Elliptic curves over finite fields and applications to cryptography

Erik Wallace

May 29, 2018

1 Introduction

These notes are not as complete or self contained as I would like. For further reading on elliptic curves, the following books are are recommended:

- “Rational points on elliptic curves” by Silverman and Tate,
- “The arithmetic of elliptic curves” by Silverman,
- “Elliptic curves” by Husemöller.

For further reading on cryptography and especially elliptic curve cryptography, the following books are recommended:

- “An Introduction to mathematical cryptography by Hoffstein, Pipher, and Silverman,
- “Handbook of elliptic and hyperelliptic curve cryptography” by Cohen, Frey et al.

though the second deserves some caution due to numerous errors. Also Bernstein’s website

- `https://safecurves.cr.yp.to/`

is indispensable. All of these resources were consulted when developing these notes.

2 Finite fields

Let p be a prime, and let $f \in \mathbb{F}_p[x]$ be a polynomial of degree $n \geq 1$ and suppose it is irreducible, i.e. f does not factor as $f = gh$ with $g, h \in \mathbb{F}_p[x]$ both of degree greater than or equal to 1. Then $\mathbb{F}_p[x]/(f)$ is a field, and is a vector space over $\mathbb{F}_p$ with basis

$$1 + (f),\ x + (f),\ \ldots x^{n-1} + (f),$$

hence $\mathbb{F}_p[x]/(f)$ is a field of p^n elements. From now on we will identify $x + (f)$ with α, where $f(\alpha) = 0$, thus an arbitrary element in a finite field in characteristic p will look like

$$a_0 + a_1\alpha + \cdots a_{n-1}\alpha^{n-1}$$

where $a_i \in \mathbb{F}_p$. Also, we will often denote a finite field with $q = p^n$ elements as $\mathbb{F}_q$. Since the multiplicative group $\mathbb{F}_q^\times$ has order $p^{n-1} - 1$, then for all $\beta \in \mathbb{F}_q^\times$ we have $\beta^{q-1} = 1$, which means that

$$\beta^q = \beta \tag{1}$$

for all $\beta \in \mathbb{F}_q$. As an immediate consequence it follows that all of the elements of $\mathbb{F}_q$ are roots of the polynomial $x^q - x$, and since $\mathbb{F}_q$ has exactly q elements. The polynomial $g(x) = x^q - x$ is *separable*, i.e. it has distinct roots, since by differentiating:

$$g'(x) = qx^{q-1} - 1 = -1$$

over $\mathbb{F}_p$, hence $g(x)$ cannot share any roots with its derivative, since its derivative does not have roots. Therefore, by counting, we see that

$$x^q - x = \prod_{\beta \in \mathbb{F}_q} (x - \beta).$$

What is more, since this is true for any finite field $\mathbb{F}_q$ of $q = p^n$ elements, and since any irreducible polynomial of degree n leads to the construction of such a field, then all irreducible polynomials of degree n over $\mathbb{F}_p$ are factors of $x^{p^n} - x$.

As an additive group $\mathbb{F}_q$ is isomorphic to

$$\underbrace{\mathbb{Z}/p\mathbb{Z} + \mathbb{Z}/p\mathbb{Z} + \cdots \mathbb{Z}/p\mathbb{Z}}_{n \text{ times}},$$

which is clear from the fact that $\mathbb{F}_q$ is an $\mathbb{F}_p$-vector space of dimension n. Meanwhile, the multiplicative group $\mathbb{F}_q^\times$ is cyclic, and thus is isomorphic to

$$\mathbb{Z}/(q-1)\mathbb{Z}.$$

Example 1. The only irreducible polynomial of degree 2 over $\mathbb{F}_2$ is $f(x) = x^2 + x + 1$. Let α denote a root of f. Then $\alpha^2 = \alpha + 1$ (negative signs can be replaced by positive signs in characteristic 2), and

$$\begin{aligned} \alpha(\alpha + 1) &= \alpha^2 + \alpha = 2\alpha + 1 = 1, \\ (\alpha + 1)^2 &= \alpha^2 + 2\alpha + 1 = \alpha^2 + 1 = \alpha + 2 = \alpha, \end{aligned}$$

which completely determines the structure of the multiplicative group, hence we see that $\mathbb{F}_4^\times$ is a cyclic group of order 3, and either α or $\alpha + 1$ are generators. It is also not hard to see that $\alpha + 1$ is the other root of f, and that

$$x^4 - x = x(x+1)(x^2 + x + 1)$$

over $\mathbb{F}_2$.

The field $\mathbb{F}_q$ is a Galois extension of $\mathbb{F}_p$ of degree n. The map

$$\sigma : \beta \mapsto \beta^p \tag{2}$$

is an automorphism of $\mathbb{F}_q$, called the *Frobenius automorphism*, because

$$(\beta_1 + \beta_2)^p = \sum_{i=0}^{p} \binom{p}{i} \beta_1^i \beta_2^{p-i} = \beta_1^p + \beta_2^p$$

over $\mathbb{F}_p$; the other properties being easier to check. From equation (1), we see that σ^n is trivial on $\mathbb{F}_q$, and it is not possible for σ^m to be on $\mathbb{F}_q$ for any $m < n$, since $x^m - x$ is separable giving us a contradiction with the number of elements in $\mathbb{F}_q$. It follows that $\mathbb{F}_q^\times$ is cyclic as claimed above, since there must exist some $\beta \in \mathbb{F}_q$ such that $\sigma^n(\beta) = \beta$ but $\sigma^m(\beta) \neq \beta$ for $m < n$. Furthermore, by the Galois correspondence, the degree of a finite Galois extension must match the order its Galois group, thus by counting, we conclude that

$$\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p) = \langle \sigma \rangle \cong \mathbb{Z}/n\mathbb{Z}.$$

Also, the normal subgroups of a Galois group are in one-to-one correspondence with the Galois subextensions. Since $\mathbb{Z}/n\mathbb{Z}$ is abelian, all of its subgroups are normal and have order $m| \mid n$, and since any finite extension of $\mathbb{F}_p$ is Galois, then it follows that the only subfields of $\mathbb{F}_q$ are those of degree $m \mid n$.

Example 2. A reducible polynomial of degree 4 over $\mathbb{F}_2$ is either

$$(x^2 + x + 1)^2 = x^4 + x^2 + 1$$

or has a root in $\mathbb{F}_2$. A polynomial in $\mathbb{F}_2[x]$ has a root in F_2 if an only if the constant term is zero, or the sum of the coefficients is even. It follows that $f(x) = x^4 + x + 1$ is irreducible. If α is a root of f, then α cannot have order 1 or 3, since it is clearly not in $\mathbb{F}_2$, and having order 3 would imply that α belongs to the subfield $\mathbb{F}_4$, which is not possible since α generates $\mathbb{F}_{16}$. We have

$$\alpha^5 = \alpha^4 \alpha = (\alpha + 1)\alpha = \alpha^2 + \alpha,$$

so α does not have order 5, hence α generates $\mathbb{F}_{16}^\times$. Since α^3 has order 5 in $\mathbb{F}_{16}^\times$ it does not generate $\mathbb{F}_{16}^\times$ and cannot belong to $\mathbb{F}_4$, since that would require it to have order dividing 3. Thus it is true that $\mathbb{F}_{16}$ is generated by α^3 over $\mathbb{F}_2$ *in the sense of fields*, while $\mathbb{F}_{16}^\times$ is not generated by α^3 in the sense of groups. The action of Frobenius on α is as follows

$$\sigma(\alpha) = \alpha^2, \quad \sigma^2(\alpha) = \alpha^4 = \alpha + 1, \quad \sigma^3(\alpha) = \alpha^8 = (\alpha + 1)^2 = \alpha^2 + 1, \quad \sigma^4(\alpha) = \alpha^{16} = \alpha.$$

In particular, since field automorphisms induce permutations on the roots of polynomials that split over the field, then we have all roots of $x^4 + x + 1$ accounted for:

$$x^4 + x + 1 = (x + \alpha)(x + \alpha^2)(x + \alpha + 1)(x + \alpha^2 + 1).$$

Note that, once again we are using the fact that minus signs can be replaced by plus signs, which is possible in characteristic 2, but not in odd characteristic. Finally, note that since $3 \nmid 4$, then $\mathbb{F}_8$ is not a subfield of $\mathbb{F}_{16}$.

Exercises

1. Compute all other irreducible polynomials of degree 4 over $\mathbb{F}_2$.
2. Construct $\mathbb{F}_9$, and find a generator for $\mathbb{F}_9^\times$. Use the Frobenius automorphism to compute the other roots of the irreducible polynomial for the chosen generator.

3 Projective Geometry

Let $f(x, y)$ be a polynomial of degree greater than or equal to 1 (the degree being the highest combined power among the terms of the polynomial). Then the equation $f(x, y) = 0$ defines a plane curve. Such an equation is often called an "affine equation." Degree 1 polynomials define lines, and degree 2 polynomials define conic sections (possibly degenerate). If two lines are drawn randomly, then we expect them to intersect, but parallel lines do not. Or so it would seem, but one of the virtues of projective geometry is that even parallel lines intersect. We accomplish this feat by including extra points, which are often called "points at infinity." If f has degree d, then we can projectivize by substituting $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$ into $f(x, y)$ and multiplying through by Z^d. The result is a "homogeneous" polynomial $F(X, Y, Z)$ of degree d, meaning that all terms have degree d: there are no higher or lower order terms. Points in the projective plane are defined by triples $(X : Y : Z)$ such that X, Y, and Z. At least two things should be clear from the substitution we have done:

1. the "points at infinity" previously mentioned are those with $Z = 0$,
2. Since $x = \frac{X}{Z} = \frac{rX}{rZ}$ and $y = \frac{Y}{Z} = \frac{rY}{rZ}$ for any $r \neq 0$, then it makes sense to consider $(X : Y : Z)$ and $(rX : rY : rZ)$ as representing the same point. In fact, that is precisely what the colons in the notation $(X : Y : Z)$, since it is a standard notation for ratios. In this sense, a point in the projective plane is really an equivalence class of triples.

Taking these facts together, we see that the points at infinity can be brought into the form $(1 : m : 0)$ or $(0 : 1 : 0)$. In particular, any two parallel lines with slope m will intersect at $(1 : m : 0)$ and any two vertical lines will intersect $(0 : 1 : 0)$. The points at infinity taken together form a line with equation $Z = 0$. We do not allow $(0 : 0 : 0)$ to be a projective point: at least one of X, Y, or Z must be non-zero.

Example 3. Consider the hyperbola $f(x, y) = x^2 - xy - 1$. Then

$$Z^2 \cdot f\left(\frac{X}{Z}, \frac{Y}{Z}\right) = X^2 - XY - Z^2$$

So $X^2 - XY - Z^2 = 0$ gives us a projective equation defining the hyperbola. Substituting $Z = 0$ give us

$$X(X - Y) = 0$$

so the points at infinity are $(0 : 1 : 0)$ and $(1 : 1 : 0)$. The lines $X = 0$ and $X - Y = 0$ are the asymptotes of the hyperbola and the points $(0 : 1 : 0)$ and $(1 : 1 : 0)$ respectively lie on them. Thus the asymptotes of the hyperbola intersect the hyperbola at infinity.

If F is a homogeneous polynomial degree d and $r \neq 0$, then

$$F(rX, rY, rZ) = r^d F(X, Y, Z),$$

so $F(rX, rY, rZ) = 0$ if and only if $F(X, Y, Z) = 0$, which allows us to speak unambiguously about a projective point P lying on a curve C defined by a homogeneous equation. As such it is often convenient to write $F(P) = 0$ to mean that $F(X, Y, Z) = 0$ for any triple representing P. It is also convenient to use a representative of a point P that is as simple as possible, to check whether or not $F(P)$ is satisfied. By "as simple as possible" we mean "over the smallest field." The point $(\sqrt{2} : 0 : \sqrt{2})$ is equivalent to $(1 : 0 : 1)$ by rescaling, and so we consider the projective point P to be defined over $\mathbb{Q}$ since the ratios of X, Y, and Z are all rational. On the other hand the square root in $(\sqrt{2} : 0 : 1)$ cannot be eliminated by any choice of r, and if we look at the ratios, we see that the smallest field this point belongs to is $\mathbb{Q}(\sqrt{2})$.

Definition 1. Let C be a curve defined by the homogeneous equation $F(X,Y,Z)=0$. If F has coefficients in K then we say that C is *defined over* K. A K*-rational point* on C is a projective point P such that $F(P)=0$. The set of all K-rational points on a plane curve is denoted by $C(K)$.

If $K=\mathbb{C}$, then $C(K)$ will be infinite by the fundamental theorem of algebra, but for a different choice of K, then $C(K)$ could be finite, and possibly even empty.

Example 4. Let C be the curve defined by $X^2+Y^2+Z^2=0$. Then C is defined over $\mathbb{Q}$, but $C(\mathbb{Q})=\emptyset$. In fact, $C(\mathbb{R})=\emptyset$, because the square of a non-zero real number must be positive, and a sum of a positive number with non-negative numbers must be positive, hence $X=Y=Z=0$, which does not yield a valid projective point.

It turns out that the asymptotes not only intersect the hyperbola, but are also tangent to the hyperbola at the points of intersection. We can see this by computing partial derivatives. Let F be a homogeneous polynomial, and let C be the curve defined by F, and let P be point on the curve, (i.e $F(P)=0$). *If*

$$\frac{\partial F}{\partial X}(P)X+\frac{\partial F}{\partial Y}(P)Y+\frac{\partial F}{\partial Z}(P)Z=0$$

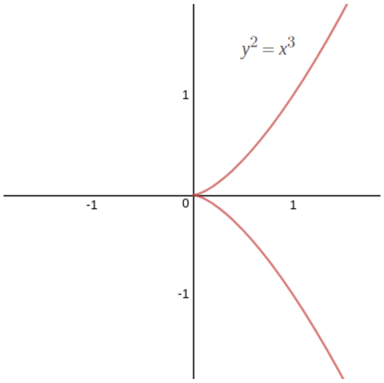
defines a line L, then L is tangent to C at the point P. Why would this equation not define a line? It does not define a line if it vanishes identically.

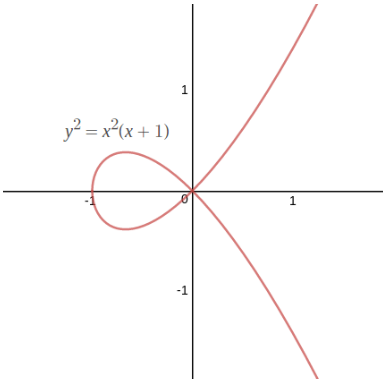
Definition 2. Let C be a curve defined by a homogeneous polynomial F. A point P such that

$$F(P)=\frac{\partial F}{\partial X}(P)=\frac{\partial F}{\partial Y}(P)=\frac{\partial F}{\partial Z}(P)=0$$

is a *singularity* of the curve C. A curve C with no singularities will be called smooth.

Example 5. The two curves in the graphs below are singular at the origin:





The one on the left is called a “cusp” and the one on the right is called a “node.” The two are clearly distinguished by the root multiplicity of the function $f(x)$ on the right hand side. Since the singularities do not occur at infinity it is possible to check that they are singularities without even projectivizing.

For the cusp if $F(x,y)=y^2-x^3$, then

$$\frac{\partial}{\partial x}F(x,y)=-3x^2 \quad \text{and} \quad \frac{\partial}{\partial y}F(x,y)=2y.$$

Clearly all of the functions y^2-x^3, $-3x^3$, and $2y$ evaluate to zero at the origin, hence the curve defined by $F(x,y)=0$ is singular at the origin. See the exercises for the other case.

Example 6. We have seen in example 3 that $(0:1:0)$ and $(1:1:0)$ are points on the hyperbola defined by $F(X,Y,Z) = X^2 - XY - Z^2$. The partial derivatives of F are

$$\frac{\partial F}{\partial X}(X,Y,Z) = 2X - Y, \quad \frac{\partial F}{\partial Y}(X,Y,Z) = -X, \quad \frac{\partial F}{\partial Z}(X,Y,Z) = -2Z,$$

By evaluating at $(0:1:0)$, we obtain the line $-X = 0$, and by evaluating at $(1:1:0)$ we obtain the line $X - Y = 0$. These lines are precisely the asymptotes. Furthermore, the hyperbola is not singular at $(0:1:0)$ and $(1:1:0)$. In fact the hyperbola is not singular at any point, but to see this, we must try to solve the system

$$2X - Y = 0, \quad -X = 0, \quad -2Z = 0, \quad X^2 - XY - Z^2 = 0.$$

The middle two equations imply that $X = 0$ and $Z = 0$. Then by the first equation we must also have $Y = 0$, but $(0:0:0)$ is not a valid projective point.

If a line L is tangent to a curve C at a point P, then we can think of L as intersecting with C at P more than once. This can be seen algebraically in terms of a polynomial having a factor more than once, or analytically in terms of a polynomial having derivatives vanishing up to a certain order.

Example 7. The curve C given by $y = (x-a)^m$ intersects the line $y = 0$ (the x-axis) exactly m times at $(a, 0)$. In terms of projective geometry the homogeneous equations corresponding to C and the x-axis are $(X - aZ)^m - YZ^{m-1} = 0$ and $Y = 0$ respectively. By eliminating the Y variable we see that $(X - aZ)^m = 0$, hence is satisfied by $(a:0:1)$ a total of m times. By eliminating the Z variable instead we also see that C intersects the line at infinity m times at the point $(0:1:0)$.

Definition 3. The number of times that a curve C intersects with a line L at a point P is called the *multiplicity*. In particular if C intersects L with multiplicity 2 at P, then L is tangent to C at P, and if C intersects L with multiplicity 3 at P, then P is an inflection point.

Naturally this definition also extends to intersections between curves, an in particular we have the following result:

Theorem 1 (Bézout's theorem)**.** *If $F_1(X,Y,Z) = 0$ and $F_2(X,Y,Z) = 0$ are homogeneous equations of degree d_1 and d_2 respectively, then the number of points of intersection counted with multiplicity is $d_1 d_2$.*

We have already observed the fact that there is a line at infinity with equation $Z = 0$. The y-axis corresponds to the line $X = 0$, and the x-axis corresponds to the line $Y = 0$. There is no particular reason why we need to choose these three lines. Generally speaking, three randomly drawn lines will determine a projectivization. More precisely, we have the following

Proposition 1. *Suppose we are given three lines L_1, L_2, L_3, where L_j is defined by the equation*

$$a_{1j}X + a_{2j}Y + a_{3j}Z = 0$$

in $(X:Y:Z)$ coordinates. Let $\mathbf{A} = (a_{ij})$ and suppose furthermore that $\det \mathbf{A} \neq 0$. *Then*

$$(U:V:W) = (X:Y:Z)\mathbf{A}$$

defines a change of variables of the projective plane.

Example 8. The asymptotes of the hyperbola in example 3 are the lines $L_1 : X = 0$ and $L_2 : Y - X = 0$. If we take $L_3 : Z = 0$ to be the third line, then $\det \mathbf{A} = 1$. The corresponding change of variables gives us $G(U, V, W) = W^2 - UV$ and by specializing, $G(1, v, w) = 0$ leads to the affine equation $v = w^2$, which is the equation of a parabola. We can also obtain a transformation to the unit circle by completing the square:

$$X^2 - XY - Z^2 = \left(X - \frac{1}{2}Y\right)^2 - \frac{1}{4}Y^2 - Z^2.$$

So, take $L_1 : \frac{1}{2}Y = 0$, $L_2 : Z = 0$, and $L_3 : X - \frac{1}{2}Y = 0$. Then under the change of variables we obtain $G(U, V, W) = W^2 - U^2 - V^2$ and so $G(u, v, 1) = 0$ gives us $u^2 + v^2 = 1$.

Both of the methods in example 8 can be generalized:

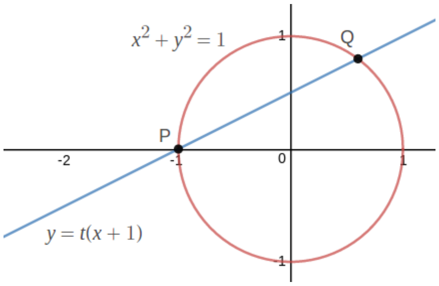
1. Let C be a plane curve defined over K, and suppose that $P \in C(K)$. Then proposition 1 can be applied to obtain an equation $G(U, V, W) = 0$ for C in which the U^d term vanishes. Vanishing of other terms can be seen in terms of the vanishing of partial derivatives.

2. If $f(x, y) = 0$ is an equation for a non-degenerate conic section in characteristic different from 2, then we can always obtain the equation $u^2 + v^2 = 1$ by a change of variables.

The maps in proposition 1 are called "projective linear maps." Projective linear maps are a special case of birational maps, the difference being that we are not limited by degree 1 maps.

Definition 4. Two curves C_1 and C_2 defined over a field K are said to be *birationally equivalent over* K if there exists a rational map $f : C_1 \to C_2$ defined over K with a rational inverse $g : C_2 \to C_1$.

Example 9. We show that the unit circle $x^2 + y^2 = 1$ is birationally equivalent to the projective line $\mathbb{P}^1$ parametrized by the variable t. By using the equation $y = t(x + 1)$ to eliminate the y-coordinate from $x^2 - 1 + y^2 = 0$, we obtain a quadratic in x, whose roots are the x-coordinates of the points P and Q:



$$x^2 - 1 + t^2(x + 1)^2 = 0 \implies (x + 1)(x - 1 + t^2(x + 1)) = 0.$$

The two factors give the x-coordinates of P and Q respectively, thus by solving $x - 1 + t^2(x + 1) = 0$ for x and plugging back into $y = t(x + 1)$ to find y, we obtain the coordinates of Q:

$$x = \frac{1 - t^2}{1 + t^2} \quad \text{and} \quad y = \frac{2t}{1 + t^2}$$

Since $t = \frac{y}{x+1}$, the map is birational.

Exercises

1. Show that $y^2 = x^2(x + 1)$ is singular at the origin.

2. Determine all singularities of the curve $x^2 + y^2 = 1 + dx^2y^2$, where $d \neq 0$, or 1.

3. Let $f(x)$ be a separable polynomial of degree d. Prove that the curve $C : y^2 = f(x)$ is smooth for $d = 2$ and 3, but that it is singular only at infinity for $d \geq 4$.

4. Let $C_1 : y^2 = f(x) = x^4 + ax^3 + bx^2 + cx$, where f is separable (in particular $c \neq 0$). Show that $(x, y) = (\frac{1}{cu}, \frac{v}{u^2})$ defines a birational equivalence, between C_1 and C_2, where $C_2 : v^2 = g(u)$ and $g(u)$ is a monic cubic polynomial.

4 Geometry of elliptic curves

Let C be a smooth curve defined by a cubic equation, and suppose that $C(K) \neq \emptyset$. Then, there is a very nice geometrical way to define addition on C:

1. Choose a point $\mathcal{O}$ in $E(K)$ as the identity.
2. Let P and Q be any two points in $E(K)$, and let L_1, be the line through them (the line tangent to E at P if $P = Q$). Then L_1 intersects E at a third point R.
3. Let L_2 be the line through $\mathcal{O}$ and R. Then the third point of intersection of L_2 and E is $P + Q$.

Smoothness is required here because otherwise the tangent line at a point P may not be well defined.

Example 10. Consider the curve $C : y^2 = x^3 + x^2 - 2x + 1$. Take $\mathcal{O}$ to be the "point at infinity." It is easy to check that $P = (-2, 1)$ and $Q = (0, -1)$ are points on the curve. We can compute $P + Q$ as follows. The line L_1 through P and Q, has equation $y = -x - 1$. We can find the third point of intersection by eliminating the y variable in the equation for C

$$0 = x^3 + x^2 - 2x + 1 - (x + 1)^2 = x^3 - 4x = x(x - 2)(x + 2)$$

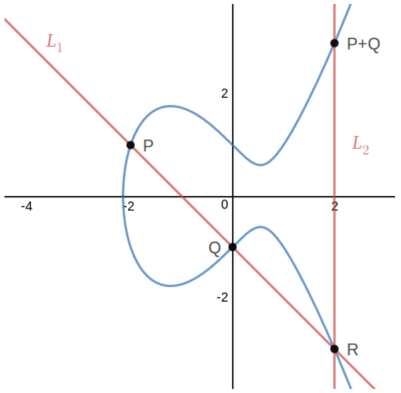
Even though we are left with a cubic, it is easy to factor since two of its roots are known, namely the x-coordinates of P and Q (the two known points of intersection of C and L_1). The third root is the x-coordinate of the unknown point of intersection, R, and to get the y-coordinate, we plug back into the equation for L_1. Thus we find $R = (2, -3)$. Since $\mathcal{O}$ is the point at infinity, the line L_2 is vertical and $P + Q$ is the reflection of R across the x-axis. Thus $P + Q = (2, 3)$. If we want to compute $2P$, we first need to compute the slope at P. By implicit differentiation:

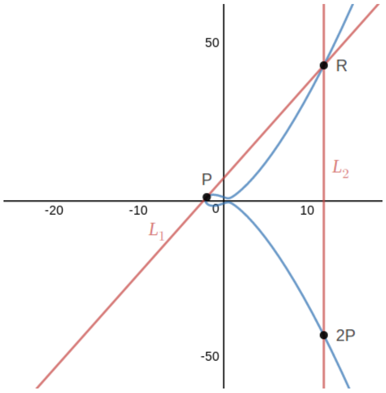
$$2y\frac{dy}{dx} = 3x^2 + 2x - 2.$$

Hence, if $P = (x_0, y_0)$ is a point on C, then the slope at P is $\frac{3x_0^2-2x_0+1}{2y_0}$. In particular, for $P = (-2, 1)$ we find that the slope is 3, thus L_1 is given by $y = 3x + 7$. We can find the x-coordinate of R the same way as before

$$0 = x^3 + x^2 - 2x + 1 - (3x + 7)^2 = x^3 - 8x^2 - 44x - 48 = (x - 12)(x + 2)^2.$$

So $R = (12, 43)$, and after reflecting across the x-axis, we find $2P = (12, -43)$.





It can be shown that a cubic curve C defined over K with a point $\mathcal{O} \in C(K)$ is birrationally equivalent to a curve of the form

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \tag{3}$$

where the point $\mathcal{O}$ has been moved to $(0 : 1 : 0)$. These days, such an equation is called a "Weierstrass equation," even though he preferred the form

$$y^2 = 4x^3 - g_2x - g_3. \tag{4}$$

Equation (4) makes a great deal of sense analytically if you are working over $\mathbb{C}$, but $\mathbb{C}$ has characteristic 0, and it is not possible to get from the form (3) to the form (4) if K has characteristic 2 or 3, which can be seen as follows.

If $\mathrm{char}(K) \neq 2$, then we can complete the square

$$y^2 + a_1xy + a_3y = \Big(y + \frac{a_1}{2}x - \frac{a_3}{2}\Big)^2 - \frac{a_1^2}{4}x^2 - \frac{a_1a_3}{2}x - \frac{a_3}{4},$$

hence we can bring (4) into the form

$$v^2 = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4} \tag{5}$$

where

$$v = y + \frac{a_1}{2}x - \frac{a_3}{2}, \quad b_2 = a_1^2 + 4a_2, \quad b_4 = a_1a_3 + 2a_4, \quad b_6 = a_3^2 + 4a_6.$$

If $\mathrm{char}(K) \neq 3$ either, then we can eliminate the x^2 term from (5), by completing the cube. Let

$$u = x + \frac{b_2}{12}, \quad c_4 = b_2^2 - 24b_4, \quad c_6 = b_2^3 - 36b_2b_4 + 216b_6$$

then

$$v^2 = u^3 - \frac{c_4}{48}u - \frac{c_6}{864}, \tag{6}$$

which differs from (4) only by rescaling. The correct values of c_4 and c_6 are easy to obtain with sage using the following code:

```
R.<b2,b4,b6,x,u>=QQ[]
f=x^3+(b2/4)*x^2+(b4/2)*x+(b6/4)
f(x=u-b2/12)(u=0).factor()
f(x=u-b2/12).derivative(u,1)(u=0).factor()
```

A plane curve C defined by a cubic equation and possessing a rational point is an elliptic curve if and only if it is non-singular. What does that mean for equations (3), (5), and (6)? If K is a field with $\mathrm{char}(K) \neq 2$ and $a, b, c \in K$, then an equation of the type

$$y^2 = x^3 + ax^2 + bx + c$$

is singular if and only if the discriminant of the right hand side is zero. If we apply this to equations (5) and (6), then we obtain the following discriminants with the help of sage:

```
R.<b2,b4,b6,c4,c6,x,u>=QQ[]
f1=x^3+(b2/4)*x^2+(b4/2)*x+(b6/4)
```

```
f1.discriminant(x).factor()
```

Output:

$$-\frac{1}{64}(-b_2^2b_4^2 + b_2^3b_6 + 32b_4^3 - 36b_2b_4b_6 + 108b_6^2) \tag{7}$$

```
f2=x^3+(c4/48)*x+(c6/864)
f2.discriminant(x).factor()
```

Output:

$$-\frac{1}{27648}(-c_4^3 + c_6^2) \tag{8}$$

These formulas are not helpful if $\mathrm{char}(K) = 2$. We cannot divide by 2, and the even coefficients that show up inside the parentheses in equation (7) end up being zero, which gets rid of useful information. But we note that 32, 36 and 108 are all divisible by 4, and

$$b_2^2b_4^2 - b_2^3b_6 = b_2^2(b_4^2 - b_2b_6) = 4b_2^2(a_1^2a_6 - a_1a_3a_4 + a_2a_3^2 + 4a_2a_6 - a_4^2),$$

so if we define

$$b_8 = a_1^2a_6 - a_1a_3a_4 + a_2a_3^2 + 4a_2a_6 - a_4^2,$$

then (7) reduces to $\frac{\Delta}{16}$, where

$$\Delta = b_2^2b_8 + 9b_2b_4b_6 - 8b_4^3 - 27b_6^2. \tag{9}$$

It turns out that by defining Δ in this way, (3) is non-singular if and only if $\Delta \neq 0$. Since the discriminants (7) and (8) are equal, then in the case where $\mathrm{char}(K) \neq 2$ or 3 we also have

$$\Delta = \frac{1}{1728}(c_4^3 - c_6^2).$$

The advantage of (9) is that it holds even when $\mathrm{char}(K) = 2$ or 3, that is if we compute Δ using (9), and we discover that $2|\Delta$ or that $3|\Delta$, then the curve (3) is not an elliptic curve in those characteristics respectively.

It is also useful to have a way of classifying elliptic curves up to birational equivalence over an algebraically closed field. The *j-invariant*

$$j = \frac{c_4^3}{\Delta} \tag{10}$$

does exactly this, but if two elliptic curves E_1 and E_2 are defined over a field K that is *not* algebraically closed, then we may need to pass to an algebraic extension of K to find a birational equivalence.

Example 11. Consider the elliptic curves

$$E_1 : y^2 = x^3 + bx + c \quad \text{and} \quad E_2 : ds^2 = r^3 + br + cr,$$

defined over $\mathbb{Q}$, where d is not a square. First, we need to bring E_2 into Weierstrass form. If we multiply though by d^3, and substitute $d^2s = v$ and $dr = u$, then we obtain the equation

$$v^2 = u^3 + bd^2u + cd^3$$

for E_2. If we had $d = 1$, then we would have the equation for E_1, but d is not a square.

We compute the j-invariant as follows:

$$\begin{gathered}a_1 = a_2 = a_3 = 0, \quad a_4 = bd^2, \quad a_6 = cd^3, \\ b_2 = 0, \quad b_4 = 2bd^2, \quad b_6 = 4cd^3, \quad b_8 = -b_2d^4, \\ c_4 = -48bd^2, \quad c_6 = 864cd^3, \\ \Delta = -16(4b^3 + 27c^2)d^6, \quad j = 1728 \cdot \frac{4b^3}{4b^3 + 27c^2}.\end{gathered}$$

Since j does not depend on d, then E_1 and E_2 have the same j-invariant. The change of variables $x = r$ and $y = s\sqrt{d}$ is a birational map between E_1 and E_2 defined over $\mathbb{Q}(\sqrt{d})$ but not over $\mathbb{Q}$. So long as d is not a square, we cannot do better than this.

Definition 5. Two elliptic curves E_1 and E_2 defined over K, are said to be *quadratic twists* if they are birationally equivalent over a quadratic extension of K, but not over K.

In the case where K is a number field (i.e. an algebraic extension of $\mathbb{Q}$) there are infinitely many quadratic twists up to birational equivalence over K itself. In the example above, simply take prime values for d. But if K is a finite field, there is essentially only one quadratic twist up to birational equivalence, thus over a finite field we will usually refer to *the* quadratic twist of an elliptic curve. In one of the exercises in this section, it is shown that

$$E_1 : y^2 = x^3 + Ax^2 + x \quad \text{and} \quad E_2 : By^2 = x^3 + Ax^2 + x, \tag{11}$$

where $A, B \in \mathbb{F}_p$ and $\left(\frac{B}{p}\right) = -1$ are quadratic twists. In elliptic curve cryptography, such curves are called *Montgomery curves.* Clearly, a different choice of a quadratic non-residue amounts only to rescaling the y-coordinate, so there is the birational equivalence right there. Furthermore, for any $x \in \mathbb{F}_p$, if $f(x) \neq 0$ in $\mathbb{F}_p$, then $\left(\frac{f(x)}{p}\right) = \pm 1$, meaning that each $x \in \mathbb{F}_p$ is an x-coordinate of a point on either E_1 or its twist E_2, and if there are two such points on the same curve (i.e. if $y \neq 0$), then that x value cannot be the x-coordinate of a point on the other curve.

We now discuss the general algebraic formulas for adding points on an elliptic curve E with equation (3), along with some special cases. First, if $P = (x_0, y_0)$ is a point on E, then the vertical line through P has equation $x = x_0$. By plugging in, we obtain

$$y^2 + a_1x_0y + a_3y = x_0^3 + a_2x_0^2 + a_4x_0 + a6.$$

Since P is a point on E, then the right hand side is equal to $y_0^2 + a_1x_0y_0 + a_3y_0$. Subtracting this quantity from both sides and factoring out $y - y_0$ gives us

$$(y - y_0)(y + y_0 + a_1x_0 + a_3).$$

The solution $y = y_0$ corresponds to the point $P = (x_0, y_0)$, which was already known to us. The other solution is

$$-P = (x_0, -y_0 - a_1x_0 - a_3). \tag{12}$$

Now suppose we have two points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ on E. If $x_1 = x_2$, then the line through P and Q is vertical, and we have $P + Q = \mathcal{O}$, otherwise the line through P and Q has a well defined slope given by

$$\lambda = \begin{cases} \frac{3x_1^2+2a_2x_1+a_4-a_1y_1}{2y_1+a_1x_1+a_3} & \text{if } P = Q \\ \frac{y_2-y_1}{x_2-x_1} & \text{otherwise.} \end{cases} \tag{13}$$

In the case $P = Q$, λ is the slope of the tangent line at P. When $P \neq Q$, the calculation of λ is clear, and if $P = Q$ then implicit differentiation can be applied to equation (3); see the exercises. The y-intercept of the line through P and Q is given by

$$\nu = \begin{cases} \frac{-x_1^3+a_4x_1+2a_6-a_3y_1}{2y_1+a_1x_1+a_3} & \text{if } P = Q \\ \frac{y_1x_2-y_2x_1}{x_2-x_1} & \text{otherwise.} \end{cases} \tag{14}$$

In the case $P \neq Q$, we have $y_1 = \lambda x_1 + \nu$ and $y_2 = \lambda x_2 + \nu$, so

$$x_2y_1 - y_2x_2 = x_2(\lambda x_1 + \nu) - x_1(\lambda x_2 + \nu) = \lambda(x_2x_1 - x_1x_2) + \nu(x_2 - x_1) = \nu(x_2 - x_1)$$

and dividing through by $x_2 - x_1$. In the case $P = Q$, we simply solve $y_1 = \lambda x_1 + \nu$ for ν; see the exercises. The equation for the line L_1 through P and Q is therefore

$$y = \lambda x + \nu. \tag{15}$$

We use this equation to eliminate y from (3), resulting in a cubic in x, specifically

$$x^3 - (\lambda^2 + a_1\lambda - a_2)x^2 + (a_4 - a_3\lambda - a_1\nu - 2\lambda\nu)x - (\nu^2 + a_3\nu - a_6). \tag{16}$$

Usually a general cubic would require Cardano's formula to solve, however we already know two roots: if $P \neq Q$, then x_1 and x_2 are roots, and if $P = Q$, then x_1 is a double root. The remaining unknown root x_3 is the x-coordinate of R. For a cubic in monic form,

$$(x - x_1)(x - x_2)(x - x_3) = x^3 - s_1x^2 + s_2x - s_3,$$

where s_1, s_2, and s_3 are the elementary symmetric polynomials in x_1, x_2, x_3, namely

$$s_1 = x_1 + x_2 + x_3, \quad s_2 = x_1x_2 + x_2x_3 + x_3x_1, \quad s_3 = x_1x_2x_3.$$

Since x_1 and x_2 are known, then choosing any of the lower degree coefficients in (16) will yield a linear equation in x_3, which can be solved, but among the options the x^2 coefficient is the easiest to work with, thus

$$x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2. \tag{17}$$

The y-coordinate of R can then be obtained by plugging into (15), thus we have $R = (x_3, \lambda x_3 + \nu)$. Since $P + Q + R = \mathcal{O}$, then $P + Q = -R$, hence if y_3 denotes the y-coordinate of $P + Q$, then by (12) we have

$$y_3 = -(\lambda + a_1)x_3 - \nu - a_3. \tag{18}$$

In the case where $P = Q$, the denominator of λ is not just anything, but rather it is recognizably equal to $2v$ in the derivation of equation (5) above. Therefore, all of the terms in the formula for x_3 can be put over the common denominator

$$4v^2 = 4x_1^3 + b_2x_1^2 + 2b_4x_1 + b_6.$$

Remarkably, the numerator also simplifies, and so the x-coordinate of $2P$ is

$$x_3 = \frac{x_1^4 - b_4x_1^2 - 2b_6x - b_8}{4x_1^3 + b_2x_1^2 + 2b_4x_1 + b_6}; \tag{19}$$

see the exercises. In the case of Montgomery curves ($a_1 = a_3 = a_6 = 0, a_4 = 1, a_2 = A$) when $B = 1$, this formula reduces to

$$x_3 = \frac{(x_1^2 - 1)^2}{4(x_1^3 + Ax_1^2 + x_1)}$$

A formula that remains valid even for the twist (see the exercises). If we projectivize, and if $P = (X : Y : Z)$, then we compute

$$X(2P) = (X^2 - Z^2)^2 \quad Z(2P) = 4XZ(X^2 + AXZ + Z^2). \tag{20}$$

Naturally $Y(2P)$ can be computed using the tangent line at P, but the fact is, if we are only interested in computing $X(nP)$, $Y(nP)$, and $Z(nP)$, then $Y(nP)$ can be ignored completely during the calculation. When we say $X(nP)$, $Y(nP)$, and $Z(nP)$, of course, these values are only determined up to rescaling by $r \neq 0$, but that is a major virtue, because it means we do not need to get $Z(nP) = 1$ every time in the middle of a calculation, but save that step for the very end. Specifically, let $nP = (X_n : Y_n : Z_n)$ where for each n the particular choice of representative among the equivalence class of triples is irrelevant, but we may assume that it is fixed throughout the calculation until possibly the very end, where we compute Z_n^{-1} and rescale. Then we have the following recursive formula among the X's and Z's for $nm(n - m)(n + m) \neq 0$:

$$\begin{aligned} X_{m+n} &= Z_{m-n}((X_m - Z_m)(X_n + Z_n) + (X_m + Z_m)(X_n - Z_n))^2, \\ Z_{m+n} &= X_{m-n}((X_m - Z_m)(X_n + Z_n) - (X_m + Z_m)(X_n - Z_n))^2. \end{aligned}$$

It may not be immediately clear which values of the indices are needed by the recursion, and indeed there is usually more than one way, but the systematic approach is to take $m - n$ to be a power of 2. Then, if we want to compute $X(nP)$ and $Y(nP)$, the only values of the indices we need can be found from the binary expansion of n and its bitwise complement (plus one).

Example 12. Consider the elliptic curve $E : y^2 = x^3 + 71x^2 + x$ over $\mathbb{F}_{8147}$. Let $P = (3347 : 4936 : 1)$. In order to compute $1630P$, we will need all powers of 2, less than or equal to 1630. Also since $1630 = (11001011110)_2$, we will need its bit complement plus one $418 = (110100010)_2$, and all integers formed by successively dropping off leading ones in the binary expansions of 1630 and 417, that is we will need

$$\begin{aligned} &606 = (1001011110)_2,\ 94 = (1011110)_2,\ 30 = (11110)_2,\ 14 = (1110)_2,\ 6 = (110)_2,\ 2 = (10)_2 \\ &162 = (10100010)_2,\ 34 = (100010)_2,\ 1 = (1)_2 \end{aligned}$$

Then the computation of X_{1630} and Z_{1630} proceeds as follows

n	m	$n-m$	$n+m$	$(X_n : Z_n)$	$(X_m : Z_m)$	$(X_{n-m} : Z_{n-m})$	$(X_{n+m} : Z_{n+m})$
4	2	2	6	(2715:2157)	(3519:750)	(3519:750)	(5555:2775)
8	6	2	14	(5824:1119)	(5555:2775)	(3519:750)	(293:5074)
16	14	2	30	(1870:3852)	(293:5074)	(3519:750)	(2209:6486)
32	2	30	34	(5901:2181)	(3519:750)	(2209:6486)	(4065:5724)
64	30	34	94	(1208:5559)	(2209:6486)	(4065:5724)	(6800:2686)
128	34	94	162	(1681:5854)	(4065:5724)	(6800:2686)	(6848:5259)
256	162	94	418	(4774:3208)	(6848:5259)	(6800:2686)	(3875:3172)
512	94	418	606	(2017:2213)	(6800:2686)	(3875:3172)	(4068:4538)
1024	606	418	1630	(5393:254)	(4068:4538)	(3875:3172)	(3590:6354)

Then it can be checked that $6354^{-1} \equiv 986 \bmod 8147$ and $3590 \times 986 \equiv 3942 \bmod 8147$, which matches the x-coordinate of $1630 \cdot P$ as computed by sage.

In elliptic curve cryptography, an *Edwards elliptic curve* over a field K of characteristic different from 2 is given by an equation

$$x^2 + y^2 = 1 + dx^2y^2 \tag{21}$$

where $d \neq 0$ or 1. By taking the identity element to be $(0, 1)$, the addition law can be defined by

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{x_1y_2 + x_2y_1}{1 + dx_1x_2y_1y_2}, \frac{y_1y_2 - x_1x_2}{1 - dx_1x_2y_1y_2} \right). \tag{22}$$

If $P = (x_0, y_0)$, then $-P = (-x_0, y_0)$. Unlike in the case of Weierstrass equations, there is no need for a separate doubling formula. Furthermore, though there are points at infinity on the curve, they play no role in the group law.

Example 13. Consider the curve $x^2 + y^2 = 1 + 12x^2y^2$ over $\mathbb{F}_{41}$. The point $P = (13, 33)$ can be shown to have order 13. For $n = 10$, we first express 10 as a sum of powers of 2

$$10 = 2 + 8,$$

then we double repeatedly until we reach 8

$$2P = (16, 34),\ 4P = (24, 36),\ 8P = (23, 4),$$

and finally we add

$$10P = 2P + 8P = (12, 39).$$

We now show how construct a birational map from an Edwards elliptic curve to a curve in Weierstrass form. We begin by putting everything on one side of the equation, and collecting terms:

$$y^2(1 - dx^2) + x^2 - 1 = 0.$$

If both x and y to be rational over K, then as a quadratic in y it must factor over K, which occurs if and only if the discriminant is a square. The discriminant is

$$-4(1 - dx^2)(x^2 - 1) = 4(dx^4 - (d + 1)x^2 + 1),$$

so if we let

$$z^2 = dx^4 - (d + 1)x^2 + 1 \tag{23}$$

then the quadratic formula gives us $y = \pm\frac{z}{1-dx^2}$, but since y and z only occur to even powers, the sign can be absorbed into z, say, so essentially this defines a birational map between (21) and (23). Now equation (23) is closely related to the theory of 2-descent of elliptic curves, and from that theory, an elliptic curve

$$E : s^2 = r^3 + ar^2 + br \tag{24}$$

is birationally equivalent to a curve of the form

$$C : z^2 = 1 - 2ax^2 + (a^2 - 4b)x^4 \tag{25}$$

via the maps

$$\begin{array}{ccc} E \longrightarrow C & \qquad & C \longrightarrow E \\ (r,s) \longmapsto \left(\dfrac{s}{r^2+ar+b}, \dfrac{r^2-b}{r^2+ar+b} \right) & \qquad & (x,y) \longmapsto \left(\dfrac{z-ax^2+1}{2x^2}, \dfrac{z-ax^2+1}{2x^3} \right) \end{array}$$

By comparing equations (23) and (25) we see that they are the same when $2a = d+1$ and $d = a^2 - 4b$. If we then solve for a and b, then (24) can be brought into the form

$$s^2 = r^3 + \left(\frac{d+1}{2}\right) r^2 + \left(\frac{d-1}{4}\right)^2 r. \tag{26}$$

We have thus established birational equivalence to Weierstrass form, however, by just one more step we can obtain birational equivalence to Montgomery form. Specifically, if we divide through by $\left(\frac{d-1}{4}\right)^3$, then under the substitution

$$u = \frac{4r}{d-1} \quad \text{and} \quad v = \frac{4s}{d-1}$$

equation (26) becomes

$$Bv^2 = u^3 + Au^2 + u \quad \text{with} \quad A = 2 \cdot \frac{d+1}{d-1} \quad \text{and} \quad B = \frac{4}{d-1}. \tag{27}$$

Exercises

1. Compute the formulas for (13) and (14) (they are reasonable to do by hand).
2. (Optional) Finish the derivation of (19) with the help of sage or magma.
3. The point $P = (0:0:1)$ is always on the curve $Y^2Z = X^3 + AX^2Z + XZ^2$, and has order 2. Apply the doubling formula (22) to P. Why isn't the result $\mathcal{O}$? *Hint: go back to the derivation of* (19) *and read through the steps. There is an assumption made that does not apply to* $P = (0:0:1)$.
4. Compute Δ and j for $By^2 = x^3 + Ax^2 + x$. Conclude that such a curve is never an elliptic curve in characteristic 2, and that the curves E_1 and E_2 given by (11) are quadratic twists of each other.
5. Compute the j-invariant of an Edwards curve using either equation (26) or (27).
6. Prove that $(0,-1)$ is a point of order 2, and $(\pm 1, 0)$ are points of order 4 on an Edwards curve.
7. Prove that if $p \equiv 3 \bmod 4$, then $y^2 = x^3 - Ax^2 + x$ is the twist of $y^2 = x^3 + Ax^2 + x$.
8. Prove that the doubling formula 20 remains valid for $B \neq 1$. The proof should work for $\mathrm{char}(K) \neq 2$.
9. If we projectivize equation (21) in the usual way with $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$, and if we projectivize (26) with $r = \frac{R}{T}$ and $s = \frac{S}{T}$, then the map directly from (21) to (26) is given by

   $$R = X\left((Z^2 - dX^2)Y - \frac{d+1}{2}X^2Z + Z^3\right),\ S = Z\left((Z^2 - dX^2)Y - \frac{d+1}{2}X^2Z + Z^3\right),\ T = 2X^3Z.$$

   Show that the points $(1:0:0)$ and $(0:1:0)$ do not map to anything. Conclude that the rational points on (26) including $(0:1:0)$ are in bijection with the rational points on (21) except for $(1:0:0)$ and $(0:1:0)$.

5 Torsion

Since an elliptic curve E is a group, and then it makes sense to ask for the order of a point P. The order of a point P is the least m such that

$$mP = \underbrace{P + P + \cdots P}_{m \text{ times}} = \mathcal{O}. \tag{28}$$

if such m exists, and it is infinite otherwise. We will also introduce the following terminology:

Definition 6. A point P on E such that (28) is satisfied for some $m \in \mathbb{Z}+$ is called a *torsion* point. For fixed $m \in \mathbb{Z}^+$, the set of all points of E satisfying (28) is a subgroup of E called the *m-torsion* of E, which will be denoted here by $E[m]$.

Note that saying that P is an m-torsion point does not necessarily mean that P has order m, it means rather that p has order dividing m. There is quite a bit of theory involving the torsion of elliptic curves, some of which we will be exposed to in the course of these notes. We begin with the simplest cases.

Suppose E is defined over a field K, where $\operatorname{char}(K) \neq 2$, then E can be expressed by an equation of the form

$$y^2 = f(x) = x^3 + ax^2 + bx + c. \tag{29}$$

Suppose P is a 2-torsion point on this curve, so $2P = \mathcal{O}$. By the doubling formula, we see that we get $\mathcal{O}$ precisely when $y = 0$, which means that we can have a rational 2-torsion point on E if and only if the cubic on the right hand side of (29) has a rational root. If $K = \mathbb{Q}$, this is easy to check by the rational root theorem. If $K = \mathbb{F}_q$, this is easy to check by plugging in.

We also see from the doubling formula that the slope of the line tangent to E at P is vertical, a fact which can also be seen from chord tangent addition as follows. Working backwards, we see that $P * P$ must be the third point on the line tangent to the curve at $\mathcal{O}$. The line tangent to the curve at $\mathcal{O}$ is the line at infinity, which intersects triply with E at $\mathcal{O}$, hence $P * P = \mathcal{O}$, or in other words the line tangent to E at P has $\mathcal{O}$ as the third point of intersection, which means it is vertical.

From the above description of lattices, it should be clear that $E[2] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. One of the points in $E[2]$ is $\mathcal{O} = (0 : 1 : 0)$, which is always rational. The remaining 3 non-trivial points of $E[2]$ each corresponds with a root of the cubic $x^3 + ax^2 + bx + c$. Three things can happen:

1. $x^3 + ax^2 + bx + c$ has no rational roots (is therefore irreducible over $\mathbb{Q}$), and so E has trivial 2-torsion over $\mathbb{Q}$, i.e. only $\mathcal{O}$ which can be thought of as generating the trivial subgroup,

2. $x^3 + ax^2 + bx + c$ has one rational root and one irreducible quadratic factor, and so the rational 2-torsion of E is isomorphic to $\mathbb{Z}/2\mathbb{Z}$.

3. $x^3 + ax^2 + bx + c$ has three rational factors, so E has "full 2-torsion" over $\mathbb{Q}$, i.e. all points in $E[2]$ are rational and so the rational 2-torsion of E is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

If $f(x)$ is a polynomial with h as a root, then it is easy to see that $f(x + h)$ has zero as a root (plug in). By applying this idea to the polynomial on the right hand side of (29), we can obtain a new equation for E, in the same form, except with $c = 0$. As a result it is often convenient to bring an elliptic curve with a non-trivial 2-torsion point into the form

$$y^2 = x^3 + ax^2 + bx = x(x^2 + ax + b). \tag{30}$$

From here, it is easy to check whether or not E has full 2-torsion by checking whether or not the discriminant of $x^2 + ax + b$ is a square.

We have not yet seen a formula for computing $3P$, but the geometric description of the 3-torsion points is rather nice. If we rewrite $3P = \mathcal{O}$ as $2P + P = \mathcal{O}$, then we see that P is the inverse of $2P$. Since the inverse of $2P$ is $P * P$, then we have $P * P = P$, which means that the line tangent to E at P actually intersects triply at P. In other words, P is an inflection point of the curve, so the question becomes "when does an elliptic curve E defined over $\mathbb{Q}$ have rational inflection points." The inflection points of a curve can be found by doing implicit differentiation twice, so that gives one strategy of describing elliptic curves with rational 3-torsion.

There are several theorems that are useful for describing the structure of rational torsion on elliptic curves defined over $\mathbb{Q}$.

Theorem 2 (Nagell-Lutz)**.** *Let E be an elliptic curve defined by an equation of the type* (29) *with $a, b, c \in \mathbb{Z}$, and suppose that $P = (x, y)$ is a non-trivial rational torsion point on E. The discriminant of f is*

$$\operatorname{disc}(f) = a^2b^2 - 4a^3c - 4b^3 + 18abc - 27c^2$$

Then x and y are both integers, and either $y = 0$ or $y^2 | \operatorname{disc}(f)$.

Theorem 3 (Mazur)**.** *Let E be an elliptic curve defined over $\mathbb{Q}$. Then the rational torsion subgroup $E(\mathbb{Q})_{tor}$ is isomorphic to one of the following fifteen groups:*

$$\mathbb{Z}/m\mathbb{Z} \text{ for } 1 \leq m \leq 10 \text{ or } m = 2$$
$$\mathbb{Z}/m\mathbb{Z} \text{ for } 1 \leq m \leq 4.$$

furthermore, each option actually occurs.

Over $\mathbb{Q}$ it is possible to get points that have do not have finite order. Over a finite field $\mathbb{F}_q$, of course, an elliptic curve E can have only a finite number of points, and thus all points have finite order. We will not spend much time studying the points of infinite order, except to see what happens when reducing such points mod p.

We conclude this section by introducing the divisions polynomials. Let E be an elliptic curve in Weierstrass form defined over a field K with $\operatorname{char}(K) \neq 2$. Let

$$\begin{aligned}
&f_0(x) = 0, f_1(x) = 1, f_2(x) = 1\\
&f_3(x) = 3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8\\
&f_4(x) = 2x^6 + b_2x^5 + 5b_4x^4 + 10b_6x^3 + 10b_8x^2 + (b_2b_8 - b_4b_6)x + (b_4b_8 - b_6^2).
\end{aligned}$$

Let $\tilde{f}(x) = 4x^3 + b_2x^2 + 2b_4x + b_6$, and for $n \geq 5$ let $f_n(x)$ be defined by the recursive formulas

$$\begin{aligned}
f_{2n} &= f_n(f_{n+2}f_{n-1}^2 - f_{n-2}f_{n+1}^2)\\
f_{2n+1} &= \begin{cases} \tilde{f}^2 f_{n+2}f_n^3 - f_{n-1}f_{n+1}^3 & \text{if } n \text{ is even,}\\ f_{n+2}f_n^3 - \tilde{f}^2 f_{n-1}f_{n+1}^3 & \text{otherwise,} \end{cases}
\end{aligned}$$

If $P = (x, y)$ is a point on E, then

$$nP = \begin{cases} \mathcal{O} & \text{if } \psi_n(x, y) = 0\\ \left(\frac{\phi_n(x,y)}{\psi_n^2(x,y)}, \frac{\omega_n(x,y)}{\psi_n^3(x,y)}\right) & \text{otherwise,} \end{cases}$$

where

$$\psi_n = \begin{cases} (2y + a_1x + a_3)f_n & \text{if } n \text{ is even,} \\ f_n & \text{otherwise,} \end{cases}$$

and

$$\phi_n = x\psi_n^2 - \psi_{n-1}\psi_{n+1} \text{ and } 2\psi_n\omega_n = \psi_{2n} - \psi_n^2(a_1\phi_n + a_3\psi_n^2).$$

In particular, we see that $P \in E[n]$ if and only if $\psi_n(x) = 0$. These equations generally become easier to work with when $a_1 = a_3 = 0$, which will always be the case in these notes

For example, suppose $a_1 = a_3 = 0$. If $n = 2$, then we have $\psi_2 = 2y$,

$$\phi_2 = x(2y)^2 - f_3 \quad \text{and} \quad 2\omega_2 = f_4.$$

Since $4y^2 = \tilde{f}$, then

$$\phi_2(x) = x^4 - b_4x^2 - 2b_6x - b_8$$

hence we obtain the same x-coordinate from doubling formula as before, except with slightly less generality. The y-coordinates match too, though it is best to use the a_n's not the b_n's when verifying this.

Exercises

1. Use implicit differentiation to compute ψ_3. Check your answer with sage.
2. Let p be an odd prime, and consider $E : y^2 = x^3 + Ax^2 + x$ over $\mathbb{F}_p$, where $A \neq \pm 2$ in $\mathbb{F}_p$.
   (a) Prove that if $\left(\frac{A^2-4}{p}\right) = 1$, then E has full 2-torsion.
   (b) Prove that if $\left(\frac{A^2-4}{p}\right) = -1$, then has exactly one of $x \equiv \pm 1 \bmod p$ as an x-coordinate of a 4-torsion point.
   (c) Conclude that $a_p \equiv p + 1 \bmod 4$ in both cases.
3. Let $A \in \mathbb{Z} - \{-2, 2\}$. Then $E : y^2 = x^3 + Ax^2 + x$ is an elliptic curve, with non-trivial 2-torsion.
   (a) Apply the rational root theorem to ψ_3 to prove that E has trivial 3-torsion over $\mathbb{Q}$.
   (b) Factor ψ_4 over $\mathbb{Q}$. Show that while $x = \pm 1$ are roots of ψ_4, they generally do not yield corresponding y values in $\mathbb{Q}$.
   (c) Show that ψ_4 factors completely over $\mathbb{Q}(\sqrt{A-2}, \sqrt{A+2})$, and moreover that E has full 4-torsion over this field.
   (d) Use your previous observations to find conditions on A such that E does not have a point of order 4 defined over $\mathbb{Q}$. Conclude by invoking Mazur's theorem, that $E(\mathbb{Q})$ has $\mathbb{Z}/2\mathbb{Z}$ as its torsion subgroup.
   (e) Solve $y^2 = x^3 + Ax^2 + x$ for A and select x and y appropriately to construct examples of elliptic curves that having a point $P \in E(\mathbb{Q})$ such that P is not a torsion point. Note that for your previous work to hold, you must have $A \in \mathbb{Z}$.

6 Endomorphisms

Definition 7. Let E be an elliptic curve, $\mathcal{O}$ be the identity element. An *endomorphism* of E is a homomorphism

$$\varphi : E \to E.$$

When considering the kernel of an endomorphism of an elliptic curve over a field K, we consider all points in the algebraic closure of that field, i.e. not only elements in K, but also roots of any polynomial with coefficients in K. One option for an endomorphism is to send everything to $\mathcal{O}$, for which all points on the elliptic curve are in the kernel of φ, in particular the kernel is infinite, even when K is a finite field (since the algebraic closure of K is not finite). As it turns out, any other homomorphism of elliptic curves is surjective and has a finite kernel.

Definition 8. The *degree* of an endomorphism φ, denoted by $\deg \varphi$, is defined to be $|\ker \varphi|$ if this quantity is finite, and zero otherwise.

Definition 9. The set of endomorphisms of E defined over a given field L is denoted $\mathrm{End}_L(E)$. The set of all endomorphisms of E is denoted by $\mathrm{End}(E)$.

Example 1. It should be clear that for $m \in \mathbb{Z}^+$ the multiplication by m map is always an endomorphism of E. If we regard $(-1) \cdot P$ as giving the inverse of P, and $0 \cdot P$ as sending everything to $\mathcal{O}$, then we obtain an endomorphism for every integer m, i.e. we have $\mathbb{Z} \subset \mathrm{End}(E)$. Since the kernel of the multiplication by m map is the m torsion, then we have $\deg(m) = m^2$, and it should be clear that holds also for negative m and zero.

Example 2. If E is defined over $\mathbb{F}_{p^n}$, then the frobenius automorphism $\sigma^n : x \mapsto x^{p^n}$ acts trivially on the coefficients of the elliptic curve, and as an automorphism, it commutes with both addition and multiplication. It follows that

$$\begin{aligned}&\sigma^n(x^3 + a_2x^2 + a_4x + a_6 - (y^2 + a_1xy + a_3y))\\&= (\sigma^n(x))^3 + a_2(\sigma^n(x))^2 + a_4\sigma^n(x) + a_6 - ((\sigma^n(y))^2 + a_1\sigma^n(x)\sigma^n(y) + a_3\sigma^n(y))\\&= x^3 + a_2x^2 + a_4x + a_6 - (y^2 + a_1xy + a_3y),\end{aligned}$$

thus if (x, y) is a point on E, then so is $(\sigma^n(x), \sigma^n(y))$. Furthermore, since the formulas for addition on E are defined over $\mathbb{F}_{p^n}$, then the same type of calculation, shows that σ^n commutes with the addition law, i.e. $\sigma^n(P + Q) = \sigma^n(P) + \sigma^n(Q)$. In particular, if $mP = \mathcal{O}$, then $m\sigma^n(P) = \sigma^n\mathcal{O} = \mathcal{O}$, i.e. if P is in the m-torsion, then so is $\sigma^n(P)$. For these reasons, σ^n induces an endomorphism of E, which will be denoted with ϕ^n.

The Frobenius endomorphism plays a very important role for elliptic curves. Let E be an elliptic curve over $\mathbb{F}_q$, where $q = p^n$ as usual. Then there exists $a_q \in \mathbb{Z}$ such that ϕ^n satisfies

$$(\phi^n(P))^2 - a_q(\phi^n(P)) + qP = 0, \tag{31}$$

for all $P \in E(K)$. Note that a_q, q, and 0 are also treated as endomorphisms and $-$ and $+$, refer to addition on the elliptic curve. It is striking that a_q is related to the number points of E defined over $\mathbb{F}_q$ as provided by the following theorem

Theorem 4 (Hasse)**.** *Let E be an elliptic curve defined over $\mathbb{F}_q$. Then*

$$|E(\mathbb{F}_q)| = q + 1 - a_q$$

where $|a_q| \leq 2\sqrt{q}$.

If E is defined over $\mathbb{F}_q$ then it is also defined over any extension of $\mathbb{F}_q$, in particular over the finite extensions $\mathbb{F}_{q^k}$. Hasse's theorem still applies over these extensions, but as it turns out it is not necessary to compute a_q from scratch. One way of looking at equation (31) is to say that ϕ^n satisfies the polynomial $x^2 - a_q x + q$. If α, β are complex roots of this polynomial, then $a_q = \alpha + \beta$. When a_q is written this way, then we have

Theorem 5. *Let α and β be the complex roots of $x^2 - a_q x + q$. Then for all $k \in \mathbb{Z}^+$*

$$a_{q^k} = \alpha^k + \beta^k.$$

If $q = p$, then a_{p^k} can also be calculated with the recursion

$$a_{p^{k+1}} = a_{p^k} a_p - p a_{p^{k-1}} \text{ for } k \geq 1$$

where $a_1 = 2$.

The discriminant of $x^2 - a_q x + q$ is $a_q^2 - 4q$. If $q = p$, then $\sqrt{p}$ is not an integer, thus the inequality in Hasse's theorem is strict, and so $a_p^2 - 4p$ is negative. As a consequence, the roots α and β belong to an imaginary quadratic number field.

The two examples above provide the easiest examples of endomorphisms, but depending on the elliptic curve and the base field in question there may be other endomorphisms. Given any two endomorphisms ϕ and ψ, it can be shown that $\phi \circ \psi$ and $\phi + \psi$ are endomorphisms, where $+$ denotes the group law on E. It is easy to show that these operations give $\mathrm{End}_L(E)$ and $\mathrm{End}(E)$ a ring structure, hence $\mathrm{End}(E)$ is called the "endomorphism ring" of E. The following theorem describes what can happen in general.

Theorem 6. *If E is an elliptic curve, then $\mathrm{End}(E)$ is isomorphic to one of the following*

1. *$\mathbb{Z}$,*

2. *an order in an imaginary quadratic number field,*

3. *an order in a quaternion algebra.*

In the first case, E is often said to have "trivial endomorphism ring," in the second case E is said to have **complex multiplication** *(CM for short), in the third case E is said to be* **supersingular** *or* **superspecial***. The third case can only occur in positive characteristic.*

It is important to understand that the base field of E is implied. For example, if E is defined over $\mathbb{Q}$, only the first two options are possible since $\mathbb{Q}$ has characteristic zero, but upon E reducing mod p, thus changing the base field to positive characteristic, then the third option becomes possible. Indeed, Elkies has proved the following:

Theorem 7. *If E is an elliptic curve defined over $\mathbb{Q}$, then there are infinitely many primes p such that E becomes supersingular over $\mathbb{F}_p$.*

The term "supersingular" is very misleading, because elliptic curves in Weierstrass form (3) are not singular at all. Moreover, most modern definitions of elliptic curves include smoothness in the definition, in spite of the existence of singular models such as $y^2 = f(x)$ where f has degree 4. The point is that such models are nonetheless equivalent to a curve in Weierstrass form, and a curve given by (3) is an elliptic curve if and only if it is smooth. Unfortunately the term supersingular persists in the literature so it is important to know both. In these notes we will attempt to use the term "superspecial."

To determine when an elliptic curve over is superspecial, we have the following result due to Deuring.

Theorem 8. *Let p be an odd prime, and let K be a field of characteristic p.*

1. *Let E/K be an elliptic curve with Weierstrass equation $y^2 = f(x)$, where $f \in K[x]$ is a separable cubic polynomial. Then E is superspecial if and only if the coefficient of x^{p-1} in $f^{(p-1)/2}$ is zero.*

2. *Let $m = (p-1)/2$ and define a polynomial*

$$H_p(t) = \sum_{i=1}^{m} \binom{m}{i}^2 t^i.$$

Let λ be in the algebraic closure of K, and different from 0 and 1. Then the elliptic curve $E_\lambda : y^2 = x(x-1)(x-\lambda)$ is superspecial if and only if $H_p(\lambda) = 0$. Additional, if $\lambda \in \mathbb{F}_p$, then $|E_\lambda(\mathbb{F}_p)| = p + 1$.

We will not discuss the proof in full detail here, aside from pointing out a corollary. For an elliptic curve in the form

$$E : y^2 = f(x)$$

over $\mathbb{F}_q$, if we wanted to compute $|E(\mathbb{F}_q)|$ directly, we could simply plug in all possible choices of x and determine whether or not $f(x)$ is a square. If $f(x) = 0$, then there is only one corresponding point, with y coordinate zero. If $f(x)$ is a non-zero square then there are two points. After accounting for the point at infinity, we have

$$|E(\mathbb{F}_p)| = 1 + \sum_{x \in \mathbb{F}_p} \left(1 + \left(\frac{f(x)}{p}\right)\right) = 1 + p + \sum_{x \in \mathbb{F}_p} \left(\frac{f(x)}{p}\right),$$

where $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol. Then by Euler's criterion, which is really a fact about cyclic groups, we have

$$\left(\frac{f(x)}{p}\right) = f(x)^{\frac{(p-1)}{2}},$$

which is now familiar from theorem 8. If we let A_p denote the coefficient of x^{p-1} in $f^{\frac{(p-1)}{2}}$, then after some more algebra we obtain

$$|E(\mathbb{F}_p)| \equiv 1 + A_p \bmod p.$$

Then by comparing with Hasse's theorem 4, we find that $A_p \equiv -a_p$. It follows that E is superspecial if and only if $a_p \equiv 0 \bmod p$. Then, by looking once again at the bound $|a_p| \leq 2\sqrt{p}$, it follows that $a_p = 0$ if $p \geq 5$. We can generalize all of this to extensions of $\mathbb{F}_p$. Let χ be the composition of the maps

$$\mathbb{F}_q^\times \to \mathbb{F}_q^\times / (\mathbb{F}_q^\times)^2 \to \{\pm 1\},$$

and extend to zero by defining $\chi(0) = 0$. Then everything works the same as before, and we get the following corollary.

Corollary 1. *Let $q \geq 5$ be odd, and let E be an elliptic curve over $\mathbb{F}_q$ given by a Weierstrass equation $y^2 = f(x)$ where f is a separable cubic. Then E is superspecial if and only if $a_q = 0$.*

Exercises

1. Let $E_1 : y^2 = x^3 + Ax^2 + x$ over $\mathbb{F}_p$ where $\left(\frac{A^2-4}{p}\right) = -1$, and let E_2 be the twist of E_1. Prove that the trace of E_2 is the negative of the trace of E_1, so

$$|E_1(\mathbb{F}_p)| = p + 1 - a_p \quad \text{and} \quad |E_2(\mathbb{F}_p)| = p + 1 + a_p.$$

2. Let $E : y^2 = x^3 + x$, and for $p = 11$ and $p = 13$, do the following steps:

   (a) Use the Legendre symbol to compute $|E(\mathbb{F}_p)|$.

   (b) Use corollary 1 to decide whether E is superspecial over $\mathbb{F}_p$.

   (c) Using sage or magma compute roots α and β for $x^2 - a_p x + p$, then compute a_{p^2} and a_{p^3} using both of the methods in theorem 5.

3. Let $E_A : y^2 = x^3 + Ax^2 + x$ be an elliptic curve in Montgomery form, and let $f(x)$ denote the cubic polynomial on the right hand side. Use sage or magma to compute the coefficient of x^{p-1} in $f(x)^{(p-1)/2}$ for $p = 11$ and $p = 13$. Your result will be a polynomial in A over $\mathbb{F}_p$, so then factor it to determine the values of A, for which E_A is superspecial.

7 Point counting

We have seen earlier that the points on an elliptic curve over $\mathbb{F}_q$ can be counted by brute force, using the Legendre symbol in the case $p = q$, or χ more generally. However, if q is large, then an efficient algorithm is needed. The first such algorithm was presented by Schoof. First, recall that

$$|E(\mathbb{F}_q)| = q + 1 - a_q$$

where a_q is the trace of the Frobenius endomorphism ϕ^n. The goal is therefore to compute a_q directly by using the properties of ϕ_q. The characteristic polynomial of ϕ_q is $x^2 - a_q x + q$, and as a linear transformation ϕ_q satisfies its own characteristic polynomial, thus

$$(\phi^n)^2(Q) - a_q \phi^n(Q) + qQ = \mathcal{O}$$

for all $Q \in E(\overline{\mathbb{F}}_q)$. If we take $Q \in E[\ell]$ for some prime ℓ, then $\phi^n(Q) \in E[\ell]$ also, which means that a_q and q can be reduced mod ℓ without changing the validity of the above equation. If we can determine a_q mod ℓ for enough primes, so that their product is larger than $4\sqrt{q}$, then the Hasse bound $|a_q| < 2\sqrt{q}$ shows that a_q is determined completely.

Example 14. Consider the elliptic curve $E : y^2 = x^3 + 13x^2 + x$ over $\mathbb{F}_{167}$. We already know that $E(\mathbb{F}_{167})$ contains points of order 4, and $p \equiv 3 \bmod 4$, so

$$a_p = p + 1 - |E(\mathbb{F}_{167})| \equiv 0 \bmod 4.$$

We could look mod higher powers of 2 if we wanted to, but instead we will compute the action of Frobenius on the ℓ-torsion for $\ell = 3$ and 5. We first need to find the smallest field over which we have full ℓ-torsion. We can ask sage to try factoring the division polynomials.

```
EllipticCurve(GF(167),[0,13,0,1,0]).division_polynomial(3).factor()
```

The output of sage tells us that the polynomial is irreducible and has degree 4, so it splits over $\mathbb{F}_{167^4}$, and if P is a non-trivial 3-torsion point, then it's x-coordinate lies in $\mathbb{F}_{887^4}$. That does not mean the y-coordinate lies in $\mathbb{F}_{167^4}$: plugging into $x^3 + 3x^2 + x$ gives us an element of $\mathbb{F}_{167^4}$, but not necessarily a square, but if we extend to $\mathbb{F}_{167^8}$, then we are guaranteed to have a square root. It turns out that $\mathbb{F}_{167^8}$ is also sufficient for full 5-torsion.

Now we get sage to compute the action of Frobenius as follows:

```
p=167;A=13;K.<z8>=GF(p^8);R.<x,y>=K[];E=EllipticCurve(K,[0,A,0,1,0]);f=y^2-(x^3+A*x^2+x)

tor3=[E(v[0],f(x=v[0]).factor()[0][0](y=0)) for v in E.division_polynomial(3).roots()]

P3=tor3[0];Q3=tor3[1];R3=E(P3[0]^p,P3[1]^p);S3=E(Q3[0]^p,Q3[1]^p)

matrix([reduce(lambda x,y:x+y,reduce(lambda x,y:x+y,
...[[ [i,j] for i  in range(3) if i*P3+j*Q3==P] for j in range(3)])) for P in [R3,S3]])
```

Here is an explanation of each line of code.

The first line of code, defines the field $K = \mathbb{F}_{167^8}$, the elliptic curve E, the ring R of polynomials in two variables over K, and the polynomial f in that ring defining the elliptic curve.

The second line of code computes *half* of the non-trivial 3-torsion points. In particular each x-coordinate only occurs once.

The third line computes a basis. Usually we would need to check that we do not have two points that belong to the same cyclic subgroup, but we have 3-torsion points and the only way for two distinct non-trivial 3-torsion points to belong to the same cyclic subgroup is if they are inverses of each other, i.e. the x-coordinates are the same. Since we have distinct x-coordinates, we can pick any two of the points we computed. We pick the first two and call them `P3` and `Q3`. We then define `R3` and `S3` by the action of Frobenius, making sure to tell sage that these are still points on E.

Line 4, is a brute force calculation of all linear combinations of `P3` and `Q3` to see which ones give us `R3` and `S3`. The output is the following matrix:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

The trace of this matrix is 1, so we now know that $a_p \equiv 1 \bmod 3$. We do the same for the 5 torsion, being a little more careful this time when constructing the basis. We do not need to repeat the first step. The second step needs to be redone with 5 instead of 3. Then before forming a basis we check that we have a pair of points not belonging to a cyclic subgroup:

```
[tor5[1]==n*tor5[0] for n in range(5)]
```

Output: `[False, False, False, False, False]`

So `tor5[1]` and `tor5[0]` do not belong to the same cyclic subgroup and can be used as a basis. Then the rest is the same as before after changing every 3 to a 5. The output is the matrix

$$\begin{bmatrix} 0 & 1 \\ 3 & 0 \end{bmatrix}$$

The trace of this matrix is 0, so we now know that $a_p \equiv 0 \bmod 5$. Note that the matrices we found are not unique: a matrix always depends on the choice of basis. But the other matrices that can occur are conjugates of the ones we found, and they have the same characteristic polynomials, meaning in particular that both the trace and determinant remain the same.

Since $4 \cdot 3 \cdot 5 = 60 > 4\sqrt{167}$, we have gone far enough to determine the value of a_p completely. If we apply the Chinese Remainder Theorem to the system

$$a_p \equiv [0, 1, 0] \bmod [4, 3, 5]$$

we obtain $a_p \equiv -20 \bmod 60$, where we use symmetric residues to satisfy $|a_p| \leq 2\sqrt{167}$, thus $a_p = -20$. Sage also has the built-in capability of computing the trace:

```
EllipticCurve(GF(167),[0,13,0,1,0]).trace_of_frobenius()
```

Now that we have computed a_p, then what is the order of E and its twist? We have

$$p + 1 - a_p = 168 + 20 = 188 = 2^2 \cdot 47 \quad \text{and} \quad p + 1 + a_p = 168 - 20 = 148 = 2^2 \cdot 37$$

respectively.

Exercises

1. Consider the elliptic curve $E : y^2 = x^3 + 10x^2 + x$ over $\mathbb{F}_7$. Extend $\mathbb{F}_7$ to a field over which E has full 3-torsion. Compute the action of Frobenius on the 3-torsion, obtaining an explicit matrix, then determine the exact trace a_7, and the order of both E and its twist.

8 Koblitz Curves

A Koblitz curve is given by

$$E_a : y^2 + xy = x^3 + ax^2 + 1 \tag{32}$$

where $a \in \mathbb{F}_2$. Let $P = (x_1, y_1)$, $Q = (x_2, y_2)$, and $P + Q = (x_3, y_3)$, then $-P = (x_2, x_2 + y_2)$, and

$$\lambda = \begin{cases} \dfrac{x_1^2 + y_1}{x_1} & \text{if } P = Q \\ \dfrac{y_1 + y_2}{x_1 + x_2} & \text{otherwise} \end{cases} \qquad \nu = \begin{cases} x_1^2 & \text{if } P = Q \\ \dfrac{x_1 y_2 + x_2 y_1}{x_1 + x_2} & \text{otherwise} \end{cases} \tag{33}$$

$$x_3 = \lambda^2 + \lambda + a + x_1 + x_2 \qquad y_3 = (\lambda + 1)x_3 + \nu$$

In each case, the number of points in $\mathbb{F}_2$ is easy to compute by brute force, specifically we find

$$|E_0(\mathbb{F}_2)| = 4 \quad \text{and} \quad |E_1(\mathbb{F}_2)| = 2.$$

By Hasse's theorem 4, the trace of frobenius is -1 for E_0 and 1 for E_1. For E_0, ϕ satisfies $x^2 + x + 2$ and its complex roots are

$$\alpha = \frac{-1 + \sqrt{-7}}{2} \quad \text{and} \quad \beta = \frac{-1 - \sqrt{-7}}{2} \tag{34}$$

and so by theorem 5, we have

$$a_{2^k} = \alpha^k + \beta^k$$

for α and β given by (34), or we can use the recursive formula from the same theorem; see the exercises for $|E_1(\mathbb{F}_{3^k})|$. It is now an easy task to compute the trace of frobenius over arbitrary extensions of $\mathbb{F}_2$.

Example 15. For E_0, we computed $a_2 = -1$. By the recursive formula we have

$$\begin{aligned} a_4 &= a_2^2 - 2a_1 = (-1)^2 - 2 \cdot 2 = -3 \\ a_8 &= a_4 a_2 - 2a_2 = (-3)(-1) - 2 \cdot (-1) = 5 \\ a_{16} &= a_8 a_2 - 2a_4 = (5)(-1) - 2 \cdot (-3) = 1 \\ a_{32} &= a_{16} a_2 - 2a_8 = (1)(-1) - 2 \cdot (5) = -11. \end{aligned}$$

On the other hand since α and β are complex conjugates, when expanding α^k or β^k using the binomial expansion theorem, we only need to consider the real terms, since the imaginary terms cancel. Moreover the real parts of α^k and β^k are equal, thus for $k = 5$

$$\alpha^5 + \beta^5 = \frac{2}{2^5}\left(-1 + \binom{5}{2}7 - \binom{5}{4}7^2\right) = \frac{1}{16}(-1 + 10 \cdot 7 - 5 \cdot 49) = \frac{-176}{16} = -11.$$

Then by Hasse's theorem

$$|E_0(\mathbb{F}_{32})| = 32 + 1 - (-11) = 44 = 2^2 \cdot 11.$$

The ease of calculating the trace of frobenius is part of the idea behind Koblitz curves. Another part of the idea comes from the fact that we can use the frobenius endomorphism to speed up point addition. The frobenius endomorphism ϕ of the elliptic curve E_0 satisfies

$$\phi^2 + \phi + 2 = 0.$$

For $n \in \mathbb{Z}^+$, consider the binary expansion $n = (n_m n_{m-1} \cdots n_1 n_0)_2$. We replace each 2^i, by $(-\phi - \phi^2)^i$, then apply the binomial expansion theorem and repeat with the coefficients until we have

$$n = \sum_{i=0}^{m} n_i(-\phi - \phi^2)^i = \sum_{j=0}^{r} b_j \phi^j,$$

for some r where $b_j \in \{-1, 0, 1\}$ for all j. Since ϕ^d acts trivially on $\mathbb{F}_{2^r}$, then we can reduce r to a value less than d, when acting on points defined over $\mathbb{F}_{2^r}$

Example 16. Take $n = 6$. Then

$$\begin{aligned} 6 = 2 + 2^2 = -\phi - \phi^2 + (-\phi - \phi^2)^2 &= -\phi - \phi^2 + \phi^2 + 2\phi^3 + \phi^4 \\ &= -\phi + (-\phi - \phi^2)\phi^3 + \phi^4 = -\phi - \phi^5. \end{aligned}$$

As we have seen in example 15, we have $|E_0(\mathbb{F}_{32})| = 2^2 \cdot 11$. The point

$$P = (\alpha^3 + \alpha^2 + 1, \alpha^3 + \alpha + 1)$$

is a point of order 11 on E_0, where α satisfies the irreducible polynomial $x^5 + x^2 + 1$. Also, since $x^5 + x^2 + 1$ divides $x^{32} - x$ but does not divide x, then α also satisfies $\alpha^{31} = 1$ in $\mathbb{F}_{32}$. We compute

$$\begin{aligned} \phi(P) = ((\alpha^3 + \alpha^2 + 1)^2, (\alpha^3 + \alpha + 1)^2) &= (\alpha^6 + \alpha^4 + 1, \alpha^6 + \alpha^2 + 1) \\ &= (\alpha(\alpha^2 + 1) + \alpha^4 + 1, \alpha(\alpha^2 + 1) + \alpha^2 + 1) \\ &= (\alpha^4 + \alpha^3 + \alpha + 1, \alpha^3 + \alpha^2 + \alpha + 1) \end{aligned}$$

and since ϕ^5 acts trivially on $\mathbb{F}_{32}$ then it acts trivially on P. Therefore

$$\begin{aligned} 6P = -\phi(P) - P &= -(\alpha^4 + \alpha^3 + \alpha + 1, \alpha^3 + \alpha^2 + \alpha + 1) - (\alpha^3 + \alpha^2 + 1, \alpha^3 + \alpha + 1) \\ &= (\alpha^4 + \alpha^3 + \alpha + 1, \alpha^4 + \alpha^2) + (\alpha^3 + \alpha^2 + 1, \alpha^2 + \alpha) = (\alpha, \alpha^4 + \alpha^3 + \alpha^2 + 1) \end{aligned}$$

where the last step is obtained with the help of

$$\lambda = \frac{\alpha + \alpha^4}{\alpha + \alpha^2 + \alpha^4} = \alpha^2 \quad \text{and} \quad \nu = \frac{\alpha^3 + \alpha^2 + \alpha + 1}{\alpha + \alpha^2 + \alpha^4} = \alpha^4 + \alpha^2 + \alpha + 1.$$

None of what we have done is really unique to characteristic 2. In odd characteristic, we could still begin with $|E(\mathbb{F}_p)|$ over a small field $\mathbb{F}_p$, compute α and β, compute the trace of frobenius for $\mathbb{F}_{p^k}$, etc. But once we get to the point of representing n in base p, we can only get down to $b_j \in \{1 - p, \ldots, -1, 0, 1, \ldots p - 1\}$. As a consequence, we still get a speedup for odd p, but we get less of one as p grows, so in the sense of giving the best speedup, Koblitz curves are optimal.

Exercises

1. Compute roots α and β of $x^2 - x + 2$. Then compute a_8 with $\alpha^3 + \beta^3$ and by the recursive formula. Let ℓ be the largest prime dividing $|E_1(\mathbb{F}_8)|$. Use sage or magma to find a point P of order ℓ on E_1 defined over $\mathbb{F}_8$. Finally, compute $5P$ both with and without the frobenius endomorphism.

9 Divisors

Let C be a curve defined over a field K, and let $f : C \to P^1_K$ be a rational function. Suppose that f has zeros at $P_1, P_2, \dots P_r$, with orders $e_1, e_2, \dots e_r$ respectively, and poles at $Q_1, Q_2, \dots Q_s$ with orders $d_1, d_2, \dots d_s$ respectively. Then the *divisor of* f is the "formal sum"

$$\operatorname{div}(f) = e_1[P_1] + e_2[P_2] \cdots + e_r[P_r] - d_1[Q_1] - d_2[Q_2] \dots - d_s[Q_s]. \tag{35}$$

We call this a formal sum because the P_i and Q_i are regarded as independent objects, somewhat like independent variables. For example if $P_1 = (1, 0)$, and $n_1 = 3$, then $n_1[P_1]$ does not mean $(3, 0)$, it means something more like we have three copies of $(1, 0)$...but we allow negatives too. Similarly addition and subtraction are not meant to suggest that we are adding or subtracting coordinates. The coefficients in the formal sum above have an important meaning, since they carry information about the zeros and poles of the function f, and so trying to combine them would be disastrous, since we would loose that information.

As an example, consider the ordinary rational function $f(x) = x^2(x - 1)$. Then f vanishes at the points $x = 0$ and $x = 1$ with orders 2 and 1 respectively, and has a pole at infinity with order 3. Why does f has a pole at infinity with order 3? This is easiest to see by projectivizing f. Let $x = \frac{X_0}{X_1}$. Then

$$f\left(\frac{X_0}{X_1}\right) = \frac{X_0^2(X_1 - X_0)}{X_1^3} = X_0^2(X_1 - X_0)X_1^{-3},$$

and now the factor $1/X_1^3$ makes sense for a pole of order 3. Furthermore, by exponent rules we see that it makes sense to treat a pole of order 3, as vanishing with order -3. Using the projective coordinates $(X_0 : X_1)$, the points $x = 0$ and $x = 1$ are $(0 : 1)$ and $(1 : 1)$ and the point at infinity is $(1 : 0)$, so the divisor of f is

$$\operatorname{div}(f) = 2[(0 : 1)] + [(1 : 1)] - 3[(1 : 0)].$$

Thus we see that the total number of zeros is 3 and the total number of poles is 3 (counting with multiplicity in each case). This is a general fact of rational functions. If $f = g/h$ is a rational function where g and h are polynomials, and if we define the degree of f to be the maximum of the degrees of g and h, then we have the following

Proposition 2. *A rational function f of degree $n \geq 1$ takes on each complex value exactly n times, counting with multiplicity, and also has n poles counting with multiplicity.*

Proof. If $\lambda \in \mathbb{C}$ is arbitrary, then the numerator of $f - \lambda$ has exactly n roots by the fundamental theorem of algebra, and the fact that there are n poles as well is clear by projectivizing. □

Thinking of f as a map $\mathbb{P}^1_{\mathbb{C}} \to \mathbb{P}^1_{\mathbb{C}}$, then this proposition says that f is surjective, and if P is an arbitrary point in the image, then $|f^{-1}(P)| \leq n$. Equality holds in the case where all points in the set $f^{-1}(P)$ have multiplicity 1. As a consequence, it is often useful to think of the degree of a map in terms of the cardinality of pre-images of individual points. This proposition remains true in the case where C is a non-singular algebraic curve. [1]

[1] From an analytic perspective this can be proved by Stokes' theorem (Calc III). The complex algebraic curve C is viewed as a real surface (by ignoring the complex structure). The surface is compact so its boundary is empty and when computing the integral of the function $f - \lambda$ over the surface, Stokes' theorem gives equality with an integral over the boundary. The boundary integral vanishes since the domain is empty.

As a consequence of this proposition we see that given a rational function f, the sum of the coefficients of its divisor $\mathrm{div}(f)$ must be zero. We often consider more general divisors

$$D = n_1[P_1] + n_2[P_2] + \cdots + n_k[P_k] \tag{36}$$

with prescribed multiplicities n_i, but such a divisor cannot possibly belong to a rational function f, unless

$$n_1 + n_2 + \cdots n_k = 0.$$

This provides the motivation for the *degree* of a divisor D of the form (36), which is defined as

$$\deg(D) = n_1 + n_2 + \cdots n_k.$$

Thus if $f : C \to \mathbb{P}^1_K$ is a rational function, then $\deg \mathrm{div}(f) = 0$. The converse, however is not necessarily true. It is true if C is $\mathbb{P}^1_K$, but if C is an elliptic curve, then there are no rational functions that vanish at only one point and have only one pole, each with multiplicity 1. For example, consider the elliptic curve

$$E : y^2 = x(x-1)(x-\lambda)$$

where λ is different from zero and 1. Then the function defined $f(P) = x(P) - a$, by taking the x-coordinate of the point P on C, is a rational function. To determine the corresponding divisor, we consider the intersection of $x = a$ with the curve E. Generally, there are two points on E with x-coordinate equal to a, which are additive inverses of each other say Q_1 and Q_2. The exceptional case occurs when $a = 0$, 1, or λ which case $x = a$ is tangent to the curve, in other words f vanishes at one of the 2-torsion points say Q with multiplicity 2. By projectivizing, we see also that f has a pole of order 2 at $\mathcal{O} = (0 : 1 : 0)$, thus the divisor is

$$\mathrm{div}(f) = \begin{cases} [Q_1] + [Q_2] - 2[\mathcal{O}] & \text{if } a \text{ is different from 0, 1, or } \lambda \\ 2[Q] - 2[\mathcal{O}] & \text{if } a \text{ is different from 0, 1, or } \lambda \end{cases}$$

so even though $f(P) = x(P) - a$ is linear as a function from $\mathbb{P}^2$ to $\mathbb{P}^1$, if we take an arbitrary point in $\mathbb{P}^1$, its pre-image is infinite and we do not get a finite pre-image until we restrict f to a curve. When restricting f, to the elliptic curve E, we see that the pre-images points generally have a cardinality of 2 and thus $f : E \to \mathbb{P}^1$ has degree 2.

On the other hand, if we take a line that is not vertical, then we will get three points of intersection with E, which defines a rational function $f : E \to \mathbb{P}^1$ of degree 3. In particular, the horizontal line $y = 0$ intersects with the three 2-torsion points. Since the 2-torsion points add to $\mathcal{O}$, then we see another example where the points on E at which a given rational function vanishes add up to $\mathcal{O}$ when added with their multiplicities (using the addition law on the elliptic curve), and this generalizes:

Proposition 3. *Let E be an elliptic curve defined over K, and let $P_1, P_2 \ldots P_k$ be points on E, such that*

$$n_1P_1 + n_2P_2 + \cdots n_kP_k = \mathcal{O},$$

then there exists a rational function $f : E \to \mathbb{P}^1_K$ such that

$$\mathrm{div}(f) = n_1[P_1] + n_2[P_2] + \cdots n_k[P_k] - (n_1 + n_2 + \cdots n_k)[\mathcal{O}].$$

10 Weil pairing

The Weil pairing is defined as follows.

Definition 10. Let P, Q be m-torsion points of E, so $mP = mQ = \mathcal{O}$, hence by proposition 3, there exist rational functions f_P and f_Q, with divisors

$$m[P] - m[\mathcal{O}] \quad \text{and} \quad m[Q] - m[\mathcal{O}]$$

respectively. Let S be any point of E such that $S \notin \{\mathcal{O}, P, -Q, P - Q\}$. Then the *Weil pairing of P and Q* is

$$e_m(P, Q) = \frac{f_P(Q + S) f_Q(-S)}{f_P(S) f_Q(P - S)}.$$

The Weil pairing has some remarkable properties, summarized in the following theorem

Theorem 9. *1. $e_m(P, Q)$ is always an m-th root of unity.*

2. $e_m(P, Q)$ is bilinear (i.e. it is linear in both components).

3. e_m is alternating (i.e. $e_m(P, P) = 1$ for all $P \in E[m]$)

4. e_m is non-degenerate (i.e. if $e_m(P, Q) = 1$ for all $Q \in E[m]$, then $P = \mathcal{O}$)

As a consequence, if the Weil pairing can be constructed, then the discrete log problem on E can be reduced to the discrete log problem on $\mathbb{F}_q^\times$ (which is much easier to solve). We show now how this can be done by an algorithm due to Victor Miller.

Algorithm 1. *Let E be an elliptic curve, and let $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ be non-trivial points of E. Let λ be the slope of the line L connecting P and Q (the slope of the tangent line at P if $P = Q$). Define the rational function*

$$g_{P,Q} = \begin{cases} \frac{y - y_P - \lambda(x - x_P)}{x + x_P + x_Q - \lambda^2} & \text{if } L \text{ is not vertical} \\ x - x_p & \text{if } L \text{ is vertical,} \end{cases}$$

which has divisor $[P] + [Q] - [P + Q] - [\mathcal{O}]$. Then a rational function f_P with divisor $m[P] - [mP] - (m-1)[\mathcal{O}]$ can be constructed as follows.

```
Let $m = (m_{n-1} \dots m_1 m_0)_2$ be the binary expansion of $m$.
Initialization step: $T \leftarrow P,\ f \leftarrow 1,\ i \leftarrow n - 2$
While $i \geq 0$
    $f \leftarrow f^2 \cdot g_{T,T}$
    $T \leftarrow 2T$
    If $m_i = 1$
        $f \leftarrow f \cdot g_{T,P}$
        $T \leftarrow T + P$
    $i \leftarrow i - 1$
Return $f$
```

Note, when implementing this algorithm over a finite field, some care needs to be taken in the definition of $g_{P,Q}$. Intuitively we like to think of vertical lines as corresponding to $\lambda = \infty$, however a straight forward implementation of λ over finite fields will probably yield $\lambda = 0$ instead. Since we also get $\lambda = 0$ for horizontal lines, it is not practical to use the value of λ alone as a strategy to check whether L is vertical.

Example 17. The elliptic curve $E : y^2 = x^3 + x$ over $\mathbb{F}_{43}$ is superspecial, and has full 11-torsion over $E(\mathbb{F}_{43^2})$. $E[11]$ has generators

$$P = (13, 24) \quad \text{and} \quad Q = (21\alpha + 6, \alpha + 23)$$

where $\alpha^2 - \alpha + 3 \equiv 0 \bmod 43$. Since $11 = (1011)_2$, then $n = 4$ and the steps of Miller's algorithm for f_P are broken down as follows

i	m_i	f_P	T
		1	P
2	0	$g_{P,P}$	$2P$
1	1	$g^2_{P,P}g_{2P,2P}$	$4P$
		$g^2_{P,P}g_{2P,2P}g_{4P,P}$	$5P$
0	1	$g^4_{P,P}g^2_{2P,2P}g^2_{4P,P}g_{5P,5P}$	$10P$
		$g^4_{P,P}g^2_{2P,2P}g^2_{4P,P}g_{5P,5P}g_{10P,P}$	$11P$

and likewise for f_Q, simply by replacing P with Q. Note, that P and Q are not 2-torsion points, nor are the other multiples of P and Q occurring in the algorithm. However, $10P + P = \mathcal{O}$ and $10Q + Q = \mathcal{O}$, so only in that case the tangent line is vertical. The point $S = (42, 16)$ has order 4, thus is suitable for computing the Weil pairing. We will forgo the full calculation, but we will show the computation of $g_{P,P}(Q + S)$, which is one of the factors in $f_P(Q + S)$. Since

$$\lambda = \frac{3 \cdot 13^2 + 1}{2 \cdot 24} = \frac{508}{48} = 7$$

in $\mathbb{F}_{43^2}$, then

$$g_{P,P} = \frac{y - 24 - 7(x - 13)}{x + 13 + 13 - 7^2} = \frac{y - 7x - 19}{x + 20}.$$

We compute $Q + S$ using addition on the elliptic curve, and the result is $(10\alpha + 5, 18\alpha + 3)$, so by plugging in

$$g_{P,P}(Q + S) = \frac{(18\alpha + 3) - 7(10\alpha + 5) - 19}{(10\alpha + 5) + 20} = 5\alpha + 16.$$

Thankfully, sage has the facility to compute the Weil pairing in full. We do this with the following code:

```
K.<z43>=GF(43^2);E=EllipticCurve(K,[0,0,0,1,0])
P=E(13,24);Q=E(21*z43+6,z43 + 23)
P.weil_pairing(Q,11)
```

Output: `(9*z43 + 19)`

meaning $e_{11}(P, Q) = 9\alpha + 19$. If we reverse P and Q, then we get $e_{11}(Q, P) = 34\alpha + 28$. Both are 11th roots of unity in $\mathbb{F}_{43^2}$, in fact that is why we need $\mathbb{F}_{43^2}$ and not $\mathbb{F}_{43}$, because 11 needs to divide the order of the multiplicative group. It is also true that $e_{11}(P, Q)$ and $e_{11}(Q, P)$ are multiplicative inverses of each other.

Definition 11. Let E be an elliptic curve over $\mathbb{F}_q$ and let ℓ be the largest prime dividing $E(\mathbb{F}_q)$. Then the *embedding degree* is the smallest integer k such that $\ell | q^k - 1$. That is to say, k is the order of $q \bmod \ell$.

Exercises

1. Using the same data above, compute $g_{2P,2P}(S)$
2. Use sage to construct the Weil pairing of $y^2 = x^3 + 6x^2 + x$ over $\mathbb{F}_{83^2}$ with $\ell = 7$.

11 Elliptic Curve Diffie-Hellman

The general setup is this. Let E be an elliptic curve defined over a finite field $\mathbb{F}_q$. Let ℓ be the largest prime dividing $|E(\mathbb{F}_q)|$, and let Q be a point of order ℓ on E. To generate a public and private key pair, Alice must must choose a random number K_A from $\{2, \ldots, \ell - 1\}$ used for encryption. The point $P_A = K_A \cdot Q$ is then the public key. The decryption key is just the inverse of K_A mod ℓ, and is computed by an extended gcd algorithm.[2]

Example 18. Alice and Bob wish to communicate using the elliptic curve $E : y^2 = x^3 + 71x^2 + x$ over $\mathbb{F}_{8147}$ from example 12. The point $P = (3270 : 7691 : 1)$ has order $\ell = 2081$, which is the largest prime dividing $E(\mathbb{F}_{8147})$. Alice and Bob each go onto random.org to pick a random number.

1. Alice gets $K_A = 1779$ which is her secret key. She computes $P_A = 1779P = (6472 : 3278 : 1)$, which is her public key.

2. Bob gets $K_B = 1425$ which is his secret key. He computes $P_B = 1425P = (2516 : 5975 : 1)$, which is his public key.

Alice and Bob do not reveal K_A and K_B to each other, but they do reveal P_A and P_B.

1. Using Bob's public key and her secret key, Alice computes $1779P_B = (2718 : 5748 : 1)$.

2. Using Alice's public Key and his secret key, Bob computes $1425P_A = (2718 : 5748 : 1)$.

Of course in view of example 12, Alice and Bob really only do this calculation using X and Z. In particular Alice only needs to send 6472, Bob, and he only needs to send 2516 to Alice. But this example was computed with sage. While there are protocols for encryption and decryption in Diffie-Hellman key exchange, Alice and Bob already agree on a common secret number, namely 2718. This number can be used to set up a key for a symmetric encryption algorithm such as AES, which is significantly faster.

Considering the heavy load that internet servers experience, and the general expectation that online communication should be practically instant, a great deal of effort has been spent to speed up the elliptic curve point addition algorithms and also to limit the amount of data that is sent. For example, As we have seen above, Koblitz curves are designed to have fast addition with the help of the frobenius endomorphism, and indeed he was one of the people to initiate the study of elliptic curve cryptography. At one time people thought that using the extra endomorphisms available to superspecial curves would be helpful, but eventually it was recognized that the low embedding degree of such curves made it easier to construct the Weil pairing. Once the weil pairing is constructed, the problem of breaking the encryption can be transfered to $\mathbb{F}_{q^k}^{\times}$ where it is easier to solve (see the MOV attack below).

Since P_A is a point on E, then we generally would expect that both the x and y-coordinates need to be sent. Numerous "point compression" algorithms have been published and patented, which allow for only a small number of bits of the y-coordinate to be sent. Many of the patents apply only in characteristic 2. However, as Bernstein has observed, if one wishes to avoid the patents, the easiest approach is simply to work with elliptic curves over $\mathbb{F}_p$, with p a sufficiently large odd prime, and avoid sending the y-coordinate altogether (for curves in Weierstrass form such as Montgomery curves). For Edwards curves with the addition law as specified by (22), we saw that if $P = (x_0, y_0)$ then $-P = (-x_0, y_0)$, so it would be the x-coordinate that does not need to be sent instead.

[2] Probably Schönhage's extended binary gcd, which has subquadratic efficiency.

For concreteness, consider the NIST recommendations for key strength:

AES	RSA	DH	ECDH	Hash (SHA or SHA3)
128	3072	3072	256	256
192	7680	7680	384	384
256	15360	15360	512	512

All values in bits.

If we are doing ordinary Diffie Hellman with an N bit prime, the process of modular exponentiation (in uniform time) requires $4(N-1)$ multiplications.[3] The fastest known multiplication algorithm has time complexity $O(N \log N \log\log N)$, thus the total time complexity of exponentiation is $O(N^2 \log N \log\log N)$, with an additional factor of 4 incurred. In the case of elliptic curves, the amount of multiplication and addition required depends largely on the model of the elliptic curve, and on the size of the coefficients. For Montgomery curves with small coefficients it is possible to get down to 15 big number multiplications, and again for an N bit prime there are $N-1$ steps, so we get the same time complexity, and for fixed N the ratio of the factors is $4/15$. However, the number of bits required for DH and ECDH are different, $3072/256 = 12$, and because of the N^2 we get a speed improvement of 12^2, so at the chosen level of security the elliptic curve algorithms are roughly $(4 \cdot 12^2)/15 = 38.4$ times faster. The higher levels yield even better improvement factors. Furthermore, by sending only the x coordinate, only 256 bits of data are sent instead of 3072 bits, which is also a significant improvement in terms of data transmission. We conclude, therefore, that the patented point compression algorithms are totally beside the point.

The basic criteria for safety for elliptic curves in Bernstein's view, and mine, are as follows.

Safety criteria. Choose a large odd prime p and an elliptic curve E over $\mathbb{F}_p$, satisfying all of the following conditions

1. The largest prime ℓ dividing $E(\mathbb{F}_p)$ is greater than some specified bound. This is computed by first computing a_p using the SEA algorithm, then factoring $p + 1 - a_p$.

2. The embedding degree is greater than some specified bound. This is computed as the multiplicative order of $p \bmod \ell$.

3. The CM discriminant is greater than some specified bound in absolute value. This can be check by verifying that the square-free part of $4p - a_p^2$ is greater than that bound.[4]

4. The addition law on E must be implemented in "uniform time," meaning that on average the computation time is the same, even when comparing with trivial or pathological cases like adding the identity.

5. There should be no way to manipulate the constants in the elliptic curve making it easier to attack.

[3] We first square repeatedly $N-1$ times, and then multiply $N-1$ times, each time we square or multiply we must also reduce mod p, which takes an extra multiplication (by $1/p$) and subtraction, for a total of $4(N-1)$ multiplications and $N-1$ subtractions). Since subtraction is asymptotically faster than multiplication, it can be ignored, and the computation of $1/p$ is universal within the problem, and thus can be precomputed once and for all.

[4] The exact value of the discriminant depends on whether or $a_p^2 - 4p \equiv 1 \bmod 4$. If this congruence holds then D is the squarefree part of $a_p^2 - 4p$, otherwise we multiply the squarefree part by 4. But D is always negative, and multiplying by 4 only makes it easier for the chosen bound to be satisfied.

The last two criteria deserve some explanation. If addition for an elliptic curve is implemented in a way that has detectable biases in speed, then an attacker could observe how long the calculations take and use the biases to determine the secret key. This is a universal problem in cryptography not just for elliptic curves. Say that you compute nP, by writing n in base 2, double repeatedly to get 2^kP for all k, and then add *only* the points needed. Then

$$15P = 8P + 4P + 2P + P \quad \text{and} \quad 9P = 8P + P$$

take a different number of operations, and therefore a different amount of time. Thus an attacker can tell how many zeros occur in the expansion for n. To get the same number of operations, you need to throw in the identity for each occurrence of 0 in the binary expansion of n, like this:

$$9P = 8P + \mathcal{O} + \mathcal{O} + P.$$

So when it was said above that modular exponentiation takes $4(N-1)$ multiplications, this is really exact in uniform time, since it must always take the same number of operations. Uniformity in time is possible to achieve for all elliptic curves, but it is easier for some than for others. In the case of Edwards curves, we are helped by the fact that there is no need for a separate doubling formula. For curves where the doubling formula is separate from the general formula, we must take exquisite care that amount of doubling and addition used is the same every time. As example 12 shows, this is true for Montgomery curves, so long as n is chosen so that 1 is the first and last bit in the binary expansion of n.

The issue of constants that can be manipulated is somewhat of a tricky question to pin down. We will illustrate the issue with an example of how the constant b can be manipulated for curves of the type

$$y^2 = x^3 - 3x + b.$$

What is alarming is that it is precisely curves of this type that NIST has included in their list. Granted their list is quite old and outdated by now. Montgomery curves do have some threat of manipulation if only one of E and its twist is safe (see the exercises).

When looking at the remaining criteria, it would seem that Koblitz curves are ruled out by the condition that p is large. If we were to drop the condition of large p, then Koblitz curves do very well with large ℓ and large embedding degree, but they still fail the CM discriminant criterion. Is this really such a problem? After all, *every* elliptic curve over a finite field has CM, because of the Frobenius endomorphism, so then why should the size of the discriminant matter? The speedups of the Frobenius endomorphism benefit both the person trying to secure their communications as well as the attacker, and in practice it would seem that the attacker gets somewhat of an edge. The speed advantage diminishes as the prime p increases in size, as pointed out already at the end of the section on Koblitz curves. If we are increasing p anyway, we might as well make it large enough so that all arithmetic is done over $\mathbb{F}_p$. Since the Frobenius endomorphism acts trivially over $\mathbb{F}_p$, then that would seem to take them completely out of the picture. But that if the Frobenius endomorphism is completely out of the picture over $\mathbb{F}_p$, then why should the discriminant matter? Sure, large p makes sense, but again, why the discriminant? The only reason that seems to make sense would be the existence of an efficient attack by lifting to a number field. No such attack is currently known, but there is also no known proof that such an attack can't exist. As such, the safety criteria take the cautious approach.

By now it should be clear that the design of a secure cryptosystem requires an understanding of everything that could possibly go wrong. For this reason we turn our attention now to attacks on elliptic curves, then we conclude the section by describing how to construct curves meeting the safety criteria.

11.1 Manipulating constants

It was mentioned above that the constant b in $E : y^2 = x^3 - 3x + b$ can be manipulated in an attack. The reason is that the constant b does not occur in the addition algorithms at all. If $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ are points on E the slope of the line through P and Q (tangent to the curve if $P = Q$) is

$$\lambda = \begin{cases} \frac{3(x_1^2-1)}{2y_2} & \text{if } P = Q \\ \frac{y_1-y_2}{x_1-x_2} & \text{otherwise.} \end{cases}$$

Consider the equation for the line through P and Q in point slope form: $y - y_1 = \lambda(x - x_1)$. By subtracting the equations $y^2 = x^3 - 3x + b$ and $y_1^2 = x_2^3 - 3x_1 + b$, b is completely eliminated:

$$\begin{aligned} y^2 - y_1^2 &= (x^3 - 3x + b) - (x_1^3 - 3x_1 + b) \\ (y - y_1)(y + y_1) &= (x - x_1)(x^2 + xx_1 + x_1^2 - 3(x + x_1)) \\ \lambda(y + y_1) &= x^2 + xx_1 + x_1^2 - 3(x + x_1). \end{aligned}$$

We then finish the calculation by eliminating y, and using $(x-x_2)(x-x_3) = x^2-s_1x+s_2$, where $s_1 = x_2+x_3$ and $s_2 = x_2x_3$ are the elementary symmetric polynomials in two variables. We thus obtain

$$x_3 = \lambda^2 - x_1 - x_2 + 3 \quad \text{and} \quad y_3 = -\lambda^3 + (2x_1 + x_2 - 3)\lambda - y_1$$

Surely an addition law not depending on b sounded like a good idea for improving computation speed, and the choice of $a_4 = -3$ makes sense from this perspective as well, since it makes the slope λ particularly simple. But the trouble is that an attacker could change the value of b while keeping the base field $\mathbb{F}_q$ the same, and get enough information to determine the key just by sending a few messages to the victim.

Example 19. Let's say that Carol chooses to use the curve $E : y^2 = x^3 - 3x + 10$ over the finite field $\mathbb{F}_{1153}$. Since $|E(\mathbb{F}_q)| = 1123$ is a prime also, this curve looks pretty safe. $Q = (933, 788)$ is a generator. Carol goes onto random.org and chooses $K_C = 633$. Seeing that Carol is using the finite field $\mathbb{F}_{1153}$, Eve searches for curves have points of smaller order over this field. She discovers:

1. The curve $E : y^2 = x^3-3x+3$ has order $2^2 \cdot 3^3 \cdot 11$; the points $P_3 = (24 : 107 : 1)$ and $P_{11} = (848 : 77 : 1)$ have order 3 and 11 respectively.
2. The curve $E : y^2 = x^3 - 3x + 8$ has order $2 \cdot 3^2 \cdot 5 \cdot 13$; the point $P_5 = (443 : 232 : 1)$ has order 5
3. The curve $E : y^2 = x^3 - 3x + 12$ has order $2^3 \cdot 3 \cdot 7^2$; the point $P_7 = (796 : 922 : 1)$ has order 7.

Eve pretends that each of these is a public key, and sends them to Carol to establish communication. Carol unwittingly computes

$$633 \cdot P_3 = (0 : 1 : 0), \quad 633 \cdot P_5 = (724 : 401 : 1), \quad P_7 = (33 : 369 : 1), \quad 633 \cdot P_{11} = (219 : 722 : 1),$$

and uses them to establish communication with Eve. So far Eve does not even know what points Carol computed, but there are only a small number of cases to check, $3 + 5 + 7 + 11 = 26$ (which is much smaller than 1123) so she systematically goes through each case comparing with the results that Carol sent, and stopping immediately when she gets a match.

As you can see $(0:1:0)$ is the identity, so in all likelihood Carol sent Eve that message in the clear, Eve recognized it immediately, and didn't need to compute anything. Since P_3 has order 3, then Even already knows $K_C \equiv 0 \bmod 3$. For the rest, Eve finds

$$K_C \equiv 3 \bmod 5, \quad K_C \equiv 3 \bmod 7, \quad K_C \equiv 6 \bmod 11,$$

which probably took her $3+3+6=12$ tries, which more or less makes sense, since on average we expect it to take half of the total number of cases. From these four congruences Eve can then recover $K_C = 633$ by the Chinese remainder theorem.

Why was this attack possible? First, because the addition law for $E : y^2 = x^3 - 3x + b$ does not depend on b, and second because Carol failed to check whether the points she was adding were actually on her curve. Since the addition law does not depend on b it is really better thought of as a group law on the projective plane over $\mathbb{F}_q$ itself. A point P has an order n with respect to this group law, irrespective of which curves it lies on. Any elliptic curve E on which P does lie will have n dividing $|E(\mathbb{F}_q)|$. Therefore since Carol's curve has prime order greater than 11, none of the points computed by Eve can lie on it. Indeed, sage will not let you add the points above on Carol's curve, because sage recognizes that they are not on the curve, and sage will complain accordingly. To prevent this attack on her curve Carol is forced to check whether the points sent to her are valid, which increases the computation time. On the other hand, if the group law depends uniquely on the choice of the elliptic curve, then this attack is not possible. As such, it seems that the wiser decision is not to use curves of the type $E : y^2 = x^3 - 3x + b$ at all, but to search for other models that don't have this peculiar feature. If we send only the x-coordinate, then Eve cannot tamper with the y-coordinate. For an Elliptic curve E over a finite field, an x-coordinate is always on E or its quadratic twist. We could check to see if the x-coordinate is valid, but Bernstein points out that even this may be skipped in the case of Montgomery curves. The reason is that the group law depends on A, and is the same both for E and its twist, so if both curves are safe then Eve doesn't get very far.

11.2 The MOV attack

The MOV attack is named after the people who discovered it (Menezes, Okamoto, and Vanstone), and it relies on the elliptic curve having low embedding degree. Since superspecial elliptic curves have low embedding degree, this is what makes them unsuitable.

Example 20. Consider the elliptic curve from example 17, namely $E : y^2 = x^3 + x$. If it is considered over $\mathbb{F}_{43}$, then $P = (13, 24)$ has order 11, and indeed $|E(\mathbb{F}_{43})| = 2^2 \cdot 11$, so $\ell = 11$ is the largest prime factor. If David goes onto random.org and picks $K_D = 6$ as his secret key, then $6P = (4, 5)$ is his public key. Eve knows that E has embedding degree 2, and so she sets up the Weil pairing, with $Q = (21\alpha + 6, \alpha + 23)$. She finds

$$e_{11}(P, Q) = 9\alpha + 19 \quad \text{and} \quad e_{11}(6P, Q) = 11\alpha + 34,$$

the first of which is familiar to us from 17. She now needs to find k such that

$$(9\alpha + 19)^k = 11\alpha + 34$$

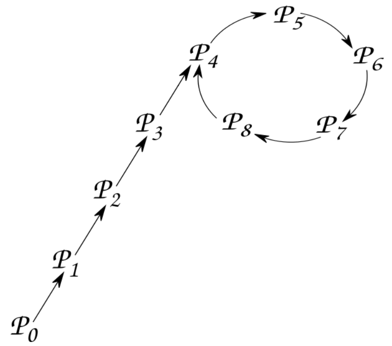
in $\mathbb{F}_{43^2}^\times$. As indicated by the NIST table above, 256 bits is considered acceptable for elliptic curves, but it is basically broken for ordinary Diffie-Hellman. The fastest method known for breaking Diffie-Hellman in $\mathbb{F}_q^\times$ is called "index calculus." In any case, it can be checked that

$$(9\alpha + 19)^6 = 11\alpha + 34$$

in $\mathbb{F}_{43^2}^\times$, so Eve wins again.

11.3 Pollard rho

The main idea behind the Pollard rho algorithm is to have a function $f : E \to E$ that is iterated, i.e. if P is a point on E then we consider the points $P_n = f(P)^n$, constructed by repeatedly applying f to P. What we are looking for is a pair of distinct points P_i and P_j such that $f(P_i) = f(P_j)$. Such a pair is called a "collision," see the diagram below.



The shape of the diagram looks like the Greek letter ρ, hence the name. The average length of time that it takes to find a collision can be estimated in terms of "waiting time," in the sense of probability. Let P be the base point used in the Diffe-Hellman setup with order ℓ, and let P_C be Carol's public key. Eve must find n such that $nP = P_C$. To do this, she uses the function

$$f(Q) = \begin{cases} P + Q & \text{if } 0 \leq x(Q) < \frac{p}{3}, \\ 2Q & \text{if } \frac{p}{3} \leq x(Q) < \frac{2p}{3}, \\ P_C + Q & \text{if } \frac{2p}{3} \leq x(Q) < p. \end{cases}$$

We can apply f repeatedly to any point in the group generated by P, but for our purposes we will define

$$P_i = f^i(P + P_C) = \alpha_i P + \beta_i P_C$$

for some $\alpha_i, \beta_i \in 0, 1, \ldots \ell - 1$. Then $\alpha_0 = \beta_0 = 1$, and the definition of f leads to the recursive formulas

$$\alpha_{i+1} = \begin{cases} \alpha_i + 1 & \text{if } 0 \leq x(Q) < \frac{p}{3}, \\ 2\alpha_i & \text{if } \frac{p}{3} \leq x(Q) < \frac{2p}{3}, \\ \alpha_i & \text{if } \frac{2p}{3} \leq x(Q) < p, \end{cases} \quad \text{and} \quad \beta_{i+1} = \begin{cases} \beta_i & \text{if } 0 \leq x(Q) < \frac{p}{3}, \\ 2\beta_i & \text{if } \frac{p}{3} \leq x(Q) < \frac{2p}{3}, \\ \beta_i + 1 & \text{if } \frac{2p}{3} \leq x(Q) < p. \end{cases}$$

If a collision is found, then

$$\alpha_i P + \beta_i P_C = \alpha_j P + \beta_j P_C \implies (\alpha_i - \alpha_j)P = (\beta_j - \beta_i)P_C.$$

Furthermore, if $\beta_j - \beta_i \not\equiv 0 \bmod \ell$, then it has an inverse and thus Eve obtains

$$K_C \equiv (\beta_j - \beta_i)^{-1}(\alpha_i - \alpha_j) \bmod \ell.$$

If we look only at the sequence $P_0, P_1, P_2 \ldots$, then all previous results must be kept when searching for a collision. However, Pollard observed that by looking at two sequences in parallel, a collision could still be found, while keeping only the most recent point in each sequence. Specifically, we define

$$Q_0 = P_0 = P + P_C, \quad P_{i+1} = f(P_i) \quad \text{and} \quad Q_{i+1} = f^2(Q_i).$$

Example 21. Carol is using the elliptic curve $E : y^2 = x^3 + 13x^2 + x$ over $\mathbb{F}_{67}$. The point $P = (21 : 73 : 1)$ has order $\ell = 47$. She goes on random.org and gets $K_C = 37$ as her secret key. She computes $P_C = (7 : 73 : 1)$ and sends it to David to establish communication. Eve intercepts P_C, and while Carol's choices are about the best that can be hoped for, for a prime p of this size, the prime p is rather small. Eve therefore sets to work with Pollard rho. For $p = 167$, the cut-offs in the definition of f are

$$\frac{p}{3} = 55\frac{2}{3} \quad \text{and} \quad \frac{2p}{3} = 111\frac{1}{3},$$

thus Eve obtains the following table

i	P_i	Q_i	$\alpha_i(P_i)$	$\beta_i(P_i)$	$\alpha_i(Q_i)$	$\beta_i(Q_i)$
0	$(126 : 94 : 1)$	$(126 : 94 : 1)$	1	1	1	1
1	$(121 : 123 : 1)$	$(54 : 6 : 1)$	1	2	1	3
2	$(54 : 6 : 1)$	$(126 : 73 : 1)$	1	3	2	4
3	$(121 : 44 : 1)$	$(0 : 1 : 0)$	2	3	3	5
4	$(126 : 73 : 1)$	$(2 : 88 : 1)$	2	4	5	5
5	$(21 : 94 : 1)$	$(130 : 138 : 1)$	2	5	7	5
6	$(0 : 1 : 0)$	$(19 : 81 : 1)$	3	5	8	6
7	$(21 : 73 : 1)$	$(18 : 83 : 1)$	4	5	9	7
8	$(2 : 88 : 1)$	$(33 : 148 : 1)$	5	5	20	14
9	$(28 : 38 : 1)$	$(28 : 129 : 1)$	6	5	42	28
10	$(130 : 138 : 1)$	$(21 : 94 : 1)$	7	5	44	28
11	$(31 : 8 : 1)$	$(21 : 73 : 1)$	7	6	46	28
12	$(19 : 81 : 1)$	$(28 : 38 : 1)$	8	6	48	28
13	$(130 : 29 : 1)$	$(31 : 8 : 1)$	9	6	49	29
14	$(18 : 83 : 1)$	$(130 : 29 : 1)$	9	7	51	29
15	$(89 : 98 : 1)$	$(89 : 98 : 1)$	10	7	52	30

From the last row, we see that

$$52P + 30P_C = 10P + 7P_C \implies 42P = -23P_C \implies 27P = 2P_C$$

and $(-23)^{-1} \equiv 2 \bmod 47$, so $2 \cdot 42 \equiv 37 \bmod 47$, thus Eve recovers Carol's secret key. As can be seen from the algorithm above, there are multiple ways we may be allowed to terminate earlier. Certainly once we obtain the identity, we should know that we have a solution:

$$3P + 5P_C = \mathcal{O} \implies 3P = -5P_C,$$

and since $(-5)^{-1} \equiv 28 \bmod 47$, Eve again gets $3 \cdot 28 \equiv 37 \bmod 47$. We also see both $(28 : 38 : 1)$ and $(28 : 129 : 1)$ occurring in the table, which we know are additive inverses, giving Eve another opportunity to stop early. There are other improvements that could be made as well, but none of them change the overall asymptotics. The number of steps expected before a collision occurs is proportional to $\sqrt{\ell}$ as ℓ grows.

11.4 Safe curves

In practice, it seems that designing safe cryposystems with Edwards curves is a bit easier than with curves in Weierstrass form. The group law has a nice symmetric form, which makes uniformity in time easy to achieve. Even though the formulas depend on both x and y it is possible to send only the y-coordinate. If an invalid y-coordinate is sent, then it will not be possible to reconstruct the x-coordinate meaning Eve has no hope of sending invalid data. The downside is that the formulas and the construction of the x-coordinate may be a little slower. To meet the safety criteria for and Edwards curve, we can take one of two approaches:

1. Choose d and then vary p until the safety criteria are met, or
2. Choose p and then vary d until the safety criteria are met.

The first choice (whether d or p) could be done at random to remove suspicion that the numbers were chosen with a hidden agenda. While p must be large, there does not seem to be a particular reason why d needs to be, in particular all of the examples of Edwards curves on Bernstein's website have small values of d and he lists the embedding degree, CM discriminant, and size of ℓ. In any case, once the curve E and prime p are set, a point P on E is chosen at random. Then if $|E(\mathbb{F}_p)| = m \cdot \ell$, where ℓ is the largest prime factor, we compute mP. If $mP \neq \mathcal{O}$, then we take mP to be the base point, otherwise we pick another random point and try again. Empirical tests for small primes p, suggests that $m = 4$ is common. Whether much control can be gained over the prime ℓ is less clear.

For Montgomery curves, both the curve E and its twist should be checked, with the idea that only the x-coordinate is sent and we do not check which of the two curves it is on. If the usual formulas for addition are used then we also require that $A^2 - 4$ is a quadratic non-residue mod p. Bernstein shows how to drop this condition in another article by changing the addition law. Finally, in order to get true uniformity, both the first and last bit of any encryption key must be 1. So, clearly there are extra steps in checking the safety of a Montgomery curve, but since addition formulas are quite fast. Again, while p must be large, it is still possible to meet all of the safety criteria with a small value of A. On the other hand, the data for small primes provide some strong hints about what optimal design might look like (See the problems).

Exercises

1. (optional) Write a program in sage or magma to compute Pollard Rho for the Koblitz curve E_1 over $\mathbb{F}_{2^7}$ with

   $$P = (\alpha^6 + \alpha^2 + 1 : \alpha^6 + \alpha^5 + \alpha^4 + \alpha^2 + \alpha + 1 : 1) \quad \text{and} \quad P_C = (\alpha^3 + \alpha + 1 : \alpha^6 + \alpha^4 + \alpha^3 + \alpha^2 + \alpha : 1)$$

   where α satisfies $x^7 + x + 1$.

2. Use sage or magma to execute the MOV attack on $y^2 = x^3 + 6x^2 + x$ over $\mathbb{F}_{43}$ with

   $$P = (4 : 32 : 1) \quad \text{and} \quad P_D = (16 : 31 : 1).$$

3. Given the elliptic curve

12 Problems

1. In this problem we will study curves in montgomery form $E_A : y^2 = x^3 + Ax^2 + x$ in an attempt to determine what effects rank might have on security.

1. Find a value of $A \in \mathbb{Z}$ for which E_A has rank 0 over $\mathbb{Q}$. Let
$$S_A(N) = \left\{ p : p \text{ is prime, } p \equiv 1 \bmod 4 \text{ and } \left(\frac{A^2 - 4}{p} \right) = -1 \right\},$$
Reduce mod p, for all $p \in S(N)$, and check the CM discriminant, embedding degree, and largest prime factor ℓ for both E_A and its twist.
2. Starting with the curve you found in part 1, choose x_0 at random such that $\alpha = \sqrt{x_0^3 + Ax_0^2 + x_0}$ is irrational, then construct the field extension $K = \mathbb{Q}(\alpha)$. Check that the new point (x_0, α) obtained on E_A is not a torsion point, so that E_A has positive rank over K. Determine which p in $S_A(N)$ split over K, which are inert, and which ramify. Divide the data found in the previous part into these cases, and detect any biases that occur.
3. Repeat the above steps with E_A having rank 1 over $\mathbb{Q}$.

2. Let $2n$ be an even number. Even though the Goldbach conjecture has not been proved, the emperical data for it is quite strong, and in fact suggests that as n grows, the number of ways that $2n = \ell_1 + \ell_2$ grows. Let $p = 4n - 1 = 2(\ell_1 + \ell_2) - 1$ and $a_p = 2(\ell_2 - \ell_1)$. Then
$$p + 1 - a_p = 4\ell_1 \quad \text{and} \quad p + 1 + a_p = 4\ell_2 \tag{37}$$
is a possible option for the number of points on an elliptic curve E and its twist over $\mathbb{F}_p$, so long as $|a_p| < 2\sqrt{p}$ or equivalently if $a_p^2 - 4p$ is negative. The discriminant is related to the squarefree part. By pluging in, we find
$$a_p^2 - 4p = 4((\ell_1 - \ell_2)^2 - 2(\ell_1 + \ell_2) + 1).$$
Using sage or magma obtain data for the following questions, and if possible try to formulate the asymptotics suggested by the data. Proving the asymptotics may be hard.

1. How often is $D = ((\ell_1 - \ell_2)^2 - 2(\ell_1 + \ell_2) + 1)$ squarefree?
2. For fixed p and all ℓ allowed by equation (37) and $|a_p| < 2\sqrt{p}$, how often does p mod ℓ have order $\ell - 1$?
3. If possible describe the relationship between the previous two results. Are the probabilities of D squarefree and maximal embedding degree independent? Does one imply the other? Is there correlation?
4. For fixed $p \equiv -1 \bmod 4$, look at all elliptic curves $E_1 : y^2 = x^3 + Ax^2 + x$, and their twists $E_2 : By^2 = x^3 + Ax^2 + x$, satisfying $\left(\frac{A^2-4}{p} \right) = -1$.
   - (a) How many have the property that $a_p^2 - 4p = 4D$ where D is squarefree?
   - (b) How many have maximal embedding degree for both curves?
   - (c) How many have the property that $|E_1(\mathbb{F}_p)| = 4\ell_1$ and $|E_2(\mathbb{F}_p)| = 4\ell_2$ for odd primes ℓ_1 and ℓ_2?
   - (d) As p grows, can you describe asymptotics for the upper and lower bounds of $|A|$? I.e. is it possible to achieve these conditions with small A even when p is large?
   - (e) As p grows, can you describe the probability that all of the above conditions are met?
   - (f) Is the any benefit to picking $\frac{\ell_i - 1}{2}$ to be a Sophie Germain prime? Are the probabilities any better?