ELLIPTIC CURVE CRYPTOGRAPHY

MATH 195

For a reference on elliptic curves and their cryptographic applications, see:

- Alfred J. Menezes, *Elliptic curve public key cryptosystems*, 1993.
- Joseph H. Silverman and John Tate, *Rational points on elliptic curves*, 1992. (An undergraduate mathematics text on elliptic curves.)
- J.W.S. Cassels, *Lectures on elliptic curves*, 1991. (Informal and mathematical.)

An elliptic curve is not an ellipse! An ellipse is a degree 2 equation of the form $x^2 + ay^2 = b$. (However, given such an ellipse, you could try to compute the arc length of a certain portion of the curve; the integral which arises can be associated to an elliptic curve.)

Elliptic curves

Let k be a field, e.g. $k = \mathbb{R}, \mathbb{C}, \mathbb{Q}, \mathbb{F}_q$. An *elliptic curve* E over k is defined by an equation of the form

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

with $a_1, \dots, a_6 \in k$, satisfying a weak condition $\Delta = \Delta(a_1, \dots, a_6) \neq 0$. We number the coefficients this way because we give x degree 2, y degree 3, so that y^2 and x^3 both have degree 6, and then the "degree" of the coefficient records the difference to make every term have degree 6.

We define

$$E(k) = \{(x, y) \in k \times k : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6\} \cup \{O\},$$

where O is the 'point at infinity'. Some explanations are to follow.

Example. If $k = \mathbb{R}$, we take the honest elliptic curve $E : y^2 = x^3 - x$. You can graph this equation by graphing $f(x) = x^3 - x$ and then plotting the squareroot of this function (omitting when $x^3 - x < 0$). Notice that we have two 'loops' (over $\mathbb{R}$), but we may also have only one, as in the case $E : y^2 = x^3 + x$.

The condition $\Delta \neq 0$ is equivalent to every point being a "smooth" point, i.e. there is a uniquely defined tangent line at each point. For example, the curve $y^2 = x^3 + x^2$ has a "node" at $(x, y) = (0, 0)$, and is *not* an elliptic curve as $\Delta = 0$. Similarly, the curve $y^2 = x^3$ has a "cusp" at the origin, and again is not allowed. Do not worry about this condition: any homework problem will be given satisfying this condition.

The 'point at infinity' is the point 'connecting' the second loop in the example. The problem: if you consider the plane, $\mathbb{R}^2$, two lines "usually" intersect, but sometimes lines are parallel. This is an unsatisfactory situation, because too many

---

This is some of the material covered April 30–May 2, in Math 195: Cryptography, taught by Hendrik Lenstra, prepared by John Voight jvoight@math.berkeley.edu.

exceptional circumstances can arise when trying to prove theorems. Instead of this *affine plane*, which we denote $\mathbb{A}^2(\mathbb{R})$, we let

$$\mathbb{P}^2(\mathbb{R}) = \mathbb{A}^2(\mathbb{R}) \cup \{\text{line at infinity}\}.$$

where the line at infinity corresponds to all of the possible slopes. We call this the *projective plane*. Two lines which are parallel in the affine plane have the same slope and hence 'intersect' at one of these additional points in the projective plane. This has other advantages: in most cases, two curves of degree d and e intersect in de points.

Example. We may take $k = \mathbb{Q}$: $E : y^2 = x^3 + 1$. Euler (1706-1783) proved that $E(\mathbb{Q}) = \{(0, \pm 1), (-1, 0), (2, \pm 3)\} \cup \{O\}$. This is all of the points which are rational numbers!

For $E : y^2 = x^3 + x$, we have $E(\mathbb{Q}) = \{O\} \cup \{(0, 0)\}$. A proof: if $x = a/b$, $\gcd(a, b) = 1$, say $b > 0$. Then

$$x^3 + x = \frac{a(a^2 + b^2)a}{b^3} = \frac{c^2}{d^2} = y^2$$

is the square of a rational number, again with $\gcd(c, d) = 1$. Now the numerator and denominator of $a(a^2 + b^2)/b^3$ are also relatively prime: if $p \mid b^3$ and $p \mid a(a^2 + b^2)$, then $p \mid b$ so $p \nmid a$ so $p \mid (a^2 + b^2)$ so $p \mid a^2$ so $p \mid a$, a contradiction. This representation as fractions is unique, so $a(a^2 + b^2) = c^2$, and $b^3 = d^2$. Therefore b must be a square by the second equation, say $b = e^2$. We have $\gcd(a, a^2 + b^2) = 1$, and their product is a square, so they must each be squares: $a = f^2$, $a^2 + b^2 = g^2$. Then $e^4 + f^4 = g^2$. Now by Fermat, this has no solutions in strictly positive integers. This does it! Geez!

Elliptic Curves over $\mathbb{F}_q$

Now let $k = \mathbb{F}_q$, a finite field.

Example. Now consider the curve $E : y^2 = x^3 + x$ over $\mathbb{F}_3$. Then what is $E(\mathbb{F}_3)$? We make a table (using Fermat's little theorem: $x^3 \equiv x \pmod 3$:

x	$x^3 + x = -x$	y
0	0	0
1	-1	-
-1	1	± 1

Notice that -1 is not a square in $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$, so the second line has no corresponding y value. Therefore $E(\mathbb{F}_3) = \{O\} \cup \{(0, 0), (-1, \pm 1)\}$, a group of order 4.

Now consider $k = \mathbb{F}_9 = \mathbb{F}_3[X]/(X^2 + 1)$. What is $E(\mathbb{F}_9)$? Notice that we may write $X = i$, since then $i^2 = X^2 = -1$. (This is purely formal!) We again make a table:

x	x^3+x	y
0	0	0
1	-1	$\pm i$
-1	1	± 1
i	0	0
$i+1$	$(i^3+1^3)+(i+1)=-1$	$\pm i$
$i-1$	1	± 1
$-i$	0	0
$-i+1$	$-(-1)=1$	± 1
$-i-1$	-1	$\pm i$

This gives

$$E(\mathbb{F}_9) = \{O, (1, \pm i), (-1, \pm 1), (\pm i, 0), (\pm(i-1), \pm 1), (\pm(i+1), \pm i), (0,0)\}$$

so $\#E(\mathbb{F}_9) = 16$.

What can we say about $\#E(\mathbb{F}_q)$? There is always the point at infinity, and at most q possibilities for both x and y, so

$$1 \leq E(\mathbb{F}_q) \leq q^2 + 1.$$

(Don't forget the point at infinity!)

In fact, $E(\mathbb{F}_q) \leq 2q+1$: for x there are at most q choices, and then y satisfies a polynomial of degree 2, hence there are at most 2 choices for y given x.

There is a more precise statement, due to Hasse:

Theorem (Hasse, first version)**.** *If E is an elliptic curve over $\mathbb{F}_q$, then*

$$|\#E(\mathbb{F}_q) - (q+1)| \leq 2\sqrt{q}.$$

This is a probablistic calculation: try out every $x \in \mathbb{F}_q$, together with the point at infinity, half of the time there will be a squareroot (and hence two values of y) and half the time not: the average is then $q+1$, and the variance (like flipping a coin q times) is $2\sqrt{q}$. Phrased in another way, we have

$$(\sqrt{q}-1)^2 = q + 1 - 2\sqrt{q} \leq \#E(\mathbb{F}_q) \leq q + 1 + 2\sqrt{q} = (\sqrt{q}+1)^2.$$

In fact, for almost all values in this range, there exists an elliptic curve with that many points. (There are exceptions, but they are completely under control.)

Note that for the example $E : y^2 = x^3 + x$ over $\mathbb{F}_3$, we have $\#E(\mathbb{F}_9) = 16 = (\sqrt{9}+1)^2$, hence this is the richest elliptic curve over $\mathbb{F}_9$, no other can have more points.

Theorem (Hasse, second version)**.** *For every elliptic curve E over $\mathbb{F}_q$ there is a complex number π with $|\pi| = \sqrt{q}$ ($\pi \neq 3.141592\ldots$!) such that for all $n \geq 1$, one has*

$$E(\mathbb{F}_{q^n}) = (\pi^n - 1)(\overline{\pi}^n - 1) = q^n + 1 - (\pi^n + \overline{\pi}^n).$$

We often write $t_n = \pi^n + \overline{\pi}^n$, this is the "trace of Frobenius". Note that

$$|t_n| \leq |\pi^n| + |\overline{\pi}^n| \leq 2\sqrt{q^n}.$$

Taking $n = 1$, we obtain the inequality in the first version of the theorem.

Returning to our example, we see that for $E : y^2 = x^3 + x$, we had

$$\#E(\mathbb{F}_3) = 4 = 3 + 1 - (\pi + \overline{\pi}),$$

and $\pi\overline{\pi} = 3$, so we see that $\pi = \pm\sqrt{3}i$. We check:

$$\#E(\mathbb{F}_9) = (\pi^2 - 1)(\overline{\pi}^2 - 1) = (-4)(-4) = 16.$$

Group Structure on an Elliptic Curve

There is an abelian group law on the set of points of an elliptic curve over a field. Given a field k and an elliptic curve E defined over k, and $P, Q \in E(k)$, how do we compute the "addition" $P + Q$? Property 1: The zero element is the point O at infinity.

Property 2: $-O = O$, $-(x, y) = (x, -y - a_1x - a_3)$. This is the 'other' solution to the equation

$$y^2 + (a_1x + a_3)y - (x^3 + a_2x^2 + a_4x + a_6) = 0.$$

This quadratic has two roots, y_1, y_2, and $y_1+y_2 = -a_1x-a_3$, so $y_2 = -a_1x-a_3-y_1$. If $a_1 = a_3 = 0$, then $-(x, y) = (x, -y)$.

Now take the example $y^2 = x^3 + x$ over $\mathbb{F}_3$. Let us build a table of the points. We have

$+$	O	$(0,0)$	$(-1,1)$	$(-1,-1)$
O	O	$(0,0)$	$(-1,1)$	$(-1,-1)$
$(0,0)$	$(0,0)$			
$(-1,1)$	$(-1,1)$			
$(-1,-1)$	$(-1,-1)$			

Now we see that $(-1, 1) = -(-1, -1)$ by Property 2, so $(-1, 1) + (-1, -1) = O$, and therefore $(0, 0) + (0, 0) = O$. This gives

$+$	O	$(0,0)$	$(-1,1)$	$(-1,-1)$
O	O	$(0,0)$	$(-1,1)$	$(-1,-1)$
$(0,0)$	$(0,0)$	O		
$(-1,1)$	$(-1,1)$			O
$(-1,-1)$	$(-1,-1)$		O	

Finally, this table must represent a group, so you cannot repeat elements in any row or column. Therefore $(-1, 1) + (0, 0) = (-1, -1)$, since it cannot be any of the other three. This allows us to fill out the table:

$+$	O	$(0,0)$	$(-1,1)$	$(-1,-1)$
O	O	$(0,0)$	$(-1,1)$	$(-1,-1)$
$(0,0)$	$(0,0)$	O	$(-1,-1)$	$(-1,1)$
$(-1,1)$	$(-1,1)$	$(-1,-1)$	$(0,0)$	O
$(-1,-1)$	$(-1,-1)$	$(-1,1)$	O	$(0,0)$

In the other cases, we have formulae to compute $P + Q$:

- If $P = O$, then $P + Q = O + Q = Q$. Otherwise $P = (x_1, y_1)$.
- If $Q = O$, then $P + Q = P$. Otherwise $Q = (x_2, y_2)$.
- If $x_1 = x_2 = x$ and $y_1 + y_2 = -a_1x - a_3$, then $P = -Q$ and $P + Q = O$. Otherwise $y_1 + y_2 + a_1x + a_3 \neq 0$.
- If $P \neq Q$ ($x_1 \neq x_2$), compute the (unique) line through the points P, Q, $y = \lambda x + \nu$:

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1};$$

otherwise, if $P = Q$, then we compute the tangent line to the curve at the point P:

$$\lambda = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3}.$$

(Note that we have taken derivatives in a formal sense: use implicit differentiation on the equation for E, and solve for dy/dx.) We always have

$$\nu = y_1 - \lambda x_1.$$

Now intersect $y = \lambda x + \nu$ with the curve

$$x^3 + a_2x^2 + a_4x + a_6 - y^2 - (a_1x + a_3)y = 0.$$

This intersects a degree 3 equation with a line (degree 1), so we expect $3 \cdot 1 = 3$ solutions: two of these are P and Q, so we will obtain another, R. We solve by substitution:

$$x^3 + (a_2 - a_1\lambda - \lambda^2)x^2 + (a_4 - 2\lambda\nu - a_1\nu - a_3\lambda) + (-\nu^2 + a_6 - a_3\nu) = 0.$$

The three roots add up to the negative of the coefficient on x^2:

$$x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2.$$

We let $y_4 = \lambda x_3 + \nu$, and then $y_3 = -y_4 - a_1x_3 - a_3$. The point $P + Q = R = (x_3, y_3)$.

Property 3: If a straight line intersects the curve in points P, Q, R, then $P + Q + R = O$. This is why $P + Q = -R$ above.

Example. Take $y^2 = x^3 + 1$ over $k = \mathbb{R}$. We add the points $P = (-1, 0)$ and $Q = (0, 1)$. We see the line $y = \lambda x + \nu$ that goes through these points is $\lambda = 1$, $\nu = 1$. We see then that

$$x_3 = 1 - x_1 - x_2 = 1 - (-1) = 2$$

and

$$y_3 = -(\lambda x_3 + \nu) = -(2 + 1) = -3.$$

Therefore $(-1, 0) + (0, 1) = (2, -3)$.

Example. Let us verify an entry in our table above, $E : y^2 = x^3 + x$ over $\mathbb{F}_3$, $P = (-1, 1)$, $Q = (0, 0)$. We then compute that $\lambda = -1$, $\nu = 0$, and $x_3 = 1 - (-1) + 0 = -1$, $y_3 = -((-1)(-1)+0) = -1$. So we have verified that $(-1, 1)+(0, 0) = (-1, -1)$.