Chapter 16
Elliptic Curves

In the mid-1980s, Miller and Koblitz introduced elliptic curves into cryptography, and Lenstra showed how to use elliptic curves to factor integers. Since that time, elliptic curves have played an increasingly important role in many cryptographic situations. One of their advantages is that they seem to offer a level of security comparable to classical cryptosystems that use much larger key sizes. For example, it is estimated in [Blake et al.] that certain conventional systems with a 4096-bit key size can be replaced by 313-bit elliptic curve systems. Using much shorter numbers can represent a considerable savings in hardware implementations.

In this chapter, we present some of the highlights. For more details on elliptic curves and their cryptologic uses, see [Blake et al.], [Hankerson et al.], or [Washington]. For a list of elliptic curves recommended by NIST for cryptographic uses, see [FIPS 186-2].

16.1 The Addition Law

An elliptic curve E is the graph of an equation

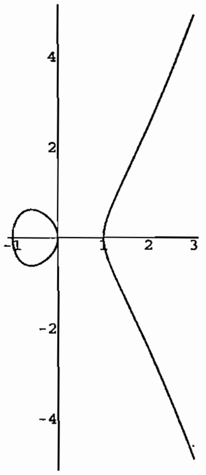
$$E: \quad y^2 = x^3 + ax^2 + bx + c,$$

where a, b, c are in whatever is the appropriate set (rational numbers, real numbers, integers mod p, etc.). In other words, let K be the rational numbers, the real numbers, or the integers mod a prime p (or, for those who know what this means, any field of characteristic not 2). Then we assume $a, b, c \in K$ and take E to be

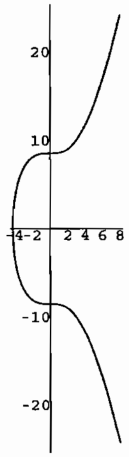
$$\{(x, y) \mid x, y \in K, y^2 = x^3 + ax^2 + bx + c\}.$$

As will be discussed below, it is also convenient to include a point (∞, ∞), which often will be denoted simply by ∞.

Let's consider the case of real numbers first, since this case allows us to work with pictures. The graph E has two possible forms, depending on whether the cubic polynomial has one real root or three real roots. For example, the graphs of $y^2 = x(x+1)(x-1)$ and $y^2 = x^3 + 73$ are the following:



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



$y^2 = x^3 + 73$

The case of two components (for example, $y^2 = x(x+1)(x-1)$) occurs when the cubic polynomial has 3 real roots. The case of one component (for example, $y^2 = x^3 + 73$) occurs when the cubic polynomial has only one real root.

For technical reasons that will become clear later, we also include a **"point at infinity,"** denoted ∞, which is most easily regarded as sitting at the top of the y-axis. It can be treated rigorously in the context of projective geometry (see [Washington]), but this intuitive notion suffices for what we need. The bottom of the y-axis is identified with the top, so ∞ also sits at the bottom of the y-axis.

Now let's look at elliptic curves mod p, where p is a prime. For example, let E be given by

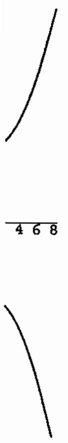
$$y^2 \equiv x^3 + 2x - 1 \pmod 5.$$

Then we assume

}.

e a point (∞,∞),

ıis case allows us
ms, depending on
e real roots. For
$x^3 + 73$ are the



$^3 + 73$

$1)(x-1))$ occurs
ıe component (for
has only one real

ve also include a
ırded as sitting at
ıtext of projective
fices for what we
so ∞ also sits at

me. For example,

We can list the points on E by letting x run through the values $0, 1, 2, 3, 4$ and solving for y:

$$(0,2), (0,3), (2,1), (2,4), (4,1), (4,4), \infty.$$

Note that we again include a point ∞.

Elliptic curves mod p are finite sets of points. It is these elliptic curves that are useful in cryptography.

Technical point: We assume that the cubic polynomial $x^3 + ax^2 + bx + c$ has no multiple roots. This means we exclude, for example, the graph of $y^2 = (x-1)^2(x+2)$. Such curves will be discussed in Section 16.3.

Technical point: For most situations, equations of the form $y^2 = x^3 + bx + c$ suffice for elliptic curves. In fact, in situations where we can divide by 3, a change of variables changes an equation $y^2 = x^3 + ax^2 + bx + c$ into an equation of the form $y^2 = x^3 + b'x + c'$. See Exercise 1. However, sometimes it is necessary to consider elliptic curves given by equations of the form

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

where $a_1, \ldots, a_6$ are constants. If we are working mod p, where $p > 3$ is prime, or if we are working with real, rational, or complex numbers, then simple changes of variables transform the present equation into the form $y^2 = x^3 + bx + c$. However, if we are working mod 2 or mod 3, or with a finite field of characteristic 2 or 3 (that is, $1 + 1 = 0$ or $1 + 1 + 1 = 0$), then we need to use the more general form. Elliptic curves over fields of characteristic 2 will be mentioned briefly in Section 16.4.

Historical point: Elliptic curves are not ellipses. They received their name from their relation to *elliptic integrals* such as

$$\int_{z_1}^{z_2} \frac{dx}{\sqrt{x^3 + bx + c}} \quad \text{and} \quad \int_{z_1}^{z_2} \frac{x\,dx}{\sqrt{x^3 + bx + c}}$$

that arise in the computation of the arc length of ellipses.

The main reason elliptic curves are important is that we can use any two points on the curve to produce a third point on the curve. Given points P_1 and P_2 on E, we obtain a third point P_3 on E as follows (see Figure 16.1): Draw the line L through P_1 and P_2 (if $P_1 = P_2$, take the tangent line to E at P_1). The line L intersects E in a third point Q. Reflect Q through the x-axis (i.e., change y to $-y$) to get P_3. Define a law of addition on E by

$$P_1 + P_2 = P_3.$$

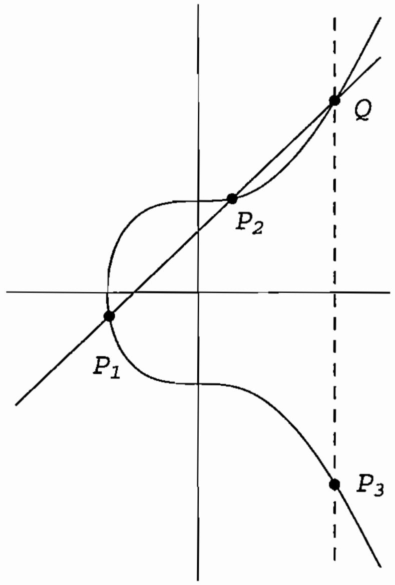


Figure 16.1: Adding Points on an Elliptic Curve.

Note that this is not the same as adding points in the plane.

Example. Suppose E is defined by $y^2 = x^3 + 73$. Let $P_1 = (2, 9)$ and $P_2 = (3, 10)$. The line L through P_1 and P_2 is

$$y = x + 7.$$

Substituting into the equation for E yields

$$(x + 7)^2 = x^3 + 73,$$

which yields $x^3 - x^2 - 14x + 24 = 0$. Since L intersects E in P_1 and P_2, we already know two roots, namely $x = 2$ and $x = 3$. Moreover, the sum of the three roots is minus the coefficient of x^2 (Exercise 1) and therefore equals 1. If x is the third root, then

$$2 + 3 + x = 1,$$

so the third point of intersection has $x = -4$. Since $y = x + 7$, we have $y = 3$, and $Q = (-4, 3)$. Reflect across the x-axis to obtain

$$(2, 0) + (3, 10) = P_3 = (-4, -3).$$

$\iota = (2, 9)$ and

P_1 and P_2, we
the sum of the
ıerefore equals

$: + 7$, we have

Now suppose we want to add P_3 to itself. The slope of the tangent line to E at P_3 is obtained by implicitly differentiating the equation for E:

$$2y\,dy = 3x^2\,dx, \quad \text{so} \quad \frac{dy}{dx} = \frac{3x^2}{2y} = -8,$$

where we have substituted $(x, y) = (-4, -3)$ from P_3. In this case, the line L is $y = -8(x + 4) - 3$. Substituting into the equation for E yields

$$(-8(x + 4) - 3)^2 = x^3 + 73,$$

hence $x^3 - (-8)^2x^2 + \cdots = 0$. The sum of the three roots is 64 (= minus the coefficient of x^2). Because the line L is tangent to E, it follows that $x = -4$ is a double root. Therefore,

$$(-4) + (-4) + x = 64,$$

so the third root is $x = 72$. The corresponding value of y (use the equation of L) is -611. Changing y to $-y$ yields

$$P_3 + P_3 = (72, 611). \qquad \blacksquare$$

What happens if we try to compute $P + \infty$? We make the convention that the lines through ∞ are vertical. Therefore, the line through $P = (x, y)$ and ∞ intersects E in P and also in $(x, -y)$. When we reflect $(x, -y)$ across the x-axis, we get back $P = (x, y)$. Therefore,

$$P + \infty = P.$$

We can also subtract points. First, observe that the line through (x, y) and $(x, -y)$ is vertical, so the third point of intersection with E is ∞. The reflection across the x-axis is still ∞ (that's what we meant when we said ∞ sits at the top and at the bottom of the y-axis). Therefore,

$$(x, y) + (x, -y) = \infty.$$

Since ∞ plays the role of an additive identity (in the same way that 0 is the identity for addition), we define

$$-(x, y) = (x, -y).$$

To subtract points $P - Q$, simply add P and $-Q$.

Another way to express the addition law is to say that

$$P + Q + R = \infty \iff P, Q, R \text{ are collinear.}$$

(see Exercise 10).

For computations, we can ignore the geometrical interpretation and work only with formulas, which are as follows:

Addition Law. *Let E be given by $y^2 = x^3 + bx + c$ and let*

$$P_1 = (x_1, y_2), \quad P_2 = (x_2, y_2).$$

Then

$$P_1 + P_2 = P_3 = (x_3, y_3),$$

where

$$\begin{aligned} x_3 &= m^2 - x_1 - x_2 \\ y_3 &= m(x_1 - x_3) - y_1 \end{aligned}$$

and

$$m = \begin{cases} (y_2 - y_1)/(x_2 - x_1) & \text{if } P_1 \neq P_2 \\ (3x_1^2 + b)/(2y_1) & \text{if } P_1 = P_2. \end{cases}$$

If the slope m is infinite, then $P_3 = \infty$. There is one additional law: $\infty + P = P$ for all points P.

It can be shown that the addition law is associative:

$$(P + Q) + R = P + (Q + R).$$

It is also commutative:

$$P + Q = Q + P.$$

When adding several points, it therefore doesn't matter in what order the points are added nor how they are grouped together. In technical terms, we have found that the points of E form an abelian group. The point ∞ is the identity element of this group.

16.2 Elliptic Curves Mod p

If p is a prime, we can work with elliptic curves mod p using the aforementioned ideas. For example, consider

$$E : y^2 \equiv x^3 + 4x + 4 \pmod{5}.$$

The points on E are the pairs (x, y) mod 5 that satisfy the equation, along with the point at infinity. These can be listed as follows. The possibilities for x mod 5 are 0, 1, 2, 3, 4. Substitute each of these into the equation and find the values of y that solve the equation:

$$\begin{array}{lll} x \equiv 0 \Longrightarrow y^2 \equiv 4 & \Longrightarrow y \equiv 2, 3 & \pmod{5} \\ x \equiv 1 \Longrightarrow y^2 \equiv 9 \equiv 4 & \Longrightarrow y \equiv 2, 3 & \pmod{5} \\ x \equiv 2 \Longrightarrow y^2 \equiv 20 \equiv 0 & \Longrightarrow y \equiv 0 & \pmod{5} \\ x \equiv 3 \Longrightarrow y^2 \equiv 43 \equiv 3 & \Longrightarrow \text{no solutions} & \\ x \equiv 4 \Longrightarrow y^2 \equiv 84 \equiv 4 & \Longrightarrow y \equiv 2, 3 & \pmod{5} \\ x = \infty \Longrightarrow y = \infty. & & \end{array}$$

The points on E are $(0,2),(0,3),(1,2),(1,3),(2,0),(4,2),(4,3),(\infty,\infty)$.

The addition of points on an elliptic curve mod p is done via the same formulas as given previously, except that a rational number a/b must be treated as ab^{-1}, where $b^{-1}b = 1 \pmod{p}$. This requires that $\gcd(b,p) = 1$.

More generally, it is possible to develop a theory of elliptic curves mod n for any integer n. In this case, when we encounter a fraction a/b, we need to have $\gcd(b,n) = 1$. The situations where this fails form the key to using elliptic curves for factorization, as we'll see in Section 16.3. There are various technical problems in the general theory that arise when $1 < \gcd(b,n) < n$, but the method to overcome these will not be needed in the following. For details on how to treat this case, see [Washington]. For our purposes, when we encounter an elliptic curve mod a composite n, we can pretend n is prime. If something goes wrong, we usually obtain useful information about n, for example its factorization.

Example. Let's compute $(1,2)+(4,3)$ on the curve just considered. The slope is

$$m \equiv \frac{3-2}{4-1} \equiv 2 \pmod{5}.$$

Therefore,

$$x_3 \equiv m^2 - x_1 - x_2 \equiv 2^2 - 1 - 4 \equiv 4 \pmod{5}$$
$$y_3 \equiv m(x_1 - x_3) - y_1 \equiv 2(1-4) - 2 \equiv 2 \pmod{5}.$$

This means that

$$(1,2) + (4,3) = (4,2).$$

∎

Example. Here is a somewhat larger example. Let $n = 2773$. Let

$$E: y^2 \equiv x^3 + 4x + 4 \pmod{2773}, \text{ and } P = (1,3).$$

Let's compute $2P = P + P$. To get the slope of the tangent line, we differentiate implicitly and evaluate at $(1,3)$:

$$2y\,dy = (3x^2+4)\,dx \Rightarrow \frac{dy}{dx} = \frac{7}{6}.$$

But we are working mod 2773. Using the extended Euclidean algorithm (see Section 3.2), we find that $2311 \cdot 6 \equiv 1 \pmod{2773}$, so we can replace $1/6$ by 2311. Therefore,

$$m \equiv \frac{7}{6} \equiv 7 \times 2311 \equiv 2312 \pmod{2773}.$$

The formulas yield

$$x_3 \equiv 2312^2 - 1 - 1 \equiv 1771 \pmod{2773}$$
$$y_3 \equiv 2312(1 - 1771) - 3 \equiv 705 \pmod{2773}.$$

The final answer is

$$2P = P + P = (1771, 705).$$

Now that we're done with the example, we mention that 2773 is not prime. When we try to calculate $3P$ in Section 16.3, we'll obtain the factorization of 2773. ∎

16.2.1 Number of Points Mod p

Let $E : y^2 \equiv x^3 + bx + c \pmod{p}$ be an elliptic curve, where $p \geq 5$ is prime. We can list the points on E by letting $x = 0, 1, \ldots, p - 1$ and seeing when $x^3 + bx + c$ is a square mod p. Since half of the nonzero numbers are squares mod p, we expect that $x^3 + bx + c$ will be a square approximately half the time. When it is a nonzero square, there are two square roots: y and $-y$. Therefore, approximately half the time we get two values of y and half the time we get no y. Therefore, we expect around p points. Including the point ∞, we expect a total of approximately $p + 1$ points. In the 1930s, H. Hasse made this estimate more precise.

Hasse's Theorem. *Suppose $E \pmod{p}$ has N points. Then*

$$|N - p - 1| < 2\sqrt{p}.$$

The proof of this theorem is well beyond the scope of this book (for a proof, see [Washington]). It can also be shown that whenever N and p satisfy the inequality of the theorem, there is an elliptic curve E mod p with exactly N points.

If p is large, say around 10^{20}, it is infeasible to count the points on an elliptic curve by listing them. More sophisticated algorithms have been developed by Schoof, Atkin, Elkies, and others to deal with this problem.

16.2.2 Discrete Logarithms on Elliptic Curves

Recall the classical discrete logarithm problem: We know that $x \equiv g^k \pmod{p}$ for some k, and we want to find k. There is an elliptic curve version: Suppose we have points A, B on an elliptic curve E and we know that $B = kA (= A + A + \cdots + A)$ for some integer k. We want to find k. This might not look like a logarithm problem, but it is clearly the analog of the classical discrete logarithm problem. Therefore, it is called the **discrete logarithm problem** for elliptic curves.

There is no good general attack on the discrete logarithm problem for elliptic curves. There is an analog of the Pohlig-Hellman attack that works in some situations. Let E be an elliptic curve mod a prime p and let n be the smallest integer such that $nA = \infty$. If n has only small prime factors, then it is possible to calculate the discrete logarithm k mod the prime powers dividing n and then use the Chinese remainder theorem to find k (see Exercise 15). The Pohlig-Hellman attack can be thwarted by choosing E and A so that n has a large prime factor.

There is no replacement for the index calculus attack described in Section 7.2. This is because there is no good analog of "small." You might try to use points with small coordinates in place of the "small primes," but this doesn't work. When you factor a number by dividing off the prime factors one by one, the quotients get smaller and smaller until you finish. On an elliptic curve, you could have a point with fairly small coordinates, subtract off a small point, and end up with a point with large coordinates (see Computer Problem 5). So there is no good way to know when you are making progress toward expressing a point in terms of the factor base of small points.

The Baby Step, Giant Step attack on discrete logarithms works for elliptic curves (Exercise 9), although it requires too much memory to be practical in most situations. For other attacks, see [Blake et al.] and [Washington].

16.2.3 Representing Plaintext

In most cryptographic systems, we must have a method for mapping our original message into a numerical value upon which we can perform mathematical operations. In order to use elliptic curves, we need a method for mapping a message onto a point on an elliptic curve. Elliptic curve cryptosystems then use elliptic curve operations on that point to yield a new point that will serve as the ciphertext.

The problem of encoding plaintext messages as points on an elliptic curve is not as simple as it was in the conventional case. In particular, there is no known polynomial time, deterministic algorithm for writing down points on an arbitrary elliptic curve $E \pmod{p}$. However, there are fast probabilistic methods for finding points, and these can be used for encoding messages. These methods have the property that with small probability they will fail to produce a point. By appropriately choosing parameters, this probability can be made arbitrarily small, say on the order of $1/2^{30}$.

Here is one method, due to Koblitz. The idea is the following. Let $E : y^2 \equiv x^3 + bx + c \pmod{p}$ be the elliptic curve. The message m (already represented as a number) will be embedded in the x-coordinate of a point. However, the probability is only about $1/2$ that $m^3 + bm + c$ is a square mod p. Therefore, we adjoin a few bits at the end of m and adjust them until we get a number x such that $x^3 + bx + c$ is a square mod p.

More precisely, let K be a large integer so that a failure rate of $1/2^K$ is acceptable when trying to encode a message as a point. Assume that m satisfies $(m+1)K < p$. The message m will be represented by a number $x = mK + j$, where $0 \leq j < K$. For $j = 0, 1, \ldots, K-1$, compute $x^3 + bx + c$ and try to calculate the square root of $x^3 + bx + c \pmod{p}$. For example, if $p \equiv 3 \pmod{4}$, the method of Section 3.9 can be used. If there is a square root y, then we take $P_m = (x, y)$; otherwise, we increment j by one and try again with the new x. We repeat this until either we find a square root or $j = K$. If j ever equals K, then we fail to map a message to a point. Since $x^3 + bx + c$ is a square approximately half of the time, we have about a $1/2^K$ chance of failure.

In order to recover the message from the point $P_m = (x, y)$ we simply calculate m by

$$m = [x/K],$$

where $[x/K]$ denotes the greatest integer less than or equal to x/K.

Example. Let $p = 179$ and suppose that our elliptic curve is $y^2 = x^3 + 2x + 7$. If we are satisfied with a failure rate of $1/2^{10}$, then we may take $K = 10$. Since we need $mK + K < 179$, we need $0 \leq m \leq 16$. Suppose our message is $m = 5$. We consider x of the form $mK + j = 50 + j$. The possible choices for x are $50, 51, \ldots, 59$. For $x = 51$ we get $x^3 + 2x + 7 \equiv 121 \pmod{179}$, and $11^2 \equiv 121 \pmod{179}$. Thus, we represent the message $m = 5$ by the point $P_m = (51, 11)$. The message m can be recovered by $m = [51/10] = 5$. ∎

16.3 Factoring with Elliptic Curves

Suppose $n = pq$ is a number we wish to factor. Choose a random elliptic curve mod n and a point on the curve. In practice, one chooses several (around 14 for numbers around 50 digits; more for larger integers) curves with points and runs the algorithm in parallel.

How do we choose the curve? First, choose a point P and a coefficient b. Then choose c so that P lies on the curve $y^2 = x^3 + bx + c$. This is much more efficient than choosing b and c and then trying to find a point.

For example, let $n = 2773$. Take $P = (1, 3)$ and $b = 4$. Since we want $3^2 \equiv 1^3 + 4 \cdot 1 + c$, we take $c = 4$. Therefore, our curve is

$$E: \quad y^2 \equiv x^3 + 4x + 4 \pmod{2773}.$$

We calculated $2P = (1771, 705)$ in a previous example. Note that during the calculation, we needed to find $6^{-1} \pmod{2773}$. This required that $\gcd(6, 2773) = 1$ and used the extended Euclidean algorithm, which was essentially a gcd calculation.

Now let's calculate $3P = 2P + P$. The line through the points $2P = (1771, 705)$ and $P = (1, 3)$ has slope 702/1770. When we try to invert 1770 mod 2773, we find that $\gcd(1770, 2773) = 59$, so we cannot do this. So what do we do? Our original goal was to factor 2773, so we don't need to do anything more. We have found the factor 59, which yields the factorization $2773 = 59 \cdot 47$.

Here's what happened. Using the Chinese remainder theorem, we can regard E as a pair of elliptic curves, one mod 59 and the other mod 47. It turns out that $3P = \infty \pmod{59}$, while $4P = \infty \pmod{47}$. Therefore, when we tried to compute $3P$, we had a slope that was infinite mod 59 but finite mod 47. In other words, we had a denominator that was 0 mod 59 but nonzero mod 47. Taking the gcd allowed us to isolate the factor 59.

The same type of idea is the basis for many factoring algorithms. If $n = pq$, you cannot separate p and q as long as they behave identically. But if you can find something that makes them behave slightly differently, then they can be found. In the example, the multiples of P reached ∞ faster mod 59 than mod 47. Since in general the primes p and q should act fairly independently of each other, one would expect that for most curves $E \pmod{pq}$ and points P, the multiples of P would reach ∞ mod p and mod q at different times. This will cause the gcd to find either p or q.

Usually, it takes several more steps than 3 or 4 to reach ∞ mod p or mod q. In practice, one multiplies P by a large number with many small prime factors, for example, 10000!. This can be done via successive doubling (the additive analog of successive squaring; see Exercise 13). The hope is that this multiple of P is ∞ either mod p or mod q. This is very much the analog of the $p-1$ method of factoring. However, recall that the $p-1$ method (see Section 6.4) usually doesn't work when $p-1$ has a large prime factor. The same type of problem could occur in the elliptic curve method just outlined when the number m such that mP equals ∞ has a large prime factor. If this happens (so the method fails to produce a factor after a while), we simply change to a new curve E. This curve will be independent of the previous curve and the value of m such that $mP = \infty$ should have essentially no relation to the previous m. After several tries (or if several curves are treated in parallel), a good curve is often found, and the number $n = pq$ is factored. In contrast, if the $p-1$ method fails, there is nothing that can be changed other than using a different factorization method.

Example. We want to factor $n = 455839$. Choose

$$E : y^2 \equiv x^3 + 5x - 5, \quad P = (1, 1).$$

Suppose we try to compute $10!P$. There are many ways to do this. One is to compute $2!P, 3!P = 3(2!P), 4!P = 4(3!P), \ldots$. If we do this, everything is fine through $7!P$, but $8!P$ requires inverting 599 (mod n). Since $\gcd(599, n) = 599$, we can factor n as 599×761.

Let's examine this more closely. A computation shows that $E \pmod{599}$ has $640 = 2^7 \times 5$ points and $E \pmod{761}$ has $777 = 3 \times 7 \times 37$ points. Moreover, 640 is the smallest positive m such that $mP = \infty$ on $E \pmod{599}$, and 777 is the smallest positive m such that $mP = \infty$ on $E \pmod{761}$. Since $8!$ is a multiple of 640, it is easy to see that $8!P = \infty$ on $E \pmod{599}$, as we calculated. Since $8!$ is not a multiple of 777, it follows that $8!P \neq \infty$ on $E \pmod{761}$. Recall that we obtain ∞ when we divide by 0, so calculating $8!P$ asked us to divide by 0 (mod 599). This is why we found the factor 599. ∎

In general, consider an elliptic curve $E \pmod{p}$ for some prime p. The smallest positive m such that $mP = \infty$ on this curve divides the number N of points on $E \pmod{p}$ (if you know group theory, you'll recognize this as a corollary of Lagrange's theorem), so $NP = \infty$. Quite often, m will be N or a large divisor of N. In any case, if N is a product of small primes, then $B!$ will be a multiple of N for a reasonably small value of B. Therefore, $B!P = \infty$.

A number that has only small prime factors is called **smooth**. More precisely, if all the prime factors of an integer are less than or equal to B, then it is called **B-smooth**. This concept played a role in the quadratic sieve (Section 6.4), the $p - 1$ factoring method (Section 6.4), and the index calculus attack on discrete logarithms (Section 7.2).

Recall from Hasse's theorem that N is an integer near p. It is possible to show that the density of smooth integers is large enough (we'll leave *small* and *large* undefined here) that if we choose a random elliptic curve $E \pmod{p}$, then there is a reasonable chance that the number N is smooth. This means that the elliptic curve factorization method should find p for this choice of the curve. If we try several curves $E \pmod{n}$, where $n = pq$, then it is likely that at least one of the curves $E \pmod{p}$ or $E \pmod{q}$ will have its number of points being smooth.

In summary, the advantage of the elliptic curve factorization method over the $p-1$ method is the following. The $p-1$ method requires that $p-1$ is smooth. The elliptic curve method requires only that there are enough smooth numbers near p so that at least one of some randomly chosen integers near p is smooth. This means that elliptic curve factorization succeeds much more often than the $p - 1$ method.

The elliptic curve method seems to be best suited for factoring numbers of medium size, say around 40 or 50 digits. These numbers are no longer used for the security of factoring-based systems such as RSA, but it is sometimes useful in other situations to have a fast factorization method for such numbers. Also, the elliptic curve method is effective when a large number has a small prime factor, say of 10 or 20 decimal digits. For large numbers where the prime factors are large, the quadratic sieve and number field sieve are superior (see Section 6.4).

16.3.1 Singular Curves

In practice, the case where the cubic polynomial $x^3 + bx + c$ has multiple roots rarely arises. But what happens if it does? Does the factorization algorithm still work? The discriminant $4b^3 + 27c^2$ is zero if and only if there is a multiple root (this is the cubic analog of the fact that $ax^2 + bx + c$ has a double root if and only if $b^2 - 4ac = 0$). Since we are working mod $n = pq$, the result says that there is a multiple root mod n if and only if the discriminant is 0 mod n. Since n is composite, there is also the intermediate case where the gcd of n and the discriminant is neither 1 nor n. But this gives a nontrivial factor of n, so we can stop immediately in this case.

Example. Let's look at an example:

$$y^2 = x^3 - 3x + 2 = (x-1)^2(x+2).$$

Given a point $P = (x, y)$ on this curve, we associate the number

$$(y + \sqrt{3}(x-1))/(y - \sqrt{3}(x-1)).$$

It can be shown that adding the points on the curve corresponds to multiplying the corresponding numbers. The formulas still work, as long as we don't use the point $(1,0)$. Where does this come from? The two lines tangent to the curve at $(1,0)$ are $y + \sqrt{3}(x-1) = 0$ and $y - \sqrt{3}(x-1) = 0$. This number is simply the ratio of these two expressions.

Since we need to work mod n, we give an example mod 143. We choose 143 since 3 is a square mod 143; in fact, $82^2 \equiv 3 \pmod{143}$. If this were not the case, things would become more technical with this curve. We could easily rectify the situation by choosing a new curve.

Consider the point $P = (-1, 2)$ on $y^2 = x^3 - 3x + 2 \pmod{143}$. Look at its multiples:

$$P = (-1, 2), \quad 2P = (2, 141), \quad 3P = (112, 101), \quad 4P = (10, 20).$$

When trying to compute $5P$, we find the factor 11 of 143.

Recall that we are assigning numbers to each point on the curve, other than (1,1). Since we are working mod 143, we use 82 in place of $\sqrt{3}$. Therefore, the number corresponding to $(-1, 2)$ is $(2 + 82(-1-1))/(2 - 82(-1-1)) = 80 \pmod{143}$. We can compute the numbers for all the points above:

$$P \leftrightarrow 80, \quad 2P \leftrightarrow 108, \quad 3P \leftrightarrow 60, \quad 4P \leftrightarrow 81.$$

Let's compare with the powers of 80 mod 143:

$$80^1 \equiv 80, \quad 80^2 \equiv 108, \quad 80^3 \equiv 60, \quad 80^4 \equiv 81, \quad 80^5 \equiv 45.$$

We get the same numbers. This is simply the fact mentioned previously that the addition of points on the curve corresponds to multiplication of the corresponding numbers. Moreover, note that $45 \equiv 1 \pmod{11}$, but not mod 13. This corresponds to the fact that 5 times the point $(-1, 2)$ is ∞ mod 11 but not mod 13. Note that 1 is the multiplicative identity for multiplication mod 11, while ∞ is the additive identity for addition on the curve.

It is easy to see from the preceding that factorization using the curve $y^2 = x^3 - 3x + 2$ is essentially the same as using the classical $p-1$ factorization method (see Section 6.4). ∎

In the preceding example, the cubic equation had a double root. An even worse possibility is the cubic having a triple root. Consider the curve

$$y^2 = x^3.$$

To a point $(x, y) \neq (0, 0)$ on this curve, associate the number x/y. Let's start with the point $P = (1, 1)$ and compute its multiples:

$$P = (1,1), \quad 2P = (\frac{1}{4}, \frac{1}{8}), \quad 3P = (\frac{1}{9}, \frac{1}{27}), \ldots, \quad mP = (\frac{1}{m^2}, \frac{1}{m^3}).$$

Note that the corresponding numbers x/y are $1, 2, 3, \ldots, m$. Adding the points on the curve corresponds to adding the numbers x/y.

If we are using the curve $y^2 = x^3$ to factor n, we need to change the points mP to integers mod n, which requires finding inverses for m^2 and m^3 mod n. This is done by the extended Euclidean algorithm, which is essentially a gcd computation. We find a factor of n when $\gcd(m, n) \neq 1$. Therefore, this method is essentially the same as computing in succession $\gcd(2, n), \gcd(3, n), \gcd(4, n), \ldots$ until a factor is found. This is a slow version of trial division, the oldest factorization technique known. Of course, in the elliptic curve factorization algorithm, a large multiple $(B!)P$ of P is usually computed. This is equivalent to factoring by computing $\gcd(B!, n)$, a method that is often used to test for prime factors up to B.

In summary, we see that the $p-1$ method and trial division are included in the elliptic curve factorization algorithm if we allow singular curves.

16.4 Elliptic Curves in Characteristic 2

Many applications use elliptic curves mod 2, or elliptic curves defined over the finite fields $GF(2^n)$ (these are described in Section 3.11). This is often because mod 2 adapts well to computers. In 1999, NIST recommended 15 elliptic curves for cryptographic uses (see [FIPS 186-2]). Of these, 10 are over finite fields $GF(2^n)$.

If we're working mod 2, the equations for elliptic curves need to be modified slightly. There are many reasons for this. For example, the derivative of y^2 is $2yy' = 0$, since 2 is the same as 0. This means that the tangent lines we compute are vertical, so $2P = \infty$ for all points P. A more sophisticated explanation is that the curve $y^2 \equiv x^3 + bx + c \pmod 2$ has singularities (points where the partial derivatives with respect to x and y simultaneously vanish).

The equations we need are of the form

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

where $a_1, \dots, a_6$ are constants. The addition law is slightly more complicated. We still have three points adding to infinity if and only if they lie on a line. Also, the lines through ∞ are vertical. But, as we'll see in the following example, finding $-P$ from P is not the same as before.

Example. Let $E : y^2 + y \equiv x^3 + x \pmod 2$. As before, we can list the points on E:

$$(0,0), \quad (0,1), \quad (1,0), \quad (1,1), \quad \infty.$$

Let's compute $(0,0) + (1,1)$. The line through these two points is $y = x$. Substituting into the equation for E yields $x^2 + x \equiv x^3 + x$, which can rewritten as $x^2(x+1) \equiv 0$. The roots are $x = 0, 0, 1 \pmod 2$. Therefore, the third point of intersection also has $x = 0$. Since it lies on the line $y = x$, it must be $(0,0)$. (This might be puzzling. What is happening is that the line is tangent to E at $(0,0)$ and also intersects E in the point $(1,1)$.) As before, we now have

$$(0,0) + (0,0) + (1,1) = \infty.$$

To get $(0,0) + (1,1)$ we need to compute $\infty - (0,0)$. This means we need to find P such that $P + (0,0) = \infty$. A line through ∞ is still a vertical line. In this case, we need one through $(0,0)$, so we take $x = 0$. This intersects E in the point $P = (0,1)$. We conclude that $(0,0) + (0,1) = \infty$. Putting everything together, we see that

$$(0,0) + (1,1) = (0,1). \qquad \blacksquare$$

In most applications, elliptic curves mod 2 are not large enough. Therefore, elliptic curves over finite fields are used. For an introduction to finite fields, see Section 3.11. However, in the present section, we only need the field $GF(4)$, which we now describe.

Let

$$GF(4) = \{0, 1, \omega, \omega^2\},$$

with the following laws:

1. $0 + x = x$ for all x.
2. $x + x = 0$ for all x.
3. $1 \cdot x = x$ for all x.
4. $1 + \omega = \omega^2$.
5. Addition and multiplication are commutative and associative, and the distributive law holds: $x(y+z) = xy + xz$ for all x, y, z.

Since

$$\omega^3 = \omega \cdot \omega^2 = \omega \cdot (1 + \omega) = \omega + \omega^2 = \omega + (1 + \omega) = 1,$$

we see that ω^2 is the multiplicative inverse of ω. Therefore, every nonzero element of $GF(4)$ has a multiplicative inverse.

Elliptic curves with coefficients in finite fields are treated just like elliptic curves with integer coefficients.

Example. Consider

$$E : y^2 + xy = x^3 + \omega,$$

where $\omega \in GF(4)$ is as before. Let's list the points of E with coordinates in $GF(4)$:

$$\begin{aligned}
&x = 0 \Longrightarrow y^2 = \omega \Longrightarrow y = \omega^2\\
&x = 1 \Longrightarrow y^2 + y = 1 + \omega = \omega^2 \Longrightarrow \text{ no solutions}\\
&x = \omega \Longrightarrow y^2 + \omega y = \omega^2 \Longrightarrow y = 1,\, \omega^2\\
&x = \omega^2 \Longrightarrow y^2 + \omega^2 y = 1 + \omega = \omega^2 \Longrightarrow \text{ no solutions}\\
&x = \infty \Longrightarrow y = \infty.
\end{aligned}$$

The points on E are therefore

$$(0, \omega^2), \quad (\omega, 1), \quad (\omega, \omega^2), \quad \infty.$$

Let's compute $(0, \omega^2) + (\omega, \omega^2)$. The line through these two points is $y = \omega^2$. Substitute this into the equation for E:

$$\omega^4 + \omega^2 x = x^3 + \omega,$$

which becomes $x^3 + \omega^2 x = 0$. This has the roots $x = 0, \omega, \omega$. The third point of intersection of the line and E is therefore (ω, ω^2), so

$$(0, \omega^2) + (\omega, \omega^2) + (\omega, \omega^2) = \infty.$$

We need $-(\omega, \omega^2)$, namely the point P with $P + (\omega, \omega^2) = \infty$. The vertical line $x = \omega$ intersects E in $P = (\omega, 1)$, so

$$(0, \omega^2) + (\omega, \omega^2) = (\omega, 1). \qquad \blacksquare$$

For cryptographic purposes, elliptic curves are used over fields $GF(2^n)$ with n large, say at least 150.

16.5 Elliptic Curve Cryptosystems

Elliptic curve versions exist for many cryptosystems, in particular those involving discrete logarithms. An advantage of elliptic curves over working with integers mod p is the following. In the integers, it is possible to use the factorization of integers into primes (especially small primes) to attack the discrete logarithm problem. This is known as the index calculus and is described in Section 7.2. There seems to be no good analog of this method for elliptic curves. Therefore, it is possible to use smaller primes, or smaller finite fields, with elliptic curves and achieve a level of security comparable to that for much larger integers mod p. This allows great savings in hardware implementations, for example.

In the following, we describe three elliptic curve versions of classical algorithms. As we'll see, there is a general procedure for changing a classical system based on discrete logarithms into one using elliptic curves:

1. Change modular multiplication to addition of points on an elliptic curve.

2. Change modular exponentiation to multiplying a point on an elliptic curve by an integer.

Of course, the second situation above is really a special case of the first, since exponentiation consists of multiplying a number by itself several times, and multiplying a point by an integer is adding the point to itself several times.

16.5.1 An Elliptic Curve ElGamal Cryptosystem

We recall the non-elliptic curve version. Alice wants to send a message x to Bob, so Bob chooses a large prime p and an integer α mod p. He also chooses a secret integer a and computes $\beta \equiv \alpha^a \pmod{p}$. Bob makes p, α, β public and keeps a secret. Alice chooses a random k and computes y_1 and y_2, where

$$y_1 \equiv \alpha^k \text{ and } y_2 \equiv x\beta^k \pmod{p}.$$

She sends (y_1, y_2) to Bob, who then decrypts by calculating

$$x \equiv y_2 y_1^{-a} \pmod{p}.$$

Now we describe the elliptic curve version. Bob chooses an elliptic curve $E \pmod{p}$, where p is a large prime. He chooses a point α on E and a secret integer a. He computes

$$\beta = a\alpha \quad (= \alpha + \alpha + \cdots + \alpha).$$

The points α and β are made public, while a is kept secret. Alice expresses her message as a point x on E (see Section 16.2). She chooses a random integer k, computes

$$y_1 = k\alpha \text{ and } y_2 = x + k\beta,$$

and sends the pair y_1, y_2 to Bob. Bob decrypts by calculating

$$x = y_2 - ay_1.$$

A more workable version of this system is due to Menezes and Vanstone. It is described in [Stinson1, p. 189].

Example. We must first generate a curve. Let's use the prime $p = 8831$, the point $G = (x, y) = (4, 11)$, and $a = 3$. To make G lie on the curve $y^2 \equiv x^3 + bx + c \pmod{p}$, we take $b = 45$. Alice has a message, represented as a point $P_m = (5, 1743)$, that she wishes to send to Bob. Here is how she does it.

Bob has chosen a random number $a_B = 3$ and has published the point $a_B G = (413, 1808)$.

Alice downloads this and chooses a random number $k = 8$. She sends Bob $kG = (5415, 6321)$ and $P_m + k(a_B G) = (6626, 3576)$. He first calculates $a_B(kG) = 3(5415, 6321) = (673, 146)$. He now subtracts this from $(6626, 3576)$:

$$(6626, 3576) - (673, 146) = (6626, 3576) + (673, -146) = (5, 1743).$$

Note that we subtracted points by using the rule $P - Q = P + (-Q)$ from Section 16.1. ∎

16.5.2 Elliptic Curve Diffie-Hellman Key Exchange

Alice and Bob want to exchange a key. In order to do so, they agree on a public basepoint G on an elliptic curve $E: y^2 \equiv x^3 + bx + c \pmod{p}$. Let's choose $p = 7211$ and $a = 1$ and $G = (3, 5)$. This gives us $b = 7206$. Alice chooses N_A randomly and Bob chooses N_B randomly. Let's suppose $N_A = 12$ and $N_B = 23$. They keep these private to themselves but publish N_AG and N_BG. In our case, we have

$$N_AG = (1794, 6375) \text{ and } N_BG = (3861, 1242).$$

Alice now takes N_BG and multiplies by N_A to get the key:

$$N_A(N_BG) = 12(3861, 1242) = (1472, 2098).$$

Similarly, Bob takes N_AG and multiplies by N_B to get the key:

$$N_B(N_AG) = 23(1794, 6375) = (1472, 2098).$$

Notice that they have the same key.

16.5.3 ElGamal Digital Signatures

There is an elliptic curve analog of the procedure described in Section 9.2. A few modifications are needed to account for the fact that we are working with both integers and points on an elliptic curve.

Alice wants to sign a message m (which might actually be the hash of a long message). We assume m is an integer. She fixes an elliptic curve $E \pmod{p}$, where p is a large prime, and a point A on E. We assume that the number of points n on E has been calculated and assume $0 \le m < n$ (if not, choose a larger p). Alice also chooses a private integer a and computes $B = aA$. The prime p, the curve E, the integer n, and the points A and B are made public. To sign the message, Alice does the following:

1. Chooses a random integer k with $1 \le k < n$ and $\gcd(k, n) = 1$, and computes $R = kA = (x, y)$
2. Computes $s \equiv k^{-1}(m - ax) \pmod{n}$
3. Sends the signed message (m, R, s) to Bob

Note that R is a point on E, and m and s are integers.

Bob verifies the signature as follows:

1. Downloads Alice's public information p, E, n, A, B
2. Computes $V_1 = xB + sR$ and $V_2 = mA$

3. Declares the signature valid if $V_1 = V_2$

The verification procedure works because

$$V_1 = xB + sR = xaA + k^{-1}(m - ax)(kA) = xaA + (m - ax)A = mA = V_2.$$

There is a subtle point that should be mentioned. We have used k^{-1} in this verification equation as the integer mod n satisfying $k^{-1}k \equiv 1 \pmod{n}$. Therefore, $k^{-1}k$ is not 1, but rather an integer congruent to 1 mod n. So $k^{-1}k = 1 + tn$ for some integer t. It can be shown that $nA = \infty$. Therefore,

$$k^{-1}kA = (1 + tn)A = A + t(nA) = A + t\infty = A.$$

This shows that k^{-1} and k cancel each other in the verification equation, as we implicitly assumed above.

The classical ElGamal scheme and the present elliptic curve version are analogs of each other. The integers mod p are replaced with the elliptic curve E, and the number $p-1$ becomes n. Note that the calculations in the classical scheme work with integers that are nonzero mod p, and there are $p-1$ such congruence classes. The elliptic curve version works with points on the elliptic curve that are multiples of A, and the number of such points is a divisor of n.

The use of the x-coordinate of R in the elliptic version is somewhat arbitrary. Any method of assigning integers to points on the curve would work. Using the x-coordinate is an easy choice. Similarly, in the classical ElGamal scheme, the use of the integer r in the mod $p-1$ equation for s might seem a little unnatural, since r was originally defined mod p. However, any method of assigning integers to the integers mod p would work (see Exercise 10 in Chapter 9). The use of r itself is an easy choice.

There is an elliptic curve version of the Digital Signature Algorithm that is similar to the preceding (Exercise 14).

16.6 Identity-Based Encryption

In most public key systems, when Alice wants to send a message to Bob, she looks up his public key in a directory and then encrypts her message. However, she needs some type of authentication – perhaps the directory has been modified by Eve, and the public key listed for Bob was actually created by Eve. Alice wants to avoid this situation. It was suggested by Shamir in 1984 that it would be nice to have an identity-based system, where Bob's public identification information (for example, his email address) serves as his public key. Such a system was finally designed in 2001 by Boneh and Franklin.

Of course, some type of authentication of each user is still needed. In the present system, this occurs in the initial setup of the system during the communications between the Trusted Authority and the User. In the following, we give the basic idea of the system. For more details and improvements, see [Boneh-Franklin].

Before describing the system, we need some preliminary information. Let p be a prime of the form $6q-1$, where q is also prime. Let E be the elliptic curve $y^2 \equiv x^3 + 1 \bmod p$. We need the following facts about E.

1. There are exactly $p+1 = 6q$ points on E.

2. There is a point $P_0 \neq \infty$ such that $qP_0 = \infty$. In fact, if we take a random point P, then, with very high probability, $6P \neq \infty$ and $6P$ is a multiple of P_0.

3. There is a function $\tilde{e}$ that maps pairs of points (aP_0, bP_0) to qth roots of unity for all integers a, b. It satisfies the **bilinearity property**

$$\tilde{e}(aP_0, bP_0) = \tilde{e}(P_0, P_0)^{ab}$$

   for all a, b.

4. If we are given two points P and Q that are multiples of P_0, then $\tilde{e}(P, Q)$ can be computed quickly from the coordinates of P and Q.

5. $\tilde{e}(P_0, P_0) \neq 1$, so it is a non-trivial qth root of unity.

Remarks. Properties (1) and (2) are fairly easy to verify (see Exercises 16 and 17). The existence of $\tilde{e}$ satisfying (3), (4), (5) is deep. In fact, $\tilde{e}$ is a modification of what is known as the Weil pairing in the theory of elliptic curves. The usual Weil pairing e satisfies $e(P_0, P_0) = 1$, but the present version is modified using special properties of E to obtain (5).

The fact that $\tilde{e}(P, Q)$ can be computed quickly needs some more explanation. The two points P, Q satisfy $P = aP_0$ and $Q = bP_0$ for some a, b. However, to find a and b requires solving a discrete log problem, which could take a long time. Therefore, the obvious solution of choosing a random qth root of unity for $\tilde{e}(P_0, P_0)$ and then using the bilinearity property to define $\tilde{e}$ does not work, since it cannot be computed quickly. Instead, $\tilde{e}(P, Q)$ is computed directly in terms of the coordinates of the points P, Q.

Although we will not need to know this, the qth roots of unity lie in the finite field with p^2 elements (see Section 3.11).

For more about the definition of $\tilde{e}$, see [Boneh-Franklin] or [Washington].

The curve E is an example of a **supersingular** elliptic curve, namely one where the number of points is congruent to 1 mod p. (See Exercise 16.) For a while, these curves were regarded as desirable for cryptographic purposes, because computations can be done quickly on them. But then it was shown that the discrete logarithm problem for them was only slightly more difficult than the classical discrete logarithm mod p (see Exercise 20), so they fell out of favor (after all, they are slower computationally than simple multiplication mod p, and they provide no security advantage). Because of the existence of the pairing $\tilde{e}$, they have become popular again.

To set up the cryptosystem, we'll need two public hash functions:

1. H_1 maps arbitrary length binary strings to multiples of P_0. A little care is needed in defining H_1, since no one should be able, given a binary string b, to find k with $H_1(b) = kP_0$. See Exercise 18.

2. H_2 maps qth roots of unity to binary strings of length n, where n is the length of the messages that will be sent. Since H_2 must be specified before the system is set up, this limits the lengths of the messages that can be sent. However, the message could be, for example, a DES key that is used to encrypt the remainder of a much longer message, so this length requirement is not a severe restriction.

To set up the system we need a Trusted Authority. Let's call him Arthur. Arthur does the following.

1. He chooses, once and for all, a secret integer s. He computes $P_1 = sP_0$, which is made public.

2. For each User, Arthur finds the user's identification ID (written as a binary string) and computes

$$D_{\text{User}} = s\,H_1(ID).$$

   Recall that $H_1(ID)$ is a point on E, so D_{User} is s times this point.

3. Arthur sends D_{User} to the user, who keeps it secret. Arthur does not need to store D_{User}, so he discards it.

The system is now ready to operate, but first let's review what is known:

Public: E, p, P_0, P_1, H_1, H_2
Secret: s (known only to Arthur), D_{User} (one for each User; it is known only by that User)

Alice wants to send an email message m (of binary length n) to Bob, who is one of the Users. She knows Bob's address, which is `bob@computer.com`. This is his ID. Alice does the following.

1. She computes $g = \tilde{e}(H_1(\texttt{bob@computer.com}), P_1)$. This is a qth root of unity.

2. She chobses a random $r \not\equiv 0 \pmod{q}$ and computes

$$t = m \oplus H_2(g^r).$$

3. She sends Bob the ciphertext

$$c = (rP_0,\, t).$$

Note that rP_0 is a point on E, and t is a binary string of length n.

If Bob receives a pair (U, v), where U is a point on E and v is a binary string of length n, then he does the following.

1. He computes $h = \tilde{e}(D_{\text{Bob}},\, U)$, which is a qth root of unity.

2. He recovers the message as

$$m = v \oplus H_2(h).$$

Why does this yield the message? If the encryption is performed correctly, Bob receives $U = rP_0$ and $v = t = m \oplus H_2(g^r)$. Since $D_{\text{Bob}} = s\,H_1(\texttt{bob@computer.com})$,

$$h = \tilde{e}(D_{\text{Bob}},\, rP_0) = \tilde{e}(H_1, P_0)^{sr} = \tilde{e}(H_1,\, sP_0)^r = g^r. \tag{16.1}$$

Therefore,

$$t \oplus H_2(h) = t \oplus H_2(g^r) = m \oplus H_2(g^r) \oplus H_2(g^r) = m,$$

as desired. Note that the main step is Equation 16.1, which removes the secret s from the D_{Bob} in the first argument of $\tilde{e}$ and puts it on the P_0 in the second argument. This follows from the bilinearity property of the function $\tilde{e}$.

It is very important that s be kept secret. If Eve obtains s, then she can compute the points D_{User} for each user and read every email. Since $P_1 = sP_0$, the security of s is compromised if Eve can compute discrete logs on the elliptic curve. Moreover, the ciphertext contains rP_0. If Eve can compute a discrete log and find r, then she can compute g^r and use this to find $H_2(g^r)$ and also m. Therefore, for the security of the system, it is vital that p be chosen large enough that discrete logs are computationally infeasible.

16.7 Exercises

1. **(a)** Let $x^3 + ax^2 + bx + c$ be a cubic polynomial with roots r_1, r_2, r_3. Show that $r_1 + r_2 + r_3 = -a$.

(b) Write $x = x_1 - a/3$. Show that

$$x^3 + ax^2 + bx + c = x_1^3 + b'x_1 + c',$$

with $b' = b - (1/3)a^2$ and $c' = c - (1/3)ab + (2/27)a^3$. (*Remark:* This shows that a simple change of variables allows us to consider the case where the coefficient of x^2 is 0.)

2. **(a)** List the points on the elliptic curve E: $y^2 \equiv x^3 - 2 \pmod 7$.

(b) Find the sum $(3, 2) + (5, 5)$ on E.

(c) Find the sum $(3, 2) + (3, 2)$ on E.

3. Show that if $P = (x, 0)$ is a point on an elliptic curve, then $2P = \infty$.

4. The points $(3, \pm 5)$ lie on the elliptic curve $y^2 = x^3 - 2$ defined over the rational numbers. Find another point with rational coordinates that lies on this curve.

5. **(a)** Show that $Q = (2, 3)$ on $y^2 = x^3 + 1$ satisfies $6Q = \infty$. (*Hint:* Compute $3Q$, then use Exercise 3.)

(b) Your computations in (a) probably have shown that $2Q \neq \infty$ and $3Q \neq \infty$. Use this to show that the points $\infty, Q, 2Q, 3Q, 4Q, 5Q$ are distinct.

6. **(a)** Factor $n = 35$ by the elliptic curve method by using the elliptic curve $y^2 \equiv x^3 + 26$ and calculating 3 times the point $P = (10, 9)$.

(b) Factor $n = 35$ by the elliptic curve method by using the elliptic curve $y^2 \equiv x^3 + 5x + 8$ and the point $P = (1, 28)$.

7. Suppose you want to factor a composite integer n by using the elliptic curve method. You start with the curve $y^2 = x^3 - 4x \pmod n$ and the point $(2, 0)$. Why will this not yield the factorization of n?

8. Devise an analog of the procedure in Exercise 8(a) in Chapter 7 that uses elliptic curves.

9. Show how to use a Baby Step, Giant Step attack (see Section 7.2) to attack the discrete log problem on elliptic curves.

10. Show that if P, Q, R are points on an elliptic curve, then

$$P + Q + R = \infty \iff P, Q, R \text{ are collinear.}$$

roots r_1, r_2, r_3.

7)a^3. (*Remark:*
s us to consider

$\cdot$ 2 (mod 7).

then $2P = \infty$.

lefined over the
oordinates that

$Q = \infty$. (*Hint:*

at $2Q \neq \infty$ and
$2Q, 3Q, 4Q, 5Q$

ing the elliptic
int $P = (10, 9)$.
ing the elliptic

sing the elliptic
x (mod n) and
on of n?

Chapter 7 that

Section 7.2) to

en

ar.

11. Let P be a point on the elliptic curve E mod n.

(a) Show that there are only finitely many points on E, so P has only finitely many distinct multiples.

(b) Show that there are integers i, j with $i > j$ such that $iP = jP$. Conclude that $(i - j)P = \infty$.

(c) The smallest positive integer k such that $kP = \infty$ is called the **order** of P. Let m be an integer such that $mP = \infty$. Show that k divides m. (*Hint:* Imitate the proof of Exercise 20(c, d) in Chapter 3.)

(d) (for those who know some group theory) Use Lagrange's theorem from group theory to show that the number of points on E is a multiple of the order of P. (Combined with Hasse's theorem, this gives a way of finding the number of points on E. See Computer Problems 1 and 4.)

12. Let P be a point on the elliptic curve E mod n. Suppose you know a positive integer k such that $kP = \infty$. You want to prove (or disprove) that k is the order of P.

(a) Show that if $(k/p)P = \infty$ for some prime factor p of k, then k is not the order of P.

(b) Suppose $m|k$ and $1 \leq m < k$. Show that $m|(k/p)$ for some prime divisor p of k.

(c) Suppose that $(k/p)P \neq \infty$ for each prime factor of k. Use Exercise 11(c) to show that the order of P is k. (Compare with Exercise 21 in Chapter 3. For an example, see Computer Problem 4.)

13. (a) Let $x = b_1 b_2 \ldots b_w$ be an integer written in binary. Let P be a point on the elliptic curve E. Perform the following procedure:

```
1. Start with $k = 1$ and $S_1 = \infty$.
2. If $b_k = 1$, let $R_k = S_k + P$. If $b_k = 0$, let $R_k = S_k$.
3. Let $S_{k+1} = 2R_k$.
4. If $k = w$, stop. If $k < w$, add 1 to $k$ and go to step 2.
```

Show that $R_w = xP$. (Compare with Exercise 23(a) in Chapter 3.)

(b) Let x be a positive integer and let P be a point on an elliptic curve. Show that the following procedure computes xP.

```
1. Start with $a = x, B = \infty, C = P$.
2. If $a$ is even, let $a = a/2$, and let $B = B, C = 2C$.
```

3. If a is odd, let $a = a - 1$, and let $B = B + C, C = C$.
4. If $a \neq 0$, go to step 2.
5. Output B.

(Compare with Exercise 23(b) in Chapter 3.)

14. Here is an elliptic curve version of the Digital Signature Algorithm. Alice wants to sign a message m, which is an integer. She chooses a prime p and an elliptic curve $E \pmod{p}$. The number of points n on E is computed and a large prime factor q of n is found. A point $A(\neq \infty)$ is chosen such that $qA = \infty$. (In fact, n is not needed. Choose a point A' on E and find an integer n' with $n'A' = \infty$. There are ways of doing this, though it is not easy. Let q be a large prime factor of n', if it exists, and let $A = (n'/q)A'$. Then $qA = \infty$.) It is assumed that the message satisfies $0 \leq m < q$. Alice chooses her secret integer a and computes $B = aA$. The public information is p, E, q, A, B. Alice does the following:

1. Chooses a random integer k with $1 \leq k < q$ and computes $R = kA = (x, y)$
2. Computes $s \equiv k^{-1}(m + ax) \pmod{q}$
3. Sends the signed message (m, R, s) to Bob

Bob verifies the signature as follows:

1. Computes $u_1 \equiv s^{-1}m \pmod{q}$ and $u_2 \equiv s^{-1}x \pmod{q}$
2. Computes $V = u_1A + u_2B$
3. Declares the signature valid if $V = R$

(a) Show that the verification equation holds for a correctly signed message. Where is the fact that $qA = \infty$ used (see the "subtle point" mentioned in the ElGamal scheme in Section 16.5)?

(b) Why does $k^{-1} \pmod{q}$ exist?

(c) If q is large, why is there very little chance that s^{-1} does not exist mod q? How do we recognize the case when it doesn't exist? (Of course, in this case, Alice should start over by choosing a new k.)

(d) How many computations "(large integer)×(point on E)" are made in the verification process here? How many are made in the verification process for the elliptic ElGamal scheme described in the text? (Compare with the end of Section 9.5.)

15. Let A and B be points on an elliptic curve and suppose $B = kA$ for some integer k. Suppose also that $2^nA = \infty$ for some integer n, but $T = 2^{n-1}A \neq \infty$.

(a) Show that if $k \equiv k' \pmod{2^n}$, then $B = k'A$. Therefore, we may assume that $0 \le k < 2^n$.

(b) Let j be an integer. Show that $jT = \infty$ when j is even and $jT \ne \infty$ when j is odd.

(c) Write $k = x_0 + 2x_1 + 4x_2 + \cdots + 2^{n-1}x_{n-1}$, where each x_i is 0 or 1 (binary expansion of k). Show that $x_0 = 0$ if and only if $2^{n-1}B = \infty$.

(d) Suppose that for some $m < n$ we know $x_0, \dots, x_{m-1}$. Let $Q_m = B - (x_0 + \cdots + 2^{m-1}x_{m-1})A$. Show that $2^{n-m-1}Q_m = \infty$ if and only if $x_m = 0$. This allows us to find x_m. Continuing in this way, we obtain $x_0, \dots, x_{n-1}$, and therefore we can compute k. This technique can be extended to the case where $sA = \infty$, where s is an integer with only small prime factors. This is the analog of the Pohlig-Hellman algorithm (see Section 7.2).

16. Let $p \equiv -1 \pmod 3$ be prime.

(a) Show that there exists d with $3d \equiv 1 \pmod{p-1}$.

(b) Show that if $a^3 \equiv b \pmod p$ if and only if $a \equiv b^d \pmod p$. This shows that every integer mod p has a unique cube root.

(c) Show that $y^2 \equiv x^3 + 1 \pmod p$ has exactly $p+1$ points (including the point ∞). (*Hint:* Apply part (b) to $y^2 - 1$.) (*Remark:* A curve mod p whose number of points is congruent to 1 mod p is called *supersingular.*)

17. (for those who know some group theory)

(a) In the situation of Exercise 16, suppose that $p = 6q - 1$ with q also prime. Show that there exists a point $P_0 \ne \infty$ such that $qP_0 = \infty$.

(b) Let $Q = (2, 3)$, as in Exercise 5. Show that if $P \notin \{\infty, Q, 2Q, 3Q, 4Q, 5Q\}$, then $6P \ne \infty$ and $6P$ is a multiple of P_0. (For simplicity, assume that $q > 3$.)

18. In the identity-based system of Section 16.6, suppose Eve can compute k such that $H_1(\texttt{bob@computer.edu}) = k\,P_0$. Show that Eve can compute g^r and therefore read Bob's messages.

19. Let H_0 be a hash function that takes a binary string of arbitrary length as input and then outputs an integer mod p. Let $p = 6q - 1$ be prime with q also prime. Show how to use H_0 to construct a hash function H_1 that takes a binary string of arbitrary length as input and outputs a point on the elliptic curve $y^2 \equiv x^3 + 1 \pmod p$ that is a multiple of the point P_0 os Exercise 17. (*Hint:* Use the technique of Exercise 16 to find y, then x. Then use Exercise 17(b).)

20. **(a)** Using the function $\tilde{e}$ of Section 16.6, show that an analogue of the Decision Diffie-Hellman problem can be solved for the curve $y^2 \equiv x^3 + 1 \pmod{p}$, where $p = 6q - 1$ is prime with q also prime. Namely, if we are given aP_0, bP_0, cP_0, show how we can decide whether $abP_0 = cP_0$.

(b) Show that the discrete logarithm problem for multiples of P_0 on E (namely, if we know kP_0, find k) can be reduced to solving a classical discrete logarithm for the qth roots of unity, hence in the field with p^2 elements. (*Remark:* This is the reason supersingular curves became unpopular.)

21. Suppose you try to set up an identity-based cryptosystem as follows. Arthur chooses large primes p and q and forms $n = pq$, which is made public. For each User, he converts the User's identification ID to a number e_{User} by some public method and then computes d with $de_{\text{User}} \equiv 1 \pmod{\phi(n)}$. Arthur gives d to the User. The integer n is the same for all users. When Alice wants to send an email to Bob, she uses the public method to convert his email address to e_{Bob} and then uses this to encrypt messages with RSA. Bob knows d, so he can decrypt. Explain why this system is not secure.

16.8 Computer Problems

1. Let E be the elliptic curve $y^2 \equiv x^3 + 2x + 3 \pmod{19}$.

(a) Find the sum $(1, 5) + (9, 3)$.

(b) Find the sum $(9, 3) + (9, -3)$.

(c) Using the result of part (b), find the difference $(1, 5) - (9, 3)$.

(d) Find an integer k such that $k(1, 5) = (9, 3)$.

(e) Show that $(1, 5)$ has exactly 20 distinct multiples, including ∞.

(f) Using (e) and Exercise 11(d), show that the number of points on E is a multiple of 20. Use Hasse's theorem to show that E has exactly 20 points.

2. You want to represent the message 12345 as a point (x, y) on the curve $y^2 \equiv x^3 + 7x + 11 \pmod{593899}$. Write $x = 12345\_$ and find a value of the missing last digit of x such that there is a point on the curve with this x-coordinate.

ique of Exercise 16

hat an analogue of
solved for the curve
prime with q also
, show how we can

multiples of P_0 on
educed to solving a
f unity, hence in the
eason supersingular

osystem as follows.
$= pq$, which is made
dentification ID to
n computes d with
ser. The integer n
d an email to Bob,
ddress to e_{Bob} and
knows d, so he can

19).

e $(1, 5) - (9, 3)$.

ples, including ∞.

umber of points on
o show that E has

t (x, y) on the curve
5_ and find a value
point on the curve

3. **(a)** Factor 3900353 using elliptic curves.

(b) Try to factor 3900353 using the $p-1$ method of Section 6.4. Using the knowledge of the prime factors obtained from part (a), explain why the $p-1$ method does not work well for this problem.

4. Let $P = (2, 3)$ be a point on the elliptic curve $y^2 \equiv x^3 - 10x + 21 \pmod{557}$.

(a) Show that $189P = \infty$, but $63P \neq \infty$ and $27P \neq \infty$.

(b) Use Exercise 12 to show that P has order 189.

(c) Use Exercise 11(d) and Hasse's theorem to show that the elliptic curve has 567 points.

5. Compute the difference $(5, 9) - (1, 1)$ on the elliptic curve $y^2 \equiv x^3 - 11x + 11 \pmod{593899}$. Note that the answer involves large integers, even though the original points have small coordinates.