IPAM Cryptography Workshop

January 11, 2002

ELLIPTIC CURVE CRYPTOGRAPHY:

WHICH CURVES TO USE?

Neal Koblitz

University of Washington

OUTLINE OF TALK:

Elliptic curve cryptography (ECC)

The Discrete Log Problem (ECDLP)

Curves over prime fields

Reducing a global curve modulo p

The Sophie Germain question

Curves over extensions of medium-size fields

Curves over $\mathbb{F}_{2^r}$ for prime r

"Magic Curves"

ELLIPTIC CURVE CRYPTOGRAPHY (ECC)

Elliptic Curves

An elliptic curve E over a field F:

$$Y^2 = X^3 + aX + b, \qquad a, b \in F.$$

(It must be *smooth*, i.e., the cubic on the right has no multiple roots; also, one needs a slightly different form in characteristic 2 or 3.)

ADDITION OF POINTS

The points $x, y \in F$ that satisfy this equation, together with a "point at infinity" denoted O, form an abelian group whose identity element is O. In other words, there is an "addition law."

Use of ECC to send a message.

Suppose we have a way to convert the message (i.e., a block of text) to a large integer, and then to a point $M \in E$.

> **Publicly known: $E/\mathbb{F}_q$, base point $P \in E$.**
>
> **Alice's keys: secret key is a random integer x; her public key is the point $Q = xP$.**

Bob wants to send the message M to Alice. He chooses a random integer k, computes kP and kQ, and sends the pair of points

$$(kP, kQ + M).$$

That is, the message is "disguised" by adding the point kQ; but since this point is equal to xkP, Alice can remove the disguise by multiplying the first point of the pair by her secret x:

$$xkP + M - x(kP) = M.$$

Elliptic Curve Digital Signature Algorithm (ECDSA)

Let E be an elliptic curve defined over $\mathbb{F}_q$ such that $\#E(\mathbb{F}_q)$ is equal to a prime ℓ of at least 160 bits (or to a small integer factor times such a prime ℓ).

Let P be an $\mathbb{F}_q$-point of E of order ℓ.

Let $f_E : \; E(\mathbb{F}_q) \to \mathbb{F}_\ell$ be a fixed, easily computable function that spreads the points over $\mathbb{F}_\ell$ fairly evenly.

For example, if q is a prime p, f_E could be the residue mod ℓ of the x-coordinate of a point (regarded as an integer between 0 and $p-1$).

Let H be a "hash function," an easily computable function from the space of message units to $\mathbb{F}_\ell$. It must have certain randomness properties and intractability properties for going backwards.

Alice's secret key: a random integer x in the range $1 < x < \ell$.

Alice's public key: the point $Q = xP \in E(\mathbb{F}_q)$.

To sign a message M, Alice does the following:

1) she selects a random integer k in the range $1 < k < \ell$;

2) she computes kP and $r = f_E(kP)$;

3) she computes $k^{-1} \in \mathbb{F}_\ell$ and

$$s = k^{-1}(H(M) + xr) \in \mathbb{F}_\ell;$$

4) her signature for the message M is the pair (r, s).

To verify the signature, Bob computes

$$u_1 = s^{-1}H(M) \in \mathbb{F}_\ell, \quad u_2 = s^{-1}r \in \mathbb{F}_\ell,$$

and then

$$u_1P + u_2Q \in E(\mathbb{F}_q).$$

If $f_E(u_1P + u_2Q) = r$, he accepts the signature.

THE DISCRETE LOG PROBLEM (DLP)

Definition. The *Discrete Logarithm Problem in the group G to the base $g \in G$* is the problem, given $y \in G$, of finding an integer x such that $g^x = y$ ($xg = y$ if the group operation in G is written additively), if such an integer exists (i.e., if y is in the subgroup generated by g).

The Discrete Log Problem in the multiplicative group $\mathbb{F}_q^\times$ is what we need to solve if we want to invert the function

$$x \mapsto g^x,$$

where $g \in \mathbb{F}_q^\times$ is a fixed element and x is an integer (modulo the order of g).

This map $x \mapsto g^x$ is the one-way function used in so-called Diffie–Hellman cryptosystems. For example, the U.S. government's first Digital Signature Standard (introduced in 1991) is such a system.

Similarly, if E is any elliptic curve defined over $\mathbb{F}_q$, then we have the following function in the elliptic curve group:

$$x \mapsto xP,$$

where P is a fixed point of E, and x is an integer.

In this case, the discrete log problem is the problem, given $Q \in E(\mathbb{F}_q)$, of finding an integer x such that $Q = xP$ if such an integer exists.

The security of all elliptic curve cryptosystems depends on the presumed intractability of the Elliptic Curve Discrete Log Problem (ECDLP).

The discrete log problem in the multiplicative group of a well chosen finite field $\mathbb{F}_q$ is hard: in practice, it seems to require about the same amount of time as factorization of an integer of approximately the same size as the finite field.

However, in the case of factorization, there are many subexponential-time* factoring algorithms. Most have running time of the form

$$\exp\Big((\log N)^{1/2+\varepsilon}\Big).$$

In the last decade, with the "number field sieve" and the "function field sieve," the heuristic asymptotic running time has been reduced to

$$\exp\Big((\log N)^{1/3+\varepsilon}\Big).$$

* **This means that the number of steps is asymptotically less than** $\exp\Big((\log N)^{\gamma}\Big)$ **for some** $\gamma < 1$**, where** N **is the number being factored.**

However, no subexponential-time algorithm is known for elliptic curves except in the following special cases:

- **there is no large prime ℓ dividing $\#E(\mathbb{F}_q)$;**

- **there is a large prime ℓ, but it also divides $q^K - 1$ for a very small value of K (in which case Menezes–Okamoto–Vanstone showed how the elliptic curve group can be imbedded in $\mathbb{F}^*_{q^K}$); or**

- **$\mathbb{F}_q$ is a prime field $\mathbb{F}_p$, and $\#E(\mathbb{F}_p) = p$ (in which case Semaev–Smart–Satoh– –Araki showed how the elliptic curve group can be imbedded in $\mathbb{F}_p^+$).**

Recall that the order of the group has to be divisible by a very large prime — perhaps even to be prime itself.

ADVANTAGES OF ELLIPTIC CURVES

- **No known subexponential-time algorithm for discrete log problem on suitably chosen curve.**

- **Therefore one can use shorter key sizes.**

- **There are many, many curves to choose from.**

RSA vs ECC:

Factoring record: N of 155 decimal digits.

ECDLP record: group size of 33 digits.

NOTE: The "Pollard ρ method" is the best one known for solving the ECDLP. It requires roughly $\sqrt{\ell}$ operations (where ℓ is the large prime factor of $\#E(\mathbb{F}_q)$).

In cryptographic applications, we would like the order of the group of points of the elliptic curve to be divisible by a very large prime — perhaps even to be prime itself.

There are several different approaches to selecting a suitable elliptic curve.

Choices to be made:

(a) What field do I want to work over?

(b) Random coefficients of the curve, or a curve with special properties?

Philosophical dispute:

HARDLINE POSITION: “All parameters for a cryptosystem must always be chosen with the maximal possible degree of randomness, because any extra structure or deviation from randomness might some day be used to attack the system.”

KINDER, GENTLER VIEWPOINT: “A user who requires the highest level of long-term security and is not so concerned about efficiency should have a cryptosystem with randomly generated parameters. On the other hand, a user who needs only short- and middle-term security and is very concerned about efficiency should have a cryptosystem that employs a special class of elliptic curve.”

By “special class” I mean a curve that has special properties that greatly improve efficiency, and for which no attack is known at present that takes advantage of these special properties to compromise the security of the system.

ONE APPROACH TO LOOKING FOR A CURVE OF PRIME (OR NEARLY PRIME) ORDER:

Let E be an elliptic curve

$$Y^2 = f(X) = X^3 + aX + b$$

defined over the field $\mathbb{Q}$ of rational numbers.

If p is any odd prime not dividing the denominators of the coefficients or the discriminant of $f(X)$, then one can consider the elliptic curve E over $\mathbb{F}_p$ that is obtained by simply reducing the coefficients modulo p.

That elliptic curve will always contain as a subgroup the image of the torsion subgroup E_{tors} (the set of points of finite order) of the curve over $\mathbb{Q}$. But one expects that in many cases the quotient will have prime order.

QUESTION. For a fixed curve E over $\mathbb{Q}$, what can be said about the probability as p varies that

$$\frac{\#E \bmod p}{\#E_{\text{tors}}}$$

is a prime number? Can one prove (for any fixed E) that there are infinitely many p for which this number is prime?

This question is analogous to a classical unsolved problem of number theory. Namely, instead of E take the multiplicative system of nonzero integers, which has torsion subgroup $\{\pm 1\}$. Then an analogous question is:

As p varies, what can be said about the probability that

$$\ell = \frac{p-1}{2} = \frac{\#\mathbb{F}_p^*}{\#\{\pm 1\}}$$

is prime? Are there infinitely many such "Sophie Germain primes" ℓ for which $p = 2\ell + 1$ is prime?

In 1823 Sophie Germain proved the 'first case' of Fermat's Last Theorem for prime exponents ℓ for which $2\ell + 1$ is prime. This was the first major result on Fermat's Last Theorem for a large class of exponents.

The question about Sophie Germain primes is of interest when using a Diffie–Hellman type cryptosystem in the multiplicative group of a prime field $\mathbb{F}_p$.

The analogous elliptic curve question given above is of interest when using an elliptic curve cryptosystem.

In both cases one needs the order of the group to be divisible by a large prime.

NOTE: The denominator $\#E_{\mathrm{tors}}$ in the elliptic curve question is often 1, and in any case it cannot be much larger than in the Sophie Germain prime question. According to a deep result of Barry Mazur, there are at most 16 torsion points on an elliptic curve over $\mathbb{Q}$.

EXAMPLE: Consider the elliptic curve

$$E: \quad Y^2 = X^3 - M^2 X,$$

and let

$$p \equiv 1 \pmod 4, \quad p \not| M.$$

Gauss found a formula for the number of points on E modulo p. Namely, one must first write p in the form $A^2 + B^2$ (a very easy computational task even if p is very large). Choose A odd, and choose its sign so that

$$A + B \equiv \left(\frac{M}{p}\right) \pmod 4.$$

Then

$$\#E \bmod p \quad = \quad p + 1 - 2A.$$

This number is always divisible by $\#E(\mathbf{Q})_{\text{tors}} = 4$, but often it is equal to 4 times a prime number. If the prime number has more than 50 digits, then the group of this elliptic curve modulo p is suitable for elliptic curve cryptography.

Remark: There's some efficiency advantage in computing multiples kP on such a curve because of the simple form of the coefficients and because of the presence of extra easily-computable endomorphisms of a CM-curve (see Gallant–Lambert–Vanstone).

Another way to get a curve over a prime field:

Generate the coefficients $a, b \in \mathbb{F}_p$ randomly.

Finding $\#E(\mathbb{F}_p)$:

First polynomial time algorithm due to Schoof in 1985;

many speed-ups since then (e.g., Atkins–Elkies using modular forms);

recently a new method was devised by Gaudry–Harley–Mestre that can count points even faster.

It is known that, as the coefficients vary, $\#E(\mathbb{F}_p)$ is distributed more-or-less uniformly throughout the Hasse interval:

$$p+1-2\sqrt{p} \le \#E(\mathbb{F}_p) \le p+1+2\sqrt{p}.$$

OPEN CONJECTURE: $\#E(\mathbb{F}_p)$ has $\approx 1/\log p$ probability of being prime.

It has not even been proved that the Hasse interval around $p+1$ always contains another prime besides p.

Hardliners prefer randomly generated curves.

Extreme example of hardliners:

German Information Security Agency

They insist that, after finding $\#E(\mathbb{F}_p)$, you find the class number h of the CM-field of the curve, which is

$$\mathbb{Q}\left(\sqrt{4p-(p+1-\#E(\mathbb{F}_p))^2}\right),$$

and discard the curve if $h \le 200$.

Why does GISA want this? No reason, really, except that a famous German professor once remarked that he feels a little uneasy about the ECDLP on curves whose CM-field has small class group. However, he didn't mention any ideas for an attack in such a case.

But in Germany there's tremendous reverence toward professors. (There's no danger of this happening in America.)

A final remark about prime fields: It can speed up arithmetic in the field a little to choose the prime p near to a power of 2.

FIELDS HAVING MEDIUM-SIZE SUBFIELDS

By this we mean any field except for a prime field or a prime-degree extension of a prime field of very small characteristic. That is, any field except $\mathbb{F}_p$ or $\mathbb{F}_{2^r}$, $\mathbb{F}_{3^r}$, etc. (with r prime).

Examples that have been proposed at various times:

1. $\mathbb{F}_{p^5}$, where $p = 2^{32} - c$ is a prime close to a power of 2.

> **As mentioned before, arithmetic is efficient in the ground field $\mathbb{F}_p$.**

2. $\mathbb{F}_{2^{8r}}$, where $r \geq 23$ is a prime.

> **Here the ground field $\mathbb{F}_{2^8}$ has size 1 byte.**

3. $\mathbb{F}_{2^N}$ where $N+1$ is a prime having 2 as a primitive root. (J. Silverman)

In this case, since $X^N + X^{N-1} + \cdots + X + 1$ is irreducible, one has

$$\mathbb{F}_2[X]/(X^{N+1}-1) \approx \mathbb{F}_2 \oplus \mathbb{F}_{2^N},$$

and the simple form of the ring on the left allows one to speed up the arithmetic in $\mathbb{F}_{2^N}$.

A threat when there's an intermediate field:

Weil Descent (ideas due to G. Frey)

Suppose that $\mathbb{F}_q$ has an intermediate field $\mathbb{F}_{q_0}$, where $q = q_0^r$. Let E be an elliptic curve over $\mathbb{F}_q$.

In the Weil descent attack one transforms the ECDLP to the Discrete Log Problem (DLP) in more dimensions but over a smaller field.

One tries to transform the ECDLP to the DLP on the Jacobian of a hyperelliptic curve over $\mathbb{F}_{q_0}$ of genus r or not much greater than r.

Then one can use algorithms for the hyperelliptic curve DLP that are significantly faster than the best available ones for the ECDLP.

See the talk to follow by E. Teske.

CURVES OVER $\mathbb{F}_{q^r}$ (r prime, q tiny)

(Perhaps $q = 2$.)

One approach: random coefficients.

Note: It is not even known (for any fixed $q \geq 2$) whether there exist infinitely many primes in the Hasse intervals around $q^r + 1$ for all r.

In the case $q = 2$ this question asks whether there exist infinitely many primes the first half of whose bits after the initial 1 are either all 0 or all 1.

CURVES DEFINED OVER SMALL FIELDS

When a curve E defined over $\mathbb{F}_q$ (with q small) is considered over $\mathbb{F}_{q^r}$, it's very easy to compute $\#E(\mathbb{F}_{q^r})$ once you trivially count $\#E(\mathbb{F}_q)$. Namely, let

$$T^2 + (\#E(\mathbb{F}_q) - q - 1)T + q = (T - \alpha)(T - \overline{\alpha})$$

be the quadratic polynomial associated to E. Then we have

$$\begin{aligned} \#E(\mathbb{F}_{q^r}) &= q^r + 1 - \alpha^r - \overline{\alpha}^r \\ &= |\alpha^r - 1|^2. \end{aligned}$$

Since

$$\alpha^{r+1} + \overline{\alpha}^{r+1} = (\alpha + \overline{\alpha})(\alpha^r + \overline{\alpha}^r) - q(\alpha^{r-1} + \overline{\alpha}^{r-1}),$$

we get a simple recursion for the sequence $\#E(\mathbb{F}_{q^r})$ that's about as easy as the recursion for the r-th Fibonacci number.

Given a curve over a small field $\mathbb{F}_q$, we want to work in an extension $\mathbb{F}_{q^r}$ such that $\#E(\mathbb{F}_{q^r})$ is equal to a prime times a small cofactor. Since $\#E(\mathbb{F}_{q^k})$ divides $\#E(\mathbb{F}_{q^\ell})$ whenever $k|\ell$, we must choose r prime. The best we can hope for is that

$$\frac{\#E(\mathbb{F}_{q^r})}{\#E(\mathbb{F}_q)} = \left|\frac{\alpha^r - 1}{\alpha - 1}\right|^2$$

is a prime.

The question of primality of this ratio is analogous to the Mersenne prime question.

CONJECTURE. For fixed E over $\mathbb{F}_q$ and variable prime r, the probability that

$$\frac{\#E(\mathbb{F}_{q^r})}{\#E(\mathbb{F}_q)} = \left|\frac{\alpha^r - 1}{\alpha - 1}\right|^2$$

is prime is about

$$\frac{e^\gamma}{\log q}\,\frac{\log r}{r},$$

where γ is Euler's constant and $\log$ denotes natural logarithm.

Let's start with the smallest field.

DEFINITION. An anomalous binary curve (ABC) is an elliptic curve

$$Y^2 + XY = X^3 + aX^2 + 1,$$

where $a = 0$ or 1, defined over $\mathbb{F}_2$.

The associated quadratic is

$$T^2 + (-1)^a T + 2 = (T - \alpha)(T - \overline{\alpha}),$$

where

$$\alpha = \frac{-(-1)^a + \sqrt{-7}}{2}.$$

This was originally proposed for ECC in my Crypto '91 paper because I saw some efficiency advantages. J. Solinas greatly improved upon my observations, making a detailed study of expansions in the number ring $\mathbb{Z}[\frac{1+\sqrt{-7}}{2}]$. (See the March 2000 issue of ***Designs, Codes, and Cryptography*****.) Called "magic curves" within the NSA.**

There are 2 basic ideas in speeding up point multiples on ABC curves:

1. You can apply the Frobenius map on points

$$\Phi : \ (x, y) \mapsto (x^2, y^2)$$

for free (if you're using a normal basis of $\mathbb{F}_{2^r}$). This map Φ is the action of the element $\alpha = \frac{-(-1)^a + \sqrt{-7}}{2}$ of the CM-ring $\mathbb{Z}[\alpha]$.

2. Any integer k has an efficient α-adic expansion, called its Non-Adjacent Form (NAF). The NAF allows you to compute kP with a small number of point additions and no point doublings.

Solinas' NAF gives a speedup of "modular exponention" (i.e., the elliptic curve analogue) that would make RSA people green with envy. (Their only successful strategy to get fast modular exponentiation is to use exponent 3. Admittedly, even ABC curves can't compete with that!)

Since $\#E(\mathbb{F}_2) = 3 + (-1)^a$, we work over an extension of prime degree r for which

$$\#E(\mathbb{F}_{2^r}) = \begin{cases} 2 \cdot \textbf{prime} & \textbf{when } a = 1; \\ 4 \cdot \textbf{prime} & \textbf{when } a = 0. \end{cases}$$

For example, with $r = 233$ and $a = 0$ we find that $\#E(\mathbb{F}_{2^r})$ is equal to 4 times a 70-digit prime. That elliptic curve group is very good for practical cryptography, for both security and efficiency reasons.

ABC curves are the <u>only</u> special class that's approved by the standards bodies. (And this took a concerted effort by people at Certicom and at NSA.)