AN ELEMENTARY PROOF OF MORDELL'S THEOREM

SUHAS V. GONDI

Abstract. Mordell's theorem states that the group of rational points on an elliptic curve is finitely generated. In this expository paper, we present an elementary proof of this theorem. We first define elliptic curves. We then prove that every elliptic curve can be brought to a general form known as the Weierstrass normal form. In doing so we define an operation on the set of rational points on an elliptic curve and prove that it forms an abelian group.

Contents

1. Introduction

The problem of finding rational solutions to polynomial equations has been a fascinating subject in mathematics since Diophantus's era. It involves three principal fields in mathematics: algebra, geometry and number theory.

For a given polynomial, the first question we would like answered is whether there exists a rational root or not. If there exists a root, we would like to know whether there are finitely or infinitely many roots. If it is the latter, we then try to search for a method (or prove no such method exists) of generating all the roots from a finite set of intial points. These questions may be answered to polynomials in one variable with relative ease, but treating a polynomial in three or more variables is quite involved. Hence, we focus on the two-variable case. We shall soon see that a satisfactory description of the rational solutions may be given to two-variable polynomials of degrees one and two, but things start getting complicated for polynomials of degree three.

Suppose $ax+by=c$ is a linear polynomial with integer coefficients, with a and b not equal to zero. Then, x is rational if and only if y is rational. Further, for every rational x, we get a unique rational y. Thus, we may list all the rational solutions of a linear polynomial through parameterization.

Date: 1st October 2018.

We now try to find the rational points on a quadric with rational coefficients. First, we note that there exists a method that determines whether a conic has a rational solution or not. This follows from a theorem due to Legendre, which we will not be stating. Now, suppose that we are given a rational conic C and a rational point O on it. Let us fix an arbitrary rational line l. Suppose we draw a line through O that intersects the line l at a rational point. Then, the other point of intersection of this line with C is also a rational point. Further, a line through O and a rational point on C intersects l at a rational point. These statements follow from the following observations:

- A line passing through two rational points has rational coefficients.
- If an intersection point of a rational line and a rational quadric is rational, then the other intersection point is also rational.
- The intersection point of two distinct rational lines is rational.

Thus, we get a one-one correspondence between the rational points of l and C. Hence, our problem of finding the rational points on a conic is reduced to the linear polynomial case.

Obtaining a one-one correspondence between rational points on a cubic and line is a lot harder for a cubic because a line intersects a cubic at three points. But we do know a useful property on the rational points of a cubic. In this paper, we look at a theorem, proposed by Poincare in 1901 and proven by Mordell in 1923, which states that for an elliptic curve, we can generate all of its rational points from a finite number of rational points.

This paper focuses on giving an elementary proof of Mordell's theorem for a wide range of cubics. In Section 2, we give a general form for an elliptic curve, called the Weierstass normal form, which is easy to work with. In Sections 3 and 4, we define and describe the group structure on the set of rational points on an elliptic curve. We also give a group theoretic statement of Mordell's theorem at the end of Section 4. The proof of the theorem is spread over Sections 5, 6, and 7. Section 5 involves some subtle arguements while Section 6 involves a lot of computations. Section 7 uses results proven in Sections 5 and 6 to give a proof of Mordell's theorem. The main idea of the proof presented in this paper is taken from chapter 3 of Silverman and Tate's Rational Points on Elliptic Curves [1].

2. Weierstrass normal form of an elliptic curve

Definition 2.1. A cubic curve is a plane curve defined by a polynomial of degree three, whose general form is given by the equation

$$ax^3 + by^3 + cx^2y + dxy^2 + exy + fx^2 + gy^2 + hx + iy + j = 0 \tag{2.2}$$

A rational cubic curve has all its coefficients coming from the field of rationals.

Definition 2.3. We call a cubic curve non-singular if it has no points at which both the partial deriatives vanish.

Geometrically, it means that you can draw a unique tangent to each point of a non-singular cubic curve.

Definition 2.4. An elliptic curve is a non-singular rational cubic curve containing a rational point.

Remark 2.5. There is no known method that determines the existence of a rational point on a cubic in a finite number of steps. It is still an open problem.

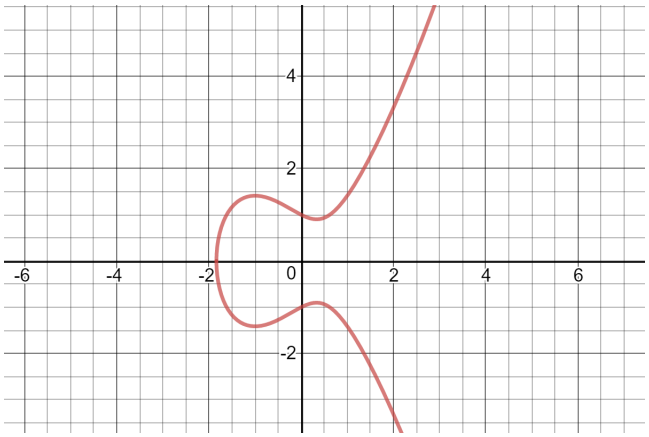

FIGURE 1. The graph of an elliptic curve given by the equation $y^2 = x^3 + x^2 - x + 1$

The general equation of a cubic has ten terms and hence ten coefficients to specify. However, the general equation of an elliptic curve can be brought down to a simpler form that is easier to work with. We prove in this section that any elliptic curve can be brought into the following form through a suitable change of coordinates that sends rational points to rational points.

$$y^2 = x^3 + ax^2 + bx + c \tag{2.6}$$

(2.6) is called the Weierstrass normal form of an elliptic curve (see figure 1). Before proving that any elliptic curve can be expressed in Weierstrass normal form, we state without proof a theorem regarding the number of intersection points of two curves. It is a widely used theorem in algebraic geometry, known as Bezout's theorem.

Theorem 2.7. *Let C_1 and C_2 be two curves defined on the projective space such that they do not have a common component. The total number of intersection points of C_1 and C_2 with coordinates coming from the complex field, taking into account multiplicities, is equal to the product of the degrees of the two curves.*

Remark 2.8. While counting the intersection points, keep in mind that we must also include the points at infinity. Thus, Theorem 2.7 also holds for a pair of parallel lines. We must count repeated roots as many times as they appear. For example, the curves $y = 0$ and $y = x^2$ intersect twice at $(0, 0)$, as it is a double root of the equation $x^2 = 0$. Finally, we cannot let the two curves have a common component as otherwise, the two curves intersect at infinitely many points.

Now, in the projective plane $\mathbb{P}^3$, the elliptic curve takes the form

$$aX^3+bY^3+cX^2Y+dXY^2+eXYZ+fX^2Z+gY^2Z+hXZ^2+iYZ^2+jZ^3 = 0 \tag{2.9}$$

Recall that an elliptic curve must contain a rational point. Let us fix a rational point on the curve and denote it by O. We take the line tangent to the curve at O to be the line $Z = 0$ (see figure 2). Let P denote the other point of intersection of this line with the curve. We take the line $X = 0$ to be the line tangent to the curve at P. We let $Y = 0$ be any line passing through O that is distinct from the line $Z = 0$.

We remark that under the projective transformation just described, rational points are sent to rational points. This follows from the fact that the tangent line at a rational point has rational coefficients and the line joining two rational points also has rational coefficients. We now show that under this projective transformation, (2.9) can be simplified substantially.

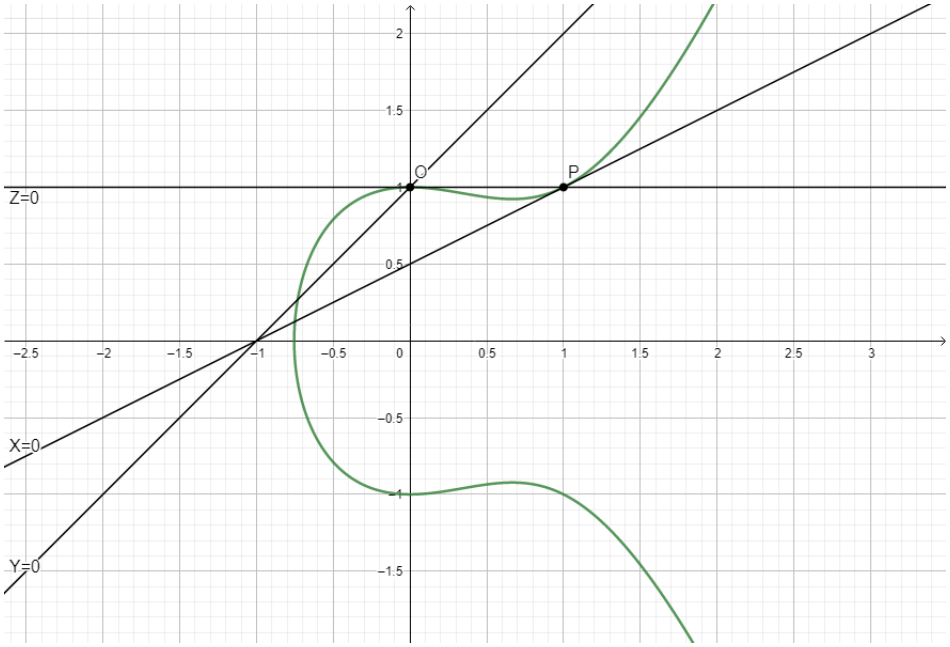

FIGURE 2. Choosing axes to put the elliptic curve in Weierstrass form

First, observe that the points O and P have coordinates $[1:0:0]$ and $[0:1:0]$, respectively. By substituting these values in (2.9), we get $a = 0$ and $b = 0$. Further, the line $Z = 0$ meets the curve at two points; twice at O (as it is tangent at that point) and once at P. The points that satisfy both the line $Z = 0$ and the elliptic curve are the points that satisfy the curve $cX^2Y + dXY^2 = 0$. Note that $[1:0:0]$ is a double root of this polynomial, as the line $Z = 0$ is tangent to the curve at O. Thus, $[1:0:0]$ is a root of $cX^2 + dXY = 0$. From this, we see that $c = 0$ holds. Thus, we now get an equation whose affine form is given by the equation

$$(dx + g)y^2 + (ex + i)y = fx^2 + hx + j. \tag{2.10}$$

By applying the transformation that sends $dx + g$ to x, we can simplify (2.10) further to give

$$xy^2 + (\alpha x + \beta)y = f(x) \tag{2.11}$$

where $f(x)$ is a polynomial of degree two. We then multiply x on both sides of (2.11) and replace xy with y to get

$$y^2 + (\alpha x + \beta)y = g(x) \tag{2.12}$$

where $g(x)$ is a cubic in x. Note that the final transformation we made is not an affine transformation, but it preserves the set of rational points, which is all we

need. We finally complete the square on the left-hand side to get the equation in the form of (2.6).

We shall henceforth use the Weierstrass normal form to represent the general form of an elliptic curve. In the next section, we define a group operation on the set of rational points on an elliptic curve. We state beforehand, without justification, that the transformation that sends the general form of a cubic to its Weierstrass normal form is a homomorphism under this group operation. Thus, we may define the group operation on an elliptic curve in its Weierstrass normal form, instead of the general form.

3. The group structure on elliptic curves

Before defining the group law on the set of rational points of an elliptic curve, we give a few comments on the properties of the curve. First, we find all the points at infinity that lie on the curve. We do this by looking at the points of intersection of the curve defined by (2.6) with the line $Z = 0$. This gives us the equation $X^3 = 0$. Thus, we see that the curve has just one point at infinity lying on it, namely $[0, 1, 0]$. This corresponds to the equivalence class of lines parallel to the y-axis, and so, intuitively, our elliptic curve consists of all points present on the affine plane plus a point that is infinitely far away along the vertical direction. We consider this point to be a rational point. In fact, we shall make it the identity element of our group. Now, we comment on the symmetries in the affine curve. Note that (2.6) is invariant under the transformation that sends y to $-y$. Thus, we see that the curve must be symmetric about the x-axis. Finally, we state and prove a corollary of Theorem 2.7 based on which we define our group law.

Corollary 3.1. *A line intersects an elliptic curve at exactly three points. Further, if two of the three points are rational, then the third is also rational.*

Proof. The first part is a direct consequence of Theorem 2.7 because the degree of an elliptic curve is three and that of a line is one. The second part follows from the fact that the sum of the x-coordinates of the three points of intersection equals the coefficient of x^2, which is rational, and hence, if two of them are rational, then the third is also rational. □

Now, we have all the tools required to define our group operation on the set of rational points that we denote by $E(\mathbb{Q})$. Let P and Q be two points in this set. The point $P + Q$ is obtained by drawing the line joining P and Q and finding its third point of intersection with the curve, and then reflecting this point about the x-axis. The fact that this operation is well defined and closed under the set $E(\mathbb{Q})$ follows from Corollary 3.1 and the fact that the curve is symmetric about the x-axis. To add two points that are the same, we draw the line tangent to the curve at this point and take the third point of intersection of this line with the curve, and then reflect this point about the x-axis.

We now show that the tuple $(E(\mathbb{Q}), +)$ forms a group. Unlike most cases, proving that the group is associative is the hardest part and hence shall be proved last.

- Existence of identity: The identity element is the point at infinity, which we denote by O. To check this, let P be an arbitrary element of $E(\mathbb{Q})$, distinct from O. To obtain $P + O$, we draw the line through P parallel to the y-axis and find the other intersection point, say P', of the line with the curve. The point $P + O$ is the reflection of the point P' about the x-axis.

But then, as the curve is symmetric about the x-axis, P' is the reflection of P. Thus, we have $P + O = P$ to be true. If $P = O$, then $O + O$ is obtained by considering the line tangent to O, which is the line $Z = 0$. But we know that this line intersects the curve thrice at O. Further, reflection of O gives back O. Thus, we get $O + O = O$ to also be true.

- Existence of inverse: Now, the inverse of the identity element is itself. So, let P be an element on the affine plane. Then the inverse of P, say P', is the reflection of P with respect to the x-axis. This is true because the line joining P and P' is parallel to the y-axis, and hence intersects the curve at O. As a result, $P + P' = O$ is satisfied.
- Associativity: We will first prove the following theorem, due to Cayley and Bacharach.

Theorem 3.2. *Suppose C, C_1 and C_2 are three cubic curves in the projective plane such that C_1 and C_2 intersect at nine points. If C passes through eight of the nine intersection points of C_1 and C_2, then it must pass through the ninth as well.*

Proof. (This proof is based on Terence Tao's proof that can be found in [3]). The proof involves using Theorem 2.7 multiple times. We first prove that any cubic passing through eight of the nine intersection points must be a linear combination of C_1 and C_2 using proof by contradiction.

Assume that C is not a linear combination of the two other curves but passes through eight of the intersection points, say A_i where $1 \leq i \leq 8$. Suppose that four of these points are collinear. Then, the line joining these four points intersects C_1 and C_2 at four points, which contradicts Bezout's theorem unless the line is part of both C_1 and C_2. But in this case, the number of intersection points between the two curves is infinite, which contradicts our assumption about C_1 and C_2. Thus, no four of these points are collinear. A similar arguement can be given as to why no seven of these points lie on a quadric. We are now restricted to three possible arrangements of points, that we treat seperately.

- Case 1: There exist three points, say A_1, A_2 and A_3, that lie on the same line, that we denote by l.
  Then, there exists a unique quadric passing through the remaining five points, say e. Existence follows from the fact that the number of parameters that must be specified to describe a quadric is five. Uniqueness follows from Theorem 2.7, which says that two quadrics can intersect at most at four points. Now, let P be a point lying on the line l, distinct from the A_is. Let Q be a point that lies neither on l nor e. Suppose p_1, p_2 and p_3 and q_1, q_2 and q_3 denote the values the cubics C_1, C_2 and C take at the points P and Q respectively. The system of equations
  $$p_1x + p_2y + p_3z = 0$$
  $$q_1x + q_2y + q_3z = 0$$
  has a non-trivial solution as the columns of the corresponding coefficient matrix are not linearly independent. This is true because the number of rows is less than the number of columns. Thus, there exists a non-trivial combination $D = kC_1 + lC_2 + mC$ such that D vanishes

at the points P and Q. Further, D is not the constant zero function because we assumed C to not be a linear combination of C_1 and C_2. Now, as l intersects D at four points, namely A_1, A_2, A_3 and P, Bezout's theorem forces D to contain the line. Thus, D is the cubic containing the line l and the quadric passing through the remaining five points, which is precisely e. But then, D also vanishes at Q which lies neither on l nor e, which is a contradiction.

- Case 2: There exist six points among the $A_i s$ that lie on a quadric. The arguement as to why this case leads to a contradiction is very similar to the previous case and hence shall be ommitted.
- Case 3: No three points among the $A_i s$ are collinear and no six points lie on a quadric.
  Now, let l denote the line passing through A_1 and A_2. Let e denote the conic passing through A_3, A_4, A_5, A_6 and A_7. As no three points are collinear and no six lie on the same quadric, none of the $A_i s$ lie on both l and e. Further, A_8 does not lie on either of the two. Now, let P and Q be two points lying on l, that are distinct from A_1 and A_2. We can find a non-zero curve $D = kC_1 + lC_2 + mC$ such that D vanishes at P and Q. D contains the line l and the conic e. But then, as A_8 is not part of either of the two, we arrive at a contradiction.

We got a contradiction for all possible arrangements of $A_i s$. Thus, we must have C to be a linear combination of C_1 and C_2. Now, as C_1 and C_2 vanish at A_9, C also vanishes at A_9. □

We now prove associativity. Let P, Q and R be three arbitrary points belonging to $E(\mathbb{Q})$. Consider the points P, Q, R, $P+Q$, $(P+Q)^{-1}$, $Q+R$, $(Q+R)^{-1}$, $((P+Q)+R)^{-1}$ and $(P+(Q+R))^{-1}$(see figure 3). If any of these elements are equal to the identity, associativity follows immediately. Thus, we may assume that none of these points correspond to O. Now, note that there exists a line, say l_1, passing through P, Q and $(P+Q)^{-1}$. Let l_2 denote the line passing through Q, R and $(Q+R)^{-1}$, l_3 the line passing though P, $Q+R$ and $(P+(Q+R))^{-1}$, l_4 the line passing through $P+Q$, R and $((P+Q)+R)^{-1}$, l_5 the line through O, $Q+R$ and $(Q+R)^{-1}$ and l_6 the line passing through O, $P+Q$ and $(P+Q)^{-1}$. Recall that a set of three distinct lines form a cubic. Thus, let C denote the cubic consisting of the lines l_1, l_4 and l_5. Let D be the cubic containing the lines l_2, l_3 and l_6. Observe that C, D and E (the elliptic curve on which we have defined our group structure) intersect at eight points. Thus, using Theorem 3.2 we see that the three curves must meet at nine points. Hence, the points $((P+Q)+R)^{-1}$ and $(P+(Q+R))^{-1}$ must coincide. From this, we find that $((P+Q)+R) = (P+(Q+R))$ holds.

Note that the group is abelian. This follows from the fact that the line passing through two points P and Q is same as the line passing through Q and P.
In this section, we gave a geometric way of operating two points in the group. In the next section, we will give an explicit description of the group operation in terms of the coordinates of the points of the group.

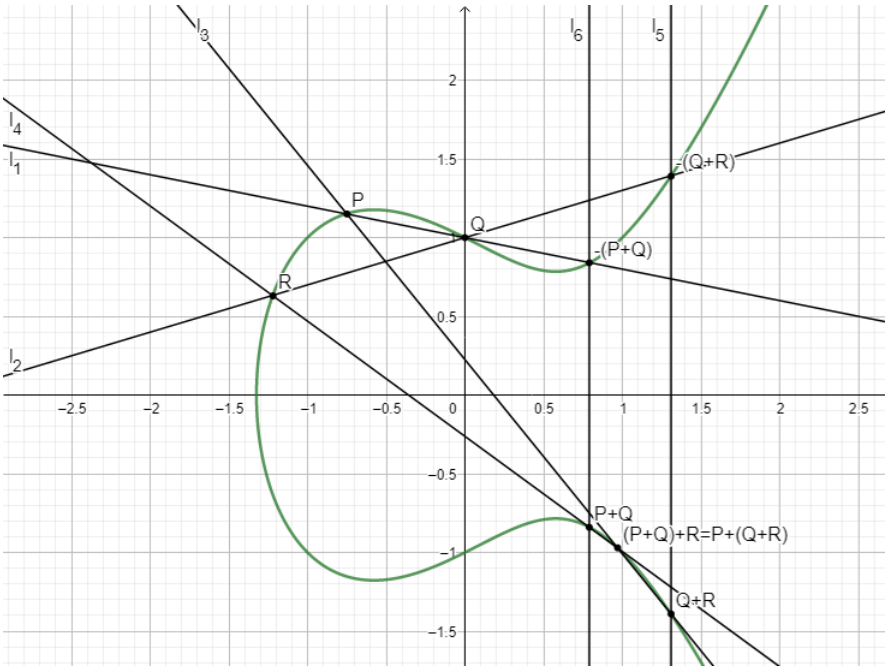

FIGURE 3. Graphical representation of the proof of associativity of the group operation

4. Some properties of the group $E(\mathbb{Q})$

Lemma 4.1. *Let P, Q and R be three points in $E(\mathbb{Q})$. Then $P+Q+R=O$ if and only if P, Q and R are collinear.*

Proof. Suppose that $P+Q+R=O$ holds. Then we get $R=(P+Q)^{-1}$ to be true. From the way we defined the group operation in the previous section, we see that P, Q and $(P+Q)^{-1}$ lie on the same line. Hence, P, Q and R are collinear. Conversely, assume that P, Q and R are collinear. From Theorem 2.7, we see that the line through P and Q intersects the curve E at exactly one more point. Thus R must coincide with $(P+Q)^{-1}$. As a result, we get $P+Q+R=O$. □

We now prove a lemma that restricts the number of possibilities for the coordinates of the elements of $E(\mathbb{Q})$.

Lemma 4.2. *Let P be a point in $E(\mathbb{Q})$ distinct from O. Then the coordinates of P are of the form $(\frac{k}{j^2}, \frac{l}{j^3})$ where k, l and j are integers with $gcd(k,j)=1$ and $gcd(l,j)=1$.*

Proof. Suppose the coordinates of the point P are $(\frac{k}{n}, \frac{l}{t})$ where k, n, l, and t are integers with $gcd(k,n)=1=gcd(l,t)$. By substituting these values in (2.6) and multiplying with t^2n^3 throughout, we get

$$l^2n^3 = k^3t^2 + ak^2nt^2 + bkn^2t^2 + cn^3t^2 \tag{4.3}$$

Thus, t^2 divides l^2n^3. As $gcd(l,t)=1$, t^2 must divide n^3. Similarly, we arrive at the conclusion that n divides t^2. By using this fact back in (4.3), we see that n^2

divides t^2 and thus n divides t. By using this fact in (4.3) once again, we conclude that n^3 divides t^2. As n^3 divides t^2 and t^2 divides n^3, we get $n^3 = t^2$. Now, if we denote the quotient $\frac{t}{n}$ by j and write x and y in terms of k, l and j, we get x and y in the desired form. □

We now provide explicit formulae for calculating the coordinates of the sum of two points, in terms of the coordinates of the points and the coefficients of the curve.

Lemma 4.4. *Suppose $P = (k, l)$ and $Q = (m, n)$ are distinct points of $E(\mathbb{Q})$. Further assume that $P + Q \neq O$. Then, the coordinates of $P + Q$ are given by the equations*

$$x = \lambda^2 - a - k - m \qquad y = -(\lambda x + p) \tag{4.5}$$

where $\lambda = \frac{n-l}{m-k}$ and $p = l - \lambda k$ hold.

Proof. Consider the line passing through P and Q. Let us denote the slope of this line by λ. Then we have $\lambda = \frac{n-l}{m-k}$. Suppose the equation of the line is $y = \lambda x + p$. Substituting the coordinates of P in the equation of the line gives $p = l - \lambda k$. Now, the intersection of this line with the curve E is given by

$$(\lambda x + p)^2 = x^3 + ax^2 + bx + c$$

By bringing all the terms to one side, we get

$$x^3 + (a - \lambda^2)x^2 + (b - 2\lambda p)x + (c - p^2) = 0 \tag{4.6}$$

Now, the sum of the roots of (4.6) is equal to $\lambda^2 - a$. Note that the roots of (4.6) are the x-coordinates of the intersection points of the curve with the line. Thus, two of its roots are k and m. As a result, the third root is given by the expression $\lambda^2 - a - m - k$. This is the x-coordinate of the inverse of $P + Q$, which is equal to the x-coordinate of $P + Q$, because inverses are obtained by reflecting about the x-axis. The y-coordinate of $(P + Q)^{-1}$ is obtained using the equation of the line. Hence, it is given by the expression $\lambda x + p$. As reflection about the x-axis changes the sign of the y-coordinate, we see that the y-coordinate of $P + Q$ is $-(\lambda x + p)$. □

Lemma 4.7. *Let $P = (k, l)$ be an element of $E(\mathbb{Q})$. Assume $P + P \neq O$. Then the coordinates of $2P$ are given by the equations*

$$x = \lambda^2 - a - 2k \qquad y = -(\lambda x + p) \tag{4.8}$$

where $\lambda = \frac{f'(k)}{2l}$ and $p = l - \lambda k$ hold. Note that $f(x)$ denotes the polynomial on the right-hand side of (2.6).

Proof. The proof is similar to the case when P and Q were distinct, except for the part of finding the slope of the line. Thus, we will only prove $\lambda = \frac{f'(k)}{2l}$ holds. Now, recall that the point $P + P$ is obtained by drawing the line tangent to the curve at P. Thus, the slope of the required line is equal to the derivative of the curve at P. Now, differentiating both sides of (2.6) with respect to x gives

$$2y\frac{dy}{dx} = f'(x) \tag{4.9}$$

From (4.9), we see that $\lambda = \frac{f'(k)}{2l}$ holds. □

We are now ready to formally state Mordell's theorem.

Theorem 4.10. *The group of rational points on an elliptic curve is finitely generated.*

The rest of the paper concentrates on the proof of the theorem.

5. A subgroup with finite index

The proof of Mordell's theorem is divided into two parts. First, we prove a particular subset of $E(\mathbb{Q})$ has finite index. Then, we define a function to provide an ordering on the set of points and use the first fact to prove the theorem. The first part of the theorem involves a few subtleties and is the tough part while the second part involves proving some inequalities. The intuition behind the proof might not seem obvious at the beginning. There is a theorem that says that all elliptic curves are isomorphic to the complex torus. If one works on the complex torus, this proof can be obtained more naturally. But we will not be going into this any further. We now state and prove the following lemma.

Lemma 5.1. *The subgroup $2E(\mathbb{Q})$, defined to be the set of all points that can be written as $P + P$ for some P belonging to $E(\mathbb{Q})$, has finite index in $E(\mathbb{Q})$.*

We introduce a useful homomorphism to prove this lemma. Before that, we make an additional assumption on the elliptic curve E. We assume that the polynomial on the right-hand side of (2.6) has a rational root. This is the same as saying that the group $E(\mathbb{Q})$ has at least one point lying on the x-axis. Let us call this point T. Now, we make a parallel transport along the x-axis that sends T to the origin. This transformation modifies (2.6) to give an equation of the form

$$y^2 = x^3 + ax^2 + bx \tag{5.2}$$

Recall that a and b are rationals. Suppose d is the lcm of their denominators. Then, we multiply (5.2) with d^6 throughout and replace x with x/d^2 and y with y/d^3 to get an equation in the form of (5.2), but with integers as coefficients. Thus, whenever we use (5.2), we may assume a and b are integers.

We remark, without elaborating further, that the proof can be extended to any elliptic curve, by working in the field generated by a root of the right-hand side of (2.6) over the rationals.

Now, suppose C is the group of rational points of the curve defined by (5.2), for some integers a and b. Then, define $\bar{C}$ to be the group of rational points for the equation

$$y^2 = x^3 + \bar{a}x^2 + \bar{b}x \tag{5.3}$$

where $\bar{a} = -2a$ and $\bar{b} = a^2 - 4b$.

Observe that the corresponding equation of $\bar{\bar{C}}$ is

$$y^2 = x^3 + 4ax^2 + 16bx \tag{5.4}$$

which can be brought to the form of C by substituting x with $4x$ and y with $8y$. Thus, we see that C is isomorphic to $\bar{\bar{C}}$.

Consider the map $\phi : C \to \bar{C}$ defined by

$$\phi(x, y) = \begin{cases} (\frac{y^2}{x^2}, y(\frac{x^2 - b}{x^2})) & x \neq 0 \\ O & x = 0 \end{cases}$$

The only points in C with x-coordinate zero are the points T and O. Note that Theorem 2.7 is not violated here because the multiplicity of T in this case is two. We define a similar map

$$\bar{\phi}(\bar{x},\bar{y}) = \begin{cases} (\frac{\bar{y}^2}{\bar{x}^2}, \bar{y}(\frac{\bar{x}^2-b}{\bar{x}^2})) & \bar{x} \neq 0 \\ O & \bar{x} = 0 \end{cases}$$

As $\bar{\bar{C}} \cong C$, $\bar{\phi}$ can be interpreted as a map ψ from $\bar{C} \to C$. We now give a series of statements regarding the maps just defined. The proofs of these statements involve using the formulae derived in Section 4. It involves tedious computation and shall be skipped. The reader can find the proof in Section 3.4 of [1].

Proposition 5.5. *The maps ϕ and ψ are well-defined homomorphisms with kernels $\{O,T\}$ and $\{\bar{O},\bar{T}\}$ respectively. The composition map $\psi \circ \phi$ corresponds to the multiplication by two map, which is defined as the map that sends P to $P+P$.*

Proposition 5.5 describes the nature of the map and its kernel. We now try to give a good description of the image of ϕ. First of all, we know that $\bar{O} \in Im(\phi)$ as $\phi(O) = \bar{O}$. The origin $\bar{T}$ belongs to $Im(\phi)$ if there exists a point in C with y-coordinate zero apart from T. For y to be zero, we must have $x^2 + ax + b = 0$. This implies that $\bar{T} \in Im(\phi)$ if and only if $a^2 - 4b$ is a perfect square. We now state and prove a proposition that gives a necessary and sufficient condition for a general point in $\bar{C}$ to be in $Im(\phi)$.

Proposition 5.6. *Let $\bar{P} = (\bar{x},\bar{y})$ be a point in $\bar{C}$ with $\bar{x} \neq 0$. Then $\bar{P}$ belongs to $Im(\phi)$ if and only if $\bar{x}$ is the square of a rational number.*

Proof. Suppose $\bar{P} \in Im(\phi)$. Then $\bar{x}$ is equal to $\frac{y^2}{x^2}$ for some (x,y) belonging to C. Thus, it follows that $\bar{x}$ is a perfect square. Conversely, suppose that we have $\bar{x} = p^2$ for some rational p. Consider the point P with coordinates

$$x = \frac{1}{2}(p^2 - a + \frac{\bar{y}}{p}) \qquad y = xp$$

It can be verified using direct substitution that P belongs to C, and $\phi(P) = \bar{P}$. Thus, we see that $\bar{P}$ belongs to $Im(P)$. □

This description of the image of ϕ should be sufficient to work with. Note that the image of ψ is also similar to this.

We now prove that the index of $Im(\psi)$ in C is finite. A similar arguement would work to prove that the index of $Im(\phi)$ in $\bar{C}$ is also finite. To prove the index is finite, we define a homomorphism from C to some group such that the kernel of the homomorphism is $Im(\psi)$ and the image is finite. To find such a homomorphism, we only need to look at the defining property of elements in $Im(\psi)$, that they contain all the points with x-coordinate that is the square of a rational number. Thus, we take the codomain set of our mapping to be the quotient group of rationals with operation being multiplication, modulo the subgroup of the squares of rationals $(\mathbb{Q}^*/\mathbb{Q}^{*2})$. Define $\Gamma : C \to \mathbb{Q}^*/\mathbb{Q}^{*2}$ as follows:

$$\Gamma(P) = \begin{cases} 1 \ (mod \mathbb{Q}^{*2}) & P = O \\ b \ (mod \mathbb{Q}^{*2}) & P = T \\ x \ (mod \mathbb{Q}^{*2}) & P = (x,y),\ x \neq 0 \end{cases}$$

Proposition 5.7. *The map Γ is a homomorphism.*

Proof. First, as the inverse of a point with coordinates (x, y) is $(x, -y)$, inverses are mapped to the same element in $\mathbb{Q}^*/\mathbb{Q}^{*2}$. As every element of $\mathbb{Q}^*/\mathbb{Q}^{*2}$ is equal to its inverse, we can say that Γ maps inverses to inverses. Thus, to prove $\Gamma(P) * \Gamma(Q) = \Gamma(P+Q)$ it is enough to show that if $P+Q+R = O$, then $\Gamma(P) * \Gamma(Q) * \Gamma(R) = 1$. According to Lemma 4.1, $P + Q + R = O$ is equivalent to saying the points P, Q and R lie on a line. Let the equation of this line be $y = mx + n$ where m and n are rationals. let the coordinates of the points P, Q and R be (x_1, y_1), (x_2, y_2) and (x_3, y_3), respectively. Then, x_1, x_2 and x_3 satisfy the cubic

$$x^3 + (a - m^2)x^2 + (b - 2mn)x - n^2 = 0$$

Thus, $x_1x_2x_3 = n^2$ holds. As n is rational, we have proved our claim. □

The fact that the kernel of this homomorphism is $Im(\psi)$ is a direct consequence of Proposition 5.6. It is left to show that the image is finite.

Proposition 5.8. *The image of* Γ *is contained in the finite subset of* $\mathbb{Q}^*/\mathbb{Q}^{*2}$ *that contains all the divisors of* b.

Proof. Now, the points O and T are mapped to 1 and b respectively. Thus, the images of O and T are contained in the subset of $\mathbb{Q}^*/\mathbb{Q}^{*2}$ that contains all the divisors of b. Let us now prove the proposition for other points. For this, we write the coordinates of the points in the form given in Lemma 4.2, and substitute them in (5.2) to get

$$l^2 = k(k^2 + aj^2k + bj^4)$$

Now, if p is a prime dividing k such that p^2 does not divide k, then p divides b. Otherwise p would not divide $(k^2 + aj^2k + bj^4)$, implying that p^2 does not divide l^2, which is a contradiction because p divides l. Thus, we see that any prime appearing in the prime factorization of k with degree 1 must also divide b. Thus, k/j^2 is congruent to one of the divisors of b modulo the subgroup $\mathbb{Q}^{*2}$. As a result, we see that all the points in C are mapped to a divisor of b. □

We have shown that the image of Γ is finite. From the first isomorphism theorem, we see that $C/Im(\psi) \cong Im(\Gamma)$. Thus, $Im(\psi)$ has finite index in C. We now complete the proof of Lemma 5.1 by using this fact and the fact that $Im(\phi)$ has finite index in $\bar{C}$.

Let $\{P_1, P_2, ...P_k\}$ be the representatives of distinct cosets of $Im(\psi)$ in C and $\{\bar{P_1}, \bar{P_2}, ...\bar{P_l}\}$ be the representatives of distinct cosets of $Im(\phi)$ in $\bar{C}$. Let P be an aribitrary point in C. Now, there exists an i such that $1 \leq i \leq k$ holds and $P - P_i$ belongs to $Im(\psi)$. Let $P - P_i = \psi(\bar{P})$ for some $\bar{P}$ in $\bar{C}$. We can find a j such that $1 \leq j \leq l$ holds and $\bar{P} - \bar{P_j} = \phi(Q)$ holds for some Q in C. Combining the two equations gives us $P = P_i + \psi(\bar{P_j}) + 2Q$ to be true. Thus, we see that any point in C can be written as the sum of an element in $2C$, P_i and $\psi(\bar{P_j})$ for some i and j satisfying the inequalities $1 \leq i \leq k$ and $1 \leq j \leq l$. Thus, the index of $2C$ is at most kl in C, and hence finite. With this, we have completed the proof of Lemma 5.1.

6. Heights of the elements of $E(\mathbb{Q})$

The proof of Mordell's theorem is essentially a proof by descent. Thus, we must obtain an ordering on the elements of the group. We do this by associating with

every element, an integer. Let h be a map from $E(\mathbb{Q})$ to the set of positive integers such that

$$h((\frac{m}{n}, y)) = \begin{cases} \log(max\{|m|, |n|\}) & P \neq O \\ 1 & P = O \end{cases}$$

We call $h(P)$ the *height* of P. Note that the x-coordinate must be written in its lowest terms, that is, $gcd(m, n) = 1$. If the x-coordinate is zero, we map the point to 0 by convention. We now state an important property of the height.

Lemma 6.1. *If M is a fixed positive integer, then the set $A = \{P \in E(\mathbb{Q}) \mid h(P) \leq M\}$ has finite cardinality.*

Proof. From the way we defined $h(P)$, we see that $h(P) \leq M$ implies $\log(max\{|m|, |n|\}) \leq M$, which in turn gives $max\{|m|, |n|\} \leq e^M$. Observe that there is a finite number of positive integers less than or equal to e^M. Thus, the number of possibilites for the numerator and denominator of a rational number $\frac{m}{n}$ satisfying $max\{|m|, |n|\} \leq e^M$ is also finite. As a result, we see that the number of possibilities for the x-coordinate of P is finite. Now, for a fixed x, there are at most two possibilities for the y-coordinate. This follows from the fact that P must satisfy the equation of the curve $y^2 = f(x)$. Thus, there is a finite number of possible coordinates for P. □

Let us recall that the main idea in the final proof is proof by descent. Thus, from Lemma 6.1 we see that if we are able to bring all the elements of $E(\mathbb{Q})$ to some element in A in a finite number of steps, we will be able to prove that the group is finitely generated. We now state and prove two important inequalities to help us find a way to reduce systematically the heights of elements of the group $E(\mathbb{Q})$. The proofs involve a lot of computation.

Lemma 6.2. *Let us fix an arbitrary point $P \in E(\mathbb{Q})$. Then there exists a constant k that depends on P such that*

$$h(Q + P) \leq 2h(Q) + k \tag{6.3}$$

for all $Q \in E(\mathbb{Q})$.

Proof. Suppose that $P = O$. Then, we have $h(Q + O) = h(Q) \leq 2h(Q)$ to be true. Thus, we can take the constant k to be zero for this case. Now assume that $P \neq O$. Let the coordinates of P be (x_0, y_0). Let Q be an arbitrary element in $E(\mathbb{Q})$ with coordinates (x, y). We further assume that Q is distinct from P, $-P$ and O. Using Lemma 4.4, we see that the x-coordinate of $P + Q$ is given by $\lambda^2 - a - x - x_0$, where $\lambda = \frac{y-y_0}{x-x_0}$. By expanding λ, we get

$$\frac{(y - y_0)^2 - (x - x_0)^2(x + x_0 + a)}{(x - x_0)^2} \tag{6.4}$$

By expanding the terms and replacing y^2 with $x^3 + ax^2 + bx + c$, we get the x-coordinate of $P + Q$ to be in the following form

$$\frac{Ay + Bx^2 + Cx + D}{Ex^2 + Fx + G} \tag{6.5}$$

where A, B, C, D, E, F and G are rational numbers that depend on a, b, c, x_0 and y_0. We can multiply all these coefficients by the lcm of their common denominators to make them integers; hence we may assume them to be integers. Recall Lemma

4.2, which states that every rational point on an elliptic curve has coordinates of the form $(m/e^2, n/e^3)$, where e, m and n are integers with e coprime to m and n. By using these expressions in place of x and y in (6.5), we get

$$\frac{Ane + Bm^2 + Cme^2 + De^4}{Em^2 + Fme^2 + Ge^4} \tag{6.6}$$

Define H to be the map $\exp h$. Observe that the inequalities $H(Q) \geq |m|$ and $H(Q) \geq |e^2|$ hold. We now prove that there exists a constant K such that $n \leq Kh(Q)^{3/2}$. From (2.6), we get

$$n^2 = m^3 + ae^2m^2 + be^4m + ce^6 \tag{6.7}$$

As $H(Q) \geq |m|$ and $H(Q) \geq |e^2|$ hold, we find that

$$n^2 \leq H(Q)^3 + aH(Q)^3 + bH(Q)^3 + cH(Q)^3 \tag{6.8}$$

Now, by taking absolute values of all the terms and using the triangular identity, we get

$$|n^2| \leq |H(Q)^3|(1 + |a| + |b| + |c|) \tag{6.9}$$

Thus, by setting $K = \sqrt{1 + |a| + |b| + |c|}$, we get the required inequality.
Now, using the triangular inequality on the numerator and denominator of (6.6), we get the inequalities

$$|Ane + Bm^2 + Cme^2 + De^4| \leq |Ane| + |Bm^2| + |Cme^2| + |De^4|$$

$$|Em^2 + Fme^2 + Ge^4| \leq |Em^2| + |Fme^2| + |Ge^4|$$

By using the bounds on m, n and e we just proved, we further get

$$|Ane + Bm^2 + Cme^2 + De^4| \leq H(Q)^2(|A| + |B| + |C| + |D|)$$

$$|Em^2 + Fme^2 + Ge^4| \leq H(Q)^2(|E| + |F| + |G|)$$

Thus, we see that

$$H(P + Q) \leq max\{(|A| + |B| + |C| + |D|), (|E| + |F| + |G|)\}H(Q)^2$$

If we denote k to be $max\{(|A|+|B|+|C|+|D|), (|E|+|F|+|G|)\}$ and take logarithm on both sides, we get (6.3). Recall that we assumed Q to be distinct from P, $-P$ and O. Let k_1, k_2 and k_3 be numbers satisfying the following inequalities.

$$h(P + P) \leq 2h(P) + k_1$$

$$h(P + -P) \leq 2h(-P) + k_2$$

$$h(P + O) \leq 2h(O) + k_3$$

To ensure that these points also satisfy (6.3), we redefine the constant k to be the maximum of k, k_1, k_2, and k_3. We thus see that for any Q belonging to $E(\mathbb{Q})$, (6.3) is satisfied. □

Lemma 6.10. *There exists a constant k such that*

$$h(2P) \geq 4h(P) - k \tag{6.11}$$

holds for all $P \in E(\mathbb{Q})$.

Proof. Let $P = (x, y)$ be an arbitrary element of $E(\mathbb{Q})$. Using Lemma 4.7, we see that the x-coordinate of $2P$ is equal to $\lambda^2 - a - 2x$ where $\lambda = \frac{f'(x)}{2y}$. By expanding λ, we get the x-coordinate of $2P$ to be equal to

$$\frac{f'(x)^2 - 4f(x)(2x + a)}{4f(x)} \tag{6.12}$$

We denote the numerator of (6.12) by $\psi(x)$ and the denominator by $\phi(x)$. Note that $\psi(x)$ is a polynomial of degree four, while $\phi(x)$ is of degree three. Observe that in (6.11), we are required to prove $H(2P)$ is greater than some quantity. Thus, unlike in Lemma 6.2, we cannot ignore the fact that $\psi(x)$ and $\phi(x)$ might have common factors. To resolve this problem, we state the following lemma.

Lemma 6.13. *Let $h(x)$ and $g(x)$ be polynomials with integer coefficients and no common roots. Suppose that d is the maximum of the degrees of the two polynomials. Then, there exists an integer Z such that $gcd(n^d h(m/n), n^d g(m/n))$ divides Z for all rationals $m/n \in \mathbb{Q}$.*

The proof of this lemma involves ring theory. We will not prove it here. Interested readers can find a proof of this lemma in Section 3.3 of [1]. Before applying Lemma 6.13 to our case, we first check whether it satisfies all the necessary conditions. We can make the coefficients of the polynomials $\phi(x)$ and $\psi(x)$ integers by multiplying all the rational coefficients with the lcm of their denominators. Further, as there are no singular points on the curve, $f(x)$ and $f'(x)$ do not share a common root. Thus, $\psi(x)$ and $\phi(x)$ also do not share a common root. Hence, Lemma 6.13 is applicable in this case. Let Z be such that $gcd(n^4\psi(m/n), (n^4\phi(m/n))$ divides Z for all $m/n \in \mathbb{Q}$. Suppose that we write $x = \frac{x_1}{x_0}$, where x_0 and x_1 are integers that are coprime to each other. Then, we have

$$gcd(x_0^4\psi(x), (x_0^4\phi(x))|Z$$

By writing the expression of the x-coordinate of $2P$ in the form $\frac{x_0^4\psi(x)}{x_0^4\phi(x)}$, we get

$$H(2P) \geq \frac{1}{Z} max\{|x_0^4\psi(x)|, |x_0^4\phi(x)|\} \tag{6.14}$$

Using the trivial inequality $max\{a, b\} \geq \frac{1}{2}(a + b)$ in (5.14), we get

$$H(2P) \geq \frac{1}{2Z}(|x_0^4\psi(x)| + |x_0^4\phi(x)|) \tag{6.15}$$

Recall that (6.11) compares $H(2P)$ and $H(P)$. Thus, we divide both sides of (6.15) by $H(P)^4 = max\{|x_1|^4, |x_0|^4\}$ to get

$$\frac{H(2P)}{H(P)^4} \geq \frac{1}{2Z} \frac{|x_0^4\psi(x)| + |x_0^4\phi(x)|}{max\{|x_1|^4, |x_0|^4\}} \tag{6.16}$$

We now wish to find a non-zero lower bound for the function

$$t(x) = \frac{|x_0^4\psi(x)| + |x_0^4\phi(x)|}{max\{|x_1|^4, |x_0|^4\}} = \frac{|\psi(x)| + |\phi(x)|}{max\{|x|^4, 1\}}$$

Observe that $t(x)$ is a continuous function. It is trivial for all points except 1 and -1. At 1 and -1, it can be checked that the left-hand limit equals the right-hand limit and thus, t is continuous at all points. Now, $t(x)$ never takes the value zero as $\phi(x)$ and $\psi(x)$ do not share a common root. Further, as the numerator of $t(x)$ is a polynomial of degree four and the denominator is also a polynomial of degree four

for large values of $|x|$, we see that the limits of $t(x)$ as x tends to infinity and minus infinity are finite and non-zero. To be precise, the limit is equal to the leading coefficient of $\psi(x)$, which is non-zero. Thus, we can find a positive real number W' such that every x in $\mathbb{R}$, barring a closed interval say I, satisfies the inequality $t(x) \geq W'$. As $t(x)$ is continuous, it assumes a minimum value, say W'', in I as I is closed. Therefore by setting $W = min\{W', W''\}$, we get

$$t(x) \geq W$$

for all $x \in \mathbb{R}$. Thus, we get

$$\frac{H(2P)}{H(P)^4} \geq \frac{W}{2Z} \tag{6.17}$$

By multiplying $H(P)^4$ on both sides of (6.17) and applying log, we get

$$h(2P) \geq 4h(P) + \log(\frac{W}{2Z}) \tag{6.18}$$

If we denote k to be $-\log(\frac{W}{2Z})$, we get (6.11). Recall that we assumed $2P$ to not be O. Now, observe that any P that has order two lies on the x-axis. Hence, there can be at most three such points. In order to include these points in (6.11), we find real numbers, say k_1, k_2 and k_3, satisfying (6.11) for each of these points. Then, we redefine k to be the minimum of the numbers k, k_1, k_2 and k_3. Then, we get (6.11) to be true for all P belonging to $E(\mathbb{Q})$. □

We now have all the tools needed to prove Mordell's theorem. We will prove it in the next section.

7. Proof of Mordell's theorem

Let P be an arbitrary element of $E(\mathbb{Q})$. Lemma 5.1 says that the subgroup $2E(\mathbb{Q})$ has finite index in $E(\mathbb{Q})$. We denote this index by l. Let Q_1, Q_2, ... Q_l denote the representatives of the distinct cosets of $2E(\mathbb{Q})$ in $E(\mathbb{Q})$. Then, there exists an i_0 such that $P - Q_{i_0} = 2P_1$ holds, for some $P_1 \in E(\mathbb{Q})$. Similarly, we can find an i_1 such that $P_1 - Q_{i_1} = 2P_2$ holds, for some $P_2 \in E(\mathbb{Q})$. In general, we can find an i_j such that $P_j - Q_{i_j} = 2P_{j+1}$ holds, for some $P_{j+1} \in E(\mathbb{Q})$. Our goal now is to find an upper bound M such that for any element P in $E(\mathbb{Q})$, we can find an $n \in \mathbb{N}$ such that $h(P_n) \leq M$.

Now, for all i such that $1 \leq i \leq l$ holds, let the constant k_i be such that the inequality

$$h(-Q_i + Q) \leq 2h(Q) + k_i$$

holds, for all $Q \in E(\mathbb{Q})$. Lemma 6.2 confirms the existence of such constants. Define k to be the maximum of the $k_i s$. Then, we have

$$h(Q_i + Q) \leq 2h(Q) + k \qquad 1 \leq i \leq l \tag{7.1}$$

Further, let t be such that

$$h(2Q) \geq 4h(Q) - t \tag{7.2}$$

holds, for all $Q \in E(\mathbb{Q})$. Lemma 6.10 confirms the existence of such a constant t. Now, $P_j - Q_{i_j} = 2P_{j+1}$ implies $h(P_j - Q_{i_j}) = h(2P_{j+1})$ holds. Thus, we must have

$$h(2P_{j+1}) \leq 2h(P_j) + k$$

to hold from (7.1). Further, we get

$$4h(P_{j+1}) - t \leq 2h(P_j) + k$$

to be true using (7.2). Thus, we have

$$h(P_{j+1}) \leq \frac{3}{4}h(P_j) - \frac{1}{4}(h(P_j) - (k+t)) \tag{7.3}$$

Now, if $h(P_j) \geq (k+t)$ holds, then $h(P_{j+1}) \leq \frac{3}{4}h(P_j)$ holds. Thus, as long as $h(P_j) \geq (k+t)$ holds, the sequence $h(P)$, $h(P_1)$, $h(P_2)$, ... goes to zero, as each term is less than or equal to the terms of the sequence $h(P)$, $\frac{3}{4}h(P)$, $\frac{3^2}{4^2}h(P)$, ..., which clearly goes to zero. Thus, at some point we get $h(P_j) \leq (t+k)$. Now, we have shown that for any given $P \in E(\mathbb{Q})$, we can find an n such that P_n belongs to the set $A = \{P \in E(\mathbb{Q}) \mid h(P) \leq (k+t)\}$. Note that P can be derived from P_n using the following relation

$$P = Q_{i_1} + 2Q_{i_2} + 2^2Q_{i_3} + \cdots + 2^{n-1}Q_n + P_n \tag{7.4}$$

Thus, we see that the point P belongs to the set generated by the the elements of $B = A \cup \{Q_1, Q_2, \ldots, Q_l\}$. As P is an arbitrary element of $E(\mathbb{Q})$, and k and t are constants independent of P, we find that set B generates whole of $E(\mathbb{Q})$. Lemma 6.1 tells us A is finite. Thus, B is finite. Hence, we have proved Mordell's theorem.

Acknowledgments

I would like to thank my mentor, Bingjin Liu, for her guidance throughout the project and for explaining the intuition behind various steps of the proof. Her support is greatly appreciated. I would also like to thank Peter May for giving me the opportunity to take part in the Apprentice Program of Research Experience for Undergraduates (REU), 2018, at the University of Chicago.

References

[1] Joseph H. Silverman and John Tate. Rational Points on Elliptic Curves. Springer.
[2] Alvaro Lozano-Robledo. Elliptic Curves, Modular Forms and Their L-functions.
[3] Terence Tao. https://terrytao.wordpress.com/2011/07/15/pappuss-theorem-and-elliptic-curves/