Elliptic Curves and Their Applications in Cryptography

Sijie Ma

October 14, 2023

1 Introduction

My first encounter with cryptography was through a podcast regarding the use of mathematics in WW2. Nazi Germany developed the enigma machine to encipher their secrete messages. However, through the ingenious endeavours of Ally mathematicians, a method of deciphering was invented. This spiked my interest into the world of cryptography.

Due to the invention of internet, the use of cryptography broadens and seeps into every aspect of our daily life. In this modern era, to cope with the rapidly developing technology and exponentially growing computing power, some preexisting encryption method are displaying their weaknesses and limits. Mathematicians therefore need to devise new means of encryption so that information can be communicated via public channels without getting intercepted. Furthermore, there is a growing need for a crypto-system which has small key sizes so that small mobile devices can have enough computing power to generate them. In 1985, two American mathematicians independently suggested to employ elliptic curves, a class of algebraic curves at the frontier of mathematical research, in cryptography (Sullivan). This system of cryptography is by far one of the most secure and efficient one in the world (Sullivan). Elliptic Curve Cryptography (ECC) is used for key exchange, encryption and digital signature by numerous governments as well as technology firms (Sullivan).

The first time I heard of elliptic curves was when I watched a documentary about Fermat's last theorem. As an aspiring mathematician, I am deeply fascinated by the mathematical elegance of elliptic curves, especially by their application in a variety of fields in pure mathematics such as number theory. Being able to further relate this to cryptography made this the perfect topic to study. This paper aims to investigate the fundamental group and algebraic properties of the elliptic curves, subsequently exploring their application in cryptography.

2 Elliptic Curves

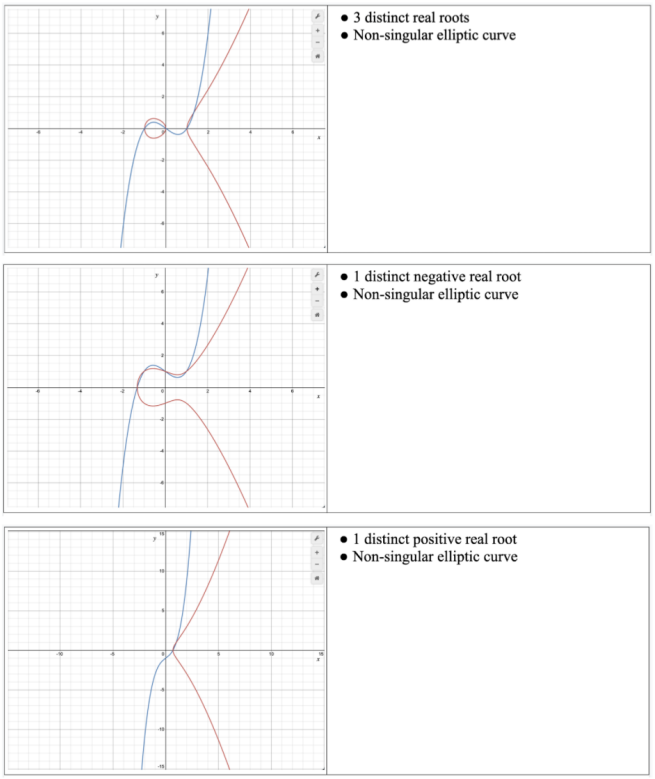
It is interesting to note that contrary to their name, elliptic curves bear no resemblance with ellipses at all. The name arose due to its origin in elliptic integrals, which are used to calculate the arc length of an ellipse (Why are They). An elliptic curve is defined as

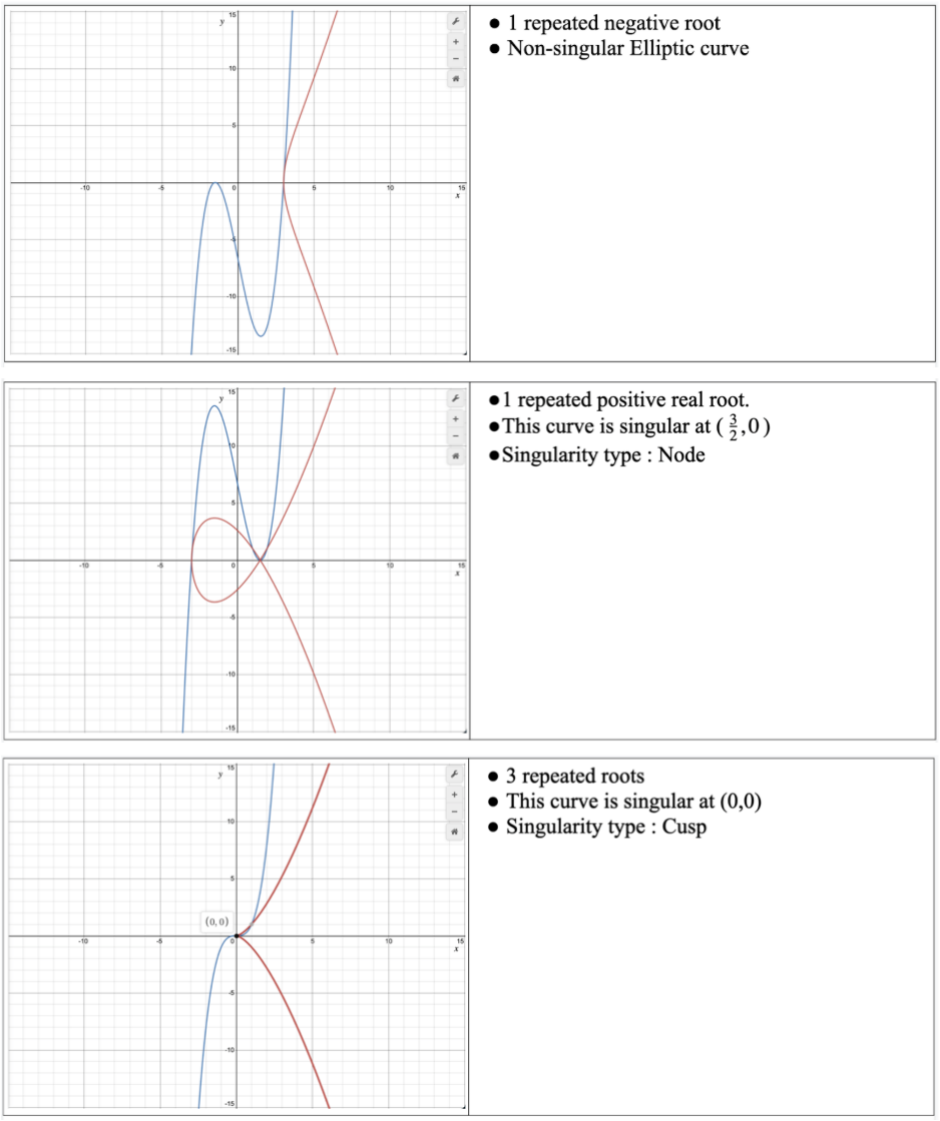
$$E : y^2 = \mathfrak{C} \tag{1}$$

$$\text{Where } \mathfrak{C} = Ax^3 + Bx + C \text{ and } \{A, B, C\} \subset \mathbb{R} \tag{2}$$

More precisely, elliptic curve is the set of points $(x, y) \in \mathbb{R}^2$ such that $E(x, y) = y^2 - Ax^3 - Bx - C = 0$. We also require that an elliptic curve to be *non-singular*, a condition that we will define later. Let us now consider different cases of $\mathfrak{C}$ and how it effects the shape and properties of E.

Figure 1: Comparison between E and $\mathfrak{C}$ in $\mathbb{R}^2$





There are several interesting properties of the elliptic curves that are associated with the shape and behavior of the cubic, $\mathfrak{C}(x)$. First of all, since $y^2 = \mathfrak{C}$ implies that $y = \pm\sqrt{\mathfrak{C}}$, all elliptic curves are symmetrical along the x-axis. Secondly, for regions such that $\mathfrak{C} < 0$, there is no real solution for y. Therefore, those intervals are not defined in E. Lastly, the elliptic curves intersect the x-axis at the exact location where $\mathfrak{C}$ has its roots. This because $y = \pm\sqrt{0} = 0$. Furthermore, note that for equation (1) to be an elliptic curve, E must not be singular. In another words, it mustn't have any singular points.

Definition 2.1 (Singular point)**.** Let $\mathcal{S}$ denote a point on E with coordinates (a,b). $\mathcal{S}$ is defined to be a *singular point* if and only if $E(a,b) = 0$ and there is no algebraically defined tangent line to E at $\mathcal{S}$ (Ash and Gross 80). In another words, more than one distinct tangential line can be drawn at this point.

There are only two types of singularities on the algebraic curve $y^2 = \mathfrak{C}$, Cusp and Node (Ash and Gross 77). Therefore, as demonstrated in Figure 1, row 5 and 6, singularity arises from repeated $\mathbb{R}_0^+$ roots of $\mathfrak{C}$. Hence, we can use the discriminant of an cubic equation to distinguish whether or not E has any singular points. Note that since there is no x^2 term in $\mathfrak{C}$, it is a "depressed" cubic, in which case the discriminant can be simplified as

$$\Delta = -16(4A^3 + 27B^2) \text{ (Ash and Gross 122).}$$

Note that for $\Delta = 0$, $\mathfrak{C}$ has repeated roots. Otherwise, it has 3 distinct roots, real or complex. To prove that singularity occurs at those repeated roots, let us first consider a new definition of tangent line.

Definition 2.2 (Tangent line (Ash and Gross 79))**.** "Suppose $f(x,y) = 0$ defines an algebraic curve C. Let $P = (a,b)$ be a point on this curve, i.e. $f(a,b) = 0$. We define the *tangent line* to C at P to be the line given by the following equation

$$f_x(a,b) \times x + f_y(a,b) \times y = f_x(a,b) \times a + f_y(a,b) \times b.$$

Note that this tangent line is undefined if both coefficients of x and y are 0. i.e. $f_x(a,b) = f_y(a,b) = 0$. We can compute the partial derivatives of equation (1) by holding the other variables as a constant. e.g. For $E(x,y) = y^2 - Ax^3 - Bx - C = 0$,

$$\frac{\partial E}{\partial x} = -3Ax^2 - B$$
$$\frac{\partial E}{\partial y} = 2y.$$

Singular points only occur when $y = 0$ because they are the repeated roots of $\mathfrak{C}$. Furthermore, since $\mathfrak{C}$ is tangent to the x-axis, the partial derivative must be equal to 0 as well at $y = 0$. If we substitute $y = 0$ into the equation $\frac{\partial E}{\partial y} = 2y$, we can quickly see that the partial derivative of E with respect to y is always 0 at any roots of $\mathfrak{C}$. Therefore, E has singular points if and only if $-3Ax^2 - B = 0$.

Lemma 2.1. Vieta's formula, which states that the sum of the roots of a polynomial of degree n in one variable, $\Sigma_i^n x_i$, is equals to $-\frac{a_{n-1}}{a_n}$. Where a_{n-1} is the coefficient of the second highest power of a polynomial and a_n denotes the coefficient of its highest power.

Since there is no x^2 term in $\mathfrak{C}$, i.e. $a_{n-1} = 0$, Vieta's formula implies that the sum of the three roots of $\mathfrak{C}$ must be 0. We can therefore consider the two different cases where E is singular.

The first scenario is for $\mathfrak{C}$ to have three repeated roots. Note that $E(x, y) = y^2 - x^3$ is the only polynomial that fits this description. To see why, let us assume a cubic equation with three repeated roots, represented by a. i.e. $\mathfrak{C} = (x - a)^3$. The sum of the three roots of $\mathfrak{C}$ must thus be $3a$.

$$\therefore 3a = 0 \Leftrightarrow a = 0$$

.

Note that throughout this exploration, assume without loss of generality that the coefficient of x^3 is 1 for the sake of simplicity. However, all statements hold true for all positive real coefficients of x^3 and the inverse holds true when the coefficient is negative.

Coming back to $E(x, y) = y^2 - x^3$, the partial derivative of E with respect to x is $-3x^2$. By substituting $x = 0$ into this expression, we obtain $\frac{\partial E}{\partial x} = 0$. Hence the tangent line is not defined at the triple repeated root of $\mathfrak{C}$.

For the second scenario, $\mathfrak{C}$ has a repeated positive real root. Note that it has to be positive because E is only defined for intervals such that $\mathfrak{C} \geq 0$. However, in Figure 1, row 4, when $\mathfrak{C}$ has a repeated negative root, $\mathfrak{C} < 0$, $\forall x < x_{\text{distinct root}}$.

Let us now consider the second scenario algebraically. Since the sum of the roots must be 0. Let $\mathfrak{C} = (x + a)(x - \frac{a}{2})^2$, where $a \in \mathbb{R}^-$. Hence the three roots of $\mathfrak{C}$ are $-a$, $\frac{a}{2}$ and $\frac{a}{2}$ respectively. Thus the sum of the roots of this cubic is always 0. Now, by expanding $\mathfrak{C}$,

$$\begin{aligned}\mathfrak{C} &= (x + a)\big(x - \frac{a}{2}\big)^2 \\ &= (x + a)\big(x^2 - ax + \frac{a^2}{4}\big) \\ &= x^3 - \frac{3a^2}{4}x + \frac{a^3}{4},\end{aligned}$$

we can define a new function to be $E(x, y) = y^2 - x^3 + \frac{3a^2}{4}x - \frac{a^3}{4}$. Its partial derivative with respect to x is therefore $\frac{\partial E}{\partial x} = -3x^2 + \frac{3a^2}{4}$. By substituting the value of x at the repeated positive root, i.e. $x = \frac{a}{2}$ into this derivative, $\frac{\partial E}{\partial x} = -\frac{3a^2}{4} + \frac{3a^2}{4} = 0$ □.

3 The Group Law

Definition 3.1 (Abelian group (Ash and Gross 117))**.** A group, specifically an *Abelian group* is a set G and a group operation, $+ : G + G \to G$, that combines two elements in G to form a third element in G. Further, any abelian -group must satisfy the following axioms under group operation.

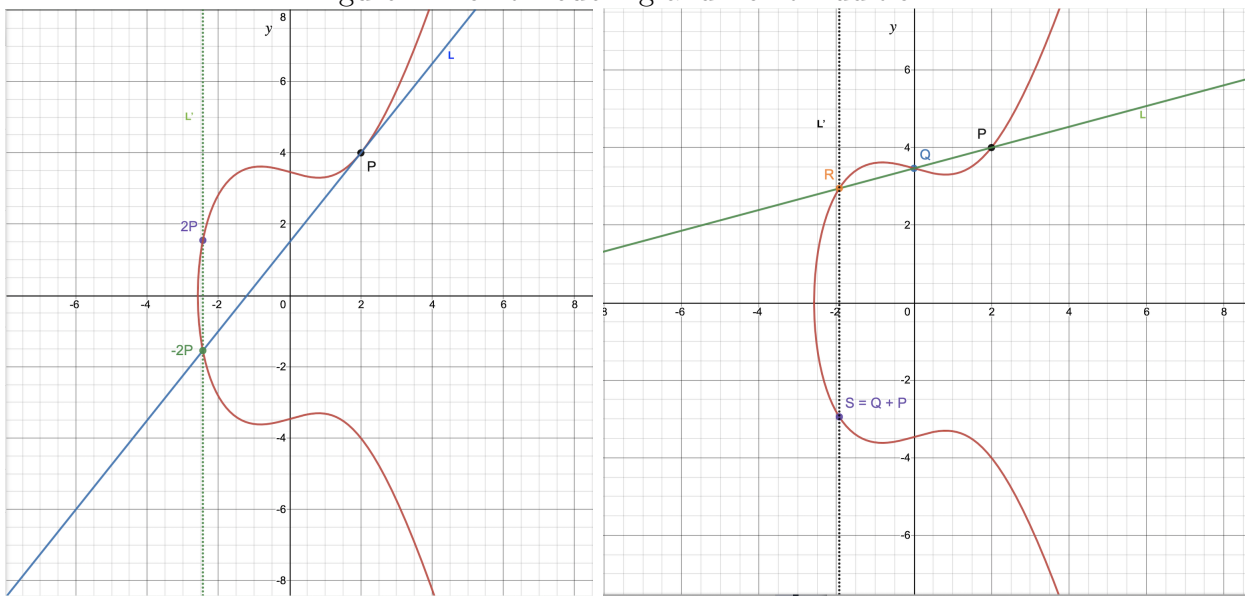
- Commutative property : $P + Q = Q + P$ for any P and Q in G
- Associative property : $(P + Q) + T = P + (Q + T)$ for any P, Q and T in G
- Identity element : There is an element in G, denoted 0 such that $P + 0 = 0 + P = P$ for any P in G
- Invertibility : For any element P in G, there is an inverse element denoted $-P$ in G such that $P + (-P) = (-P) + P = 0$.

Note that in an elliptic curve group over a number field K, denoted by $E(K)$, points can be added by using the defined group operation $+$. Let us first sketch a definition for an algebraic *number field*.

Definition 3.2 (Number Field)**.** A *field* $F \subseteq \mathbb{C}$ is a non empty collection of numbers that is *closed* under multiplication and addition. This simply means that whenever the operation is performed on any two elements in a field, the result is again an element in the same field. Furthermore, all elements in the field should have additive and multiplicative ($\neq 0$) inverses.

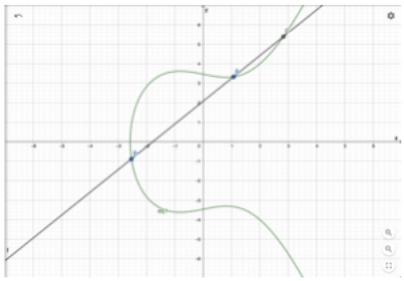
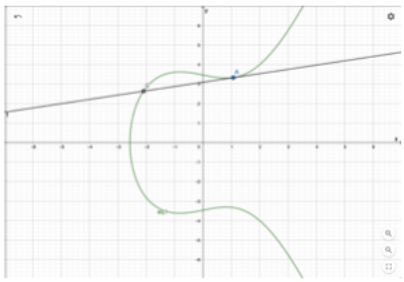
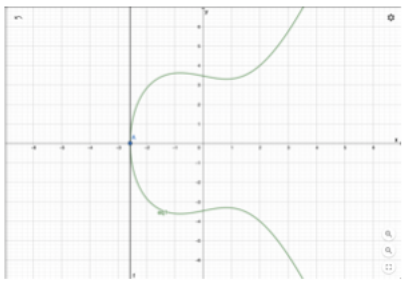
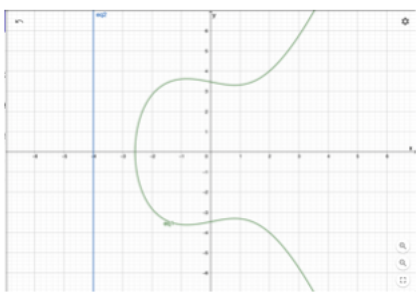
Let P and Q denote 2 arbitrary points on the curve, in another words, 2 elements in $E(K)$. If we want to add point P with point Q, a geometric approach is to first connect P and Q with a straight line L. (Note that in the case where $P = Q$, a tangent line through P is drawn). Since the algebraic degree of the polynomial $E(x, y)$ is 3, L will always intersect the E at a third point, denoted by R. Lastly, draw a line L' from R to $\mathcal{O}$, let S denote its second intersection with E. The last step, simply put, is to reflect R over the x-axis. Note that the commutative property of point addition is true because there is only 1 distinct line joining two points. Therefore, the third intersection of L and E will always be the same. This property is essential in cryptography because it allows us to ignore the order of which the private keys are applied.

Figure 2: Point Doubling and Point Addition



We can now define a few properties of $E(K)$ such that the third and forth axioms of the group law are satisfied. Let us first define a point at infinity, denoted by $\mathcal{O}$, to be the identity element of the group $E(K)$. Strictly speaking, $\mathcal{O}$ has a projective coordinate of $(0:1:0)$ in the projective plane (Ash and Gross 66). However, due to word limit, the paper will not go into the details of projective geometry. Therefore, loosely speaking, the point $\mathcal{O}$ is on the y-axis, however it is infinitely far away from the origin. Geometrically, if we were to connect an arbitrary point W on the curve with some point that is infinitely high up on the y-axis, we obtain a perfectly vertical line through W. Due to the symmetry of E along the x axis, the third intersection with the elliptic curve will occur at the reflection of W over the x axis, denoted by $-W$. Therefore, by reflecting $-W$ again over the x-axis, we obtain the original point, W. Therefore, $W + \mathcal{O} = W$, satisfying the property of the identity element. We can now define the inverse of any element in $E(K)$. Note that under point addition, the inverse of point W must satisfy $W + (-W) = (-W) + W = \mathcal{O}$. Therefore the inverse of any point on an elliptic curve is simply its reflection over the x-axis. This is because the line joining two such points would be perfectly vertical and thereby intersecting the point of infinity.

Figure 3: Different scenarios of intersections between L and E

	3 real intersection
	2 distinct real intersection, +1 due to tangency.
	1 distinct real intersection, +1 due to tangency, +1 due to intersection at O
	0 real intersection, +3 at O

By using the definition of $\mathcal{O}$, I could explain why I stated previously that L will always intersect with E at three points. This is because when there is no intersection points at real, finite coordinates as illustrated in row 3 and 4, Figure 3, we can say that the straight line L intersects E at $\mathcal{O}$ because any vertical line, L will intersect at $\mathcal{O}$.

Let (x_1, y_1), (x_2, y_2) and (x_3, y_3) denote the coordinates P, Q, R respectively. Let (x_4, y_4) denote S where $x_4 = x_3$ and $y_4 = -y_3$ due to S being the horizontal reflection of R. To derive a formula for x_3 and y_3, we first separate the two situations, point addition and point doubling (the latter when P = Q). Let λ denote the gradient of the straight line L. Note that to calculate λ under point addition, simply find the change in y with respect to x.

$$\therefore \lambda = \frac{y_2 - y_1}{x_2 - x_1}, \text{ for point addition} \tag{3}$$

To obtain λ for point doubling. Apply implicit differentiation to the equation (1),

$$\begin{aligned} \frac{dy}{dx}(y^2) &= \frac{dy}{dx}(Ax^3 + Bx + C) \\ 2y\frac{dy}{dx} &= 3Ax^2 + B \\ \therefore \lambda = \frac{dy}{dx} &= \frac{3Ax^2 + B}{2y}, \text{ for point doubling.} \end{aligned}$$

We can now look back at the seemingly arbitrary definition of tangent line which we have established in Definition 2.2, $f_x(a, b) \times x + f_y(a, b) \times y = f_x(a, b) \times a + f_y(a, b) \times b$. By rearranging the equation and factorise out the like terms, we obtain

$$\begin{aligned} \frac{\partial f}{\partial x} \times (x - a) &= -\frac{\partial f}{\partial y} \times (y - b) \\ -\frac{\partial f}{\partial x} \times \frac{\partial y}{\partial f} &= \frac{y - b}{x - a}, \end{aligned}$$

which is the gradient of a straight line. For elliptic curves specifically, since $\frac{\partial E}{\partial x} = -3Ax^2 - B$ and $\frac{\partial E}{\partial y} = 2y$, the equation above can be written as

$$\frac{\partial y}{\partial x} = \frac{3Ax^2 + B}{2y} = \frac{dy}{dx}.$$

To find out the the coordinates of the third intersection, R, we first calculate the equation of the line L and substitute for $L = y$ into E. For example, let $E_1(\mathbb{R}) = \{(x, y)|x, y \in \mathbb{R},\ y^2 = x^3 - 2x + 12\}$. Note that the candidate has chosen this equation because it has small coefficients, thus it is relatively easy to solve analytically. Given that (2, 4) is a point on the curve, denoted by $\mathbb{P}$. If we want to calculate (2, 4) + (2, 4), we first need to find out λ using equation (3).

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

By substituting $x = 2$ and $y = 4$,

$$\lambda = \frac{3 \times 4 - 2}{2 \times 4} = \frac{5}{4}.$$

The equation of the line, L is therefore

$$y - 4 = \frac{5}{4}(x - 2) \Leftrightarrow y = \frac{5}{4}x + \frac{3}{2}$$

Now, by substituting y into E_1, we obtain : $16x^3 - 25x^2 - 92x + 156 = 0$. Since L is tangent to E at $x = 2$, we know that $x = 2$ is a double root of this cubic $\therefore x^2 - 4x + 4$ must be a factor of $16x^3 - 25x^2 - 92x + 156$.

$$\begin{array}{rr|l} & & 16x \ + 39 \\ \hline x^2 - 4x + 4) & 16x^3 - 25x^2 \ - 92x + 156 & \\ & \underline{- 16x^3 + 64x^2 \ - 64x} & \\ & 39x^2 - 156x + 156 & \\ & \underline{- 39x^2 + 156x - 156} & \\ & 0 & \end{array}$$

Hence through long division, x_3 must occur when $16x + 39 = 0$

$\therefore x_3 = -\frac{39}{16}$; $y_3 = \frac{5}{4} \times -\frac{39}{16} + \frac{3}{2} = -\frac{99}{64}$ and $y_4 = \frac{99}{64}$

$$\therefore 2\mathbb{P} = (-\frac{39}{16}, \frac{99}{64})$$

However, if I were to use this method for $E_1(\mathbb{F}_p)$, where p is a prime, this method breaks down and quickly becomes too complicated because I have to take the modulo form of very big fractions (details will be explained in the next section). Therefore, it is worth noting that there is a general formula for x_3 and y_3 due to Vieta's Formulas *(Lemma 2.1)*. Let $L = \lambda x + c$, where λ and c are constants which depend on the equation of the curve as well as the points we are adding. By substituting $L = y$ into E_1, we obtain $(\lambda x + c)^2 = Ax^3 + Bx + C$. Then we expand and rearrange this equation to obtain $Ax^3 - \lambda x^2 + (B - 2\lambda c)x + C - c^2 = 0$, where the coefficient of the x^2 term is $-\lambda^2$.

$$\therefore x_1 + x_2 + x_3 = -\frac{-\lambda^2}{1}$$

$$\Leftrightarrow x_3 = \lambda^2 - x_1 - x_2 \tag{4}$$

Lastly, to find y_3, apply equation (2) and substitute x_2, y_2 with x_3, y_3. i.e.

$$\lambda = \frac{y_3 - y_1}{x_3 - x_1}$$

$$\Leftrightarrow y_3 = \lambda(x_3 - x_1) + y_1$$

$$\therefore y_4 = -\lambda(x_3 - x_1) - y_1 \tag{5}$$

Now, let us can employ equation (3), (4) and (5) to calculate $3\mathbb{P}$ in $E_1(\mathbb{R})$.

i.e. $(2, 4) + (-\frac{39}{16}, \frac{99}{64})$

$$\lambda = \frac{4 - \frac{99}{64}}{2 + \frac{39}{16}} = \frac{157}{284}$$

$$
\begin{aligned}
x_3 &= \frac{157}{284}^2 + \frac{39}{16} - 2 = \frac{3746}{5041} \\
y_3 &= \frac{157}{284} \times (\frac{3746}{5041} - 2) + 4 \approx 3.31 \\
y_4 &\approx -3.31 \\
\therefore 3\mathbb{P} &\approx (\frac{3746}{5041}, -3.31) \ \square
\end{aligned}
$$

I chose to approximate the value of y_3 because the numerator and denominator of the fraction is simply too big and if I want to find higher scalar multiples of the point (2,4), I can use other integer multiples to calculate it because $E(\mathbb{R})$ has commutative properties.
e.g. $4 \times (2,4) \equiv 3\mathbb{P} + (2,4) \equiv 2 \times (-\frac{39}{16}, \frac{99}{64})$

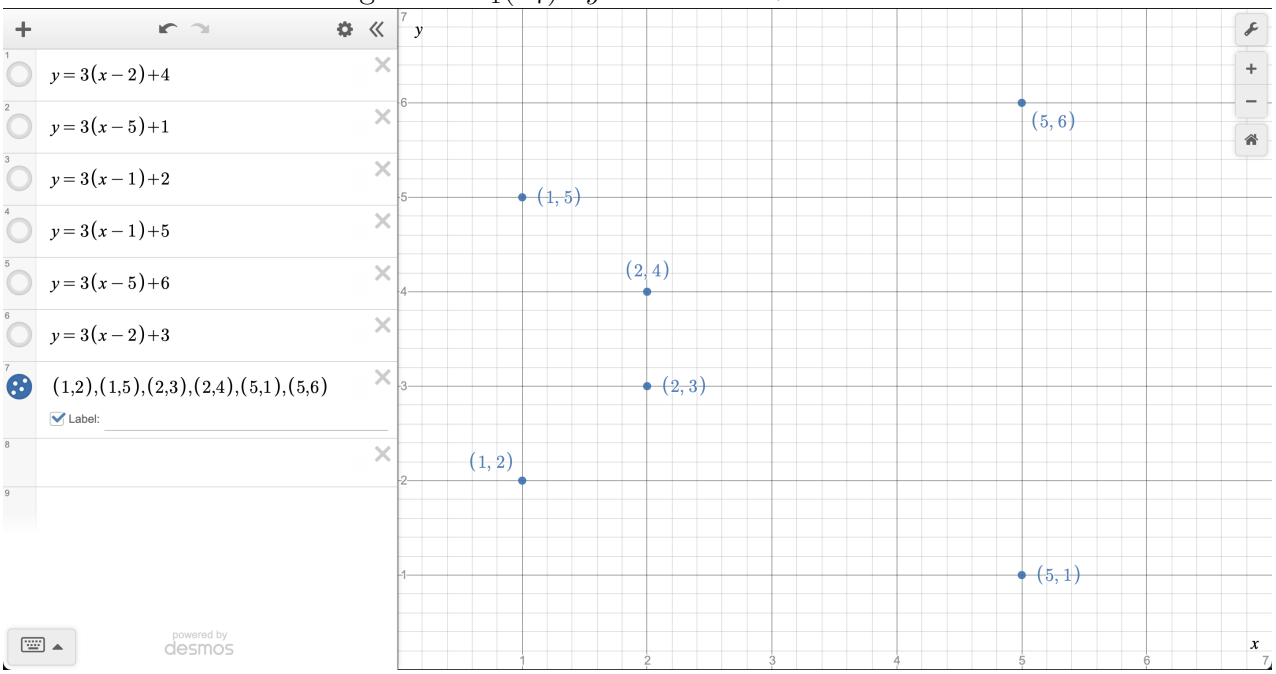
4 Elliptic curves over an integer modulo prime field

Definition 4.1 (Congruence (Ash and Gross 35))**.** $a \equiv b \bmod n$ signifies that a and b have the same remainder when divided by n

Definition 4.2 (Integer modulo prime field)**.** Similar to any other fields, e.g. $\mathbb{R}$ or $\mathbb{Z}$, an *Integer modulo p field*, where p is a prime number, consists of a set of positive integers n, such that $1 \leq n \leq p-1$. Common notations for such a field are $(\mathbb{Z}/p\mathbb{Z})$ or simply $\mathbb{F}_p$, where p is a prime number. Note that a ring - an non-empty collection of numbers - $(\mathbb{Z}/n\mathbb{Z})$, where n is any positive integer, is a field if and only if n is a prime.

Proof. Assume for the sake of contradiction that $n = ab$, where $a, b \in \mathbb{Z}/n\mathbb{Z}$ and are non-zero elements. By definition of a field, a and b must both have a multiplicative inverse. i.e. a^{-1} and b^{-1}. Note that $ab = n \equiv 0 \bmod n$. However, $0 = 0 \times a^{-1}b^{-1} = ab \times a^{-1}b^{-1} = (a \times a^{-1})(b \times b^{-1}) = 1$. This is a contradiction. Since a and b are non-zero elements, at least of of which doesn't have an inverse. Hence $\mathbb{Z}/n\mathbb{Z}$ is not a field if n is not prime. $\square$

Figure 4: $E_1(\mathbb{F}_7) : y^2 = x^3 - 2x + 12 \bmod 7$



Now let's consider the same elliptic curve over an integer modulo 7 finite field. i.e. $E_1(\mathbb{F}_7) : y^2 = x^3 - 2x + 12 \bmod 7$. Figure 3. It is interesting to note that the coordinates of the elements in $\mathbb{F}_7$ are symmetrical along the line $y = \frac{7}{2}$. This is because in $E(\mathbb{R})$, the points on the curve are symmetrical along the x-axis. Let (a, b) denote the coordinates of an arbitrary point P on $E(\mathbb{R})$. Its inverse must have the coordinate $(a, -b)$. Since there is no negative number in an $\mathbb{F}_7$ field, we must take the congruence of this negative number. i.e. $-b \equiv 7 - b \bmod 7$. Note that by definition, all coordinates as well as λ need to be in $\mathbb{F}_7$. e.g. for 2 $\times$ (2,4), it is shown previously that $\lambda = \frac{5}{4}$. Hence to calculate x_3 using equation (4), we need to first calculate the value of $\lambda_{\mathbb{F}_7} \equiv \frac{5}{4} \bmod 7$ to facilitate future calculations. This is a problem because integers cannot have a fractional remainder when divided by 7. However, we can "add, subtract and multiple congruence if they have the same modulus" (Elliptic tale). Therefore, $\frac{5}{4} \bmod 7$ can be written as follows,

$$\lambda_{\mathbb{F}_7} = 5 \bmod 7 \times 4^{-1} \bmod 7.$$

Note that 4^{-1} mod 7 is the modulo multiplicative inverse of 4 mod 7. To find its value, we need find a positive integer, n, which is smaller than 7, such that $4 \times n \equiv 1 \bmod 7$. This is because by definition of inverse, under group operation between an element in the group and its inverse would obtain the identity element, which in this case is 1 ("Modular Multiplicative"). By simple trial and error, we can quickly notice that for $n = 2$, $4 \times 2 = 8 \equiv 1 \bmod 7$.

$$\begin{aligned} \therefore \lambda_{\mathbb{F}_7} &\equiv 5 \bmod 7 \times 2 \bmod 7 \\ &\equiv 10 \bmod 7 \\ &\equiv 3 \bmod 7 \end{aligned}$$

We first have to redefine the divisibility rules for rational numbers to obtain an intuitive interpretation of the result. Suppose a number $n = \frac{p}{q} \in \mathbb{Q}$ where $p, q \in \mathbb{Z}$ and are coprimes. An integer d divides n if d divides p. Since $\frac{5}{4} - 3 = -\frac{7}{4}$, which is divisible by 7, $\frac{5}{4} \bmod 7 \equiv 3 \bmod 7$. Using equation (4), $x_3 = 3^2 - 4 = 5$. Since $5 < 7$, no changes need to be made to x_3. Now we can use equation (5) to calculate y_4,

$$\begin{aligned} y_4 &= -(3 \times (5 - 2) + 4) \bmod 7 \\ &= -13 \bmod 7 \\ &= -6 \bmod 7 \\ &= 1 \bmod 7 \end{aligned}$$

$\therefore$ $2 \times \mathbb{P} = (5, 1)$. Similarly, $3 \times \mathbb{P}$ = (1,2) ; $4 \times \mathbb{P}$ = (1,5) ; $5 \times \mathbb{P}$ = (5,6); $6 \times \mathbb{P}$ = (2,3). Note that $7 \times \mathbb{P}$ = (2,4) + (2,3) = $\mathbb{P} + (-\mathbb{P}) = \mathcal{O}$ because (2,3) is the inverse of $\mathbb{P}$. Therefore, 8 $\mathbb{P}$ = (2,4) = $\mathbb{P}$.(Appendix 1) We can hence state that the group carndinality, $\#E$, equals to 7. In another words, there are seven elements in this cyclic subgroup generated by $\mathbb{P}$.

Definition 4.3 (Cyclic group (Ash and Gross 104)). A *cyclic group* is a group that can be entirely produced by performing the group operation on a single element. All elements in a *cyclic subgroup* is a subset of the original group and undergoes the same group operation.

Note that for $E_1(\mathbb{F}_7)$, all its cyclic subgroup is equal to itself. In another words, we can choose any point on the curve as a generator, such that by repeating the group operation on itself, all elements in $E(\mathbb{F}_7)$ can be generated. However, it is not always the case. For example, let us consider the curve $E_2(\mathbb{F}_{17}) = \{(x, y)|x, y \in \mathbb{F}_{17}, y^2 = x^3 - 4x + 5 \bmod 17\}$ which has exactly 20 points, including $\mathcal{O}$. If we take the point (4,6) on the curve and multiply it by some integer scalar constant, we can obtain all elements in $E_2(\mathbb{F}_{17})$. On the other hand, if we take the point (9,1) as our generator, $2\times(9,1) = (12,6)$; $3 \times (9,1) = (12,11)$; $4 \times (9,1) = (9,16)$ and lastly, $4 \times (9,1) = \mathcal{O}$. Therefore, the $\#E = 5$. In cryptography, it is important to choose a generator which can generate a cyclic subgroup with a large group cardinality so that there are more possible private key choices (explained in the next section). Hence it is more difficult to decrypt. Furthermore, if we increase the value of p for any elliptic curves, $E(\mathbb{F}_p)$, there will be more points on the curve and therefore higher $\#E$. It is interesting to note that the number of points on the curve, N, can be estimated using Hasse's theorem.

Theorem 4.1 (Hasse's theorem). Given an elliptic curve E modulo p, the number of points on the curve is denoted by N and is bounded by :

$$p + 1 - 2\sqrt{p} \leq N \leq p + 1 + 2\sqrt{p}$$ ("Lecture 17: Elliptic" [HH:MM:SS]).

Hence, the number of points in $E_2(\mathbb{F}_{17})$, denoted by N_2 is bounded by

$$17 + 1 - 2\sqrt{17} \leq N_2 \leq 17 + 1 + 2\sqrt{17}$$

$$9 < N_2 < 27$$

Since $9 < 20 < 27$, the theorem holds true.

5 Application of Elliptic Curves in Diffie-Helman Key Exchange Protocol (ECDH)

Let us now consider the previous elliptic curve over $\mathbb{F}_{17}$. Due to its cyclic property, an elliptic curve over a finite field can be used in a Discrete Logarithm Problem(DLP). It is a problem that is very easy to obtain the answer, however it is extremely difficult to work backwards. Hence this type of "trapdoor" problem is the cornerstone of modern cryptography.

Definition 5.1 (Discrete Logarithm Problem)**.** Let G denote a group whose group operation is $\times$ and identity element is 1. Let a and b denote 2 elements in the group G, we are essentially solving for a value of k, such that $b^k = a$. In another words, we are performing the group operation, $\times$, on b itself k times to obtain a. It is called a discrete logarithm problem because by rearranging the previous expression, we obtain $k = \log_b a$.

Consider the following problem : $2^k = 16$. Note that $k = 4$ is trivial to obtain. However, the solution to $2^n \bmod p \equiv 16$ is more difficult to compute both analytically and numerically because there can be multiple solutions. By employing the same technique to an elliptic curve over a finite field, we can choose a generator point, $\mathcal{G}$ and multiply it by a scalar constant k to obtain a second point on the curve, $k\mathcal{G}$. This step is easy to calculate. However, if we were only given $k\mathcal{G}$ and $\mathcal{G}$, it takes a long computing time to obtain k.

We can now apply DLP with elliptic curves in cryptography through Diffie-Helman Key Exchange Protocol. In cryptography, there are mainly two types of keys, private and public keys which are both used to encrypt and decrypt data that are shared between the sender and receiver (Parahar). As suggested by their names, private keys are usually kept secrete and public keys are visible to everyone on the internet.

Let $\mathcal{A}$ and $\mathcal{B}$ denote the sender and the receiver respectively. Let $P_{\mathcal{A}} = n_{\mathcal{A}} \times \mathcal{G}$ and $P_{\mathcal{B}} = n_{\mathcal{B}} \times \mathcal{G}$. In the case of ECDH, the public keys are the following : $P_{\mathcal{A}}$; $P_{\mathcal{B}}$; $\mathcal{G}$; equation of the elliptic curve and lastly the cardinality of the cyclic subgroup generated by $\mathcal{G}$. The private keys are two arbitrary positive integer constants denoted by $n_{\mathcal{A}}$ and $n_{\mathcal{B}}$. They have to be smaller than $\#E$ otherwise, there can be infinitely many possible private keys. For example, suppose $\#E = 5$, $3\mathcal{G} = 8\mathcal{G} = (3 + 5n)\mathcal{G}$, where n is any positive integer number. Therefore, a larger group cardinatlity offers more possibilities for the private keys. The DLP is therefore harder and it would require more computing time to decipher.

The mechanism behind ECDH is as follows. Both party first multiply their private key by the generator to obtain $P_{\mathcal{A}}$ and $P_{\mathcal{B}}$. Then $\mathcal{A}$ will multiply its private key by $P_{\mathcal{B}}$ and vise versa. Since an elliptic curve is an Abelian group, $n_{\mathcal{A}} \times (n_{\mathcal{B}} \times \mathcal{G}) = n_{\mathcal{B}} \times (n_{\mathcal{A}} \times \mathcal{G})$. Therefore, $\mathcal{A}$ and $\mathcal{B}$ will arrive at the same point on the curve. Note that one application of ECDH is the transmission of encrypted message. This application is widely used in a range of social media platforms, including but not limited to Whatsapp and Facebook.

Consider the following scenario, $\mathcal{A}$ would like to share a plain text message, $\mathcal{M}$ with $\mathcal{B}$. $\mathcal{M}$ is first encrypted into a sequence of numbers represented by a point $P_{\mathcal{M}}$ on $E(\mathbb{F}_p)$ by using some other encryption system. Then, we obtain $P_{\mathcal{A}}$ and $P_{\mathcal{B}}$ as stated above. After that, $\mathcal{A}$ would multiply $P_{\mathcal{B}}$ by its private key, $n_{\mathcal{A}}$ and add the product to the message point $P_{\mathcal{M}}$. i.e. Let $P_{\mathcal{C}} = P_{\mathcal{M}} + n_{\mathcal{A}} \times P_{\mathcal{B}}$. Lastly, $\mathcal{B}$ would multiply $P_{\mathcal{A}}$ with its private key, $n_{\mathcal{B}}$ and subtract it from $P_{\mathcal{C}}$ to obtain the original message point, $P_{\mathcal{M}}$. This is because

$$(P_{\mathcal{M}} + n_{\mathcal{A}} \times P_{\mathcal{B}}) - n_{\mathcal{B}} \times (n_{\mathcal{A}} \times \mathcal{G}) = P_{\mathcal{M}} \ \square$$

Let us consider the curve $E_2(\mathbb{F}_{17})$ again. Let $\mathcal{G} = (4, 6)$; $n_{\mathcal{A}} = 7$; $n_{\mathcal{B}} = 12$ and $P_{\mathcal{M}} = (10, 8)$ for the message $\mathcal{A}$ wants to share with $\mathcal{B}$. To avoid repetition in calculation, I have used "Elliptic Curves over Finite Fields" by Sascha Grau to compute the values. Note that $P_{\mathcal{A}} = 7 \times (4, 6) = (16, 5)$ and $P_{\mathcal{B}} = 12 \times (4, 6) = (12, 11)$. (Appendix 2) $\mathcal{A}$ can therefore calculate $P_{\mathcal{C}}$. i.e. $P_{\mathcal{C}} = (10, 8) + 7 \times (12, 11) = (13, 5)$. Now, $\mathcal{B}$ can work reversely to calculate $P_{\mathcal{M}}$. i.e. $P_{\mathcal{M}} = (13, 5) - 12 \times (16, 5)$. Note that $12 \times (16, 5) = (9, 1)$. Furthermore, the inverse of $(9, 1)$ is $(9, 16)$ because $-1 \bmod 17 \equiv 16 \bmod 17$. Under the group law of elliptic curves, subtracting a point is equivilant to adding the inverse of the point.

$$\therefore P_{\mathcal{M}} = (13, 5) + (9, 16) = (10, 8)$$

This calculated value from $\mathcal{B}$ is the same as the $P_{\mathcal{M}}$ made up by $\mathcal{A}$. The correct message is therefore accurately transmitted. Furthermore, this is an example of the associative property of elliptic curves. Let us now consider why this system is secure. Let $\mathcal{E}$ denote the eavesdropper who wants to intercept the message $\mathcal{M}$ by calculating $P_{\mathcal{M}}$ without knowing any of the private keys $n_{\mathcal{A}}$ and $n_{\mathcal{B}}$. However, $\mathcal{E}$ needs to know both $n_{\mathcal{A}}$ and $n_{\mathcal{B}}$ to calculate $P_{\mathcal{M}}$. This brings back the elliptic curves DLP. $\mathcal{E}$ knows $P_{\mathcal{A}}$; $P_{\mathcal{B}}$ and $\mathcal{G}$ because they are the public keys. However, calculating $n_{\mathcal{A}}$ and $n_{\mathcal{B}}$ is a discrete logarithm problem, e.g. $P_{\mathcal{A}} = n_{\mathcal{A}} \times \mathcal{G}$. It is considered to be a "hard" problem in computer science because the computation time required to solve it varies exponentially with the the complexity of the problem as opposed to an easy problem, where the time varies linearly.

6 Conclusion

In conclusion, the paper has thoroughly investigated the mathematical properties of $E(\mathbb{R})$ and $E(\mathbb{F}_p)$, including but not limited to the group laws and the shapes of the elliptic curves. The paper has also outlined one of the applications of $E(\mathbb{F}_p)$ in cryptography, specifically

in ECDH. However, due to word constraints, the paper cannot fully and rigorously explain some of the ideas behind elliptic curves, such as its associative property. Since abstract algebra topics such as *fields* and *modular arithmetic* are beyond the content of syllabus, explaining some of those concepts will require a substantial amount of definition and therefore digressing from the original aim of the paper. Furthermore, the proof of some properties such as associativity and discriminant of elliptic curves are not included in the paper because they are beyond the scope of the IB syllabus. Therefore setting them out and explaining them will deter the coherence of the paper.

This investigation has completely altered my understanding and approach to mathematics by introducing me to the fabulous world of abstraction and rigour. For further research, I could possibly explore elliptic curves over rational and complex field. The former, $E(\mathbb{Q})$ is closely related to one of the Millennium Prize Problems, the Birch and Swinnerton-Dyer conjecture, which describes the rational points on an elliptic curve. Whereas the latter, $E(\mathbb{C})$ is interestingly a torus in a 3D space. Therefore, research into topology and complex analysis is needed to fully understand $E(\mathbb{C})$.

A possible weakness of the paper is the limited application of elliptic curve in cryptography. First of all, the value of p in $E(\mathbb{F}_p)$ as well as the coefficient of the elliptic curve would practically be extremely large. For example, "it has been noted by the NSA that the encryption of a top-secret document by elliptic curve cryptography requires a key length of 384 bit" ("Elliptic Curve"). This means that $p > 2^{384}$. Furthermore, there is also a range of use of elliptic curves in cryptography which isn't explored by this paper, including but not limited to digital signature and direct encryption. I have really enjoyed this project and in the future, I will write my own encryption programme as well as an algorithm that solves elliptic curve discrete logarithm problem.

Works Cited

Ash, Avner, and Robert Gross. *Elliptic Tales: Curves, Counting, and Number Theory*. Princeton, Princeton UP, 2012.

"Elliptic Curve Cryptography." *Keycdn*, 4 Oct. 2018, www.keycdn.com/support/elliptic-curve-cryptography. Accessed 7 Apr. 2021.

Elliptic Curves over Finite Fields. Sascha Grau, 2017, graui.de/code/elliptic2/. Accessed 7 Apr. 2021.

"Lecture 17: Elliptic Curve Cryptography (ECC) by Christof Paar." *Youtube*, 30 Jan. 2014, www.youtube.com/watch?v=zTt4gvuQ6sY. Accessed 7 Apr. 2021.

"Modular Multiplicative Inverse." *GeeksforGeeks*, 31 Mar. 2021, www.geeksforgeeks.org/multiplicative-inverse-under-modulo-m/. Accessed 7 Apr. 2021.

Parahar, Mahesh. "Difference between Private Key and Public Key." *tutorialspoint*, 15 Apr. 2020, www.tutorialspoint.com/difference-between-private-key-and-public-key. Accessed 7 Apr. 2021.

Sullivan, Nick. "A (Relatively Easy To Understand) Primer on Elliptic Curve Cryptography." *The CloudFlare Blog*, 24 Oct. 2013, blog.cloudflare.com/a-relatively-easy-to-understand-primer-on-elliptic-curve-cryptography/. Accessed 7 Apr. 2021.

Why are They Called "Elliptic" Curves? University of North Florida, www.unf.edu/~ddreibel/mas4932/elliptic-integrals.pdf. Accessed 7 Apr. 2021.

Works Consulted

"The Beauty of Elliptic Curves." *Youtube*, 26 Oct. 2017, www.youtube.com/watch?v=ThWilMcOUxg. Accessed 7 Apr. 2021.

De Feo, Luca. "Elliptic Curve Cryptography." *defeo.lu*, Luca De Feo, 18 Apr. 2007, defeo.lu/talks/pisa-jun-07.pdf. Accessed 7 Apr. 2021.

"Elliptic Curve Cryptography." *School of Electrical Engineering and Computer Science (EECS)*, University of Ottawa, www.site.uottawa.ca/~chouinar/Handout_Elliptic_Curve_Crypto.pdf. Accessed 7 Apr. 2021.

"Elliptic Curve Diffie Hellman." *Youtube*, 10 Dec. 2014, www.youtube.com/watch?v=F3zzNa42-tQ&t=517s. Accessed 7 Apr. 2021.

"Lecture 17: Elliptic Curve Cryptography (ECC) by Christof Paar." *Youtube*, 30 Jan. 2014, www.youtube.com/watch?v=zTt4gvuQ6sY. Accessed 7 Apr. 2021.

"Lecture 16: Introduction to Elliptic Curves by Christof Paar." *Youtube*, 30 Jan. 2014, www.youtube.com/watch?v=vnpZXJL6QCQ. Accessed 7 Apr. 2021.

Appendix 1 : Point addition table for $E_1(\mathbb{F}_7) : y^2 = x^3 - 2x + 12 \bmod 7$ (Elliptic Curves).

+	∞	(1,2)	(1,5)	(2,3)	(2,4)	(5,1)	(5,6)
∞	∞	(1,2)	(1,5)	(2,3)	(2,4)	(5,1)	(5,6)
(1,2)	(1,2)	(2,3)	∞	(5,1)	(1,5)	(5,6)	(2,4)
(1,5)	(1,5)	∞	(2,4)	(1,2)	(5,6)	(2,3)	(5,1)
(2,3)	(2,3)	(5,1)	(1,2)	(5,6)	∞	(2,4)	(1,5)
(2,4)	(2,4)	(1,5)	(5,6)	∞	(5,1)	(1,2)	(2,3)
(5,1)	(5,1)	(5,6)	(2,3)	(2,4)	(1,2)	(1,5)	∞
(5,6)	(5,6)	(2,4)	(5,1)	(1,5)	(2,3)	∞	(1,2)

Appendix 2 : Point addition table for $E_2(\mathbb{F}_{17}) : y^2 = x^3 - 4x + 5 \bmod 17$(Elliptic Curves).

+	∞	(1,6)	(1,11)	(4,6)	(4,11)	(5,5)	(5,12)	(8,3)	(8,14)	(9,1)	(9,16)	(10,8)	(10,9)	(11,0)	(12,6)	(12,11)	(13,5)	(13,12)	(16,5)	(16,12)
∞	∞	(1,6)	(1,11)	(4,6)	(4,11)	(5,5)	(5,12)	(8,3)	(8,14)	(9,1)	(9,16)	(10,8)	(10,9)	(11,0)	(12,6)	(12,11)	(13,5)	(13,12)	(16,5)	(16,12)
(1,6)	(1,6)	(13,12)	∞	(12,11)	(11,0)	(10,9)	(9,16)	(12,6)	(10,8)	(5,5)	(16,5)	(5,12)	(8,3)	(4,6)	(4,11)	(8,14)	(1,11)	(16,12)	(13,5)	(9,1)
(1,11)	(1,11)	∞	(13,5)	(11,0)	(12,6)	(9,1)	(10,8)	(10,9)	(12,11)	(16,12)	(5,12)	(8,14)	(5,5)	(4,11)	(8,3)	(4,6)	(16,5)	(1,6)	(9,16)	(13,12)
(4,6)	(4,6)	(12,11)	(11,0)	(13,12)	∞	(9,16)	(10,9)	(13,5)	(9,1)	(5,12)	(8,3)	(5,5)	(16,5)	(1,6)	(1,11)	(16,12)	(4,11)	(8,14)	(12,6)	(10,8)
(4,11)	(4,11)	(11,0)	(12,6)	∞	(13,5)	(10,8)	(9,1)	(9,16)	(13,12)	(8,14)	(5,5)	(16,12)	(5,12)	(1,11)	(16,5)	(1,6)	(8,3)	(4,6)	(10,9)	(12,11)
(5,5)	(5,5)	(10,9)	(9,1)	(9,16)	(10,8)	(11,0)	∞	(12,11)	(13,5)	(4,11)	(1,6)	(1,11)	(4,6)	(5,12)	(8,14)	(16,5)	(16,12)	(8,3)	(13,12)	(12,6)
(5,12)	(5,12)	(9,16)	(10,8)	(10,9)	(9,1)	∞	(11,0)	(13,12)	(12,6)	(1,11)	(4,6)	(4,11)	(1,6)	(5,5)	(16,12)	(8,3)	(8,14)	(16,5)	(12,11)	(13,5)
(8,3)	(8,3)	(12,6)	(10,9)	(13,5)	(9,16)	(12,11)	(13,12)	(10,8)	∞	(4,6)	(16,12)	(1,6)	(8,14)	(16,5)	(5,12)	(1,11)	(5,5)	(4,11)	(9,1)	(11,0)
(8,14)	(8,14)	(10,8)	(12,11)	(9,1)	(13,12)	(13,5)	(12,6)	∞	(10,9)	(16,5)	(4,11)	(8,3)	(1,11)	(16,12)	(1,6)	(5,5)	(4,6)	(5,12)	(11,0)	(9,16)
(9,1)	(9,1)	(5,5)	(16,12)	(5,12)	(8,14)	(4,11)	(1,11)	(4,6)	(16,5)	(12,6)	∞	(13,5)	(11,0)	(10,8)	(12,11)	(9,16)	(13,12)	(10,9)	(1,6)	(8,3)
(9,16)	(9,16)	(16,5)	(5,12)	(8,3)	(5,5)	(1,6)	(4,6)	(16,12)	(4,11)	∞	(12,11)	(11,0)	(13,12)	(10,9)	(9,1)	(12,6)	(10,8)	(13,5)	(8,14)	(1,11)
(10,8)	(10,8)	(5,12)	(8,14)	(5,5)	(16,12)	(1,11)	(4,11)	(1,6)	(8,3)	(13,5)	(11,0)	(12,6)	∞	(9,1)	(13,12)	(10,9)	(12,11)	(9,16)	(4,6)	(16,5)
(10,9)	(10,9)	(8,3)	(5,5)	(16,5)	(5,12)	(4,6)	(1,6)	(8,14)	(1,11)	(11,0)	(13,12)	∞	(12,11)	(9,16)	(10,8)	(13,5)	(9,1)	(12,6)	(16,12)	(4,11)
(11,0)	(11,0)	(4,6)	(4,11)	(1,6)	(1,11)	(5,12)	(5,5)	(16,5)	(16,12)	(10,8)	(10,9)	(9,1)	(9,16)	∞	(13,5)	(13,12)	(12,6)	(12,11)	(8,3)	(8,14)
(12,6)	(12,6)	(4,11)	(8,3)	(1,11)	(16,5)	(8,14)	(16,12)	(5,12)	(1,6)	(12,11)	(9,1)	(13,12)	(10,8)	(13,5)	(9,16)	∞	(10,9)	(11,0)	(5,5)	(4,6)
(12,11)	(12,11)	(8,14)	(4,6)	(16,12)	(1,6)	(16,5)	(8,3)	(1,11)	(5,5)	(9,16)	(12,6)	(10,9)	(13,5)	(13,12)	∞	(9,1)	(11,0)	(10,8)	(4,11)	(5,12)
(13,5)	(13,5)	(1,11)	(16,5)	(4,11)	(8,3)	(16,12)	(8,14)	(5,5)	(4,6)	(13,12)	(10,8)	(12,11)	(9,1)	(12,6)	(10,9)	(11,0)	(9,16)	∞	(5,12)	(1,6)
(13,12)	(13,12)	(16,12)	(1,6)	(8,14)	(4,6)	(8,3)	(16,5)	(4,11)	(5,12)	(10,9)	(13,5)	(9,16)	(12,6)	(12,11)	(11,0)	(10,8)	∞	(9,1)	(1,11)	(5,5)
(16,5)	(16,5)	(13,5)	(9,16)	(12,6)	(10,9)	(13,12)	(12,11)	(9,1)	(11,0)	(1,6)	(8,14)	(4,6)	(16,12)	(8,3)	(5,5)	(4,11)	(5,12)	(1,11)	(10,8)	∞
(16,12)	(16,12)	(9,1)	(13,12)	(10,8)	(12,11)	(12,6)	(13,5)	(11,0)	(9,16)	(8,3)	(1,11)	(16,5)	(4,11)	(8,14)	(4,6)	(5,12)	(1,6)	(5,5)	∞	(10,9)