El Gamal Cryptography on an Elliptic Curve

Revised: April 28, 2009

1 Introduction

In the following $\ell$ will denote a prime greater than 2, and $\mathbb{F}_\ell \cong \mathbb{Z}/\ell\mathbb{Z}$ the field of integers modulo $\ell$. We will be talking about “addition”, as previously studied, on a cubic curve $E$ given in Weierstrass form, i.e., $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, with coefficients in $\mathbb{F}_\ell$, and points $(x, y)$ of $E$ will be pairs of elements $x, y$ in $\mathbb{F}_\ell$. A reference for this material is the book *A Course in Number Theory and Cryptography* by Neal Koblitz, published in 1987 by Springer. Some information may also be found online; for example, one might look at Wikipedia.

The basic idea is to use the El Gamal method — which makes sense for a (large) finite cyclic group. The case where the finite cyclic group is the multiplicative group $(\mathbb{Z}/p\mathbb{Z})^*$ for some odd prime $p$ was discussed previously in the course. Here the task is to use the method with a cyclic subgroup of the points on an elliptic curve over $\mathbb{F}_\ell$, when $\ell$ is an odd prime.

2 Representing characters by points on a curve

As before, characters are represented by numbers; in particular, characters and standard symbols in U.S. English may be represented by their ASCII codes, which are integers from 0 to 127. The question here is how to represent a number $N$ in this range by a point of $E$. In the first place $\ell$ must be large enough that $E$ contains at least 127 points. Since for a point $(x, y)$ of $E$ the second coordinate $y$ is the root of a quadratic polynomial in the first coordinate $x$, letting $x$ be $N$ and then solving for $y$ will not lead to a root $y$ in $\mathbb{F}_\ell$ unless the discriminant of the corresponding quadratic equation is the square of an element of $\mathbb{F}_\ell$. Precisely half of the non-zero elements of $\mathbb{F}_\ell$ are squares, so the discriminant will be a square roughly half the time. Because of that $x$ cannot simply be $N$ but rather something determined by $N$ that offers a range of possible values of $x$.

One chooses an integer $m$ so that $1/2^m$ is an acceptably small probability of failure to find a $y$ for given $x$. The idea then is, for a given value $N$, to try as many as $m$ different values of $x$ until there is found a $y$ with $(x, y)$ on $E$. The values of $x$ one tries are

$$x = mN + j, \quad 1 \leq j \leq m.$$

The event that one does not find a $y$ after trying all $m$ of these values has probability $1/2^m$. If a $y$ is found, then the point $p = (x, y)$ becomes the point of the curve representing the number $N$. There is no secrecy in this. The original number $N$ may be recovered from $p$ as the largest integer strictly smaller than $x/m$ or

$$N = \text{floor}\left(\frac{\text{lift}(x) - 1}{m}\right)$$

---

where the function lift returns the least non-negative residue of an integer mod. It is necessary that \( \ell \geq 128m \) if this method is to be viable for representing integers \( 0 \leq N \leq 127 \) by a point on a given curve \( E \) over \( \mathbb{F}_l \).

3 Encoding points on a curve

Inasmuch as the basic El Gamal technique needs a cyclic group, in order to be sure that the points on an elliptic curve obtained by the method of the previous section to represent codes all lie in a cyclic subgroup of \( E(\mathbb{F}_l) \), it is almost necessary to choose \( l \) and \( E \) so that the entire group \( E(\mathbb{F}_l) \) is cyclic. This is, in particular, the case if the size of \( E(\mathbb{F}_l) \) is square-free.

Here the question is encoding for secrecy of the points on a curve. Suppose that \( b \) is a point, regarded as the “base”, of large order relative to the arithmetic on \( E \). This applies, in particular, when the group of points of \( E \) in \( \mathbb{F}_l \) is cyclic and \( b \) is a generator. For example, if the number of all points \( |E| \) of \( E \) happens to be prime, which is far from always true, then any point \( b \) of \( E \) other than the origin has order \( |E| \). As suggested above, for any \( E \) the number \( |E| \) of its points is usually somewhere around \( \ell \) since there are two points on \( E \) for each of the roughly \( \ell/2 \) values of \( x \) for which there is a \( y \) except for the case when \( x \) leads to a quadratic equation for \( y \) having discriminant 0. In this scheme a single point \( p \) on \( E \) will be encrypted by a pair \( (q, r) \) where both \( q \) and \( r \) are points of \( E \).

The designer of the scheme picks the prime \( \ell \), the curve \( E \), a “base point” \( b \) on \( E \) of large order, all of which are to be public, and a secret element \( j \) of \( \mathbb{F}_l \). With those items fixed, the designer publishes one more point \( c \) on \( E \) that is determined by the formula \( c = j b \). For given \( \ell \) and \( E \), the scheme’s “public key” is the pair of points \( b \) and \( c \) on \( E \).

A user of this scheme may encode a point \( p \) of \( E \) as follows: (i) draw a random value \( k \) modulo the order of \( b \) and then (ii) produce the pair of points \( (q, r) \) using the formulae:

\[
q = kb \\
r = p + kc
\]

Anyone who knows the secret value \( j \) as well as the published data may recover the original point \( p \) from the pair \( (q, r) \) using the simple formula

\[
p = r - jq
\]

Security for this system relies on it being difficult to ascertain \( j \) even though \( b \) and \( c \) are both known.