in rep_quality and rep_quality[m] >= 56:
                break

# Running these commands without errors confirms
# the conclusions of Lemma 8.5.
assert all(i in rep_quality for i in range(1, n, 2))
assert all(rep_quality[i] >= 49 for i in range(3095, n, 2))
```

References

[1] R. van Bommel, E. Costa, W. Li, B. Poonen, and A. Smith, Abelian varieties of prescribed order over finite fields, arXiv:2106.13651v1 (2021).

[2] S.A. DiPippo and E.W. Howe, Real polynomials with all roots on the unit circle and abelian varieties over finite fields, *J. Number Theory* **73** (1998), no. 2, 426–450; corrigendum in [3]. (MR1657992)

[3] S.A. DiPippo and E.W. Howe, Corrigendum: "Real polynomials with all roots on the unit circle and abelian varieties over finite fields", *J. Number Theory* **83** (2000), no. 1, 182. (MR1767658)

[4] T. D'Nelly-Warady and K.S. Kedlaya, Geometric decomposition of abelian varieties of order 1, arXiv:2109.03986v2 (2022); to appear in *Women in Numbers 5*.

[5] T. Dupuy, K.S. Kedlaya, D. Roe, and C. Vincent, Isogeny classes of abelian varieties over finite fields in the LMFDB, in *Arithmetic Geometry, Number Theory, and Computation*, Simons Symposia, Springer, 2022, 375–448. (MR4427971)

[6] E.W. Howe and K.S. Kedlaya, Every positive integer is the order of an ordinary abelian variety over $\mathbb{F}_2$, *Research in Number Theory* **7** (2021). (MR4309841)

[7] B. Kadets, Estimates for the number of rational points on simple abelian varieties over finite fields, *Math. Z.* **297** (2021), 465–473. (MR4204701)

[8] K.S. Kedlaya, Search techniques for root-unitary polynomials, in K.E. Lauter and K.A. Ribet (eds.), *Computational Arithmetic Geometry*, Contemporary Math. 463, Amer. Math. Soc., 2008, 71–82.

[9] M.L. Madan and S. Pal, Abelian varieties and a conjecture of R. M. Robinson, *J. reine angew. Math.* **291** (1977), 78–91. (MR0439848)

[10] M.L. Madan and C.S. Queen, Algebraic function fields of class number one, *Acta Arith.* **20** (1972), 423–432. (MR0306160)

[11] S. Marseglia and C. Springer, Every finite abelian group is the group of rational points of an ordinary abelian variety over $\mathbb{F}_2$, $\mathbb{F}_3$, and $\mathbb{F}_5$, *Proc. Amer. Math. Soc.* **151** (2023), 501–510.

[12] P. Mercuri and C. Stirpe, Classification of algebraic function fields with class number one, *J. Number Theory* **154** (2015), 365–374. (MR3339577)

[13] J.S. Milne, The Riemann hypothesis over finite fields, from Weil to the present day, in *The Legacy of Bernhard Riemann after One Hundred and Fifty Years. Vol. II*, Int. Press, Somerville, MA, 2016, 487–565; reprinted in *ICCM Not.* **4** (2016), 14–52.

[14] G.W. Reitwiesner, Binary arithmetic, *Advances in Computers* **1** (1960), 231–308. (MR0122018)

[15] R.M. Robinson, Intervals containing infinitely many sets of conjugate algebraic units, *Annals of Math.* **80** (1964), 411–428. (MR0175881)

[16] R.M. Robinson, Conjugate algebraic units in a special interval, *Math. Z.* **154** (1977), 31–40. (MR0447175)

[17] The Sage Developers, SageMath version 9.6, 2022, `https://www.sagemath.org`.

[18] Q. Shen and S. Shi, Function fields of class number one, *J. Number Theory* **154** (2015), 375–379. (MR3339578)

[19] C. Stirpe, A counterexample to "Algebraic function fields with small class number", *J. Number Theory* **143** (2014), 402–404. (MR3227356)

[20] J. Tate, Classes d'isogénie des variétés abéliennes sur un corps fini (d'après T. Honda), Séminaire Bourbaki. Vol. 1968/69: Exposés 347–363, Lecture Notes in Math., vol. 175, Springer, Berlin, 1971, pp. 95–110, Exp. No. 352. (MR3077121)

[21] W.C. Waterhouse and J.S. Milne, Abelian varieties over finite fields, 1969 Number Theory Institute (Proc. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969), pp. 53–64. Amer. Math. Soc., Providence, R.I., 1971. (MR0314847)

Department of Mathematics, University of California San Diego, La Jolla, CA 92093, United States of America
Email address: `kedlaya@ucsd.edu`
URL: `https://kskedlaya.org`