CHAPTER 4

Elliptic Curves

Chapter Goals

- To describe what an elliptic curve is.
- To explain the basic mathematics behind elliptic curve cryptography.
- To show how projective coordinates can be used to improve computational efficiency.
- To show how point compression can be used to improve communications efficiency.

4.1. Introduction

This chapter is devoted to introducing elliptic curves. Some of the more modern public key systems make use of elliptic curves since they can offer improved efficiency and bandwidth. Since much of this book can be read with just the understanding that an elliptic curve provides another finite abelian group in which one can pose a discrete logarithm problem, you may decide to skip this chapter on an initial reading.

Let K be any field. The projective plane $\mathbb{P}^2(K)$ over K is defined as the set of triples

$$(X, Y, Z)$$

where $X, Y, Z \in K$ are not all simultaneously zero. On these triples is defined an equivalence relation

$$(X, Y, Z) \equiv (X_1, Y_1, Z_1)$$

if there exists a $\lambda \in K$ such that

$$X = \lambda \cdot X_1,\ Y = \lambda \cdot Y_1 \text{ and } Z = \lambda \cdot Z_1.$$

So, for example, if $K = \mathbb{F}_7$, the finite field of seven elements, then the two points

$$(4, 1, 1) \text{ and } (5, 3, 3)$$

are equivalent. Such a triple is called a projective point.

An *elliptic curve* over K will be defined as the set of solutions in the projective plane $\mathbb{P}^2(K)$ of a homogeneous Weierstrass equation of the form

$$E:\ Y^2 \cdot Z + a_1 \cdot X \cdot Y \cdot Z + a_3 \cdot Y \cdot Z^2 = X^3 + a_2 \cdot X^2 \cdot Z + a_4 \cdot X \cdot Z^2 + a_6 \cdot Z^3,$$

with $a_1, a_2, a_3, a_4, a_6 \in K$. This equation is also referred to as the long Weierstrass form. Such a curve should be non-singular in the sense that, if the equation is written in the form $F(X, Y, Z) = 0$, then the partial derivatives of the curve equation

$$\partial F/\partial X,\ \partial F/\partial Y \text{ and } \partial F/\partial Z$$

should not vanish simultaneously at any point on the curve, i.e. the three simultaneous equations have no zero defined over the algebraic closure $\overline{K}$.

The set of K-rational points on E, i.e. the solutions in $\mathbb{P}^2(K)$ to the above equation, is denoted by $E(K)$. Notice that the curve has exactly one rational point with coordinate Z equal to zero, namely $(0, 1, 0)$. This is called the point at infinity, which will be denoted by $\mathcal{O}$.

N.P. Smart, *Cryptography Made Simple*, Information Security and Cryptography, DOI 10.1007/978-3-319-21936-3_4

4.1.1. The Affine Form: For convenience, we will most often use the affine version of the Weierstrass equation, given by

$$E:\ Y^2 + a_1 \cdot X \cdot Y + a_3 \cdot Y = X^3 + a_2 \cdot X^2 + a_4 \cdot X + a_6, \tag{5}$$

where $a_i \in K$; this is obtained by setting $Z = 1$ in the above equation. The K-rational points in the affine case are the solutions to E in K^2, plus the point at infinity $\mathcal{O}$. Although most protocols for elliptic-curve-based cryptography make use of the affine form of a curve, it is often computationally important to be able to switch to projective coordinates. Luckily this switch is easy:

- The point at infinity always maps to the point at infinity in either direction.
- To map a projective point (X, Y, Z) which is not at infinity, so $Z \neq 0$, to an affine point we simply compute $(X/Z, Y/Z)$.
- To map an affine point (X, Y), which is not at infinity, to a projective point we take a random non-zero $Z \in K$ and compute $(X \cdot Z, Y \cdot Z, Z)$.

As we shall see later it is often more convenient to use a slightly modified form of projective point where the projective point (X, Y, Z) represents the affine point $(X/Z^2, Y/Z^3)$, which equates to using the projective equation

$$E:\ Y^2 + a_1 \cdot X \cdot Y \cdot Z + a_3 \cdot Y \cdot Z^3 = X^3 + a_2 \cdot X^2 \cdot Z^2 + a_4 \cdot X \cdot Z^4 + a_6 \cdot Z^6.$$

4.1.2. Isomorphisms of Elliptic Curves: Given an elliptic curve defined by equation (5), it is useful to define the following constants for use in later formulae:

$$\begin{aligned}
b_2 &= a_1^2 + 4 \cdot a_2, \\
b_4 &= a_1 \cdot a_3 + 2 \cdot a_4, \\
b_6 &= a_3^2 + 4 \cdot a_6, \\
b_8 &= a_1^2 \cdot a_6 + 4 \cdot a_2 \cdot a_6 - a_1 \cdot a_3 \cdot a_4 + a_2 \cdot a_3^2 - a_4^2, \\
c_4 &= b_2^2 - 24 \cdot b_4, \\
c_6 &= -b_2^3 + 36 \cdot b_2 \cdot b_4 - 216 \cdot b_6.
\end{aligned}$$

The discriminant of the curve is defined as

$$\Delta = -b_2^2 \cdot b_8 - 8 \cdot b_4^3 - 27 \cdot b_6^2 + 9 \cdot b_2 \cdot b_4 \cdot b_6.$$

When the characteristic of the field $\mathsf{char} K \neq 2, 3$ the discriminant can also be expressed as

$$\Delta = (c_4^3 - c_6^2)/1728.$$

Notice that $1728 = 2^6 \cdot 3^3$ so, if the characteristic of the underlying finite field is not equal to 2 or 3, dividing by this latter quantity makes sense. A curve is then non-singular if and only if $\Delta \neq 0$; from now on we shall assume that $\Delta \neq 0$ in all our discussions. When $\Delta \neq 0$, the j-invariant of the curve is defined as

$$j(E) = c_4^3/\Delta.$$

As an example, which we shall use throughout this chapter, we consider the elliptic curve

$$E:\ Y^2 = X^3 + X + 3$$

defined over the field $\mathbb{F}_7$. Computing the various quantities above we find that we have

$$\Delta = 3 \text{ and } j(E) = 5.$$

The j-invariant is closely related to the notion of elliptic curve isomorphism. Two elliptic curves defined by Weierstrass equations E (with variables X, Y) and E_1 (with variables X_1, Y_1) are isomorphic over K if and only if there exist constants $r, s, t \in K$ and $u \in K^*$, such that the change of variables

$$X = u^2 \cdot X_1 + r\ , \quad Y = u^3 \cdot Y_1 + s \cdot u^2 \cdot X_1 + t$$

transforms E into E_1. This transformation might look special, but arises from the requirements that the isomorphism should map the two points at infinity to themselves, and should keep unchanged the structure of the affine equation (5).

Such an isomorphism defines a bijection between the set of rational points in E and the set of rational points in E_1. Notice that isomorphism is defined relative to the field K. As an example consider again the elliptic curve

$$E:\ Y^2 = X^3 + X + 3$$

over the field $\mathbb{F}_7$. Now make the change of variables defined by $[u, r, s, t] = [2, 3, 4, 5]$, i.e.

$$X = 4 \cdot X_1 + 3 \text{ and } Y = Y_1 + 2 \cdot X_1 + 5.$$

We then obtain the isomorphic curve

$$E_1:\ Y_1^2 + 4 \cdot X_1 \cdot Y_1 + 3 \cdot Y_1 = X_1^3 + X_1 + 1,$$

and we have

$$j(E) = j(E_1) = 5.$$

Curve isomorphism is an equivalence relation. The following lemma establishes the fact that, over the algebraic closure $\overline{K}$, the j-invariant characterizes the equivalence classes in this relation.

Lemma 4.1. *Two elliptic curves that are isomorphic over K have the same j-invariant. Conversely, two curves with the same j-invariant are isomorphic over the algebraic closure $\overline{K}$.*

But curves with the same j-invariant may not necessarily be isomorphic over the ground field. For example, consider the elliptic curve, also over $\mathbb{F}_7$,

$$E_2:\ Y_2^2 = X_2^3 + 4 \cdot X_2 + 4.$$

This has j-invariant equal to 5 so it is isomorphic to E over $\overline{\mathbb{F}_7}$, but it is not isomorphic over $\mathbb{F}_7$ since the change of variable required is given by

$$X = 3 \cdot X_2 \text{ and } Y = \sqrt{6} \cdot Y_2.$$

However, $\sqrt{6} \notin \mathbb{F}_7$. Hence, we say both E and E_2 are defined over $\mathbb{F}_7$, but they are isomorphic over $\mathbb{F}_{7^2} = \mathbb{F}_7[\sqrt{6}] \subset \overline{\mathbb{F}_7}$.

4.2. The Group Law

Assume, for the moment, that $\mathsf{char}K \neq 2, 3$, and consider the change of variables given by

$$X = X_1 - \frac{b_2}{12},$$
$$Y = Y_1 - \frac{a_1}{2} \cdot \left(X_1 - \frac{b_2}{12} \right) - \frac{a_3}{2}.$$

This change of variables transforms the long Weierstrass form given in equation (5) to the equation of an isomorphic curve given in short Weierstrass form,

$$E:\ Y^2 = X^3 + a \cdot X + b,$$

for some $a, b \in K$. One can then define a group law on an elliptic curve using the chord-tangent process.

The chord process is defined as follows; see Figure 4.1 for a diagrammatic description. Let P and Q be two distinct points on E. The straight line joining P and Q must intersect the curve at one further point, say R, since we are intersecting a line with a cubic curve. The point R will also be defined over the same field of definition as the curve and the two points P and Q. If we then reflect R in the x-axis we obtain another point over the same field which we shall call $P + Q$.

The tangent process is given diagrammatically in Figure 4.2 or as follows, for a point P on the curve E. We take the tangent to the curve at P; such a line must intersect E in at most one other

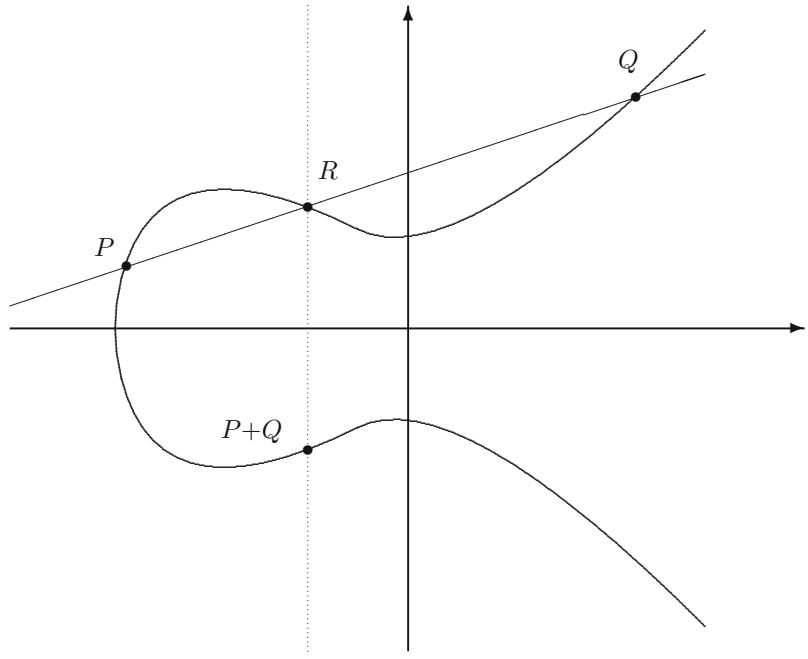

FIGURE 4.1. Adding two points on an elliptic curve

point, say R, as the elliptic curve E is defined by a cubic equation. Again we reflect R in the x-axis to obtain a point which we call $[2]P = P + P$. If the tangent to the point is vertical, it "intersects" the curve at the point at infinity and $P + P = \mathcal{O}$, and P is said to be a point of order 2.

One can show that the chord-tangent process turns E into an abelian group with the point at infinity $\mathcal{O}$ being the identity. The above definition can easily be extended to the long Weierstrass form (and so to characteristic two and three). One simply changes the definition by replacing "reflection in the x-axis" by "reflection in the line $Y = a_1 \cdot X + a_3$". In addition a little calculus will result in explicit algebraic formulae for the chord-tangent process. This is necessary since drawing diagrams as above is not really allowed in a field of finite characteristic. The algebraic formulae are summarized in the following lemma.

Lemma 4.2. *Let E denote an elliptic curve given by*

$$E: \; Y^2 + a_1 \cdot X \cdot Y + a_3 \cdot Y = X^3 + a_2 \cdot X^2 + a_4 \cdot X + a_6$$

and let $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ denote points on the curve. Then

$$-P_1 = (x_1, -y_1 - a_1 \cdot x_1 - a_3).$$

Set

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1},$$
$$\mu = \frac{y_1 \cdot x_2 - y_2 \cdot x_1}{x_2 - x_1}$$

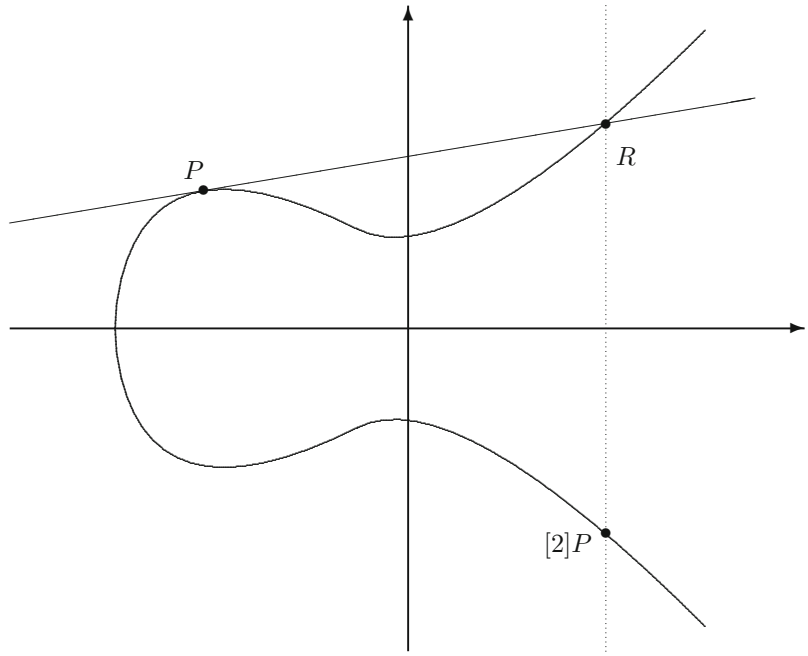

FIGURE 4.2. Doubling a point on an elliptic curve

when $x_1 \neq x_2$*, and set*

$$\lambda = \frac{3 \cdot x_1^2 + 2 \cdot a_2 \cdot x_1 + a_4 - a_1 \cdot y_1}{2 \cdot y_1 + a_1 \cdot x_1 + a_3},$$
$$\mu = \frac{-x_1^3 + a_4 \cdot x_1 + 2 \cdot a_6 - a_3 \cdot y_1}{2 \cdot y_1 + a_1 \cdot x_1 + a_3}$$

when $x_1 = x_2$ *and* $P_2 \neq -P_1$*. If*

$$P_3 = (x_3, y_3) = P_1 + P_2 \neq \mathcal{O}$$

then x_3 *and* y_3 *are given by the formulae*

$$x_3 = \lambda^2 + a_1 \cdot \lambda - a_2 - x_1 - x_2,$$
$$y_3 = -(\lambda + a_1) \cdot x_3 - \mu - a_3.$$

The elliptic curve isomorphisms described earlier then become group isomorphisms as they respect the group structure.

4.2.1. The Elliptic Curve Discrete Logarithm Problem (ECDLP): For a positive integer m we let $[m]$ denote the multiplication-by-m map from the curve to itself. This map takes a point P to

$$P + P + \cdots + P,$$

where we have m summands. This map is the basis of elliptic curve cryptography, since whilst it is easy to compute, it is believed to be hard to invert, i.e. given $P = (x, y)$ and $[m]P = (x', y')$ it is

hard to compute m. Of course this statement of hardness assumes a well-chosen elliptic curve etc., something we will return to later.

Example: We end this section with an example of the elliptic curve group law. Again we take our elliptic curve

$$E:\ Y^2 = X^3 + X + 3$$

over the field $\mathbb{F}_7$. It turns out there are six points on this curve given by

$$\mathcal{O},\ (4,1),\ (6,6),\ (5,0),\ (6,1) \text{ and } (4,6).$$

These form a group with the group law being given by the following table, which is computed using the addition formulae given above.

$+$	$\mathcal{O}$	$(4,1)$	$(6,6)$	$(5,0)$	$(6,1)$	$(4,6)$
$\mathcal{O}$	$\mathcal{O}$	$(4,1)$	$(6,6)$	$(5,0)$	$(6,1)$	$(4,6)$
$(4,1)$	$(4,1)$	$(6,6)$	$(5,0)$	$(6,1)$	$(4,6)$	$\mathcal{O}$
$(6,6)$	$(6,6)$	$(5,0)$	$(6,1)$	$(4,6)$	$\mathcal{O}$	$(4,1)$
$(5,0)$	$(5,0)$	$(6,1)$	$(4,6)$	$\mathcal{O}$	$(4,1)$	$(6,6)$
$(6,1)$	$(6,1)$	$(4,6)$	$\mathcal{O}$	$(4,1)$	$(6,6)$	$(5,0)$
$(4,6)$	$(4,6)$	$\mathcal{O}$	$(4,1)$	$(6,6)$	$(5,0)$	$(6,1)$

As an example of the multiplication-by-m map, if we let $P = (4,1)$ then we have

$$[2]P = (6,6),\ [3]P = (5,0),\ [4]P = (6,1),\ [5]P = (4,6),\ [6]P = \mathcal{O}.$$

So we see in this example that $E(\mathbb{F}_7)$ is a finite cyclic abelian group of order six generated by the point P. For all elliptic curves over finite fields the group is always finite and it is also highly likely to be cyclic (or "nearly" cyclic).

4.3. Elliptic Curves over Finite Fields

Over a finite field $\mathbb{F}_q$, the number of rational points on a curve is finite, and its size will be denoted by $\#E(\mathbb{F}_q)$. The expected number of points on the curve is around $q+1$ and if we set

$$\#E(\mathbb{F}_q) = q + 1 - t$$

then the value t is called the *trace of Frobenius* at q. A first approximation to the order of $E(\mathbb{F}_q)$ is given by the following well-known theorem of Hasse.

Theorem 4.3 (H. Hasse, 1933)**.** *The trace of Frobenius satisfies*

$$|t| \leq 2 \cdot \sqrt{q}.$$

Consider our example of

$$E:\ Y^2 = X^3 + X + 3$$

then recall this has six points over the field $\mathbb{F}_7$, and so the associated trace of Frobenius is equal to 2, which is less than $2 \cdot \sqrt{q} = 2 \cdot \sqrt{7} = 5.29$.

The qth-power Frobenius map, on an elliptic curve E defined over $\mathbb{F}_q$, is given by

$$\varphi : \begin{cases} E(\overline{\mathbb{F}}_q) \longrightarrow E(\overline{\mathbb{F}}_q) \\ (x,y) \longmapsto (x^q, y^q) \\ \mathcal{O} \longmapsto \mathcal{O}. \end{cases}$$

The map φ sends points on E to points on E, no matter what the field of definition of the point is. In addition the map φ respects the group law in that

$$\varphi(P+Q) = \varphi(P) + \varphi(Q).$$

In other words the map φ is a group endomorphism of E over the algebraic closure of $\mathbb{F}_q$, which is denoted by $\overline{\mathbb{F}}_q$, referred to as the Frobenius endomorphism. The trace of Frobenius t and the Frobenius endomorphism φ are linked by the equation

$$\varphi^2 - [t]\varphi + [q] = [0].$$

Hence, for any point $P = (x, y)$ on the curve, we have

$$(x^{q^2}, y^{q^2}) - [t](x^q, y^q) + [q](x, y) = \mathcal{O},$$

where addition and subtraction denote curve operations.

As was apparent from the earlier discussion, the cases char $K = 2, 3$ often require separate treatment. Practical implementations of elliptic curve cryptosystems are usually based on either $\mathbb{F}_{2^n}$, i.e. characteristic two, or $\mathbb{F}_p$ for large primes p. Therefore, in the remainder of this chapter we will focus on fields of characteristic two and $p > 3$, and will omit the separate treatment of the case $\mathsf{char}K = 3$. Most arguments, though, carry across easily to characteristic three, with modifications that are well documented in the literature.

4.3.1. Curves over Fields of Characteristic $p > 3$: Assume that our finite field is given by $K = \mathbb{F}_q$, where $q = p^n$ for a prime $p > 3$ and an integer $n \geq 1$. As mentioned, the curve equation in this case can be simplified to the short Weierstrass form

$$E: \ Y^2 = X^3 + a \cdot X + b.$$

The discriminant of the curve then reduces to $\Delta = -16 \cdot (4 \cdot a^3 + 27 \cdot b^2)$, and its j-invariant to $j(E) = -1728 \cdot (4 \cdot a)^3/\Delta$. The formulae for the group law in Lemma 4.2 also simplify to

$$-P_1 = (x_1, -y_1),$$

and if

$$P_3 = (x_3, y_3) = P_1 + P_2 \neq \mathcal{O},$$

then x_3 and y_3 are given by the formulae

$$x_3 = \lambda^2 - x_1 - x_2,$$
$$y_3 = (x_1 - x_3) \cdot \lambda - y_1,$$

where if $x_1 \neq x_2$ we set

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

and if $x_1 = x_2$, $y_1 \neq 0$ we set

$$\lambda = \frac{3 \cdot x_1^2 + a}{2 \cdot y_1}.$$

4.3.2. Curves over Fields of Characteristic Two: We now specialize to the case of finite fields where $q = 2^n$ with $n \geq 1$. In this case, the expression for the j-invariant reduces to $j(E) = a_1^{12}/\Delta$. In characteristic two, the condition $j(E) = 0$, i.e. $a_1 = 0$, is equivalent to the curve being supersingular. As mentioned earlier, this very special type of curve is avoided in cryptography. We assume, therefore, that $j(E) \neq 0$.

Under these assumptions, a representative for each isomorphism class of elliptic curves over $\mathbb{F}_q$ is given by

$$E: \ Y^2 + X \cdot Y = X^3 + a_2 \cdot X^2 + a_6, \tag{6}$$

where $a_6 \in \mathbb{F}_q^*$ and $a_2 \in \{0, \gamma\}$ with γ a fixed element in $\mathbb{F}_q$ such that $\mathrm{Tr}_{q|2}(\gamma) = 1$, where $\mathrm{Tr}_{q|2}$ is the absolute trace

$$\mathrm{Tr}_{2^n|2}(\alpha) = \sum_{i=0}^{n-1} \alpha^{2^i}.$$

The formulae for the group law in Lemma 4.2 then simplify to

$$-P_1 = (x_1, y_1 + x_1),$$

and if

$$P_3 = (x_3, y_3) = P_1 + P_2 \neq \mathcal{O},$$

then x_3 and y_3 are given by the formulae

$$\begin{aligned} x_3 &= \lambda^2 + \lambda + a_2 + x_1 + x_2, \\ y_3 &= (\lambda + 1) \cdot x_3 + \mu \\ &= (x_1 + x_3) \cdot \lambda + x_3 + y_1, \end{aligned}$$

where if $x_1 \neq x_2$ we set

$$\begin{aligned} \lambda &= \frac{y_2 + y_1}{x_2 + x_1}, \\ \mu &= \frac{y_1 \cdot x_2 + y_2 \cdot x_1}{x_2 + x_1} \end{aligned}$$

and if $x_1 = x_2 \neq 0$ we set

$$\begin{aligned} \lambda &= \frac{x_1^2 + y_1}{x_1}, \\ \mu &= x_1^2. \end{aligned}$$

4.4. Projective Coordinates

One of the problems with the above formulae for the group laws, given in both large and even characteristic, is that at some stage they involve a division operation. Division in finite fields is considered to be an expensive operation, since it usually involves some variant of the extended Euclidean algorithm, which although of approximately the same complexity as multiplication can usually not be implemented as efficiently.

To avoid these division operations one can use projective coordinates. Here one writes the elliptic curve using three variables (X, Y, Z) instead of just (X, Y). Instead of using the projective representation given at the start of this chapter we instead use one where the curve is written as

$$E: \ Y^2 + a_1 \cdot X \cdot Y \cdot Z + a_2 \cdot Y \cdot Z^4 = X^3 + a_2 \cdot X^2 \cdot Z^2 + a_4 \cdot X \cdot Z^4 + a_6 \cdot Z^6.$$

The point at infinity is still denoted by $(0, 1, 0)$, but now the map from projective to affine coordinates is given by

$$(X, Y, Z) \longmapsto (X/Z^2, Y/Z^3).$$

This choice of projective coordinates is made to provide a more efficient arithmetic operation.

4.4.1. Large Prime Characteristic: The formulae for point addition when our elliptic curve is written as

$$E: \ Y^2 = X^3 + a \cdot X \cdot Z^4 + b \cdot Z^6$$

are now given by the law

$$(X_3, Y_3, Z_3) = (X_1, Y_1, Z_1) + (X_2, Y_2, Z_2)$$

where (X_3, Y_3, Z_3) are derived from the formulae

$$\begin{aligned}
\lambda_1 &= X_1 \cdot Z_2^2, & \lambda_2 &= X_2 \cdot Z_1^2,\\
\lambda_3 &= \lambda_1 - \lambda_2, & \lambda_4 &= Y_1 \cdot Z_2^3,\\
\lambda_5 &= Y_2 \cdot Z_1^3, & \lambda_6 &= \lambda_4 - \lambda_5,\\
\lambda_7 &= \lambda_1 + \lambda_2, & \lambda_8 &= \lambda_4 + \lambda_5,\\
Z_3 &= Z_1 \cdot Z_2 \cdot \lambda_3, & X_3 &= \lambda_6^2 - \lambda_7 \cdot \lambda_3^2,\\
\lambda_9 &= \lambda_7 \cdot \lambda_3^2 - 2 \cdot X_3, & Y_3 &= (\lambda_9 \cdot \lambda_6 - \lambda_8 \cdot \lambda_3^3)/2.
\end{aligned}$$

Notice the avoidance of any division operation, bar division by 2 which can be easily accomplished by multiplication of the precomputed value of $2^{-1} \pmod p$. Doubling a point,

$$(X_3, Y_3, Z_3) = [2](X_1, Y_1, Z_1),$$

can be accomplished using the formulae

$$\begin{aligned}
\lambda_1 &= 3 \cdot X_1^2 + a \cdot Z_1^4, & Z_3 &= 2 \cdot Y_1 \cdot Z_1,\\
\lambda_2 &= 4 \cdot X_1 \cdot Y_1^2, & X_3 &= \lambda_1^2 - 2 \cdot \lambda_2,\\
\lambda_3 &= 8 \cdot Y_1^4, & Y_3 &= \lambda_1 \cdot (\lambda_2 - X_3) - \lambda_3.
\end{aligned}$$

4.4.2. Even Characteristic: In even characteristic we write our elliptic curve in the form

$$E: \; Y^2 + X \cdot Y \cdot Z = X^3 + a_2 \cdot X^2 \cdot Z^2 + a_6 \cdot Z^6.$$

Point addition,

$$(X_3, Y_3, Z_3) = (X_1, Y_1, Z_1) + (X_2, Y_2, Z_2)$$

is now accomplished using the recipe

$$\begin{aligned}
\lambda_1 &= X_1 \cdot Z_2^2, & \lambda_2 &= X_2 \cdot Z_1^2,\\
\lambda_3 &= \lambda_1 + \lambda_2, & \lambda_4 &= Y_1 \cdot Z_2^3,\\
\lambda_5 &= Y_2 \cdot Z_1^3, & \lambda_6 &= \lambda_4 + \lambda_5,\\
\lambda_7 &= Z_1 \cdot \lambda_3, & \lambda_8 &= \lambda_6 \cdot X_2 + \lambda_7 \cdot Y_2,\\
Z_3 &= \lambda_7 \cdot Z_2, & \lambda_9 &= \lambda_6 + Z_3,\\
X_3 &= a_2 \cdot Z_3^2 + \lambda_6 \cdot \lambda_9 + \lambda_3^3, & Y_3 &= \lambda_9 \cdot X_3 + \lambda_8 \cdot \lambda_7^2.
\end{aligned}$$

Doubling is performed using

$$\begin{aligned}
Z_3 &= X_1 \cdot Z_1^2, & X_3 &= (X_1 + d_6 \cdot Z_1^2)^4,\\
\lambda &= Z_3 + X_1^2 + Y_1 \cdot Z_1, & Y_3 &= X_1^4 \cdot Z_3 + \lambda \cdot X_3,
\end{aligned}$$

where $d_6 = \sqrt[4]{a_6}$. Notice how in both even and odd characteristic we have avoided a division operation when performing curve operations.

4.5. Point Compression

In many cryptographic protocols we need to store or transmit an elliptic curve point. Using affine coordinates this can be accomplished using two field elements, i.e. by transmitting x and then y. However, one can do better using a technique called point compression. Point compression is based on the observation that for every x-coordinate on the curve there are at most two corresponding y-coordinates. Hence, we can represent a point by storing the x-coordinate along with a bit b to say which value of the y-coordinate we should take. All that remains to decide is how to compute the bit b and how to reconstruct the y-coordinate given the x-coordinate and the bit b.

Large Prime Characteristic: For elliptic curves over fields of large prime characteristic we notice that if $\alpha \in \mathbb{F}_p^*$ is a square, then the two square roots $\pm\beta$ of α have different parities when represented as integers in the range $[1, \ldots, p-1]$. This is because $-\beta = p - \beta$. Hence, as the bit b we choose the parity of the y-coordinate. Given (x, b), we can reconstruct y by computing

$$\beta = \sqrt{x^3 + a \cdot x + b} \pmod{p}.$$

If the parity of β is equal to b we set $y = \beta$, otherwise we set $y = p - \beta$. If $\beta = 0$ then no matter which value of b we have we set $y = 0$.

As an example consider the curve

$$E: \ Y^2 = X^3 + X + 3$$

over the field $\mathbb{F}_7$. Then the points $(4, 1)$ and $(4, 6)$ which in bits we need to represent as

$$(0b100, 0b001) \text{ and } (0b100, 0b110),$$

i.e. requiring six bits for each point, can be represented as

$$(0b100, 0b1) \text{ and } (0b100, 0b0),$$

where we only use four bits for each point. In larger, cryptographically interesting, examples the advantage becomes more pronounced. For example consider the curve with the same coefficients but over the finite field $\mathbb{F}_p$ where

$$p = 1\,125\,899\,906\,842\,679 = 2^{50} + 55$$

then the point

$$(1\,125\,899\,906\,842\,675,\ 245\,132\,605\,757\,739)$$

can be represented by the integers

$$(1\,125\,899\,906\,842\,675,\ 1).$$

So instead of requiring 102 bits we only require 52 bits.

Even Characteristic: In even characteristic we need to be slightly more clever. Suppose we are given a point $P = (x, y)$ on the elliptic curve

$$E: \ Y^2 + X \cdot Y = X^3 + a_2 \cdot X + a_6.$$

If $y = 0$ then we set $b = 0$, otherwise we compute

$$z = y/x$$

and let b denote the least significant bit of z. To recover y given (x, b), for $x \neq 0$, we set

$$\alpha = x + a_2 + \frac{a_6}{x^2}$$

and let β denote a solution of

$$z^2 + z = \alpha.$$

Then if the least significant bit of β is equal to b we set $y = x \cdot \beta$, otherwise we set $y = x \cdot (\beta + 1)$. To see why this works notice that if (x, y) is a solution of

$$E: \ Y^2 + XY = X^3 + a_2 \cdot X^2 + a_6$$

then $(x, y/x)$ and $(x, 1 + y/x)$ are the two solutions of

$$Z^2 + Z = X + a_2 + \frac{a_6}{X^2}.$$

4.6. Choosing an Elliptic Curve

One of the advantages of elliptic curves is that there is a very large number of possible groups. One can choose both the finite field and the coefficients of the curve. In addition finding elliptic curves with the correct cryptographic properties to make the systems using them secure is relatively easy; we just have to know which curves to avoid.

For any elliptic curve and any finite field the group order $\#E(\mathbb{F}_q)$ can be computed in polynomial time. But this is usually done via a complicated algorithm that we cannot go into in this book. Hence, you should just remember that computing the group order is computationally easy. As we saw in Chapter 3, when considering algorithms to solve discrete logarithm problems, knowing the group order is important in understanding how secure a group is. For some elliptic curves computing the group order is easy; in particular supersingular curves. The curve $E(\mathbb{F}_q)$ is said to be supersingular if the characteristic p divides the trace of Frobenius, t. If $q = p$ then this means that $E(\mathbb{F}_p)$ has $p + 1$ points since we must have $t = 0$. For other finite fields the possible values of t corresponding to supersingular elliptic curves are given by, where $q = p^f$,

- f odd: $t = 0$, $t^2 = 2q$ and $t^2 = 3q$.
- f even: $t^2 = 4q$, $t^2 = q$ if $p = 1 \pmod 3$ and $t = 0$ if $p \neq 1 \pmod 4$.

For elliptic curves there are no known sub-exponential methods for the discrete logarithm problem, except in certain special cases. There are three particular classes of curves which, under certain conditions, will prove to be cryptographically weak:

- The curve $E(\mathbb{F}_q)$ is said to be anomalous if its trace of Frobenius is one, giving $\#E(\mathbb{F}_q) = q$. These curves are weak when $q = p$, the field characteristic. In this case there is an algorithm to solve the discrete logarithm problem which requires $O(\log p)$ elliptic curve operations.
- For any q we must choose curves for which there is no small number t such that r divides $q^t - 1$, where r is the large prime factor of $\#E(\mathbb{F}_q)$. This alas eliminates the supersingular curves and a few others. In this case there is a simple computable mapping from the elliptic curve discrete logarithm problem to the discrete logarithm problem in the finite field $\mathbb{F}_{q^t}$. Hence, in this case we obtain a sub-exponential method for solving the elliptic curve discrete logarithm problem.
- If $q = 2^n$ then we usually assume that n is prime to avoid the possibility of certain attacks based on the concept of "Weil descent".

One should treat these three special cases much like one treats the generation of large integers for the RSA algorithm. Due to the P − 1 factoring method one often makes RSA moduli $N = p \cdot q$ such that p is a so-called safe prime of the form $2p_1 + 1$. Another special RSA-based case is that we almost always use RSA with a modulus having two prime factors, rather than three or four. This is because moduli with two prime factors appear to be the hardest to factor.

It turns out that the only known practical method to solve the discrete logarithm problem in general elliptic curves is the parallel version of Pollard's Rho method given in Chapter 3. Thus we need to choose a curve such that the group order $\#E(\mathbb{F}_q)$ is divisible by a large prime number r, and for which the curve is not considered weak by the above considerations. Hence, from now on we suppose the elliptic curve E is defined over the finite field $\mathbb{F}_q$ and we have

$$\#E(\mathbb{F}_q) = h \cdot r$$

where r is a "large" prime number and h is a small number called the cofactor. By Hasse's Theorem 4.3 the value of $\#E(\mathbb{F}_q)$ is close to q so we typically choose a curve with r close to q, i.e. we choose a curve E so that $h = 1, 2$ or 4.

Since the best general algorithm known for the elliptic curve discrete logarithm problem is the parallel Pollard's Rho method, which has complexity $O(\sqrt{r})$, which is about $O(\sqrt{q})$, to achieve the same security as a 128-bit block cipher we need to take $q \approx 2^{256}$, which is a lot smaller than the

field size recommended for systems based on the discrete logarithm problems in a finite field. This results in the reduced bandwidth and computational times of elliptic curve systems.

Chapter Summary

- Elliptic curves over finite fields are another example of a finite abelian group. There are many such groups since we are free to choose both the curve and the field.
- For cryptography we need to be able to compute the number of elements in the group. Although this is done using a complicated algorithm, it can be done in polynomial time.
- One should usually avoid supersingular and anomalous curves in cryptographic applications.
- Efficient algorithms for the group law can be produced by using projective coordinates. These algorithms avoid the need for costly division operations in the underlying finite field.
- To save bandwidth and space it is possible to efficiently compress elliptic curve points (x, y) down to x and a single bit b. The uncompression can also be performed efficiently.
- For elliptic curves there are no sub-exponential algorithms known for the discrete logarithm problem, except in very special cases. Hence, the only practical general algorithm to solve the discrete logarithm problem on an elliptic curve is the parallel Pollard's Rho method.

Further Reading

Those who wish to learn more about elliptic curves in general may try the textbook by Silverman (which is really aimed at mathematics graduate students). Those who are simply interested in the cryptographic applications of elliptic curves and the associated algorithms and techniques may see the book by Blake, Seroussi and Smart and its follow-up book.

I.F. Blake, G. Seroussi and N.P. Smart. *Elliptic Curves in Cryptography.* Cambridge University Press, 1999.

I.F. Blake, G. Seroussi and N.P. Smart. *Advances in Elliptic Curve Cryptography.* Cambridge University Press, 2004.

J.H. Silverman. *The Arithmetic of Elliptic Curves.* Springer, 1985.