Cryptology Fall 2017

Chloe Martindale
TU/e

October 10, 2017

These notes are based on notes by Tanja Lange. In the last lecture we saw how to construct a group law on a circle. Today we will construct a group law on an Edwards curve, which can be used for a Diffie-Hellman key exchange and is not susceptible to index calculus-style attacks.

1 Clarification from previous lectures: arithmetic in finite fields

Following questions from a significant number of students, we first quickly go over some basics of arithmetic in finite fields of prime order ($\mathbb{F}_p$)).

- Suppose that $n, m \in \mathbb{Z}$. We say that 'n and m are the same mod p' (and we write $n \equiv m \bmod p$ if there exists $a \in \mathbb{Z}$ such that $n = m + ap$. Note that this happens if and only if $p|(n-m)$, which we can think of as 'the remainder when you divide $n - m$ by p is 0'.
- There is a finite list of numbers that are all different mod p, and we can represent $\mathbb{Z}/p\mathbb{Z}$ by any choice of p numbers are not the same mod p. Some different representations of $\mathbb{Z}/7\mathbb{Z}$ are
$$\{-3,-2,-1,0,1,2,3\},$$
$$\{0,1,2,3,4,5,6\},$$
$$\{32,33,34,35,36,37,38\},\ldots$$
- Suppose that $g \in \mathbb{F}_p^*$ has order ℓ, so that
$$\{g, g^2, \ldots, g^\ell\} = \langle g \rangle \subseteq \mathbb{F}_p^*$$
and $g^\ell = 1 = |\langle g \rangle|$. Then calculations with elements in $\mathbb{F}_p^*$ are *mod p*, giving statements like '$g^a \equiv h \bmod p$'. Suppose that there exist $a, b \in \mathbb{Z}$ such that $g^a \equiv g^b \equiv h \bmod p$. Then
$$h \equiv g^a \equiv g^b \equiv 1 \cdot g^b \equiv g^\ell \cdot g^b \equiv g^{b+\ell} \bmod p,$$

and in fact for any $n \in \mathbb{Z}$,

$$h \equiv g^a \equiv g^b \equiv 1^n \cdot g^b \equiv g^{n\ell} \cdot g^b \equiv g^{b+n\ell} \bmod p.$$

So *any* $a \in \{b, b+\ell, b-\ell, b+2\ell, b-2\ell, \cdots\}$ satisfies the equation

$$g^a \equiv h \bmod p.$$

In other words, it is enough to determine *a mod ℓ*. In particular, after we take the log of an equation, all our calculations will be mod ℓ (not mod p).

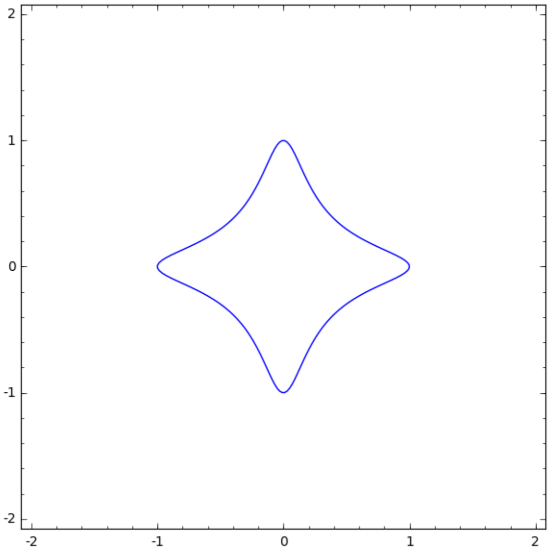
- If you're in doubt about whether to compute mod p or mod ℓ, you can always double-check with $g = -1$ (and $p \neq 2$). Then $\langle g \rangle = \{-1, 1\}$, so $\ell = |\langle g \rangle| = \text{ord}(g) = 2$. In this case, $g^a \equiv 1 \bmod p$ if and only if a is even, i.e., $a \equiv 0 \bmod \ell$.

2 Edwards curves

Having constructed a group law for points on a circle, and noting that the security level of a Diffie-Hellman key exchange is similar to that of finite fields, we try a slightly more complicated curve that will not be susceptible to such attacks: the Edwards curve.

Example. Let's try to make a group from the points on an Edwards curve. We will look first at the example

$$C : x^2 + y^2 = 1 - 30x^2y^2.$$



Note that the equation of C looks similar to the equation of a circle with a 'fudge factor', and we will see that we can construct a group law similar to that of the circle plus this 'fudge factor'. Define

$$G = \{(x, y) \in \mathbb{R} \times \mathbb{R} : x^2 + y^2 = 1 - 30x^2y^2\}.$$

Claim 1. For $(x_1, y_1), (x_2, y_2) \in G$ define

$$(x_1, y_1) \oplus (x_2, y_2) := \left(\frac{x_1y_2 + y_1x_2}{1 - 30x_1y_1x_2y_2}, \frac{y_1y_2 - x_1x_2}{1 + 30x_1y_1x_2y_2} \right).$$

Then $(G, \oplus)$ is a group with neutral element $(0, 1)$.

Proof. We first have to check that we didn't divide by zero, that is, we should check that for $(x_1, y_1), (x_2, y_2) \in G$, we never get that $1 \pm 30x_1x_2y_1y_2 = 0$. If x_1, x_2, y_1, or $y_2 = 0$ then this is clearly non-zero, so suppose that x_1, x_2, y_1, and y_2 are non-zero. Then by the curve equation, for $i = 1, 2$,

$$x_i^2 + y_i^2 + 30x_i^2y_i^2 = 1,$$

and $x_i^2 + y_i^2 > 0$ so

$$30x_i^2y_i^2 < 1,$$

hence

$$\sqrt{30}|x_iy_i| < 1.$$

Therefore

$$30|x_1x_2y_1y_2| = \sqrt{30}|x_1y_1|\sqrt{30}|x_2y_2| < 1 \cdot 1 = 1,$$

so the denominators of the operation $\oplus$ are never zero. We still need to check that it actually defines a group law, that is, that the group axioms (G1)-(G4) are satisfied.

(G1) For the axiom (G1), we have to check that $(x_1, y_1) \oplus (x_2, y_2) \in G$, that is, we have to check that

$$\left(\frac{x_1y_2 + y_1x_2}{1 - 30x_1y_1x_2y_2} \right)^2 + \left(\frac{y_1y_2 - x_1x_2}{1 + 30x_1y_1x_2y_2} \right)^2$$
$$= 1 - 30 \left(\frac{x_1y_2 + y_1x_2}{1 - 30x_1y_1x_2y_2} \right)^2 \left(\frac{y_1y_2 - x_1x_2}{1 + 30x_1y_1x_2y_2} \right)^2,$$

which we can do just by simplification.

(G2) For the axiom (G2), we have to check that if $(x_1, y_1), (x_2, y_2)$, and $(x_3, y_3) \in G$, then

$$((x_1, y_1) \oplus (x_2, y_2)) \oplus (x_3, y_3) = (x_1, y_1) \oplus ((x_2, y_2) \oplus (x_3, y_3)),$$

which we can again doing just by plugging in the formulae and simplifying.

(G3) For the axiom (G3), we have to check that for every $(x, y) \in G$, $(x, y) \oplus (0, 1) = (0, 1) \oplus (x, y) = (x, y)$. We plug $(x_1, y_1) = (x, y)$ and $(x_2, y_2) = (0, 1)$ into our formula for $\oplus$ to get

$$(x, y) \oplus (0, 1) = \left(\frac{x \cdot 1 + 0 \cdot y}{1 - 30x \cdot y \cdot 0 \cdot 1}, \frac{y \cdot 1 - x \cdot 0}{1 + 30x \cdot y \cdot 0 \cdot 1} \right) = (x, y),$$

and similarly for $(0, 1) \oplus (x, y)$.

(G4) For the axiom (G4), we have to check that for every $(x, y) \in G$, there exists $-(x, y) \in G$ such that $(x, y) + (-(x, y)) = (0, 1)$. We claim that $-(x, y) = (-x, y)$:

$$\begin{aligned}(x, y) \oplus (-x, y) &= \left(\frac{xy - xy}{1 - 30x^2y^2}, \frac{x^2 + y^2}{1 + 30x^2y^2} \right) \\ &= (0, 1),\end{aligned}$$

as by the curve equation $x^2 + y^2 = 1 + 30x^2y^2$.

□

Definition 1. Suppose that $d \in \mathbb{F}_q^*$ is a non-square (i.e., that for g a primitive element of $\mathbb{F}_q^*$, $d = g^k$ for k odd). Then the curve

$$C_d : x^2 + y^2 = 1 + dx^2y^2$$

is an *Edwards curve over* $\mathbb{F}_q$.

Note that the example we looked at was C_{-30} but over $\mathbb{R}$. In fact

$$(x_1, y_1) \oplus (x_2, y_2) := \left(\frac{x_1y_2 + y_1x_2}{1 + dx_1y_1x_2y_2}, \frac{y_1y_2 - x_1x_2}{1 - dx_1y_1x_2y_2} \right)$$

defines a group law of C_d just as before. Checking the group axioms is exactly the same process, but the proof that the denominators are non-zero is different, we will write that out. Before we do, let's remind ourselves why we want $(G, \oplus)$ where

$$G = \{(x, y) \in \mathbb{F}_q : x^2 + y^2 = 1 + dx^2y^2\}$$

to be a group in the first place: let's write down a Diffie-Hellman key exchange with our group as $(G, \oplus)$ rather than $(\mathbb{F}_q^*, \cdot)$.

- Setup: Alice and Bob agree on a curve C_d and an element $(x, y) \in G$ that generates a large subgroup of G (via $\oplus$).
- Alice chooses a secret key $a \in \mathbb{Z}$, computes her public key $a \cdot (x, y)$ and sends it to Bob.
- Bob chooses a secret key $b \in \mathbb{Z}$, computes his public key $b \cdot (x, y)$ and sends it to Alice.

- Alice computes the shared secret $b \cdot (a \cdot (x, y)) = (ab) \cdot (x, y)$.
- Bob computes the shared secret $a \cdot (b \cdot (x, y)) = (ab) \cdot (x, y)$.

In order to be able to do this, we need to able to multiply points by an integer, which is defined as adding them to themselves (via $\oplus$), for which we need $\oplus$ to be a group law, which we now prove is at least well-defined.

Claim 2. Suppose that $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ are on C_d, i.e. that for $i = 1, 2$

$$x_i^2 + y_i^2 = 1 + dx_i^2 y_i^2.$$

Then

$$1 \pm dx_1 x_2 y_1 y_2 \neq 0.$$

Proof. Suppose for a contradiction that

$$dx_1 x_2 y_1 y_2 = \pm 1. \tag{1}$$

Then

$$\begin{aligned}
dx_1^2 y_1^2 (x_2 + y_2)^2 &= dx_1^2 y_1^2 (x_2^2 + y_2^2 + 2x_2 y_2) \\
&= dx_1^2 y_1^2 (1 + dx_2^2 y_2^2 + 2x_2 y_2) \\
&= d^2 x_1^2 y_1^2 x_2^2 y_2^2 + dx_1^2 y_1^2 + 2(dx_1 x_2 y_1 y_2) x_1 y_1 \\
&= 1 + dx_1^2 y_1^2 \pm 2x_1 y_1 \\
&= (x_1 \pm y_1)^2,
\end{aligned}$$

but as d is non-square, $dx_1^2 y_1^2 (x_2 + y_2)^2$ is non-square or zero, and $(x_1 \pm y_1)^2$ is square, so we must have that

$$dx_1^2 y_1^2 (x_2 + y_2)^2 = (x_1 \pm y_1)^2 = 0.$$

By the assumption $dx_1 x_2 y_1 y_2 = \pm 1$ that x_1, y_1, x_2, and y_2 are non-zero, and by definition that $d = 0$, hence

$$x_2 + y_2 = 0.$$

But if (x_2, y_2) is on C_d, then $(x_2, -y_2)$ is also on C_d, hence the above argument with $y_2 = -y_2$ gives that

$$x_2 - y_2 = 0,$$

hence $x_2 = y_2 = 0$, which is a contradiction to (1). □

So we have a group under $\oplus$ made up of the $\mathbb{F}_q$-points on C_d, but how easy is arithmetic in this group? It is harder to break Diffie-Hellman in a group with this structure, but arithmetic still needs to be easy enough to perform encryption! Note first of all that doubling a point is actually easier than adding 2 different points:

$$\begin{aligned} 2 \cdot (x, y) &= (x, y) \oplus (x, y) \\ &= \left(\frac{2xy}{1 + dx^2y^2}, \frac{y^2 - x^2}{1 - dx^2y^2} \right) \\ &= \left(\frac{2xy}{x^2 + y^2}, \frac{y^2 - x^2}{2 - x^2 - y^2} \right). \end{aligned}$$

These equations have lower degree than the equations for adding two different points, which means faster computation (we will see later how much faster). Still, we have to do an inversion to compute the sum of 2 points or the double of a point, but we can 'delay' this inversion. So, our aim now is to compute

$$(x_3, y_3) := (x_1, y_1) \oplus (x_2, y_2)$$

with the minimum number of inversions and multiplications. To 'delay' the inversion, we introduce new variables X_i, Y_i, Z_i and substitute $x_i = X_i/Z_i$ and $y_i = Y_i/Z_i$. Then

$$x_3 = \frac{(X_1Y_2 + X_2Y_1)Z_1Z_2}{(Z_1Z_2)^2 + dX_1X_2Y_1Y_2}$$

and

$$y_3 = \frac{Z_1Z_2(Y_1Y_2 - X_1X_2}{(Z_1Z_2)^2 - dX_1X_2Y_1Y_2}.$$

Define

$$\begin{aligned} X_3 &= Z_1Z_2(X_1Y_2 + X_2Y_1)((Z_1Z_2)^2 - dX_1X_2Y_1Y_2), \\ Y_3 &= Z_1Z_2(Y_1Y_2 - X_1X_2)((Z_1Z_2)^2 + dX_1X_2Y_1Y_2), \end{aligned}$$

and

$$Z_3 = ((Z_1Z_2)^2 - dX_1X_2Y_1Y_2)((Z_1Z_2)^2 + dX_1X_2Y_1Y_2).$$

Then $x_3 = X_3/Z_3$ and $y_3 = Y_3/Z_3$, and if we just compute X_3, Y_3, and Z_3 then we don't have to do any inversions! In fact, X_3, Y_3, and Z_3 can be computed in just 10 multiplications (M), one squaring (S), and one multiplication by (D) in the following way:

1. $A = Z_1Z_2$, $B = A^2$, $C = X_1X_2$, $D = Y_1Y_2$. (3M + 1S).
2. $E = dCD$, $F = B - E$, $G = B + E$. (1M + 1D).
3. $X_3 = AF((X_1 + Y_1)(X_2 + Y_2) - C - D)$. (3M).
4. $Y_3 = AG(D - C)$. (2M).
5. $Z_3 = FG$. (1M).

Note that in step 3 we reduced the multiplications by a clever trick:

$$X_1Y_2+X_2Y_1 = X_1Y_2+X_2Y_1+X_1X_2+Y_1Y_2-X_1X_2-Y_1Y_2 = (X_1+Y_1)(X_2+Y_2)-C-D.$$

Doubling can be done in just 4S + 3M, so here we concretely that it is much faster than adding distinct points.

You can also make scalar multiplication faster by precomputing some multiplications of P, e.g., by using that

$$15P = 8P + 4P + 2P + P.$$