Summation polynomials and the discrete logarithm problem on elliptic curves

Igor Semaev
Department of Mathematics
University of Leuven,Celestijnenlaan 200B
3001 Heverlee,Belgium
Igor.Semaev@wis.kuleuven.ac.be

February 5, 2004

Abstract

The aim of the paper is the construction of the index calculus algorithm for the discrete logarithm problem on elliptic curves. The construction presented here is based on the problem of finding bounded solutions to some explicit modular multivariate polynomial equations. These equations arise from the elliptic curve summation polynomials introduced here and may be computed easily. Roughly speaking, we show that given the algorithm for solving such equations, which works in polynomial or low exponential time in the size of the input, one finds discrete logarithms faster than by means of Pollard's methods.

Keywords: elliptic curves, summation polynomials, the discrete logarithm problem

1 Introduction

Let E be the elliptic curve defined over the prime finite field F_p of p elements by the equation

$$Y^2 = X^3 + AX + B. \tag{1}$$

The discrete log problem here is given $P, Q \in E(F_p)$ find an integer number n such that $Q = nP$ in $E(F_p)$ if such an n exists. It is of great importance in cryptography, see [1] and [2]. The aim of the paper is the construction of the index calculus algorithm for the problem. The construction presented here is based on the problem of finding bounded solutions to some explicit modular multivariate polynomial equations. These equations arise from the summation polynomials introduced in the second Section of the paper. In the third Section we show, roughly speaking, that given a good algorithm for solving such equations one finds discrete logarithms in $E(F_p)$ probably faster than by means of Pollard's methods, see [3],[4],[6] for them. An index calculus for the problem, called the xendi calculus, was published by Silverman [7]. It was shown in [8] that the xendi calculus fails to improve known bounds. We stress here that the underlying idea of the present new approach is different from Silverman's.

2 Summation Polynomials

Let E be the elliptic curve given by the equation (1) over a field K of characteristic $\neq 2, 3$, which is not necessary F_p now. For any natural number $n \geq 2$ we introduce the polynomial $f_n = f_n(X_1, X_2, \ldots, X_n)$ in n variables which is related to the arithmetic operation on E. We call this polynomial summation polynomial and define it by the following property. Let $x_1, x_2, \ldots, x_n$ be any elements from $\overline{K}$, the algebraic closure of the field K, then $f_n(x_1, x_2, \ldots, x_n) = 0$ if and only if there exist $y_1, y_2, \ldots, y_n \in \overline{K}$ such that points (x_i, y_i) are on E and $(x_1, y_1) + (x_2, y_2) + \ldots + (x_n, y_n) = P_\infty$ in the group $E(\overline{K})$.

Theorem 1 *The polynomial f_n may be defined by $f_2(X_1, X_2) = X_1 - X_2$, and $f_3(X_1, X_2, X_3) =$*

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

and $f_n(X_1, X_2, \ldots, X_n) =$

$$\mathrm{Res}_X(f_{n-k}(X_1, \ldots, X_{n-k-1}, X), f_{k+2}(X_{n-k}, \ldots, X_n, X)) \tag{2}$$

for any $n \geq 4$ and $n - 3 \geq k \geq 1$.

The polynomial f_n is symmetric and of degree 2^{n-2} in each variable X_i for any $n \geq 3$.

The polynomial f_n is absolutely irreducible and

$$f_n(X_1, \ldots, X_{n-1}, X_n) = f_{n-1}^2(X_1, \ldots, X_{n-1})X_n^{2^{n-2}} + \ldots$$

for any $n \geq 3$.

Proof. First we define the polynomial f_n for $n = 2$ and $n = 3$. One sees that $f_2 = X_1 - X_2$. Now we determine f_3. Let (x_1, y_1) and (x_2, y_2) be two affine points on E such that $x_1 \neq x_2$. We denote

$$(x_3, y_3) = (x_1, y_1) + (x_2, y_2),$$
$$(x_4, y_4) = (x_1, y_1) - (x_2, y_2).$$

One can see that x_3, x_4 are roots of a quadratic polynomial, whose coefficients are symmetric functions in x_1 and x_2. Really, we derive

$$x_3 = \lambda_3^2 - (x_1 + x_2),$$
$$x_4 = \lambda_4^2 - (x_1 + x_2),$$

where $\lambda_3 = (y_1 - y_2)/(x_1 - x_2)$ and $\lambda_4 = (y_1 + y_2)/(x_1 - x_2)$. Then

$$x_3 + x_4 =$$
$$\lambda_3^2 + \lambda_4^2 - 2(x_1 + x_2) =$$
$$2\tfrac{(x_1+x_2)(x_1x_2+A)+2B}{(x_1-x_2)^2},$$

and

$$x_3x_4 =$$
$$(\lambda_3^2 - (x_1 + x_2))(\lambda_4^2 - (x_1 + x_2)) =$$
$$\tfrac{(x_1x_2-A)^2-4B(x_1+x_2)}{(x_1-x_2)^2},$$

The x-coordinates x_3 and x_4 are roots of the polynomial

$$(x_1 - x_2)^2X^2 - 2((x_1 + x_2)(x_1x_2 + A) + 2B)X + ((x_1x_2 - A)^2 - 4B(x_1 + x_2)).$$

If $x_1 = x_2$ and $(x_3, y_3) = 2(x_1, y_1)$, where (x_3, y_3) is an affine point on E, one can see that x_3 is the root of the same polynomial. It means that one can take $f_3(X_1, X_2, X_3) =$

$$(X_1-X_2)^2X_3^2-2((X_1+X_2)(X_1X_2+A)+2B)X_3+((X_1X_2-A)^2-4B(X_1+X_2)).$$

One sees that the polynomial $f_3(X_1, X_2, X_3)$ is irreducible over the field $\overline{K(X_3)}$. It follows from the fact that the equation $f_3(X_1, X_2, X_3) = 0$ is isomorphic over $\overline{K(X_3)}$ to the initial elliptic curve (1). In particular, the polynomial $f_3(X_1, X_2, X_3)$ is absolutely irreducible. So we have proved all claims when $n = 3$.

Let $n \geq 4$, and $n - 3 \geq k \geq 1$, and

$$(x_1, y_1) + (x_2, y_2) + \ldots + (x_n, y_n) = P_\infty \tag{3}$$

in the group $E(\overline{K})$. First we consider the case $(x_1, y_1)+\ldots+(x_{n-k-1}, y_{n-k-1}) = (x, y)$ for some affine point $(x, y) \in E$. So $(x_{n-k}, y_{n-k})+\ldots+(x_n, y_n) = (x, -y)$. It implies the polynomials $f_{n-k}(x_1, \ldots, x_{n-k-1}, X)$ and $f_{k+2}(x_{n-k}, \ldots, x_n, X)$ have nonzero leading coefficients and the common root x. It follows by induction that the leading coefficients of the polynomials are $f^2_{n-k-1}(x_1, \ldots, x_{n-k-1})$ and $f^2_{k+1}(x_{n-k}, \ldots, x_n)$ which are nonzero. Then $f_n(x_1, x_2, \ldots, x_n) =$

$$\mathrm{Res}_X(f_{n-k}(x_1, \ldots, x_{n-k-1}, X), f_{k+2}(x_{n-k}, \ldots, x_n, X)) = 0$$

Let $(x_1, y_1) + \ldots + (x_{n-k-1}, y_{n-k-1}) = P_\infty$ then $(x_{n-k}, y_{n-k}) + \ldots + (x_n, y_n) = P_\infty$ and the leading coefficients of the polynomials $f_{n-k}(x_1, \ldots, x_{n-k-1}, X)$ and $f_{k+2}(x_{n-k}, \ldots, x_n, X)$ are zeros. Again $f_n(x_1, x_2, \ldots, x_n) =$

$$\mathrm{Res}_X(f_{n-k}(x_1, \ldots, x_{n-k-1}, X), f_{k+2}(x_{n-k}, \ldots, x_n, X)) = 0.$$

When $f_n(x_1, x_2, \dots, x_n) = 0$ the equality (3) is true. Really, if the leading coefficients of the polynomials $f_{n-k}(X_1, \dots, X_{n-1}, X)$ and $f_{k+2}(X_{n-k}, \dots, X_n, X)$ in X are zeros at $x_1, \dots, x_n$ then

$$(x_1, y_1) + \ldots + (x_{n-k-1}, y_{n-k-1}) = P_\infty,$$
$$(x_{n-k}, y_{n-k}) + \ldots + (x_n, y_n) = P_\infty,$$

by induction for some $y_i \in \overline{K}$. So (3) is true. If one of these coefficients isn't zero then the polynomials $f_{n-k}(x_1, \dots, x_{n-1}, X)$ and $f_{k+2}(x_{n-k}, \dots, x_n, X)$ have a common root $x \in \bar{K}$. Again by induction

$$(x_1, y_1) + \ldots + (x_{n-k-1}, y_{n-k-1}) = (x, y)$$
$$(x_{n-k}, y_{n-k}) + \ldots + (x_n, y_n) = \pm(x, y)$$

and (3) is true.

By induction and using known properties of the resultant one gets $\deg_{X_n} f_n \le 2^{n-2}$. On the other hand one can always find $x_1, \dots, x_{n-1} \in \overline{K}$ such that the x-coordinates of 2^{n-2} points

$$(x_1, y_1) \pm \ldots \pm (x_{n-1}, y_{n-1})$$

are pairwise different. It means that the polynomial $f_n(x_1, \dots, x_{n-1}, X_n)$ in X_n has just 2^{n-2} different roots. That is $\deg_{X_n} f_n = 2^{n-2}$. The same is true for all other variables.

Now we prove that f_n is absolutely irreducible. Let on the contrary $f_n = G_1 G_2$ for some polynomials G_i over $\overline{K}$. It follows from the definition of the polynomial f_n that G_i is a constant or depends on all variables. From (2) it follows that

$$f_n = (X_{n-1} - X_n)^{2^{n-2}} F_1 F_2,$$

where $F_1 = f_{n-1}(X_1, \dots, X_{n-2}, X)$, and $F_2 = f_{n-1}(X_1, \dots, X_{n-2}, \overline{X})$, and $X, \overline{X} \in K_1 = \overline{K(X_{n-1}, X_n)}$ are roots of the polynomial $f_3(X_{n-1}, X_n, X)$. One proves by induction on n and using the same argument that the polynomials F_1 and F_2 are irreducible over K_1. So F_1 should divide one of G_i which is defined over $\overline{K}$. Therefore F_1 and F_2 divide the same polynomial G_i, for example G_1. So G_2 should be a constant and the polynomial f_n is absolutely irreducible.

To prove the last claim of the Theorem we observe that the coefficient at $X_n^{2^{n-2}}$ of the polynomial f_n is just

$$Z_n^{2^{n-2}} f_n(X_1, \dots, X_{n-1}, X_n/Z_n),$$

when $Z_n = 0$. One sees that

$$Z_n^{2^{n-2}} f_n(X_1, \dots, X_{n-1}, X_n/Z_n) =$$
$$\mathrm{Res}_X(f_{n-k}(X_1, \dots, X_{n-k-1}, X), Z_n^{2^k} f_{k+2}(X_{n-k}, \dots, X_n/Z_n, X)).$$

By induction the last resultant, when $Z_n = 0$, is the resultant

$$\mathrm{Res}_X(f_{n-k}(X_1, \dots, X_{n-k-1}, X), f_{k+1}^2(X_{n-k}, \dots, X_{n-1}, X))$$

which equals $f_{n-1}^2(X_1, \dots, X_{n-1})$. This finishes the proof of the Theorem.

Remark 1 *In the case of characteristic* 2 *and* 3 *the same polynomial may be introduced and computed in a similar way. So we omit this. Insted we give two first summation polynomials* f_3 *and* f_4 *for the Koblitz elliptic curves, see* [9], *defined over the finite field of two elements* F_2 *by the equation*

$$Y^2 + XY = X^3 + aX^2 + 1, \tag{4}$$

where $a = 0, 1$. *They are*

$$f_3(x_1, x_2, x_3) = (x_1x_2 + x_1x_3 + x_2x_3)^2 + x_1x_2x_3 + 1,$$

and

$$f_4(x_1, x_2, x_3, x_4) =$$
$$(x_1 + x_2 + x_3 + x_4)^4 + (x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4)^4 +$$
$$x_1x_2x_3x_4(x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4 + x_1 + x_2 + x_3 + x_4)^2$$
$$(x_1x_2x_3x_4)^2(x_1 + x_2 + x_3 + x_4)^2 + (x_1x_2x_3 + x_1x_2x_4 + x_1x_3x_4 + x_2x_3x_4)^2$$

3 The Discrete Logarithm Problem

We return now to the discrete logarithm problem in $E(F_p)$, where E is given by (1) over the field F_p. We fix any natural number $n \geq 2$. Let $R = (x, y) = l_1P + l_2Q$ in $E(F_p)$ for random integers l_1 and l_2. Let's consider the equation

$$f_{n+1}(x_1, \ldots, x_n, x) \equiv 0(\bmod p) \tag{5}$$

in variables $x_1, \ldots, x_n$. Very probably (5) has a solution $x_1^0, \ldots, x_m^0$, where x_i^0 are integer numbers bounded by $p^{1/n+\delta}$ for some small $\delta > 0$ or x_i^0 are rational numbers the numerator and the denominator of which are bounded by $p^{1/(2n)+\delta}$. Imagine we have an algorithm able to find such a solution. It would imply we are able to find the relation

$$(x_1^0, y_1^0) + \ldots + (x_n^0, y_n^0) = l_1P + l_2Q \tag{6}$$

for some $y_1^0, \ldots, y_n^0$ in F_p or in F_{p^2}. It isn't important if some $y_i^0 \in F_{p^2} \setminus F_p$, since the sum of all such points in (6) is an order 2 point on E. The relations (6) may be combined with the relations

$$(x_1, y_1) + \ldots + (x_m, y_m) = P_\infty$$

that would come from the equations

$$f_m(x_1, \ldots, x_m) \equiv 0(\bmod p) \tag{7}$$

for $m \geq n$ if one could find them bounded by $p^{1/n+\delta}$. One should avoid trivial solutions to (5) and (7) like $x_1, x_1, x_2, x_2, \ldots, x_k, x_k$ is always solution to (7) when $m = 2k$. One needs about $p^{1/n+\delta}$ nontrivial solutions to find the logarithm of Q to the base P. So if the algorithm, finding a bounded solution to (5) and (7), works in $t_{p,n}$ operations then the complexity of the discrete logarithm problem in $E(F_p)$ is essentially

$$t_{p,n}p^{1/n+\delta} + p^{2/n+2\delta}$$

operations. When $n \geq 5$, even for some exponential $t_{p,n}$, this amount may be less than $O(p^{1/2})$ provided by Pollard's methods.

There exist modular multivariate polynomial equations a bounded solution for which may be found in polynomial or low exponential time in the size of the input. The exciting question arising here is whether or not it is true for (5) and (7)?

Remark 2 *The similar approach may be developed for the Koblitz curve E defined by* (4)*. To construct the index calculus algorithm for the discrete logarithm problem in $E(F_{2^l})$ one should have an algorithm, working in polynomial or low exponential time, for findind polynomials $x_1, x_2, \ldots, x_n$ over F_2 of degree $\leq l/n + \delta$ satisfing the equation $f_{n+1}(x_1, x_2, \ldots, x_n, x) \equiv 0$ for a random polynomial x modulo an irreducible polynomial of degree l over the field F_2.*

References

[1] V.Miller, *Use of elliptic curves in cryptography. Advances in cryptology—CRYPTO '85 (Santa Barbara, Calif., 1985)*, 417–426, Lecture Notes in Comput. Sci., **218**(1986), Springer, Berlin, 417–426.

[2] N. Koblitz *Elliptic curve cryptosystems,* Math. Comp. **48** (1987), 203–209.

[3] J.Pollard *Monte-Carlo methods for index computation* mod *p,* Math.Comp. **32** (1978), 918–924.

[4] P.van Oorschot and M.Wiener, *Parallel collision search with cryptanalytic applications,* J. Cryptology **12** (1999), no. 1, 1–28.

[5] M.Wiener and R.Zuccherato, *Faster attacks on elliptic curve cryptosystems. Selected areas in cryptography (Kingston, ON, 1998)*, Lecture Notes in Comput. Sci., **1556**(1999), Springer, Berlin, 190–200.

[6] R.Gallant, R.Lambert,and S.Vanstone, *Improving the parallelized Pollard lambda search on anomalous binary curves.* Math. Comp. **69** (2000), no. 232, 1699–1705.

[7] Silverman, J. H.: The xedni calculus and the elliptic curve discrete logarithm problem. Des. Codes Cryptogr. 20 (2000), no. 1, 5–40.

[8] Jacobson, M. J., Koblitz, N., Silverman, J. H., Stein, A., Teske, E.: Analysis of the xedni calculus attack. Des. Codes Cryptogr. 20 (2000), no. 1, 41–64.

[9] *Digital Signature Standard(DSS),* FIPS PUB 186-2,2000 January 27.