Elliptic Curves and Elliptic Curve Cryptography

Amiee O'Maley

Amiee O'Maley graduated Summa Cum Laude from Ball State in May 2004 with a major in Mathematics. She is currently an Actuarial Analyst for Anthem Insurance Company in Indianapolis, IN. The content of this paper was part of her honors thesis with Dr. Michael Karls.

Introduction

Quadratic equations are studied extensively within mathematics throughout a student's high school and college careers. The standard form for these equations (in the variable x) is given by

$$ax^2 + bx + c = 0,$$

where a, b, and c are real, and $a \neq 0$. Their solutions are given by the *quadratic formula*

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

which is introduced in algebra. *Ellipses*, *parabolas*, and *hyperbolas* are studied in geometry, and surfaces such as *hyperboloids* and *paraboloids*, given by

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 - \left(\frac{z}{c}\right)^2 = 1 \text{ and } \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = z,$$

respectively, are studied in multivariable calculus. All of these are quadratic equations. What is beyond quadratics? For example, there are *elliptic curves*, which are curves of the form

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

The study of elliptic curves can be traced back to the ancient Greeks and Alexandrians, from which a deep theory has emerged. The name "elliptic curve" comes from the work of G.C. Fagnano (1682-1766) who showed that computing the arc length of an ellipse leads to an integral of the form

$$\int \frac{1}{\sqrt{(1-u^2)(1-k^2u^2)}}\, du.$$

By making the changes of variables,

$$v^2 = (1-u^2)(1-k^2u^2) = (u-\alpha)(u-\beta)(u-\gamma)(u-\delta)$$

followed by

$$x = \frac{1}{u-\alpha} \quad \text{and} \quad y = \frac{v}{(u-\alpha)^2},$$

one is led to the integral

$$\int \frac{1}{\sqrt{x^3+ax^2+bx+c}}\, dx,$$

which is why a curve of the form $y^2 = x^3 + ax^2 + bx + c$ is called an elliptic curve [1].

Elliptic curves have been used to study or solve many famous problems, such as the Congruent Number Problem and Fermat's Last Theorem. A rational number n is said to be *congruent* if there exists a right triangle with rational sides whose area is n. For example, 6 is a congruent number, since the right triangle with sides 3, 4, and 5 has area 6. Mathematicians such as Pierre de Fermat (1601-1665) and Leonhard Euler (1707-1783) studied the problem of which numbers are congruent. This problem can be turned into an investigation of points on certain elliptic curves. Fermat's Last Theorem, which states that there are no non-zero integer solutions x, y, z to the equation $x^n + y^n = z^n$ for integers $n > 2$, was proved in 1993 by Andrew Wiles. A key to Wiles' proof was to show that if Fermat's Last Theorem were false, a certain type of elliptic curve would exist that leads to a contradiction [1].

Elliptic curves can also be used as schemes to transmit information securely. In 1985, Neal Koblitz, from the University of Washington, and Victor Miller, who worked at IBM, first proposed the application of elliptic curve systems to *cryptography*, which is the science of concealing the meaning of a message [3, 8]. To *encrypt* a message, one conceals the meaning of the message using a code or cipher, and to *decrypt* the message, one turns the encrypted message back into the original message.

Many cryptosystems necessitate the use of an algebraic structure known as a group, and elliptic curves can be used to form such a structure, referred to as an elliptic curve group. To understand elliptic curve groups, a good starting point is to look at elliptic curves over the real numbers. The next step is to consider elliptic curves over finite fields such as the integers modulo p, where p is a prime number.

Properties of elliptic curves and elliptic curve groups can then be applied to cryptographic schemes, known as *elliptic curve cryptography* (ECC) schemes. We will look at one such ECC scheme, known as the Elliptic Curve ElGamal Method. Elliptic curve cryptography, used in many applications today, maintains the three objectives of information security: *confidentiality*, the concealment of data from unauthorized parties; *integrity*, the assurance that data is genuine; and *availability*, the fact that the system still functions efficiently after security provisions are in place [2]. Elliptic curve cryptography has expanded the use of public-key cryptosystems, providing systems of encryption that are easier to implement and harder to crack [6].

Elliptic curves over $\mathbb{R}$

An *elliptic curve over the real numbers* is the set of points (x, y) that satisfy an equation of the form

$$y^2 = x^3 + ax + b, \tag{1}$$

where x, y, a, and b are real numbers. There are other elliptic curves of the more general "Weierstrass" form:

$$y^2 + a_1xy + a_2y = a_3x^3 + a_4x^2 + a_5x + a_6,$$

but through a change of variable, one can put any elliptic curve over the reals into the form of Equation (1) [5, 6]. Figure 1 shows some examples of elliptic curves.

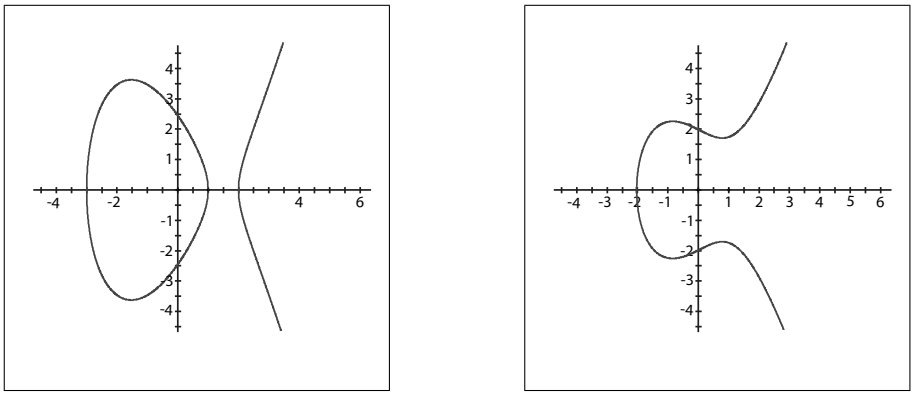


Figure 1: Elliptic curves $y^2 = x^3 - 7x + 6$ (left) and $y^2 = x^3 - 2x + 4$ (right)

Elliptic curves in the form of Equation (1) can be divided into two groups, *non-singular* and *singular* elliptic curves. A continuously differentiable curve written in the form $F(x, y) = 0$ is said to be singular if there is a point on the curve at which both partial derivatives of F are zero. Otherwise the curve is called non-singular. It follows from the Implicit Function Theorem that at every point of a non-singular curve, there is a tangent line [7]. We leave it as an exercise to the reader to show that the elliptic curve of Equation (1) is

non-singular if and only if $4a^3+27b^2 \neq 0$. We hasten to add that $4a^3+27b^2 \neq 0$ is a necessary and sufficient condition for the cubic polynomial $x^3 + ax + b$ to have three distinct roots [5]. We will only use non-singular elliptic curves, as we will need to have curves at which each point has a tangent line. The two curves pictured in Figure 1 are both non-singular, as $4(-7)^3 + 27(6)^2 = -400$ and $4(-2)^3 + 27(4)^2 = 400$.

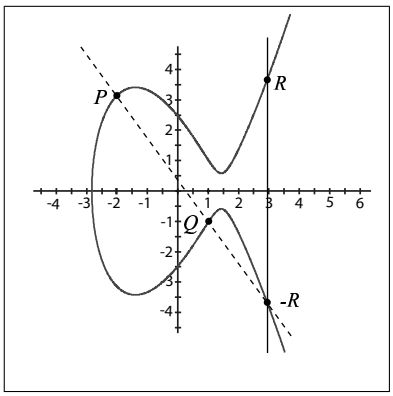
Adding points on elliptic curves over $\mathbb{R}$

A binary operation, usually denoted by addition, defined over a non-singular elliptic curve E in form of Equation (1) can be used to transform the curve into an abelian group. An *elliptic curve group* over the real numbers consists of the points on the curve, along with a special point ∞, called the *point at infinity*, which will be the identity element under this addition operation.

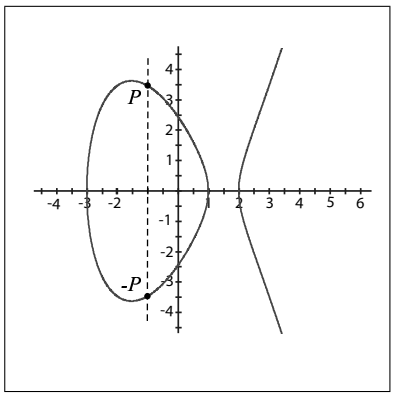
The adding of points on elliptic curves can be done using two different methods, graphical and algebraic. The key to each approach is to find the third point of intersection of the elliptic curve with the line through two given points on the curve. Any vertical line will contain the point at infinity and tangent lines contain the point of tangency twice [5].

Define the *negative of the point at infinity* to be $-\infty = \infty$ and the *negative* of any other point $P = (x_P, y_P)$ on elliptic curve E to be its reflection over the x-axis, that is $-P = (x_P, -y_P)$. Note that if $P = (x_P, y_P)$ is on the curve, then so is $-P$. The definition of addition is broken into three cases:

1. Adding two distinct points P and Q with $P \neq -Q$;
2. Adding the points P and $-P$;
3. Doubling the point P (i.e. adding the point P to itself).



$P = (-2, 3.162)$ and $Q = (1, -1)$ yield $P + Q = R = (2.925, 3.671)$ on $y^2 = x^3 - 6x + 6$



$P + (-P) = \infty$ at $P = (-1, 3.464)$ on $y^2 = x^3 - 7x + 6$

Figure 2: Adding two distinct points

Adding distinct points P and Q when P is not equal to -Q

Suppose that P and Q are distinct points on an elliptic curve with $P \neq -Q$. To add P to Q, a line is drawn between the two points and extended until it crosses the elliptic curve at the third point, $-R$. We recall that if either P or Q is a point of tangency to the curve, then $-R$ is that point of tangency. This point $-R$ is then reflected over the x-axis to its negative R. The addition of points P and Q is defined to be: $P + Q = R$. Figure 2 (left) gives an example of this case.

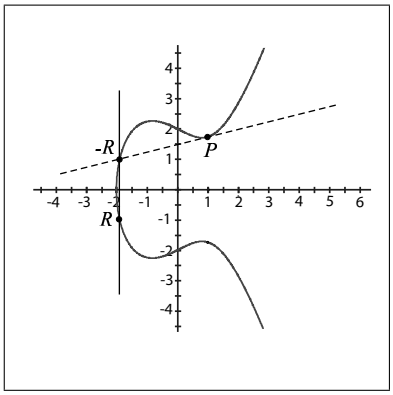
Algebraically, the coordinates of R can be calculated as $x_R = s^2 - x_P - x_Q$ and $y_R = -y_P + s(x_P - x_R)$ where $s = (y_P - y_Q)/(x_P - x_Q)$ is the slope of the line through $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$.

Adding the points P and $-P$

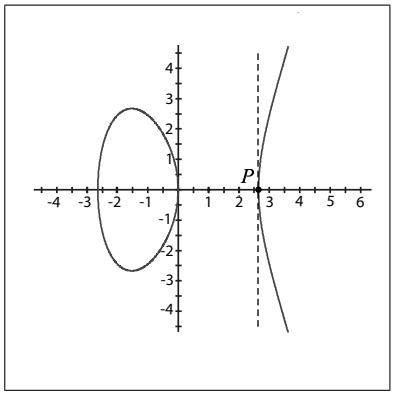
The addition of the points P and $-P$ poses a unique situation. The line through the two points is a vertical line, which will not intersect the elliptic curve at any third point, so we define $P + (-P) = \infty$, the point at infinity. Figure 2 (right) gives an example of this case.

Doubling the point P

The doubling of a point P poses yet another unique situation. Instead of drawing a line between two different points, the tangent line to the curve at the point P is drawn and extended until it crosses the elliptic curve at one other point, called $-R$. If the y-coordinate of P is zero, this tangent line will be vertical and $-R = \infty$ so that we define $2P = P + P = P + (-P) = \infty$ as in the second case. Otherwise, as in the first case, point $-R$ is reflected over the x-axis to its negative, R. Thus, the doubling of the point P is defined to be $2P = P + P = R$. Figure 3 illustrates both scenarios.



$P = (1, 1.732)$ on $y^2 = x^3 - 2x + 4$,
$P + P = R = (-1.917, -.890)$



$P = (2.646, 0)$ on $y^2 = x^3 - 7x$,
$P + P = \infty$

Figure 3: Doubling a point P

Algebraically, the coordinates of R can be calculated as $x_R = s^2 - 2x_P$ and $y_R = -y_P + s(x_P - x_R)$ where $s = (3x_P^2 + a)/(2y_P)$ is the slope of the tangent line through $P = (x_P, y_P)$.

With our definition for addition on non-singular elliptic curves, it should be clear that the group properties of closure, and commutativity are upheld. The set has an identity element, which is the point at infinity, and every point P has an inverse, as $P + (-P) = \infty$. The axiom of associativity is not as clear and is difficult to prove, but is sustained under this operation nonetheless [4]. Thus, an abelian group is formed.

Adding points on elliptic curves over $\mathbb{Z}_p$

The addition of points on elliptic curves over the real numbers is a good approach to see the underlying steps in performing the operation. However, calculations prove to be slow and inaccurate due to rounding errors, and the implementation of these calculations into cryptographic schemes requires fast and precise arithmetic. Therefore elliptic curve groups over finite fields such as $\mathbb{Z}_p$, when $p > 3$ is prime, are used in practice.

An elliptic curve with $\mathbb{Z}_p$ as its underlying field can be formed by choosing a and b within the field $\mathbb{Z}_p$. Similar to the real case, the curve includes all points (x, y) in $\mathbb{Z}_p \times \mathbb{Z}_p$ that satisfy the elliptic curve equation

$$y^2 \equiv x^3 + ax + b \mod p,$$

where x and y are numbers in $\mathbb{Z}_p$. Note that there are only finitely many points on this type of curve.

As in the real case, if $4a^3 + 27b^2 \not\equiv 0 \mod p$, then the corresponding elliptic curve forms a group [5]. This group consists of the points on the curve, along with ∞, the point at infinity. Again, we define the negative of the point at infinity to be $-\infty = \infty$ and the negative of a point $P = (x_P, y_P)$ to be $-P = (x_P, -y_P \mod p)$.

The arithmetic in an elliptic curve group over $\mathbb{Z}_p$ is very similar to that done algebraically with elliptic curve groups over the real numbers–the only difference is that all calculations are performed modulo p (see [3]).

Adding distinct points P and Q when $P \neq -Q$

Suppose $P = (x_P, y_P)$ and $Q = (x_Q, y_Q)$ and that $P \neq -Q$. Let s be given by $s \equiv (y_P - y_Q)(x_P - x_Q)^{-1} \mod p$. Then $P + Q = R$, where

$$x_R \equiv (s^2 - x_P - x_Q) \mod p \text{ and}$$

$$y_R \equiv -y_P + s(x_P - x_R) \mod p.$$

Adding the points P and $-P$

As before, we define $P + (-P) = \infty$.

Doubling the point P

If the y-coordinate of P is zero, modulo p, then $P = -P$. To double the point $P = (x_P, y_P)$ with $y_P \not\equiv 0 \bmod p$, let s be given by $s \equiv (3x_P^2 + a)(2y_P)^{-1} \bmod p$. We define $2P = P + P = R$ where

$$x_R \equiv s^2 - 2x_P \mod p \text{ and}$$

$$y_R \equiv -y_P + s(x_P - x_R) \bmod p.$$

Example. Addition table for the points on $y^2 = x^3 + 5x + 4$ over $\mathbb{Z}_{11}$.

$+$	$(0,2)$	$(0,9)$	$(2,0)$	$(4,0)$	$(5,0)$	$(10,3)$	$(10,8)$	∞
$(0,2)$	$(5,0)$	∞	$(10,8)$	$(10,3)$	$(0,9)$	$(2,0)$	$(4,0)$	$(0,2)$
$(0,9)$	∞	$(5,0)$	$(10,3)$	$(10,8)$	$(0,2)$	$(4,0)$	$(2,0)$	$(0,9)$
$(2,0)$	$(10,8)$	$(10,3)$	∞	$(5,0)$	$(4,0)$	$(0,9)$	$(0,2)$	$(2,0)$
$(4,0)$	$(10,3)$	$(10,8)$	$(5,0)$	∞	$(2,0)$	$(0,2)$	$(0,9)$	$(4,0)$
$(5,0)$	$(0,9)$	$(0,2)$	$(4,0)$	$(2,0)$	∞	$(10,8)$	$(10,3)$	$(5,0)$
$(10,3)$	$(2,0)$	$(4,0)$	$(0,9)$	$(0,2)$	$(10,8)$	$(5,0)$	∞	$(10,3)$
$(10,8)$	$(4,0)$	$(2,0)$	$(0,2)$	$(0,9)$	$(10,3)$	∞	$(5,0)$	$(10,8)$
∞	$(0,2)$	$(0,9)$	$(2,0)$	$(4,0)$	$(5,0)$	$(10,3)$	$(10,8)$	∞

Elliptic curve cryptography

Having defined the addition of points on elliptic curves over $\mathbb{Z}_p$, we now look at how to apply these ideas to the ElGamal scheme.

Elliptic curve cryptography scheme, using Alice and Bob

An ECC scheme is a form of public-key cryptosystem. Public-key cryptosystems are a relatively new technology, developed in 1976 by Whitfield Diffie and Martin Helman, both Stanford researchers. These cryptosystems involve separate encryption and decryption operations. The encryption rule uses a *public key*, while the decryption rule employs a *private key*. Knowledge of the public key allows encryption of a message but does not permit decryption of the encrypted message. The private key is kept secret so that only the intended individual can decrypt the message [2].

ECC schemes use an elliptic curve E over a finite field such as $\mathbb{Z}_p$, where p is a very large prime, and involve both an encryption and decryption operation. There are several public key schemes that can be used to encrypt and decrypt messages, such as the Diffie-Hellman scheme, the Vanstone-Menezes scheme, and the ElGamal scheme. We will look at the ElGamal encryption and decryption scheme. For more on the ElGamal or other schemes, see [5, 6, 9].

The ElGamal public-key cryptosystem is based on the Discrete Logarithm problem in $\mathbb{Z}_p^*$, the set of integers $1, 2, \ldots, p-1$, under multiplication modulo p.

The utility of the Discrete Logarithm problem in a cryptographic setting is that finding discrete logarithms is difficult, but the inverse operation of exponentiation can be computed efficiently [9]. In other words, if a person is given α, β, and $\alpha^z \equiv \beta \bmod p$, then it is very difficult to figure out the exponent z. We will use this idea in an ECC cryptosystem and perform the operations on an elliptic curve over $\mathbb{Z}_p$. Note that in an elliptic curve group, α^z is interpreted as adding α to itself z times.

This scheme will be demonstrated using Alice and Bob as sender and receiver of a secret message, respectively. Typically, the message consists of some large secret number, which is subsequently used by the two parties to open a conventional secure communication channel. The coordinates of the points on the elliptic curve itself serve as a pool of numbers to choose from.

The encryption operation

Step 1: Bob chooses a point α on an elliptic curve E over some $\mathbb{Z}_p$ and an integer z between 1 and the order of the abelian group E.

Step 2: Bob computes $\beta = z\alpha$ on the curve and publishes α, β, E, and p. He keeps his private key z secret.

Step 3: Suppose Alice wants to send a message to Bob. Alice picks an integer k between 1 and the order of E, which will be her private key.

Step 4: To encrypt a message, Alice looks up Bob's public key. As the message, she selects a point x on the elliptic curve E. Next, Alice performs the following encryption operation to encrypt the message:

$$e_k(x, k) = (k\alpha, x + k\beta) = (y_1, y_2).$$

The encrypted message is $y = (y_1, y_2)$; it includes Alice's public key y_1.

The decryption operation

Step 5: Alice sends Bob the encrypted message. To decrypt the message, Bob uses the decryption operation:

$$d_z(y_1, y_2) = y_2 - zy_1 = (x + k\beta) - z(k\alpha) = x + k(z\alpha) - z(k\alpha) = x,$$

where z is Bob's private key.

Note the interlocking of public and private keys here: Bob's private key z will decrypt this message correctly, because it matches his public key $\beta = z\alpha$, and he can be sure that it was Alice who transmitted this message, since nobody else is in possession of the private key k that matches her public key $y_1 = k\alpha$.

An example of the encryption and decryption operations

Step 1: E is the elliptic curve $y^2 = x^3 + 5x + 4$ over $\mathbb{Z}_{11}$ (see table above), $\alpha = (10, 3)$, $z = 3$. Bob's private key: $z = 3$.

Step 2: $\beta = 3(10, 3) = (10, 8)$.
Bob's public key: $\alpha = (10, 3)$, $\beta = (10, 8)$, $y^2 = x^3 + 5x + 4$ over $\mathbb{Z}_{11}$.

Step 3: Alice chooses $k = 2$.

Step 4: Alice's message is $x = (2, 0)$, which is a point on the elliptic curve E.
$y_1 = 2(10, 3) = (5, 0)$.
$y_2 = (2, 0) + 2(10, 8) = (2, 0) + (5, 0) = (4, 0)$.
The encrypted message is $y = ((5, 0), (4, 0))$.

Step 5: Beginning with $y = ((5, 0), (4, 0))$, Bob computes
$x = (4, 0) - 3(5, 0) = (4, 0) - (5, 0) = (4, 0) + (5, 0) = (2, 0)$.
The decrypted message is $x = (2, 0)$.

References

[1] E. Brown, *Three Fermat trials to elliptic curves*, The College Mathematics Journal **31** (2000) 162–172.

[2] Certicom, *The elliptic curve cryptosystem: an introduction to information security* [Retrieved October 3, 2003], `www.certicom.com`

[3] Certicom, Online ECC tutorial [Retrieved October 10, 2003], `www.certicom.com/resources/ecc_tutorial/ecc_tut_1_0.html`

[4] J. Hastad and S. Strom, *Elliptic curves*, Seminars in Theoretical Computer Science [Retrieved April 18, 2004], `www.nada.kth.se/kurser/kth/2D1441/lecturenotes/elliptic.pdf`

[5] N. Koblitz, Algebraic aspects of cryptography, Springer-Verlag(1998).

[6] M. Rosing, Implementing elliptic curve cryptography, Manning Publications (1999).

[7] W. Rudin, Principles of mathematical analysis (3rd Edition), McGraw-Hill (1964).

[8] S. Singh, The code book: the science of secrecy from ancient Egypt to quantum cryptography, Anchor Books (1999).

[9] D. Stinson, Cryptography: theory and practice (2nd Edition), Chapman & Hall/CRC (2002).