Use of Elliptic Curves in Cryptography

Victor S. Miller

Exploratory Computer Science, IBM Research, P.O. Box 218, Yorktown Heights, NY 10598

ABSTRACT

We discuss the use of elliptic curves in cryptography. In particular, we propose an analogue of the Diffie-Hellmann key exchange protocol which appears to be immune from attacks of the style of Western, Miller, and Adleman. With the current bounds for infeasible attack, it appears to be about 20% faster than the Diffie-Hellmann scheme over GF(p). As computational power grows, this disparity should get rapidly bigger.

H.C. Williams (Ed.): Advances in Cryptology - CRYPTO '85, LNCS 218, pp. 417-426, 1986.

INTRODUCTION

Elliptic curves have been objects of intense study in Number Theory for the last 90 years. To quote Lang "It is possible to write endlessly on Elliptic Curves (This is not a threat)." [1]. Recently [2], H.W. Lenstra has proposed a new integer factorization algorithm based on the arithmetic of elliptic curves, which, under reasonable hypotheses, runs at least as fast as the best known factorization algorithm, and uses a negligible amount of storage. This has obvious implications for cryptographic techniques depending on the difficulty of factoring. It is my intent to show that elliptic curves have a rich enough arithmetic structure so that they will provide a fertile ground for planting the seeds of cryptography.

NOTATION AND RESUME OF PROPERTIES OF ELLIPTIC CURVES

If S is a finite set, we denote its cardinality by $|S|$. If p is a prime number, and $n \neq 0$ is an integer, we denote by $v_p(n)$ the exact exponent of p dividing n. If $a = b/c$ is rational, then we set $v_p(a) = v_p(b) - v_p(c)$. As usual, $\mathbf{Q}$ denotes the rational numbers, and $\mathbf{Z}$ denote the integers. If $n \neq 0$ is an integer, let $\mathbf{Q}_{(n)}$ denotes the subring of $\mathbf{Q}$ consisting of elements whose denominators are relatively prime to n. If σ denotes a set of primes then let $\mathbf{Z}_\sigma$ denote those rational numbers whose denominators are divisible only by primes in σ. Note that if no prime in σ divides n, then $\mathbf{Z}_\sigma$ is a subring of $\mathbf{Q}_{(n)}$.

An (affine) **algebraic group** defined over a ring R is a set of simultaneous polynomial equations in $x_1, \ldots, x_n$, with coefficients in R :

$$f_1(P) = 0, \ldots, f_t(P) = 0$$

along with a composition law, and inverse given by n polynomial functions with coefficients in R.

$$m_i(x_1, \ldots, x_n, y_1, \ldots, y_n)$$
$$a_i(x_1, \ldots, x_n)$$

which satisfies the usual axioms for a group. If G is an algebraic group, and S is a ring which has a multiplication by elements of R defined, then $G(S)$ denotes the set of solutions to the polynomial equations with the variables having values in S. The law of composition given above then makes $G(S)$ into a group. We also may have a projective algebraic group with the same definition as above, except that the polynomials must be homogeneous of the same degree. In this case $G(S)$ denotes the set of solutions all of whose coordinates are not zero, with two solutions being considered the same if one is a scalar multiple of the other. Note that in this case, the law of composition really consists of a set of rational functions.

As an example, we have the multiplicative group:

$$\mathbf{G}_m: \quad xy = 1$$

with law of composition $((x_1, y_1),(x_2, y_2)) \to (x_1x_2, y_1y_2)$ and inverse $(x,y) \to (y,x)$.

Define the **logarithmic height** of a point: Given a point $P = (x_1, x_2, \ldots, x_n)$ with rational coordinates, let D be a common denominator for all the x_i such that, there is a j such that $(Dx_j, D) = 1$. The logarithmic height, $h(P) = \log \max(|D|, |Dx_1|, \ldots, |Dx_n|)$. This height is a measure of the number of bits needed to write down the point P. Let $H_n(K) = \{P \in \mathbf{Q}^n \mid h(P) \leq K\}$.

Let p_i denote the i-th prime number, and $G_m(r)$ be the subgroup of $G_m(\mathbf{Q})$ generated by $p_1, \ldots, p_r$. Note that $G_m(r)$ is the same as $G_m(\mathbf{Z}_\sigma)$ where σ consists of the set of primes $p_1, \ldots, p_r$. Also let $G_m(r, K) = G_m(r) \cap H_2(K)$

An **elliptic curve** defined over a field F is the curve defined by the following equation:

$$E: y^2 = x^3 + ax + b$$

where a and b are elements of F (assumed to have characteristic $\neq 2$ or 3. There is a slightly more complicated formulation in those cases.). There is a natural law of composition on the points of E obtained by the "tangent and chord method": Given two points P and Q, the straight line containing them intersects the curve in a third point R (if $P = Q$ take the tangent to the curve at P). Define $P + Q$ as being the point $(x(R), -y(R))$. This provides a commutative and associative law of composition, whose zero element is the point at infinity: (∞, ∞) We denote the set of points (including the point at infinity) of the curve E with coordinates in the field F by $E(F)$ The **discriminant** of the curve $\Delta = 16(4a^3 - 27b^2)$. The elliptic curve

$$E_\lambda: y^2 = x^3 + \lambda^4 ax + \lambda^6 b$$

is isomorphic to the curve E above by the substituion

$$(x,y) \to \left(\frac{x}{\lambda^2}, \frac{y}{\lambda^3}\right)$$

We say the E is **minimal** if a and b are integers, and there is no integer $\lambda \neq \pm 1$ such that $\lambda^4 \mid a$ and $\lambda^6 \mid b$. Clearly, every elliptic curve is isomorphic to a minimal one. We denote the discriminant of the minimal curve isomorphic to E by $\Delta_{\min}$. There is a slightly more general definition of minimal by using a more complicated model for an elliptic curve (see [1]). Its value of $\Delta_{\min}$ differs by a factor dividing 24, from the one described above.

To calculate multiples of a point $P = (x,y)$ we may use the following recurrences (see Lang [1], p. 37):

$$
\begin{aligned}
g_1 &= 1 \\
g_2 &= 1 \\
g_3 &= 3x^4 + 6ax^2 + 12bx - a^2 \\
g_4 &= x^6 + 5ax^4 + 20bx^3 - 5a^2x^2 - 4abx - 8b^2 - a^3 \\
g_{2n} &= g_n(g_{n+2}g_{n-1}^2 - g_{n-2}g_{n+1}^2) \\
g_{4n+1} &= 16y^4 g_{2n+2}g_{2n}^3 - g_{2n+1}^3 g_{2n-1} \\
g_{4n+3} &= g_{2n+3}g_{2n+1}^3 - 16y^4 g_{2n+2}^3 g_{2n} \\
f_{2n} &= 2yg_{2n} \\
f_{2n+1} &= g_{2n+1} \\
\phi_n &= xf_n^2 - f_{n+1}f_{n-1} \\
4y\omega_n &= f_{n+2}f_{n-1}^2 - f_{n-2}f_{n+1}^2
\end{aligned}
$$

Then

$$
n(x,y) = \left(\frac{\phi_n}{f_n^2}, \frac{\omega_n}{f_n^3} \right)
$$

Using the above recursions we may calculate the coordinates of the above point in $26 \log_2 n$ multiplications.

We let $\mathbf{F}_q$ denote GF(q), the finite field with q elements. We now state a few results for elliptic curves which are needed for the discussion in the next section. Two good general references for elliptic curves are Cassels [11] and Lang [1]. All of the results quoted below are contained therein, unless indicated otherwise. The number of points, $|E(\mathbf{F}_p)| = p + 1 - a_p$ where $|a_p| \leq 2\sqrt{p}$ (the "Riemann hypothesis for finite fields" proved by Hasse in 1931 for Elliptic curves). The Mordell-Weil theorem states that the rank of the free part of the group $E(\mathbf{Q})$ is finite (for any specific E). In fact it is usually quite small. Indeed, no one to date has been able to find an elliptic curve with rational coefficients whose rank is greater than 14 (this record is held by J-F Mestre, see [12] for a description of a rank 12 case).

A fundmental theorem of Neron and Tate (see [1]) is that there exists a unique positive semi-definite quadratic function $\hat{h}(P)$ such that for all $P \in E(\mathbf{Q})$ (even on $E(M)$ where M is a number field) such that

$$
h(P) = \hat{h}(P) + O(1)
$$

The O(1) is quite small, even being bounded by $\log \max(|a|, |b|)$ (see Zimmer [15]). In fact this bound always seems to be much too large (see [16]). We also have $\hat{h}(P) = 0$ if and only if P is a

point of finite order (of which there are at most 16, by a theorem of Mazur). The meaning of $\hat{h}(P)$ being a quadratic function is that

$$<P, Q> = \hat{h}(P + Q) - \hat{h}(P) - \hat{h}(Q)$$

is a positive definite inner product. If $P_1, \ldots, P_r$ is a basis for the points of $E(\mathbf{Q})$ of infinite order, we define the regulator to be

$$R = \det(<P_i, P_j>)$$

This value is independent of the basis chosen. We also define $|P| = \sqrt{<P, P>}$. In this case $<P, P> = 1/2\hat{h}(P)$.

KEY EXCHANGE, AND DISCRETE ELLIPTIC LOGARITHMS

The Diffie-Hellman key exchange protocol [3] was proposed to allow the agreement on a secret key between two parties communicating over an insecure channel. It operates as follows: A large prime p and a primitive root g of p are made public. Party A chooses an exponent a between 0 and $p - 1$ at random. Party B does the same with an exponent b Party A transmits g^a to B, and vice-versa. Both parties agree on g^{ab}. The security of this protocol rests on two unproven (but reasonable) assumptions:

1. Any method of obtaining g^{ab} from g^a and g^b would be as hard as obtaining a from g^a (taking "discrete logarithms").
2. If $p - 1$ did not have only small prime factors, that finding discrete logarithms was intractible (i.e. could not run in time polynomial in $\log p$).

Neither assumption has been disproven. However, Western and Miller [4], and Adleman [5] have come up with algorithms for the discrete logarithm problem which run in time $L(p)$, where

$$L(x) = \exp(\sqrt{\log x \log \log x})$$

In addition, Pohlig and Hellman [17], and Pollard [18] have a method for calculating discrete logarithms, depending only on the fact that we are working in a group, which runs in time $O(\sqrt{p'})$ where p' is the largest prime factor of $p - 1$. Given current processing speeds, this escalates the size of the prime p which must be used, in order to make this method more secure. A figure of $p \simeq 2^{512}$ seems to be necessary.

The above protocol really only uses the property that we are working in a group. As stated above, the points on an elliptic curve have the structure of an abelian group. Thus we may make the analogous constructions over elliptic curves. We shall briefly describe the "Index Calculus" algorithm of Adleman, and Western and Miller, and give arguments why such an algorithm is not likely to work on elliptic curves. We have the reduction map:

$$G_m(Q_{(p)}) \to G_m(F_p)$$

denoted by $x \to \bar{x}$. Now

$$|G_m(r, \log \sqrt{p}/2)| = \overline{|G_m(r, \log \sqrt{p}/2)|}$$

Let $\text{prob}(r) = |G_m(r, \log \sqrt{p}/2)|/(p-1)$. This is the probability that an element of the multiplicative group is in the above image. As r increases to $\pi(\sqrt{p}/2)$, $\text{prob}(r)$ increases to 1.

The "index calculus" method fixes a value of r, and chooses elements $a \in F_p$ at random until there is $x \in G_m(r, \log \sqrt{p}/2)$ such that $\bar{x} = g^a$. The probability of that succeeding is $\text{prob}(r)$. To each such successful test we have an equation

$$a = \nu_0 l_0 + \ldots + \nu_r l_r \bmod p - 1 \tag{1}$$

where

$$x = (-1)^{\nu_0} p_1^{\nu_1} \ldots p_r^{\nu_r} \tag{2}$$

and l_k is such that $p_k = g^{l_k} \bmod p - 1$ where $p_0 = -1$. Evidently, we have $l_0 = (p-1)/2$. We need to generate r such independent equations. Once we have accumulated them, we may solve for the l_i.

Given $z \in F_p$, we find l such that $z = g^l$ as follows: Choose $a \bmod p - 1$ at random until there exists $x \in G_m(r, \log \sqrt{p}/2)$ such that $\bar{x} = z^a$. Then

$$al = \nu_0 l_0 + \nu_1 l_1 + \ldots + \nu_r l_r$$

where

$$x = p_0^{\nu_0} p_1^{\nu_1} \ldots p_r^{\nu_r}$$

We may then solve for l. Each such test has probability $\text{prob}(r)$ of succeeding.

There is a trade-off between increasing r in order to make $\text{prob}(r)$ bigger (in order to decrease the expected number of tests to make), and in decreasing r in order to make the calculation of the decomposition (2) faster, and of solving a smaller system of equations. Fortunately, good algorithms exist for both the latter problems. Using the new factorization algorithm of Lenstra [2], we may find the decomposition (2), or signal failure, in time

$$O(L(p)^{\sqrt{2\alpha}+\varepsilon})$$

where

$$\alpha = \frac{\log p_r}{\log p} \simeq \frac{\log r + \log\log r}{\log p}$$

The equations (1) are provably sparse, namely at most $\log p$ of the v_i are $\neq 0$, because $p_1 \ldots p_r \simeq e^r$ by the prime number theorem. We may solve these equations in random time $O(r^2 \log^2 p)$ by the algorithm of Wiedemann [13]. It is evident that this last figure is the big bottleneck in trying to make r larger. It turns out that the optimum trade off is made by letting $r = L(p)^{c/2}$ for some small constant c between 3/2 and 2. The total running time turns out to be $L(p)^c$. Recently, Coppersmith, Odlyzko, and Schroeppel have devised a slightly more complicated variant of the above, which has the above running time with $c = 1$.

The reason why the above algorithm works so well, is that there are lots of free generators for the group $G_m(Q_{(p)})$, which have fairly small heights. If one tries to use an analogous method with elliptic curves, one immediately runs into the barrier of the Mordell-Weil Theorem (see above). We show below, that this finititude of the rank combined with other estimates, that it is extremely unlikely that an "index calculus" attack on the elliptic curve method will ever be able to work. We may view $E(\mathbf{Q}) \otimes \mathbf{R}$ as an r-dimensional inner product space, with the inner product given above, which contains $E(\mathbf{Q})$ as a lattice, whose fundamental domain has volume $\sqrt{R}$. Thus,

$$|E(\mathbf{Q}) \cap H(K)| = 2\frac{V_r}{\sqrt{R}} K^{r/2} + O(K^{(r-1)/2})$$

where V_r is the volume of the r dimensional sphere of radius 1. Thus, unless the rank of the curve can be made very large, and the regulator made fairly small, the probability of a point of $E(\mathbf{F}_p)$ lifting to a point on $E(\mathbf{Q})$ whose height is bounded by something reasonable (say a polynomial in $\log p$) is vanishingly small. In particular, in order to make the probality of finding a point with a specified height $< p^a$ it is necessary to make $K = \Omega(p^{(1-a)/r})$. That is we must compute points whose coordinates are represented by rational numbers whose length is exponential in $\log p$. That is rather a daunting prospect!

Despite the remarks above about it being difficult to find curves of large rank, it is widely believed that there is no bound on the rank attainable. However, it is also true that rank $(E(\mathbf{Q})) = O(\log \max(|a|, |b|))$. This shows that the size of the coefficients needs to be exponentially larger than the rank. This would seem to preclude high rank from the point of view of computational complexity. In fact, the above bound is really quite bad, which would tend to make the situation much worse from the point of view of computational complexity. As far as a lower bound on the regulator, Lang has conjectured [1], and Silverman proved [7](in some cases)

that if $\hat{h}(P) \neq 0$ that $\hat{h}(P) > c_1 \log |\Delta_{\min}| + c_2$ for some constants c_i. This estimate is even true over algebraic number fields, with the constants depending on the field. Laurent [8] gives a precise lower bound for the constant c_1, if one has $c_2 = 0$ (this only make c_1 larger), of $c_3/(D(\log\log D)^3)$ where D is the degree of the field above Q and c_3 is an absolute constant independent of the curve and the field. These estimates say that the regulator can't be too small, as long as a and b can't get too big. This remark would seem to preclude an attack which tries to look at points in $E(M)$ for some finite field extension M of Q.

Even if one could somehow get around the barrier mentioned above there is still the problem of actually lifting a point. In the original case of G_m it is trivial, or nearly so. In the case of an elliptic curve it seems to be much more difficult. If we are given a point $(x,y) \in E(\mathbf{F}_p)$ and some point $(x_1, y_1) \in E(\mathbf{Z}/p^k\mathbf{Z})$ which projects to the original point, we could find a rational point (X, Y) whose height is bounded by $k\log p - \log 8$ by an integer basis reduction algorithm (L^3 or Kannan) in the 3 dimensional lattice generated by the vectors

$$\begin{pmatrix} 1 \\ x_1 \\ y_1 \end{pmatrix} \begin{pmatrix} 0 \\ p^k \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ p^k \end{pmatrix}$$

However, there are many possible choices for (x_1, y_1), about $p^{k/3}$ of them. Furthermore, even though they are parametrizable, the parametrization is non-linear. Thus, unless there is a new idea, it would seem that this is another barrier, difficult to surmount.

Implementation and Practice

A number of details need to be addressed in order to make this scheme practical:

1. The actual algorithm for multiplication on an elliptic curve
2. The choice of the parameters A and B for the elliptic curve.
3. The choice of the prime modulus p.
4. What information needs to be transmitted.

There are two possible algorithms that one could use for multiplying a point by an integer: the recursion cited above, or repeated use of addition and doubling with the binary method for multiplication. In either algorithm, it appears to be best to represent the points on the curve in the following form: Each point is represented by the triple (x,y,z) which corresponds to the point $(x/z^2, y/z^3)$. This is a homogeneous representation with x having weight 2, y having weight 3, and z having weight 1. If this representation is used with the recursions in the first section, then it is easily checked that the only change is in the initialization. A simple induction shows that g_{2n} has weight $4n^2 - 4$, and that g_{2n+1} has weight $4n^2 + 4n$.

In order to be secure from the Pohlig-Hellmann (or Pollard) algorithm, it is necessary that N_p, the number of points of E in $\mathbf{F}_p$ have a prime factor $> p^{\alpha}$, for α as close to 1 as possible. This is made possible by the algorithm of Schoof [19], which calculates N_p in time polynomial in $\log p$. In general it is not hard to find such good p. Theoretically, the best result known is one of Fouvry [20]: For any fixed non-zero integer a, a positive proportion of primes p have the property that the largest prime factor of $p + a$ is $\geq p^{\delta}$ where $\delta = 0.6687$.

Instead of using the Schoof algorithm, when searching for a good p, I have taken the following approach: Choose the curve to be:

$$E: \ y^2 = x^3 - ax$$

where a is not a perfect square. This curve has complex multiplication by $\sqrt{-1}$, and there is an exact formula for N_p (see [10]). In the case $p = 3 \bmod 4$ we have $N_p = p + 1$. This is the so-called "supersingular" case. In this case we know even more. It is well known (see [1]) that any field containing the coordinates of all points of order l also contains the l-th roots of unity. This shows that a necessary condition for group of point over $\mathbf{F}_p$ to contain a subgroup isomorphic to $\mathbf{Z}/l\mathbf{Z} \times \mathbf{Z}/l\mathbf{Z}$ is that $l \mid p - 1$. Because the number of points in the supersingular case is $p + 1$ we have 2 as the only possibility for l. But, in our case, this happens if and only if, a is a quadratic residue modulo p. To sum up, in the case above the group of points modulo p is of order $p + 1$, cyclic in the case $(a/p) = -1$, and a product of a cyclic group of order 2 and a cyclic group of order $(p + 1)/2$ when $(a/p) = 1$.

The above choice of curve was taken for convenience in calculation. However, it may be prudent to avoid curves with complex multiplication because the extra structure of these curves might somehow be used to give a better algorithm.

Finally, it should be remarked, that even though we have phrased everything in terms of points on an elliptic curve, that, for the key exchange protocol (and other uses as one-way functions), that only the x-coordinate needs to be transmitted. The formulas for multiples of a point cited in the first section make it clear that the x-coordinate of a multiple depends only on the x-coordinate of the original point.

BIBLIOGRAPHY [1] Lang, Serge, Elliptic Curves: Diophantine Analysis, Springer-Verlag New York, 1978.

[2] Lenstra, H. W., Letter to A. M. Odlyzko.

[3] Diffie, W. and Hellman M., New Directions in Cryptography, IEEE Trans. Inform. Theory, IT-22 (1976), 644-654.

[4] Western, A. E., and Miller, J. C. P., *Table of Indices and Primitive Roots*, Royal Society Mathematical Tables, vol. 9, Cambridge Univ. Press, 1968.

[5] Adleman, L., A subexponential algorithm for the discrete logarithm problem with applications to cryptography, Proc. 20th IEEE Found. Comp. Sci. Symp. (1979), 55-60.

[6] Odlyzko, A. M., Discrete logarithms in finite fields and their cryptographic significance, preprint.

[7] Silverman, J., Lower bound for the canonical height on elliptic curves, Duke Math. J. 48, 633-648 (1981).

[8] Laurent, M., Minoration de la hauteur de Neron-Tate, Seminaire de Theorie does Nombres, Paris 1981-82, 137-151, Birkhauser (1983).

[9] Birch, B. J., Swinnerton-Dyer H.P.F., Notes on Elliptic Curves I, J. reine u. angewandte Math., 212, 7-25 (1963).

[10] Birch, B. J., Swinnerton-Dyer H.P.F., Notes on Elliptic Curves II, J. reine u. angewandte Math., 218, 79-108 (1965).

[11] Cassels, J. W. S., Diophantine Equations with special reference to elliptic curves, J. London Math. Soc., 41, 193-291 (1966).

[12] Mestre, J-F., Courbes elliptique et formule explicites, Seminaire de Theorie does Nombres, Paris 1981-82, 179-187, Birkhauser (1983).

[13] Wiedemann, D., Solving sparse linear equations over finite fields, preprint.

[14] Coppersmith, D., Odlyzko, A. M., and Schroeppel, R., Discrete logarithms in GF(p), IBM Research Report RC 10985 (1985).

[15] Zimmer, H. G., On the difference of the Weil height and the Neron-Tate height, Math. Z. 147 (1976) 35-51.

[16] Buhler, J., Gross, B., and Zagier, D., On the conjecture of Birch and Swinnerton-Dyer for an elliptic curve of rank 3, preprint.

[17] Pohlig, S. and Hellman, M., An improved algorithm for computing logarithms over GF(p) and its cryptgraphic significance, IEEE Inform. Theory IT-24 (1978), 106-110.

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

[19] Schoof, R., Elliptic Curves over finite fields and the computation of square roots mod p, Report 83-09, Math. Inst. Univ. v. Amsterdam (1983).

[20] Fouvry, E., Theoreme de Brun-Titchmarsh; application au theoreme de Fermat, Invent. Math. 79 (1985), 383-407.

[21] Bremner, A. and Cassels, J. W. S., On the Equation $Y^2 = X(X^2 + p)$, Math. Comp. 42 (1984), 257-264.