

On counting and generating curves over small finite fields

Qi Cheng^a, Ming-Deh Huang^b

^a*School of Computer Science
University of Oklahoma, Norman, OK 73019, USA*

^b*Computer Science Department
University of Southern California
Los Angeles, CA 90089, USA*

Abstract

We consider curves defined over small finite fields with points of large prime order over an extension field. Such curves are often referred to as Koblitz curves and are of considerable cryptographic interest. An interesting question is whether such curves are easy to construct as the target point order grows asymptotically. We show that under certain number theoretic conjecture, if q is a prime power, r is a prime and $\sqrt{q} > (r \log q)^{2+\epsilon}$, then there are at least $\Omega(\frac{q}{r^{1+\epsilon} \log^2 q})$ non-isomorphic elliptic curves E/\mathbf{F}_q , such that the quotient group $E(\mathbf{F}_{q^r})/E(\mathbf{F}_q)$ has prime order. We also show that under the same conjecture, if q is a prime power and r is a prime satisfying $q > (r \log q)^{2+\epsilon}$ and $\sqrt{q} = o(\frac{q}{r^{1+\epsilon} \log q})$, then there are at least $\Omega(\frac{q}{r^{1+\epsilon} \log q})$ curves H/\mathbf{F}_q of genus 2, such that the order of the quotient group $\mathbf{Jac}(H)(\mathbf{F}_{q^r})/\mathbf{Jac}(H)(\mathbf{F}_q)$ is a prime. Based on these results we present simple and efficient algorithms for generating $\Omega(\log^3 n)$ non-isomorphic elliptic curves in $\Omega(\log n)$ isogenous classes, each with a point of prime order $\Theta(n)$. The average time to generate one curve is $O(\log^2 n)$. We also present an algorithm which generates $\Omega(\log^3 n)$ curves of genus two with Jacobians whose orders have a prime factor of order $\Theta(n)$, in heuristic expected time $O(\log^4 n)$ per curve.

Keyword: Curve-based cryptography, Koblitz curve, Bateman-Horn conjecture.

Email addresses: qcheng@cs.ou.edu (Qi Cheng), huang@cs.usc.edu (Ming-Deh Huang).

1 Introduction

Elliptic curve and other curve-based cryptosystems require the construction of curves over finite fields with points of large prime order on the Jacobians of the curves. This problem can be formulated as follows: Given a natural number n , to construct a curve over some finite field and a point on the Jacobian of the curve whose order is prime and close to n .

In the case of elliptic curves, the Jacobian of an elliptic curve is the curve itself. A random elliptic curve will have prime order with roughly the same probability as finding a prime in the range between $p + 1 - 2\sqrt{p}$ and $p + 1 + 2\sqrt{p}$ if the curve is defined over \mathbf{F}_p . Counting the order of the group of rational points on an elliptic curve has been made easier due to recent improvements on Schoof's method [21,17]. Hence a reasonable approach in the case of elliptic curves, is to find a prime p around n , then randomly choose elliptic curves E over \mathbf{F}_p until $\#E(\mathbf{F}_p)$ is a prime, and find a nontrivial rational point on the curve. However, extending the approach to hyperelliptic curves is difficult since no practically efficient method is known for counting points on hyperelliptic curves and their Jacobians when p is large (say greater than 10^{25}), despite some recent progress on the problem [9,8]. An alternative approach which avoids point counting, is to apply the heuristic method of Atkin-Morain [3] and Elkies involving the CM theory of elliptic curves. Generalization of this method to hyperelliptic curves is possible but not practical [5,23].

Dramatic progress has been made in point counting when the characteristic of the field is very small. Satoh first proposed a very efficient algorithm based on the canonical p -adic lift of the elliptic curve when the characteristic of the field, p , is small but greater than 5. His idea was extended to $p = 2, 3$ subsequently [15]. In [16], experimental results were reported. It was concluded in that paper that "it is no longer necessary to use precomputed curves in cryptography since one can easily compute new curves as desired. Finding a curve with a security level comparable with RSA-1024 takes minutes or less. Curve generation for short-term security, with a level equivalent to DES, is feasible on a low-power chip." Recently Kedlaya [10] gave a $O(g^5 r^3)$ counting algorithm for hyperelliptic curve over field \mathbf{F}_{p^r} with genus g when p is fixed. The time complexity of all these algorithms depends polynomially on the characteristic of the field.

The approach considered in this paper is to start with a relatively small finite field \mathbf{F}_q , then look for curves defined over \mathbf{F}_q which, when considered as curves over an extension \mathbf{F}_{q^m} , has rational points of large prime order on their Jacobians. Such curves are sometimes called Koblitz curves [22,12,19]. One simple and natural approach is to fix an elliptic curve over a small field \mathbf{F}_q and then consider it over \mathbf{F}_{q^r} as r varies [11]. This heuristic method seems to work well in the practical range of cryptographic interest. However it does not

work as r grows asymptotically, since the probability that

$$\frac{\#E(\mathbf{F}_{q^r})}{\#E(\mathbf{F}_q)}$$

is prime when r varies is conjectured (in analogy to the the classical Mersenne prime problem) to be about $\frac{e^\gamma \log r}{r \log q}$ (where γ is the Euler's constant), which tends to 0 as r increases and q is fixed.

In this paper, we explore the possibility of determining a relatively small base field \mathbf{F}_q (say $q = (\log n)^{O(1)}$) and some extension \mathbf{F}_{q^k} (say $k = O(\log n / \log \log n)$), so that curves with \mathbf{F}_{q^k} -points of prime order $\Theta(n)$ can be constructed. The following theorems are proven based on a weak version of the Bateman-Horn conjecture concerning the density of primes when evaluating an integral polynomial. These theorems lead to methods which guarantee to generate curves of genus one and two with above-mentioned properties.

Theorem 1 *Assume that the weak Bateman-Horn conjecture is true and a Siegel zero doesn't exist. Let q be a prime power and r be a prime. If $\sqrt{q} > (r \log q)^{2+\epsilon}$, there are at least $\Omega(\frac{q}{r^{1+\epsilon} \log^2 q})$ non-isomorphic elliptic curves E/\mathbf{F}_q in $\Omega(\frac{\sqrt{q}}{r^{1+\epsilon} \log q})$ \mathbf{F}_{q^r} -isogenous classes, such that the order of the quotient group $E(\mathbf{F}_{q^r})/E(\mathbf{F}_q)$ is prime.*

It is known that the Dirichlet L-function $L(s, \chi)$ of a Dirichlet character $\chi \bmod q$ is zero free in $\sigma > 1 - c/\log(q(2 + |t|))$, where $s = \sigma + it$ and c is an absolute constant, with at most one exception. If the exception exists, then χ must be real and the zero is also real. Such a zero is called a Siegel zero [6]. Theorem 1 will be proved in Section 3. It leads to an algorithm which on input n , determines a suitable base field \mathbf{F}_q and extension degree r , where $r = \Theta(\frac{\log n}{(4+\epsilon) \log \log n})$ and $q = \Theta(\log^{4+\epsilon} n)$; then generates $\Omega(\log^3 n)$ non-isomorphic elliptic curves in $\Omega(\log n)$ isogenous classes, and a point on each curve with prime order greater than n and less than $n(1 + O(\frac{r+\sqrt{q}}{q}))$; in time $O(\log^{5+\epsilon} n)$. The average time to generate one curve is $O(\log^2 n)$. Consequently this method is particularly efficient if we want to generate a large collection of good elliptic curves while minimizing the average construction time per curve. We note that, in contrast, it is hard to overcome $\log^4 n$ per curve barrier if we first select a random curve, then do the point-counting and finally test the order for primality, since after all testing primality of a number around n takes $O(\log^3 n)$ time, and such a number is a prime with probability only $\frac{1}{\log n}$. Note that we use the fast arithmetic algorithm in primality testing, but the error probability needs to be kept below $1/n$, hence the time complexity per number is $O(\log^3 n)$. The advantage in looking for Koblitz curves defined over \mathbf{F}_q , with $q = O(\log^4 n)$, is that there are $\Theta(\sqrt{q}) = \Theta(\log^2 n)$ isogenous classes, hence $\Theta(\log^2 n)$ numbers to test for possible orders over the extension field \mathbf{F}_{q^r} . Although $O(\log^3 n)$ non-isomorphic Koblitz curves are generated, primality testing is

performed on only $O(\log^2 n)$ numbers, and this is essentially why the average construction time per good curve can be as low as $O(\log^2 n)$. We refer to Section 4 for more detailed analysis.

We also prove a similar theorem for curves of genus two under the same conjecture.

Theorem 2 *Assume that the weak Bateman-Horn conjecture is true, and a Siegel zero doesn't exist. Let q be a prime power and r be a prime. If $q > (2r \log q)^{2+\epsilon}$ and $\sqrt{q} = o(\frac{q}{r^{1+\epsilon} \log q})$, then there are at least $\Omega(\frac{q}{r^{1+\epsilon} \log q})$ curves H/\mathbf{F}_q of genus 2, such that the order of the quotient group $\mathbf{Jac}(H)(\mathbf{F}_{q^r})/\mathbf{Jac}(H)(\mathbf{F}_q)$ is prime.*

We will prove this theorem in Section 5. It leads to an algorithm which generates $\Omega(\log^3 n)$ curves of genus two with Jacobians whose orders have a prime factor greater than n and less than $n(1 + O(\frac{r+\sqrt{q}}{q}))$, in heuristic expected time $O(\log^4 n)$ per curve.

Setting $q = 103, r = 19$, as many as 400 curves of genus 2 were generated at the average rate of less than five minutes per curve, as we implement the algorithm on a PII 300Mhz computer using GP scripting. All of the group orders are about 240 bits long.

The method developed in this paper can be extended in a natural way to hyperelliptic curves of any fixed genus. However it seems to be difficult to have rigorous analysis of the method when the genus of interest is greater than two.

2 The weak Bateman-Horn conjecture

Gauss observed that the density of primes around x is $\frac{1}{\log x}$. One might predict, as a more precise estimate, that the asymptotic formula is

$$\pi(x+y) - \pi(x) = (1 + o(1)) \frac{y}{\log x}, \quad (1)$$

where $\pi(x)$ is the number of primes less than x and $y \leq x$. The formula is proved for $y > x^\alpha$, α is any constant greater than $7/12$ and is disproved [14] if y is less than any fixed power of $\log x$ in the sense that

$$\limsup_{x \rightarrow \infty} \frac{\pi(x + (\log x)^\lambda) - \pi(x)}{(\log x)^{\lambda-1}} > 1$$

and

$$\liminf_{x \rightarrow \infty} \frac{\pi(x + (\log x)^\lambda) - \pi(x)}{(\log x)^{\lambda-1}} < 1.$$

However, if the conjecture is modified to a weaker form which states that there exists an absolute constant c such that $\pi(x+y) - \pi(x) > c \frac{y}{\log x}$ for $10 \log^2 x < y < x$, no counter

example has been found. Moreover A. Selberg proved that under Riemann Hypothesis, (1) is true for almost all x if $\frac{y}{\log^2 x} \rightarrow \infty$.

More generally, it has been conjectured [4] that the number of prime values assumed by any irreducible polynomial $F(X)$ is given by the formula

$$\pi_F(x) = (C_F + o(1)) \frac{x}{\log |F(x)|},$$

for $x > \log^{2+\epsilon} |F(x)|$. The constant C_F is $\prod_{p \text{ prime}} (1 - \frac{\omega_F(p)}{p}) / (1 - \frac{1}{p})$ where $w_F(p)$ is number of distinct roots $F(x) = 0$ in \mathbf{F}_p . We refer to this conjecture as the Bateman-Horn Conjecture.

There are strong heuristic arguments [4] in support of the conjecture, at least in terms of the order of estimate implied in the conjecture. However the precise estimate predicted in the conjecture can be problematic in some cases. It was recently shown in [7] that for any given degree some polynomials can be constructed to take either significantly more or significantly less prime values than predicted by the conjecture. Such discrepancy seems to disappear if one does not insist on the precise estimate in the conjecture. If for example the conjecture is weakened to the following

$$\pi_F(x) \geq \frac{C_F}{2} \frac{x}{\log |F(x)|},$$

for $x > 10 \log^2 \max_{1 \leq y \leq x} |F(y)|$ and $x > 10 \deg F$, then no counter example has been found. We refer to this conjecture as the weak Bateman-Horn Conjecture. Note that the constant C_F depends on the polynomial F . Hence the probability that $|F(x)|$ becomes a prime at a random integer x can be very different from the probability that a random integer of size around $|F(x)|$ becomes a prime, unless that C_F is very close to a constant.

What we will need is a special case of this weaker statement where the polynomial F splits over a cyclotomic field. In this case, C_F will be shown to be bounded from below by a function on the degree of F . The function grows very slowly with the degree.

Theorem 3 *Let r be a prime. Let ξ be r -th primitive root of unity. Let F be a monic irreducible polynomial with degree $d = \phi(r) = r - 1$ (ϕ is Euler function) and splitting field $\mathbf{Q}(\xi)$. Let $\pi_F(x)$ be the number of integers $n \leq x$ for which $|F(n)|$ is prime. Then under the weak Bateman-Horn Conjecture and the conjecture of non-existence of a Siegel zero, there is an absolute constant C such that*

$$\pi_F(x) > C \frac{x}{\log |F(x)| e^{(\log \log r)^2}},$$

whenever $x > 10 \log^2 \max_{1 \leq y \leq x} |F(y)|$ and $x > 10d$.

Proof: Denote $Disc(F)/Disc(\mathbf{Q}(\xi))$ by β^2 . It is known [2] that

$$\sum_{p \equiv i \pmod r, p \leq x} \frac{d}{p} = \log \log x + A(r, i) + O\left(\frac{1}{\log x}\right)$$

Now we evaluate the constant C_F ,

$$C_F = \prod_p \frac{1 - \frac{\omega_F(p)}{p}}{1 - \frac{1}{p}}.$$

Observe that $\omega_F(p) = d$ if $p \equiv 1 \pmod r, p \nmid \beta$, $\omega_F(p) = 1$ if $p = r$. And $\omega_F(p) = 0$ if $p \not\equiv 1 \pmod r$ and $p \neq r$. Hence

$$C_F = \frac{(1 - \frac{1}{r}) \prod_{p \equiv 1 \pmod r, p \nmid \beta} (1 - \frac{d}{p}) \prod_{p \equiv 1 \pmod r, p \mid \beta} (1 - \frac{\omega_F(p)}{p})}{\prod_p (1 - \frac{1}{p})}$$

We have

$$\log C_F = - \sum_{p \equiv 1 \pmod r, p \nmid \beta} \frac{d}{p} - \sum_{p \equiv 1 \pmod r, p \mid \beta} \frac{\omega_F(p)}{p} + \sum_p \frac{1}{p} + A,$$

where $|A|$ is bounded from above by an absolute constant. Hence

$$\begin{aligned} \log C_F &= - \sum_{p \equiv 1 \pmod r} \frac{d}{p} + \sum_{p \equiv 1 \pmod r, p \mid \beta} \frac{d - \omega_F(p)}{p} + \sum_p \frac{1}{p} + A \\ &\geq - \sum_{p \equiv 1 \pmod r} \frac{d}{p} + \sum_p \frac{1}{p} + A \end{aligned}$$

We also have

$$-A(r, 1) = - \sum_{p \equiv 1 \pmod r} \frac{d}{p} + \sum_p \frac{1}{p} + D,$$

where D is an absolute constant. Hence $\log C_F \geq -A(r, 1) + A - D$. Applying the method in [18], if a Siegel zero doesnot exist, one can have

$$\frac{1}{(\log \log r)^2} \leq A(r, 1) \leq (\log \log r)^2.$$

It implies that $C_F = \Omega\left(\frac{1}{e^{(\log \log r)^2}}\right)$.

3 Special bivariate polynomials associated with elliptic curves

Let E be an elliptic curve defined over \mathbf{F}_q , where q is prime power p^d , with $p \neq 2, 3$. Then E has $q + 1 - a$ points over \mathbf{F}_q where a is the trace of the Frobenius endomorphism of the

curve over \mathbf{F}_q , and $-2\sqrt{q} \leq a \leq 2\sqrt{q}$. Let α be one root of $x^2 - ax + q = 0$. Let $\bar{\alpha}$ denote the complex conjugate of α . Then the order of abelian group $E(\mathbf{F}_{q^r})$, denoted by $\#E(\mathbf{F}_{q^r})$, is $(\alpha^r - 1)(\bar{\alpha}^r - 1)$.

Let Φ_n denote the n -th cyclotomic polynomial, the minimal polynomial of $\xi_n = e^{\frac{2\pi i}{n}}$. Then $x^r - 1 = \prod_{k>0, k|r} \Phi_k(x)$. Therefore

$$\begin{aligned} \#E(\mathbf{F}_{q^r}) &= (\alpha^r - 1)(\bar{\alpha}^r - 1) \\ &= \prod_{k>0, k|r} \Phi_k(\alpha) \prod_{k>0, k|r} \Phi_k(\bar{\alpha}) \\ &= \prod_{k>0, k|r} \Phi_k(\alpha)\Phi_k(\bar{\alpha}), \end{aligned}$$

and $\Phi_k(\alpha)\Phi_k(\bar{\alpha}) = \prod_{\gcd(i,k)=1, 0<i\leq k} (\alpha - \xi_k^i)(\bar{\alpha} - \xi_k^i) = \prod_{\gcd(i,k)=1, 0<i<k} (q - a\xi_k^i + \xi_k^{2i})$.

Denote

$$\prod_{\gcd(i,k)=1, 0<i<k} (x - y\xi_k^i + \xi_k^{2i})$$

by $\Psi_k(x, y)$. The first three Ψ_k 's are: $\Psi_1(p, a) = p + 1 - a$, $\Psi_2(p, a) = p + 1 + a$, and $\Psi_3(p, a) = p^2 + (a - 1)p + (a^2 + a + 1)$.

From the above discussion we see that for an elliptic curve E defined over \mathbf{F}_q of trace a ,

$$\#E(\mathbf{F}_{q^r}) = \prod_{k>0, k|r} \Psi_k(q, a) \tag{2}$$

The polynomial $\Psi_k(x, y)$ possesses several nice properties as shown in the following lemma.

Lemma 1 (1) $\Psi_k(x, y) \in \mathbf{Z}[x, y]$.

(2) If $k > 3$ is prime, then for any integer c , $F_1(x) = \Psi_k(c, x)$ and $F_2(x) = \Psi_k(x, c)$ are irreducible polynomial over \mathbf{Q} , and has a cyclotomic field as its splitting field.

(3) $\Psi_k(x, y)$ is irreducible over \mathbf{Q} .

(4) If r is a prime, then for any $-2\sqrt{q} \leq a \leq 2\sqrt{q}$,

$$q^r - 2q^{r/2} + 1 \leq (q + 1 - a)\Psi_r(q, a) \leq q^r + 2q^{r/2} + 1.$$

Proof: Part (1) follows directly from the definition. As for part (2), given any integer c , $F_1(x) = \Psi_k(c, x)$ is an irreducible polynomial if the only Galois element of $\mathbf{Q}(\xi_k)/\mathbf{Q}$ that fixes $c\xi_k - \xi_k^2$ is the identity. Similarly, $F_2(x) = \Psi_k(x, c)$ is irreducible if the only Galois element of $\mathbf{Q}(\xi_k)/\mathbf{Q}$ that fixes $c\xi_k^{-1} + \xi_k$ is the identity. When $k > 3$ is a prime, the minimum polynomial of ξ_k has degree $k - 1$ and has more than 5 terms. For any Galois element σ of $\mathbf{Q}(\xi_k)/\mathbf{Q}$, $\sigma(c\xi_k - \xi_k^2) - (c\xi_k - \xi_k^2)$ simplifies to a polynomial expression of ξ_k of degree less

than k with at most four terms. Hence if $k > 3$, $F_1(x)$ is irreducible and has cyclotomic fields as its splitting field. By a similar argument one can show that if $k > 3$, $F_2(y)$ is irreducible, Part (3) follows from Part (2). Part (4) follows from the equation 2.

Proof of Theorem 1: The order of the quotient group $E(\mathbf{F}_{q^r})/E(\mathbf{F}_q)$ is $F_1(x) = \Psi_r(q, x)$. The variable x will take value from $-2\sqrt{q}$ to $2\sqrt{q}$. From Lemma 1 we see that as long as $\sqrt{q} > (r \log q)^{2+\epsilon}$, we may apply Theorem 3 to the polynomial $F_1(x)$, and it will evaluate to $\Omega(\frac{\sqrt{q}}{r^{1+\epsilon} \log q})$ number of primes. Hence there are $\Omega(\frac{\sqrt{q}}{r^{1+\epsilon} \log q})$ \mathbf{F}_q -isogenous classes, such that the order of the quotient group $E(\mathbf{F}_{q^r})/E(\mathbf{F}_q)$ is prime. It is proved in [13,20] that there exist two constants c_1, c_2 such that if A is a set of integers between $q+1-\sqrt{q}$ and $q+1+\sqrt{q}$, the number of non-isomorphic classes of elliptic curves defined over \mathbf{F}_q whose number of points over \mathbf{F}_q are in A is

$$c_1\sqrt{q}(|A| - 2)/\log q \leq N \leq c_2\sqrt{q}|A|\log q(\log \log p)^2.$$

Thus there are at least $\Omega(\frac{q}{r^{1+\epsilon} \log^2 q})$ non-isomorphic elliptic curves over \mathbf{F}_q in these isogenous classes.

4 Algorithms for the case of elliptic curves

We are ready to describe an algorithm for constructing an elliptic curve whose order has prime factor bigger than a given number n .

Algorithm 1 Input: n .

Output: Two primes $q, r > 3$, and a set of elliptic curves defined over \mathbf{F}_q . If E is any of the output curves, then the quotient group $E(\mathbf{F}_{q^r})/E(\mathbf{F}_q)$ has a prime order larger than n ;

- (1) Let r be the largest prime less than $\lceil \frac{\log n}{4 \log \log n} \rceil$;
- (2) Let $Q = \lceil n^{\frac{1}{r-1}} \rceil$; Make sure that $\frac{Q^r - 2Q^{r/2} + 1}{Q + 2\sqrt{Q} + 1} \geq n$. If not, increase Q to the least integer satisfying the inequality.
- (3) Search a prime q such that $Q \leq q \leq Q + 10 \log^2 Q$;
- (4) Find a quadratic nonresidue c in \mathbf{F}_q ;
- (5) Compute polynomial $f(x) = \Psi_r(q, x) = \prod_{i>0, \gcd(i,r)=1} (q - x\xi_r^i + \xi_r^{2i}) \in Z[x]$;
- (6) Search for numbers $-2\sqrt{q} \leq a \leq 2\sqrt{q}$, such that $f(a)$ is prime; .
- (7) For all possible j -invariants $j \in \mathbf{F}_q$ of curves over \mathbf{F}_q , compute the number of points of its corresponding curves. Namely for $j = 1728$, check all the curves $y^2 = x^3 + \alpha x$, $\alpha \neq 0$; for $j = 0$, check all the curves $y^2 = x^3 + \beta$,

$\beta \neq 0$; in all the remaining cases, let

$$k = \frac{108j}{j - 1728}$$

check the curves $y^2 = x^3 - kx - 4k$ and $y^2 = x^3 - kc^2x - 4kc^3$. If any of the curves has $q + 1 - a$ points over \mathbf{F}_q for any a chosen in the previous step, output the curve.

Now we elaborate on the steps of the Algorithm 1.

Step 1 to 3 in Algorithms 1 determine a suitable base field \mathbf{F}_q and extension degree r . The output curves will be defined over \mathbf{F}_q . Note that if $r = \frac{\log n}{(4+\epsilon)\log \log n}$, then $Q = \log^{4+\epsilon} n$.

In step 4, we search a quadratic residue in field \mathbf{F}_q . The naive search method is adequate as it takes time $O(\sqrt{q})$, which is $O(\log^{2+\epsilon} n)$.

Step 5 and 6 search for traces of suitable elliptic curves. Note that if an elliptic curve E/\mathbf{F}_q has trace a , then quotient group $E(\mathbf{F}_{q^r})/E(\mathbf{F}_q)$ has order $\Psi_r(q, a)$, since r is a prime. Thus we look for those a where $\Psi_r(q, a)$ is prime as a ranges from $-2\sqrt{q}$ to $2\sqrt{q}$. Theorem 1 implies that in this range $\Psi_r(q, a)$ will evaluate to $\Omega(\frac{\sqrt{q}}{(r-1)^{1+\epsilon}\log q}) = \Omega(\log n)$ primes. We use Rabin-Miller's primality testing algorithm to see whether $\Psi_r(q, a)$ is prime. The primality testing algorithm takes time $O(\log^3 n)$ for each number. The total time complexity is $O(\log^{5+\epsilon} n)$ in these two steps.

Notice that the maximum value for $\Psi_r(q, a)$ will be

$$\begin{aligned} \frac{q^r + 2q^{r/2}}{q - 2\sqrt{q}} &= q^{r-1} \frac{1 + 2/q^{r/2}}{1 - 2/\sqrt{q}} \\ &= q^{r-1} (1 + O(1/\sqrt{q})) \\ &= (Q + \log^2 Q)^{r-1} (1 + O(1/\sqrt{q})) \\ &= Q^{r-1} (1 + \frac{\log^2 Q}{Q})^{r-1} (1 + O(1/\sqrt{q})) \\ &= n (1 + O(\frac{r + \sqrt{q}}{q})) \end{aligned}$$

This illustrates a nice property of the algorithm: it will not find a curve with prime part of the order too far away from n .

Step 7 construct elliptic curve for all possible j -invariant values and output those whose traces a are good in the sense that $f(a)$ is prime. The theory of elliptic curves assures that for every a in the search range, there is at least one elliptic curve over \mathbf{F}_q with trace a . We apply Shanks's Baby-Step-Giant-Step (BSGS) strategy to count points over \mathbf{F}_q for each curve. This takes time $O(q^{1/4+\epsilon}) = O(\log^{1+\epsilon} n)$ for each curve. Hence the total complexity

for the last step is $O(\log^{5+\epsilon} n)$.

A variant of this algorithm is searching for a curve over F_{2^a} (or \mathbf{F}_q where q is small prime power). The Theorem 3 implies that by brute-force searching through elliptic curves over \mathbf{F}_{2^a} of all possible j -invariants, we are guaranteed to find a good curve very efficiently.

Now suppose E is a curve generated by the algorithm, then group $E(\mathbf{F}_{q^r})/E(\mathbf{F}_q)$ has prime order, hence is cyclic. Any point that has coordinates in $\mathbf{F}_{q^r} - \mathbf{F}_q$ must have order containing that big prime. It is easy to generate such a point.

5 The genus two curves

Let H be hyperelliptic curve with genus 2 over \mathbf{F}_q where q is a power of a prime p . If $p > 3$, H may be given as $y^2 = f(x)$, where $f(x) \in \mathbf{F}_q[x]$ is a monic polynomial of degree 5. Let

$$P(X) = x^4 + a_1x^3 + a_2x^2 + qa_1x + q^2 \quad (3)$$

be the characteristic polynomial of the Frobenius endomorphism on $\mathbf{Jac}(H)$. $P(X)$ can be factored over \mathbf{C} as

$$P(X) = (x - \alpha_1)(x - \bar{\alpha}_1)(x - \alpha_2)(x - \bar{\alpha}_2),$$

where $\bar{\alpha}_i$ is the conjugate of α_i . The order of Jacobian group over \mathbf{F}_{q^r} is

$$\begin{aligned} & (1 - \alpha_1^r)(1 - \bar{\alpha}_1^r)(1 - \alpha_2^r)(1 - \bar{\alpha}_2^r) \\ &= \prod_{k>0, k|r} \Phi_k(\alpha_1) \prod_{k>0, k|r} \Phi_k(\bar{\alpha}_1) \prod_{k>0, k|r} \Phi_k(\alpha_2) \prod_{k>0, k|r} \Phi_k(\bar{\alpha}_2) \\ &= \prod_{k>0, k|r} \Phi_k(\alpha_1)\Phi_k(\bar{\alpha}_1)\Phi_k(\alpha_2)\Phi_k(\bar{\alpha}_2) \end{aligned}$$

We factor Φ_k over $\mathbf{Q}(\xi_k)$.

$$\begin{aligned} & \Phi_k(\alpha_1)\Phi_k(\bar{\alpha}_1)\Phi_k(\alpha_2)\Phi_k(\bar{\alpha}_2) \\ &= \prod_{0<i<k, \gcd(i,k)=1} (\xi_k^i - \alpha_1)(\xi_k^i - \bar{\alpha}_1)(\xi_k^i - \alpha_2)(\xi_k^i - \bar{\alpha}_2) \\ &= \prod_{0<i<k, \gcd(i,k)=1} (\xi_k^{4i} + a_1\xi_k^{3i} + a_2\xi_k^{2i} + qa_1\xi_k^i + q^2) \end{aligned}$$

Definition 1 *Define*

$$\Delta_k(x, y, z) = \prod_{i>0, \gcd(i,k)=1} (\xi_k^{4i} + y\xi_k^{3i} + z\xi_k^{2i} + xy\xi_k^i + x^2).$$

Lemma 2 (1) $\Delta_k(x, y, z) \in \mathbf{Z}[x, y, z]$.

(2) If H is a curve defined over \mathbf{F}_q with genus 2 and $x^4 + a_1x^3 + a_2x^2 + qa_1x + q^2$ is the minimal polynomial of its Frobenius endomorphism. Then the order of Jacobian of H over \mathbf{F}_{q^r} is $\#\mathbf{Jac}(H)(\mathbf{F}_{q^r}) = \prod_{k>0, k|r} \Delta_k(q, a_1, a_2)$.

(3) If $r > 8$ is prime, then for any integers c, d , $Z(z) = \Delta_r(c, d, z)$ is an irreducible polynomial over \mathbf{Q} , and its splitting field is cyclotomic.

(4) $\Delta_k(x, y, z)$ is irreducible over \mathbf{Q} .

Proof: Part (1) and (2) are directly from the definition of $\Delta_k(x, y, z)$. The proof of Part (3) is similar to that for Part (2) of Lemma 1, noting that for any integer c, d , $Z(z) = \Delta_k(c, d, z)$ is irreducible polynomial if the only Galois action in $\mathbf{Q}(\xi_k)/\mathbf{Q}$ that fixes $\xi_k^2 + d\xi_k + cd\xi_k^{k-1} + c^2\xi_k^{k-2}$ is the identity. Part (4) follows from part (3).

Given a polynomial P , one can ask whether P is the characteristic polynomial of the Frobenius endomorphism of a genus-2 curve. This question is considerably harder than the similar question in the elliptic curve case. For simplicity we replace a_1, a_2 by $-a, b + 2q$ respectively in (3)

$$P(X) = (X^2 + q)^2 - aX(X^2 + q) + bX^2.$$

In [1, page 54, 59], a partial answer was obtained.

Proposition 1 *If a pair of integers a, b satisfies following conditions:*

- (1) $0 < b < a^2/4 < q$,
- (2) b is not divisible by p , and
- (3) neither of $a^2 - 4b$ nor $(b + 4q)^2 - 4qa^2$ is an integer square.

Then there must exist a genus-2 curve, whose Frobenius endomorphism has minimal polynomial $(X^2 + q)^2 - aX(X^2 + q) + bX^2$.

It is easy to show

Lemma 3 *If q is a prime, a is the least prime less than $2\sqrt{q}$. For all $0 < b < \frac{a^2}{4}$, there are only $O(\sqrt{q})$ number of b 's such that one of $a^2 - 4b$ and $(b + 4q)^2 - 4qa^2$ is an integer square.*

Proof of Theorem 2: If r is a prime, the order of $\mathbf{Jac}(H)(\mathbf{F}_{q^r})/\mathbf{Jac}(H)(\mathbf{F}_q)$ is $\Delta_r(q, -a, b + 2q)$. Fix a , and let b vary from 0 to $a^2/4$. From Lemma 2 we see that if $q > (2r \log q)^{2+\epsilon}$ then we can apply Theorem 3 to the polynomial $\Delta_r(q, -a, x + 2q)$, and there will be

$$\Omega\left(\frac{a^2/4}{2(r-1)^{1+\epsilon} \log q}\right) = \Omega\left(\frac{q}{r^{1+\epsilon} \log q}\right)$$

b 's such that $\Delta_r(q, -a, b + 2q)$ is prime. Among them, there are only $O(\sqrt{q})$ number of b 's such that one of $a^2 - 4b$ and $(b + 4q)^2 - 4qa^2$ is an integer square. Since $\sqrt{q} =$

$o(\frac{q}{r^{1+\epsilon}\log q})$, there exist at least $\Omega(\frac{q}{r^{1+\epsilon}\log q})$ curves H/\mathbf{F}_q of genus 2 such that the order of $\mathbf{Jac}(H)(\mathbf{F}_{q^r})/\mathbf{Jac}(H)(\mathbf{F}_q)$ is prime.

If $q = \log^{4+\epsilon} n$, $r = \frac{\log n}{(8+\epsilon)\log\log n}$, we will get at least $\Omega(\log^3 n)$ a_2 's, such that $\Delta_r(q, -a, a_2)$ are primes, among them, at most $O(\sqrt{q}) = O(\log^{2+\epsilon} n)$ a_2 's make one of $a^2 - 4b$ and $(b + 4q)^2 - 4qa^2$ an integer square. This suggests the following strategy to set r and q .

- (1) Let r be the largest prime less than $\frac{\log n}{8\log\log n}$;
- (2) Let $Q = \lceil n^{\frac{1}{2(r-1)}} \rceil$; Increase Q if necessary to satisfy $\frac{(Q^{r/2}-1)^4}{(Q^{1/2}+1)^4} \geq n$;
- (3) Search for a prime between Q and $Q + 10\log^2 Q$, assign it to q ;

Once q and r is fixed, the algorithm then randomly selects coefficients for a degree-5 monic polynomial $f(x)$. It uses the BSGS method to count number of elements in $H(\mathbf{F}_q)$ and $H(\mathbf{F}_{q^2})$, where H is the hyperelliptic curve defined by $y^2 = f(x)$. Then we calculate $\mathbf{Jac}(H)(\mathbf{F}_{q^r})/\mathbf{Jac}(H)(\mathbf{F}_q)$ and test for the primality. It is very hard to estimate the time complexity rigorously. Heuristically, we get a prime order with probability roughly equal to $\frac{1}{\log n}$. The counting algorithm takes time $O(q^{3/4}) = O(\log^3 n)$. Hence the time complexity to generate one curve is $O(\log^4 n)$.

Although our algorithms assume number theoretic conjectures, they work very well in practice. In fact, let $n = 2^{240}$, as many as 400 curves of genus 2 were generated at the average rate of less than five minutes per curve, as we ran a casual implementation of the algorithm on a PII 300Mhz computer. This algorithm is remarkably faster and generates much more curves than the other high genus curve generating algorithms.

References

- [1] L. M. Adleman and M.A. Huang. *Primality Testing and Abelian Varieties Over Finite Fields*. Lecture Notes in Mathematics. Springer-Verlag, 1992.
- [2] T. M. Apostol. *Introduction to Analytic Number Theory*. Springer-Verlag, 1976.
- [3] A.O.L. Atkin and F. Morain. Elliptic curves and primality proving. *Mathematics of Computation*, 61:29–67, 1993.
- [4] P.T. Bateman and R.A.Horn. Primes represented by irreducible polynomials in one variable. In *Proc. Symp. Pure Math*, pages 119–132, Providence, 1965. AMS press.
- [5] Jinhui Chao, Kazuto Matsuo, Hiroto Kawashiro, and Shigeo Tsujii. Construction of hyperelliptic curves with cm and its application to cryptosystems. In *AsiaCrypto*, volume 1976 of *Lecture Notes in Computer Science*. Springer-Verlag, 2000.
- [6] H. Davenport. *Multiplicative number theory*. Springer-Verlag, 2000.

- [7] John Friedlander and Andrew Granville. Limitation to the equi-distribution of primes iv. *Proc. Roy. Soc. London Ser. A*, 435(1893):197–204, 1991.
- [8] P. Gaudry and R. Harley. Counting points on hyperelliptic curves over finite fields. In *ANTS*, 2000.
- [9] Ming-Deh Huang and D. Ierardi. Counting points on curves over finite fields. *J of Symbolic computation*, 25:1–21, 1998.
- [10] Kiran S. Kedlaya. Counting points on hyperelliptic curves using monsky-washnitzer cohomology. *J. Ramanujan Math. Soc.*, 16(4):323–338, 2001.
- [11] N. Koblitz. *Algebraic Aspects of Cryptography*. Springer-Verlag, 1998.
- [12] David R. Kohel. Rational groups of elliptic curves suitable for cryptography. Preprint, 1999.
- [13] H. W. Lenstra. Factoring integers with elliptic curves. *Annals of Mathematics*, 126:649–673, 1987.
- [14] H. Maier. Primes in short intervals. *Michigan Math. J.*, 32:221–225, 1985.
- [15] M.Fouquet, P.Gaudry, and R.Harley. An extension of Satoh’s algorithm and its implementation. *J. Ramanujan Math.Soc.*, 15:281—318, 2000.
- [16] M.Fouquet, P.Gaudry, and R.Harley. Finding secure curves with the Satoh-FGH algorithm and an early-abort strategy. In *Eurocrypt*, 2001.
- [17] N.D.Elkies. Elliptic and modular curves over finite fields and related computational issues. In D.A. Buell and J.T. Teitelbaum, editors, *Computational Perspectives on Number Theory: Proceedings of a Conference in Honor of A.O.L. Atkin*, pages 21–76. AMS/International Press, 1998.
- [18] Jan-Christoph Puchta. On the class number of p -th cyclotomic field. *Arch Math.*, 74:266–268, 2000.
- [19] Y. Sakai and K. Sakurai. Design of hyperelliptic cryptosystems in small characteristic and a software implementation over f_{2^n} . In *AsiaCrypt*, volume 1514 of *Lecture Notes in Computer Science*, 1998.
- [20] R. Schoof. Nonsingular plane cubic curves over finite fields. *J. Combin. Theory Ser. A*, 46(2):183–211, 1987.
- [21] R. Schoof. Counting points on elliptic curves over finite fields. *Journal of Theorie des Nombres de Bordeaux 7*, pages 219–254, 1995.
- [22] N.P. Smart. Elliptic curve cryptosystems over small fields of odd characteristic. *J. Cryptology*, 1999.
- [23] A. Weng. Constructing hyperelliptic curves of genus 2 suitable for cryptography. Preprint.