SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Points of finite order on an elliptic curve over the rational numbers

av

Andreas Adamsson

2018 - No K18

Points of finite order on an elliptic curve over the rational numbers

Andreas Adamsson

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Wushi Goldring

2018

Points of finite order on an elliptic curve over the rational numbers

Andreas Adamsson

May 29, 2018

Abstract

This thesis is aimed to serve as an introduction to the theory of rational points on elliptic curves over the rational numbers. The thesis starts by introducing fundamental concepts in the theory of projective geometry including the theorem of Bezout. Using this introduction as a theoretical framework, the paper defines what elliptic curves are and partially proves that elliptic curves coupled with the canonical binary operation makes them into abelian groups. After introducing additional terminology and theorems about elliptic curves, the remainder of the paper is dedicated to presenting a proof of the Nagell-Lutz Theorem. The Nagell-Lutz Theorem is a practical tool in finding all rational points of finite order on an elliptic curve over the rationals.

Acknowledgements

I would like to extend my thanks to my supervisor Wushi Goldring for helping me throughout the process of writing this thesis.

1 Introduction

The study of elliptic curves remains an important field of study for purely theoretical questions asked by mathematicians as well as a tool for practical problems facing people all over the world. This thesis aims to serve as an introduction to the basics of elliptic curves from a geometrical perspective with focus on showing an algebraically interesting result. The hope is that this will give a glimpse of the intriguing duality of the geometric and algebraic perspectives of the field.

Chapter two of this thesis gives a brief introduction to the theory and language of projective geometry. The aim of the chapter is to introduce enough terminology such that an uninitiated student may understand both the fundamentals of projective geometry as well as the very important theorem of Bezout.

Following the introduction of projective geometry is a chapter focusing on introducing elliptic curves. It defines the Weierstrass normal form and shows the role that it plays in the definition of elliptic curves. The canonical binary operation on points on an elliptic curve is defined and it is shown that an elliptic curve coupled with this operation forms an abelian group.

The following chapters introduces additional terminology and results that are of importance for proving the main result of this thesis, the Nagell-Lutz Theorem. This theorem is a computationally effective tool for finding all rational points of finite order on an elliptic curve over the rationals. The final chapter of this thesis includes a proof of the theorem as well as an application of how it may be used.

2 Projective geometry and Bezout's Theorem

A fundamental concept that underpins the theory of elliptic curves is projective geometry. In this chapter, a very brief introduction to the subject is introduced culminating in stating Bezout's Theorem as well as the Cayley-Bacharach Theorem.

Definition 2.1. Let $\sim$ be an equivalence relation on the set of triplets in $\mathbb{C}^3$ defined by the following rule:

$$[a,b,c] \sim [a',b',c'] \text{ if there is a non-zero } t \in \mathbb{C} \text{ so that } a = ta', b = tb' \text{ and } c = tc'.$$

△

Definition 2.2. *The projective plane*, denoted $\mathbb{P}^2$, is defined as

$$\mathbb{P}^2 = \{[a,b,c] \in \mathbb{C}^3 \backslash \{[0,0,0]\}\}/\sim$$

For a point $[a,b,c] \in \mathbb{P}^2$, the variables a,b,c are called the *homogeneous coordinates* for the point $[a,b,c]$. △

That is, two points $[a,b,c],[a',b',c'] \in \mathbb{P}^2$ are equal if there is some non-zero variable $t \in \mathbb{C}$ such that $a = a't, b = b't$ and $c = c't$. Hence, $\mathbb{P}^2$ consists of all equivalence classes of complex triplets $[a,b,c]$ excluding $[0,0,0]$.

Definition 2.3. A polynomial $F(X,Y,Z)$ is a *homogeneous polynomial* of degree d if it has the property that $F(tX,tY,tZ) = t^d F(X,Y,Z)$ for some integer d. △

For instance, consider

$$F_1(X,Y,Z) = \alpha X + \beta Y + \gamma Z$$

where at least one of the variables $\alpha,\beta,\gamma \in \mathbb{C}$ is non-zero. It is easy to see that $F_1(tX,tY,tZ) = tF_1(X,Y,Z)$, which means that F_1 is a homogeneous polynomial of degree 1. The set of solutions in $\mathbb{P}^2$ to the equation $F_1(X,Y,Z) = 0$ is a *line* in the projective plane. More generally, consider the following definition:

Definition 2.4. If $F_d(X,Y,Z)$ is a homogeneous polynomial of degree d, the set of solutions in $\mathbb{P}^2$ to $F_d(X,Y,Z) = 0$ is a *curve of degree* d in the projective plane. △

As each element in $\mathbb{P}^2$ is an equivalence class of complex triplets that are all equal under $\sim$, it must follow that if $[a,b,c] \in \mathbb{P}^2$ is a solution to the equation $F(X,Y,Z) = 0$ where $F(X,Y,Z)$ is some homogeneous polynomial of degree d, then $[ta,tb,tc]$ must also be a solution for all non-zero $t \in \mathbb{C}$. This follows as $F(ta,tb,tc) = t^d F(a,b,c) = 0$ by the assumption that $[a,b,c]$ is a solution and by definition 2.3. This makes this set well defined. On the other hand, if one were to look at solutions to the equation $F(X,Y,Z) = 1$, then if $[a,b,c]$ is a solution to the equation, it follows that $[ta,tb,tc]$ is not a solution in general as $F(ta,tb,tc) = t^d F(a,b,c) = t^d$. Thus it only makes sense to talk about the solutions to equations of the type $F(X,Y,Z) = 0$.

Any polynomial $f(x,y) = \sum_{i,j} a_{i,j} x^i y^j$ that is not homogeneous can easily be *homogenized*. Take d to be the degree of f, i.e. the highest value of $i+j$ where $a_{i,j} \neq 0$. Then the homogeneous polynomial $F(X,Y,Z) = \sum_{i,j} a_{i,j} X^i Y^j Z^{d-i-j}$ is the homogenization of $f(x,y)$. Any homogeneous polynomial may also be *dehomogenized* by letting $f(x,y) = F(x,y,1)$. Viewing a polynomial in both its homogeneous and dehomogeneous form is important as there are many important properties in the projective geometry that are fundamental in the theory of elliptic curves. One such important property is Bezout's Theorem.

Example 2.5 (Homogenization and dehomogenization)**.** Consider the polynomial $f(x,y) = y^3 - x^5 - 3xy^2 - 1$. The degree of f is 5. Thus, by the procedure described the homogenization of f becomes $F(X,Y,Z) = Y^3Z^2 - X^5 - 3XY^2Z^2 - Z^5$. Conversely, dehomogenizing F with respect to Z is done by letting $f(x,y) = F(x,y,1) = y^3 - x^5 - 3xy^2 - 1$. △

Let C_1 be the curve defined by the solutions to the polynomial equation $f_1(x,y) = 0$ and C_2 defined by the solutions to the polynomial equation $f_2(x,y) = 0$, both polynomials of degree 1. In an affine space such as $\mathbb{R}^2$ one is used to the concept that two lines may have no points of intersection, that is $C_1 \cap C_2 = \emptyset$. However, if $\hat{C}_1$ and $\hat{C}_2$ are the corresponding point sets in the projective space of the homogenized polynomials $F_1(X,Y,Z)$ and $F_2(X,Y,Z)$ it can be shown that $\hat{C}_1 \cap \hat{C}_2 \neq \emptyset$ [1, p. 223]. The property will not be proven, but the idea is illustrated in the the following example.

Example 2.6 (Intersection of two lines in the projective plane)**.** Let $f_1(x,y) = y - x - 1$ and $f_2(x,y) = y - x$. The lines $C_1 : f_1(x,y) = 0$ and $C_2 : f_2(x,y) = 0$ have no points of intersection

in $\mathbb{C}$ as they are parallel. Call the homogenizations of f_1 and f_2, $F_1(X,Y,Z) = Y - X - Z$ and $F_2(X,Y,Z) = Y - X$ respectively, and let $\hat{C}_1 : F_1(X,Y,Z) = 0$ and $\hat{C}_2 : F_2(X,Y,Z) = 0$. One can see that the point $[1,1,0] \in \mathbb{P}^2$ is a solution to both lines, and hence $C_1 \cap C_2 = \emptyset$ but $\hat{C}_1 \cap \hat{C}_2 = \{[1,1,0]\} \neq \emptyset$. △

When considering the intersection of a line with a curve one may similarly find situations where the line does not intersect the curve in $\mathbb{R}^2$. For instance, the line $y = -1$ and the curve $y = x^2$ have no points of intersection in $\mathbb{R}^2$, however, it has two points of intersection in $\mathbb{C}^2$ at $(\pm i, -1)$. Hence, in the following discussion, when counting points of intersection between curves, it is the solutions in $\mathbb{C}^2$ that are counted.

Another consideration that has to be made is that a line may intersect a curve with multiplicity. This can occur when the line is a tangent to some point on the curve. Another situation when this may occur is if the line passes through a *singular point* on the curve.

Definition 2.7. A *singular point* on a curve is a point on the curve where all partial derivatives of the function vanish simultaneously. A curve is called *singular* if it contains a singular point. Conversely, a curve is called *non-singular* or *smooth* if no such points exists on the curve. △

Finally, one must also consider the case when the line is a linear factor of the cubic. The line $x - y = 0$ and the curve $x^2 - y^2 = 0$ have infinitely many points of intersection as all solutions to $x - y = 0$ are also solutions to $x^2 - y^2 = 0$. This is easily seen when the curve is factorized as $x^2 - y^2 = (x - y)(x + y) = 0$. In this case it is said that the two curves share a *component*. This idea can be generalized by the following lemma.

Lemma 2.8. [2, p. 305] *If R is a Unique Factorization Domain (UFD), then a polynomial ring in an arbitrary number of variables with coefficients in R is also a Unique Factorization Domain.* □

As $\mathbb{C}$ is a UFD, the lemma shows that both $\mathbb{C}[x,y]$ and $\mathbb{C}[X,Y,Z]$ are UFDs. By definition, any element of a UFD can be written as a finite product of irreducible elements in the UFD. Also, such a decomposition will be unique up to associates [2, p. 285]. Therefore any polynomial with coefficients in $\mathbb{C}$ can be factored into irreducibles. This allows for the following definition:

Definition 2.9. Two curves, defined as the solutions to the zeros of their corresponding polynomials, share a *component* if both polynomials share the same irreducible factor, up to associates. △

For each point $P \in \mathbb{P}^2$ and for some projective curves C_1 and C_2 with no common components, let $I(C_1 \cap C_2, P)$ be the intersection multiplicity function from $\mathbb{P}^2 \times \mathbb{P}^2$ to $\mathbb{N}$, which has the following properties [1, p. 237]:
1. If $P \notin C_1 \cap C_2$, then $I(C_1 \cap C_2, P) = 0$.
2. If $P \in C_1 \cap C_2$, if P is a non-singular point of C_1 and C_2, and if C_1 and C_2 have distinct tangents at P, then $I(C_1 \cap C_2, P) = 1$. (I.e. C_1 and C_2 intersect transversally at P)
3. If $P \in C_1 \cap C_2$ and if C_1 and C_2 do not intersect transversally at P, then $I(C_1 \cap C_2, P) \geq 2$.

These informal introductions should be enough such that one can understand the following theorem of Bezout.

Theorem 2.10 (Bezout's Theorem)**.** [1, p. 237] *Let C_1 and C_2 be projective curves with no common components. Then*

$$\sum_{P \in C_1 \cap C_2} I(C_1 \cap C_2, P) = (\deg C_1)(\deg C_2)$$

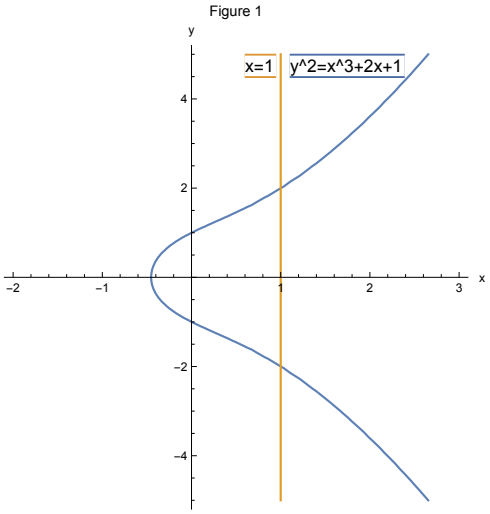
where the sum is over all points of $C_1 \cap C_2$ having complex coordinates. In particular, if C_1 and C_2 are smooth curves with only transversal intersections, then $\#(C_1 \cap C_2) = (\deg C_1)(\deg C_2)$; and in all cases there is an inequality

$$\#(C_1 \cap C_2) \leq (\deg C_1)(\deg C_2).$$

□

This is a natural generalization of the fundamental theorem of algebra, i.e. that, when counting multiplicities, a polynomial of degree n has exactly n roots. For the purposes of this thesis, this theorem will mostly be used for the idea that a line and a cubic always have three points of intersection when also considering the homogenization of the curves in the projective plane. This is illustrated in the following example.

Example 2.11 (Intersection of cubic and line)**.** Consider the curves $C_1 : y^2 = x^3 + 2x + 1$ and $C_2 : x = 1$. A section of the real points of these curves are illustrated in figure 1. By substitution of the linear expression for y into C_2, one finds that C_1 and C_2 have two points of intersection, one at the point $(1, 2)$ and one at the point $(1, -2)$.



Let $\hat{C}_1 : ZY^2 = X^3 + 2XZ^2 + Z^3$ and $\hat{C}_2 : X = Z$ be the homogenizations of C_1 and C_2 respectively. It follows from Bezout's Theorem that these two curves should have exactly three points of intersection in the projective plane. Whenever $Z = 1$, the solutions $[1, 2, 1]$ and $[1, -2, 1]$ were already found. The third and final point of intersection can be found when letting $Z = 0 (= X)$. By substitution one finds that any value of Y is a solution to the curve. This corresponds to the point $[0, 1, 0]$ in the projective plane. Therefore, when also considering this extra point $[0, 1, 0]$ and the homogenized versions of the curves, one may say that C_1 and C_2 have three points of intersection. △

In the example above it can also be seen that for any curve $x = d$, where d is any constant, this curve will have at least one point of intersection with C_1 at the point $[0, 1, 0]$ as it is clearly a point on the homogenized line equation $X = dZ$. That any curve on this form intersects C_1 at the specific point $[0, 1, 0]$ is not a coincidence, as for the particular class of curves in which C_1 belongs, this will always be the case. For that class of curves, the point $[0, 1, 0]$ is referred to as the point at infinity and often denoted $\mathcal{O}$.

The following theorem will be used for proving parts of a lemma in the next chapter. Like the theorem of Bezout it will be stated without any proof.

Theorem 2.12 (Cayley-Bacharach Theorem)**.** [1, p. 240] *Let C_1 and C_2 be curves in $\mathbb{P}^2$ without common components of respective degree d_1 and d_2, and suppose that C_1 and C_2 intersect at $d_1 d_2$ distinct points. Let D be a curve in $\mathbb{P}^2$ of degree $d_1 + d_2 - 3$. If D passes through all but one of the points of $C_1 \cap C_2$, then D must pass through the remaining point also.* □

For this thesis, the only case when this theorem will be used is when C_1 and C_2 are both of degree 3. That is, when the two curves intersect at nine distinct points and D, also of degree 3, passes through eight of the points.

3 Elliptic curves and the group law

In this chapter, a formal definition of elliptic curves is introduced. The additive operation that makes an elliptic curve into a group is also explained, and the group axioms partly proven. Before this is possible to do, a few more concepts are introduced.

Definition 3.1 (Weierstrass normal form)**.** [1, p. 22] A cubic of the form

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

for some constants $a, b, c \in \mathbb{C}$, is said to be in Weierstrass normal form. △

It is possible to show that any cubic with a rational point, i.e. a cubic containing a point (a, b) where $a, b \in \mathbb{Q}$, is *birationally equivalent* to a curve in Weierstrass normal form [1, p. 22]. Two cubics are birationally equivalent if, by some transformation, either curve can be transformed such that it coincides with the other cubic. This transformation must be of a form such that it preserves rational solutions. That is, if the transformation is applied to a rational solution on one of the cubics, then the resulting point must be a rational solution on the other cubic, and vice versa. This statement will not be proven, instead the following example will illustrate the concept.

Example 3.2 (Transformation to Weierstrass normal form)**.** Let C be the cubic $u^3+v^3 = u+v+1$, with the solution $\mathcal{O}_{uv} = [1, -1, 0]$ in the projective plane. A birational transformation will be derived such that $x' = x(u, v)$ and $y' = y(u, v)$ where x' and y' satisfies a cubic equation in Weierstrass normal form with $\mathcal{O} = [0, 1, 0]$ as a solution in the projective plane.

In order to produce this birational transformation the homogenization of C will have to be considered. The homogenization procedure yields the curve $\hat{C} = F(U, V, W) = U^3 + V^3 - UW^2 - VW^2 - W^3 = 0$. The first step will be to construct a linear transformation that maps points in the UVW-space to points in the XYZ-space. Once this is done the transformation will be modified to fit the Weierstrass normal form. Note that the tangent to the point $\mathcal{O}_{uv} = [1, -1, 0]$ is:

$$\frac{\partial F}{\partial U}(\mathcal{O}_{uv})U + \frac{\partial F}{\partial V}(\mathcal{O}_{uv})V + \frac{\partial F}{\partial W}(\mathcal{O}_{uv})W = 0.$$

This expression evaluates to the line $U+V = 0$. Under the transformation that is being constructed, this line will correspond to be the line $Z = 0$. Next, under the substitution $U = -V$ in $\hat{C}$, it follows that $U+V = 0$ intersects $\hat{C}$ with multiplicity three at $W = 0$, i.e. the tangent intersects three times at $\mathcal{O}_{uv}$. Let the line $X = 0$ be $W = 0$ and the line $Y = 0$ be $U - V = 0$ under the transformation, the transformation may then be written as the linear system:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} U \\ V \\ W \end{pmatrix}.$$

Where matrix inversion gives:

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 1 \\ 2 & 0 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

This becomes the candidate transformation. Substituting U, V and W by this linear transformation in $\hat{C}$ one gets the following expression:

$$\frac{1}{8}(Z + Y)^3 + \frac{1}{8}(Z - Y)^3 - \frac{1}{2}(Z + Y)X^2 - \frac{1}{2}(Z - Y)X^2 - X^3 = 0.$$

Which is equivalent to:

$$\frac{1}{8}(2Z^3 + 6Y^2Z) - ZX^2 - X^3 = 0.$$

Dehomogenizing the expression gives:

$$6y^2 = 8x^3 + 8x^2 - 2.$$

As this is not yet on Weierstrass normal form it is also required to multiply the expression by 6^3 and then making the change of variable $y' = 36y$ and $x' = 12x$ before the final transformation is derived. This gives the curve:

$$C_{xy} : y'^2 = x'^3 + 12x'^2 - 432.$$

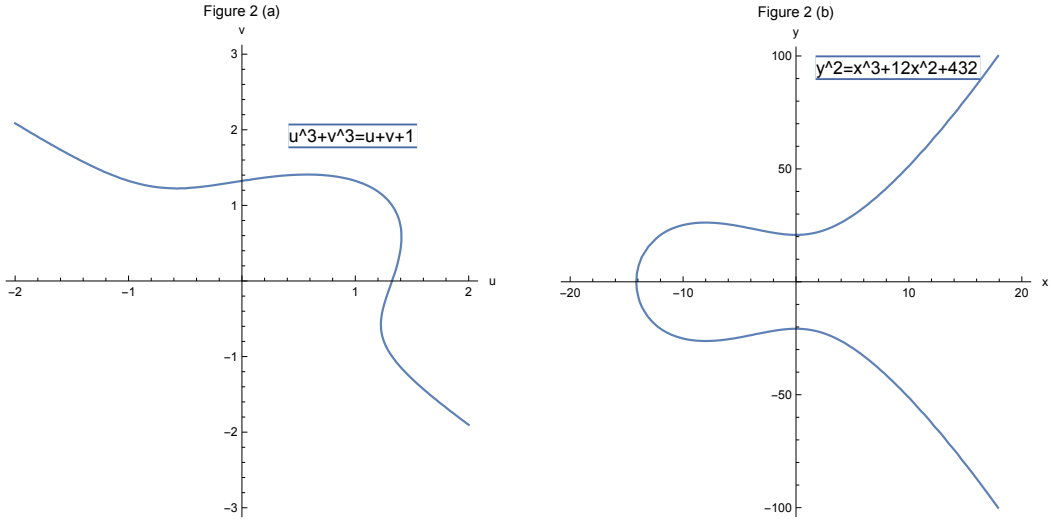
That is, the transformations

$$x' = 12x = 12\frac{X}{Z} = 12\frac{W}{U + V} = \frac{12}{u + v}$$

and

$$y' = 36y = 36\frac{Y}{Z} = 36\frac{U - V}{U + V} = 36\frac{u - v}{u + v},$$

gives an expression on the desired form. As the line $u + v = 0$ contains no solutions of C, any such points can be ignored under the transformation. As can be seen, both x' and y' are rational expressions of the points u and v, which shows that any rational point on the curve C will correspond to a rational point on C_{xy}. Hence, this is a birational transformation of C to C_{xy}. Figure 2(a) and figure 2(b) illustrates a section of the real solutions of both curves.



△

The homogenized version of a cubic on the general Weierstrass normal form becomes

$$ZY^2 = X^3 + aX^2Z + bXZ^2 + cZ^3.$$

When the cubic intersects the line $Z = 0$, the equation reduces to $X^3 = 0$. That is, it intersects $Z = 0$ with a multiplicity of three at $[0, 1, 0]$. That is, the line $Z = 0$ and the cubic has a single point of intersection, namely $[0, 1, 0]$. It also follows that that $Z = 0$ is the tangent to $[0, 1, 0]$ on the cubic. In the following lemma, an important property of this point is shown.

Lemma 3.3. *Let $F(X, Y, Z) = Y^2Z - X^3 - aX^2Z - bXZ^2 - cZ^3$ for some constants $a, b, c \in \mathbb{C}$. Let C be the cubic defined by $F(X, Y, Z) = 0$. Then the point $[0, 1, 0]$ is a non-singular point on C.*

Proof. By substitution it is trivial to see that $[0, 1, 0]$ is a point on the curve for any values of a, b, c. As was discussed in chapter 2, a non-singular point on a curve is a point on the curve where the partial derivatives do not vanish simultaneously. For this lemma it is enough to study the partial derivative with respect to Z.

$$\frac{\partial F}{\partial Z} = Y^2 - aX^2 - 2bXZ - 3cZ^2$$

Evaluating the expression at $[0, 1, 0]$ yields the value 1 regardless of the constants a, b and c. This shows that $[0, 1, 0]$ is always a non-singular point on C. □

The argument preceding lemma 3.3 also shows that on the line $Z = 0$ there is in fact only a single solution to the homogenized polynomial, i.e. the point $[0, 1, 0]$. Any solution to the homogenized polynomial where $Z \neq 0$ corresponds to a solution where $Z = 1$. This follows from the the definition of homogeneous coordinates, which states that $[X, Y, Z]$ is equivalent to $[X/Z, Y/Z, 1]$ as it is only a factor of $1/Z$ that differentiates the two coordinates. This allows for the following definition of an elliptic curve.

Definition 3.4 (Elliptic curve over $\mathbb{C}$)**.** Let

$$E(\mathbb{C}) = \{(x, y) \in \mathbb{C}^2 | y^2 = f(x) = x^3 + ax^2 + bx + c\} \cup \{\mathcal{O}\}$$

where $y^2 - f(x) = 0$ is a non-singular cubic in Weierstrass normal form and where $\mathcal{O} = [0, 1, 0]$. Then $E(\mathbb{C})$ is an elliptic curve over $\mathbb{C}$. The point $\mathcal{O}$ is referred to as the point at infinity. △

By lemma 3.3 and by definition 3.4 it follows that all points on an elliptic curve over $\mathbb{C}$ are non-singular. Also note that any cubic that is birationally equivalent to any such elliptic curve is also considered an elliptic curve [1, p. 25]. However, because of this equivalence, the study of elliptic curves can be limited to the study of curves in Weierstrass normal form.

The remainder of this chapter will focus on showing that an elliptic curve may be coupled with a binary operation so that it forms an abelian group.

Definition 3.5. For any elements $P, Q \in E(\mathbb{C})$, where $E(\mathbb{C})$ is some elliptic curve over $\mathbb{C}$. When P and Q are distinct points there exists a unique line between them. Then the third point of intersection of this line with $E(\mathbb{C})$ is denoted $P * Q$, or equivalently $Q * P$ as the line from P to Q is the same as the line from Q to P and thus intersect the same third point. If $P = Q$, then instead consider the tangent line at P and let the third point of intersection with $E(\mathbb{C})$ be denoted $P * Q$ or $Q * P$, or equivalently $P * P$ or $Q * Q$. △

By Bezout's Theorem it follows that a line in the projective plane intersects with a curve of degree 3 in the projective plane three times if one includes all special cases discussed in chapter 2. This means that there will always exist a third point of intersection, $P * Q$, that is an element in $E(\mathbb{C})$. It is important to note that $P * Q$ must not be distinct from P or Q. For instance, a line may intersect a point P on the elliptic curve with multiplicity 3, and thus $P * P = P$, this is in particular true for the point $\mathcal{O}$. This allows for the formal definition of the additive operation.

Definition 3.6 (The additive operation)**.** For any elements $P, Q \in E(\mathbb{C})$, where $E(\mathbb{C})$ is some elliptic curve over $\mathbb{C}$. Define the operation $+ : E(\mathbb{C}) \times E(\mathbb{C}) \to E(\mathbb{C})$ by $(P, Q) \mapsto (P * Q) * \mathcal{O}$. △

As was explained in definition 3.5, $P * Q \in E(\mathbb{C})$, and by definition 3.4 it follows that $\mathcal{O} \in E(\mathbb{C})$. Therefore, $(P * Q) * \mathcal{O}$ is simply the third point of intersection between the line between the two points and the elliptic curve, which will intersect the curve at a third point by Bezout's Theorem. Hence, the additive operation is well defined. It remains to show that it forms a group when coupled with an elliptic curve over $\mathbb{C}$. In order to prove that, some additional lemmas will be introduced below.

Lemma 3.7 (Commutativity)**.** *For all $P, Q \in E(\mathbb{C})$, where $E(\mathbb{C})$ is an elliptic curve over $\mathbb{C}$, the following equality holds:*

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

Proof. By definition of the additive operation, $P + Q = (P * Q) * \mathcal{O}$. On the other hand, by the definition of the $*$ operation, $P * Q = Q * P$. This means that $P + Q = (P * Q) * \mathcal{O} = (Q * P) * \mathcal{O} = Q + P$. □

Lemma 3.8 (Existence of identity element)**.** *For all $P \in E(\mathbb{C})$, where $E(\mathbb{C})$ is an elliptic curve over $\mathbb{C}$, the following holds:*

$$P + \mathcal{O} = \mathcal{O} + P = P.$$

Proof. By lemma 3.7, $P + \mathcal{O} = \mathcal{O} + P$. So it is enough to show that $P + \mathcal{O} = P$.

By definition 3.6, $P + \mathcal{O} = (P * \mathcal{O}) * \mathcal{O}$. Let $Q = P * \mathcal{O}$ and let $P \neq \mathcal{O}$. That is, Q is the third intersection with $E(\mathbb{C})$ on the unique line L_1 between P and $\mathcal{O}$. So $P + \mathcal{O} = Q * \mathcal{O}$. The line between Q and $\mathcal{O}$ also forms a unique line L_2 that intersects $E(\mathbb{C})$ at a third point. If L_1 and L_2 are the same line, then the third point of intersection must be P, which concludes the proof. Assume for contradiction that $L_1 \neq L_2$, then by Bezout's Theorem the lines should have exactly one point of intersection. However, by assumption they both share the points Q and $\mathcal{O}$. This contradicts the assumption when Q and $\mathcal{O}$ are distinct points. If $Q = \mathcal{O}$, then L_1 and L_2 intersects $E(\mathbb{C})$ at $\mathcal{O}$ with multiplicity two. This means that both lines must be tangents to $\mathcal{O}$. But the tangent at any point is unique, which is a contradiction. Hence, in both cases $L_1 \neq L_2$ leads to a contradiction.

When $P = \mathcal{O}$, it follows that $\mathcal{O} + \mathcal{O} = (\mathcal{O} * \mathcal{O}) * \mathcal{O}$. In the discussion preceding lemma 3.3 it was shown that $\mathcal{O}$ has the tangent $Z = 0$, which intersects the point with multiplicity three. Thus $\mathcal{O} * \mathcal{O} = \mathcal{O}$, which applied twice shows that $\mathcal{O} + \mathcal{O} = \mathcal{O}$. □

Lemma 3.9 (Existence of inverse)**.** *For each $P \in E(\mathbb{C})$, where $E(\mathbb{C})$ is an elliptic curve over $\mathbb{C}$, there exists an element $-P \in E(\mathbb{C})$, called an inverse, such that the following holds:*

$$P + (-P) = (-P) + P = \mathcal{O}.$$

Proof. If $P = \mathcal{O}$, then $\mathcal{O}$ is also the element $-P$, as, by the preceding lemma, $\mathcal{O} + \mathcal{O} = \mathcal{O}$.

When $P \neq \mathcal{O}$, let $P = (x', y')$ and let $y^2 = x^3 + ax^2 + bx + c$ be the equation defining the elliptic curve. Since we know that (x', y') satisfies this equation, by substitution it can easily be seen that $(x', -y')$ also satisfies the equation. The points described by these coordinates are distinct whenever $y' \neq 0$.

First consider the case when $y' \neq -y'$, call the second point $Q = (x', -y')$. The line $x = x'$ intersects both P and Q. As $\mathcal{O}$ is a solution to the homogenized line $X = x'Z$ it must also be the third and final point of intersection with the elliptic curve by Bezout's Theorem. So $P+Q = \mathcal{O}*\mathcal{O}$. The tangent of $\mathcal{O}$, $Z = 0$, only intersects the elliptic curve at $\mathcal{O}$, so $\mathcal{O}*\mathcal{O}$ must therefore correspond to the point $\mathcal{O}$. Thus, $P + Q = \mathcal{O}$, and by lemma 3.7 it follows that $P + Q = Q + P$. Thus, when $y' \neq 0$, Q satisfies the required properties of $-P$, therefore let $-P = Q$.

When $y' = 0$, consider the homogenized tangent of P, $X = x'Z$. The point $\mathcal{O}$ clearly lies on the line and by construction it intersects P with a multiplicity of two. Therefore $P + P = \mathcal{O} * \mathcal{O}$ and by the same reasoning as above, $P + P = \mathcal{O}$. So when $y' = 0$, let $-P = P$. □

Note that the lemma shows that for any point $P = (x, y)$ on an elliptic curve (when $P \neq \mathcal{O}$), $-P$ is simply the point $(x, -y)$.

Lemma 3.10 (Associativity)**.** *For all $P, Q, R \in E(\mathbb{C})$, where $E(\mathbb{C})$ is an elliptic curve over $\mathbb{C}$, the following equality holds*

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

Partial proof. This partial proof follows a discussion of Silverman and Tate [1, p. 19-20] where it is assumed that the points $\mathcal{O}, P, Q, R, P * Q, P + Q, Q * R, Q + R$ and a final point S, which is introduced in the proof, are distinct points. This is clearly a limitation in the proof as the lemma does not make any such reservations. Proving this lemma completely would require one to consider a lot of different cases. However, by introducing additional results that are not covered in this thesis, the proof can be done more efficiently, see for instance Silverman [3, p. 52, 62-63]. The case covered in this partial proof essentially corresponds to the case where three completely random points are considered.

Consider the lines L_1, L_2 and L_3 defined as follows. L_1 is the line between the points P and Q on $E(\mathbb{C})$, which by definition also contains the point $P*Q$. The line L_2 is the line between $P*Q$ and $\mathcal{O}$, which also intersects $E(\mathbb{C})$ at the point $P + Q$. Finally, L_3 is defined as the line between the points $P + Q$ and R. In addition, also consider the lines J_1, J_2 and J_3 defined similarly in the following way. J_1 is the line between Q and R, which also intersects $E(\mathbb{C})$ at $Q * R$. J_2 is the line given by the points $Q * R$ and $\mathcal{O}$ that intersects $E(\mathbb{C})$ at $Q + R$. Finally, J_3 is the line between the points P and $Q + R$.

By following the additive operation, the lines introduced are exactly the lines that would be considered when deriving the points $(P+Q)*R$ and $P*(Q+R)$ respectively. Let S be the intersection between the lines L_3 and J_3. If it can be shown that S lies on $E(\mathbb{C})$ it follows that $S = P*(Q+R)$ and $S = (P + Q) * R$. It would also follow that the line between S and $\mathcal{O}$ has a single final point of intersection on $E(\mathbb{C})$ that would correspond to both $P+(Q+R)$ and $(P+Q)+R$, which would prove the lemma given the particular assumptions.

Each line is defined as some set of points in $\mathbb{P}^2$ that evaluates a homogeneous polynomial of degree 1 to zero. Take the lines L_1, J_2 and L_3. Multiply their corresponding homogeneous polynomials so that the product forms a homogeneous polynomial of degree 3, the points in $\mathbb{P}^2$ that evaluates this polynomial to zero forms a curve of degree 3, call it C_1. It follows that any point on the lines L_1, J_2 and L_3 are also points on C_1. In fact, this makes C_1 the union of the points on L_1, J_2 and L_3 and would thus necessarily contain the points $\mathcal{O}, P, Q, R, P * Q, Q * R, P + Q, Q + R, S$. Similarly, let C_2 be constructed in the same way by the lines J_1, L_2 and J_3, then C_2 would also contain the same nine points as they are all contained in the union of these lines as well.

It is also important to note that none of the defined lines are the same. This follows from the assumption that the points considered are distinct. For instance, if one claims that L_1 and J_1 are

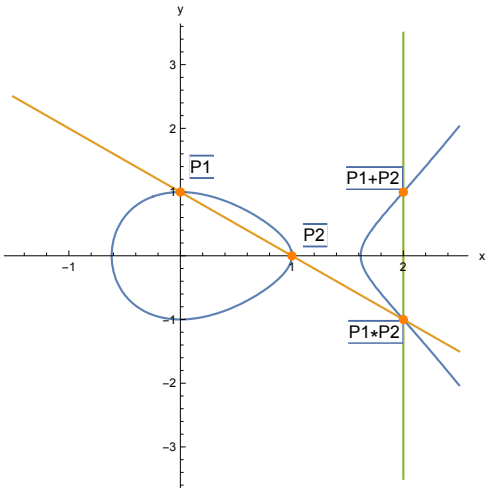
the same line, it would mean that they intersect $E(\mathbb{C})$ at the same points, but this would contradict that the points of intersection are distinct, hence the lines are different. This means that the two cubics C_1 and C_2 share nine points of intersection and have no common components. In addition, the elliptic curve $E(\mathbb{C})$ contains all points but S. However, by the Cayley-Bacharach Theorem, this means that $E(\mathbb{C})$ must also contain the ninth point of intersection, S. But then L_3 and J_3 has the same final point of intersection on the curve, i.e. $S = P * (Q + R) = (P + Q) * R$. □

Theorem 3.11. *Let $E(\mathbb{C})$ be an elliptic curve over $\mathbb{C}$ and $+$ be the additive operation defined on elliptic curves over $\mathbb{C}$. Then $(E(\mathbb{C}), +)$ is an abelian group.*

Proof. The associativity axiom follows from lemma 3.10, the existence of an identity element axiom follows from lemma 3.8, the existence of an inverse axiom follows from lemma 3.9 and the commutativity axiom follows from lemma 3.7. $(E(\mathbb{C}), +)$ therefore fulfills all the requirements for being an abelian group [2, p. 16-17]. □

Example 3.12 (Application of the group operation)**.** Let $E(\mathbb{C})$ be an elliptic curve over $\mathbb{C}$ defined by the curve $y^2 = x^3 - 2x^2 + 1$. Then $P_1 = (0, 1)$ and $P_2 = (1, 0)$ are points on $E(\mathbb{C})$. The point $P_1 + P_2$ is derived by first finding the point $P_1 * P_2$. By definition $P_1 * P_2$ is the third point where the line between P_1 and P_2, $y = 1 - x$, intersects the cubic. Hence, finding this point can be done by substituting the linear expression for y into the cubic. This yields the result $P_1 * P_2 = (2, -1)$. Finally, consider the line between $P_1 * P_2$ and $\mathcal{O}$, $x = 2$, to find the point $P_1 + P_2$. Once more, substitution of the new line into the linear equation yields the result $P_1 + P_2 = (2, 1)$. The operation explained in this example is illustrated in figure 3. △

Figure 3



4 Rational points on an elliptic curve over the rationals

As has been shown, the solutions in $\mathbb{P}^2$ of a non-singular cubic on Weierstrass normal form can be seen as an abelian group. Clearly, a subset of those solutions are solutions $[x, y, 1] \in \mathbb{P}^2$ such that $x, y \in \mathbb{R}$ or even $x, y \in \mathbb{Q}$. For the purposes of this thesis, the focus now shifts to show that the solutions in $\mathbb{Q}$, under certain conditions on the polynomial, is in fact a subgroup.

Definition 4.1. Let

$$E = \{(x, y) \in \mathbb{C}^2 | y^2 = f(x) = x^3 + ax^2 + bx + c\} \cup \{\mathcal{O}\}$$

where $y^2 - f(x) = 0$ is a non-singular cubic in Weierstrass normal form, where $a, b, c \in \mathbb{Q}$ and where $\mathcal{O} = [0, 1, 0]$. Then E is an elliptic curve over $\mathbb{Q}$. For any such E, let $E(\mathbb{Q})$ be defined as

$$E(\mathbb{Q}) = \{(x, y) \in E | x, y \in \mathbb{Q}\} \cup \{\mathcal{O}\}.$$

The set $E(\mathbb{Q})$ is then called the set of rational points on E. △

It is clear that E is just an elliptic curve over $\mathbb{C}$ but with the constraint that the polynomial coefficients are elements of $\mathbb{Q}$. As $\mathbb{Q} \subset \mathbb{C}$ it follows that E forms an abelian group. The set $E(\mathbb{Q})$ is by definition a subset of some particular elliptic curve over $\mathbb{Q}$. It is, however, not clear that $E(\mathbb{Q})$ forms subgroup of E when coupled with the additive operation that was introduced in the previous chapter. The main goal of this chapter is to show that it in fact forms a subgroup.

Noteworthy is that if the polynomial equation is multiplied by some integer such that all denominators of a, b and c are cleared and such that a birational change of variable can be formed where the Weierstrass normal form is preserved, then this gives a birational equivalence between a curve on the form introduced above and a form where the new coefficients are all in $\mathbb{Z}$. This means that it is in fact equivalent to study curves where the coefficients are integers, rather than rationals.

Lemma 4.2. *$\forall P, Q \in E(\mathbb{Q})$ where E is an elliptic curve over $\mathbb{Q}$, $P * Q \in E(\mathbb{Q})$.*

Proof. Note that $E(\mathbb{Q})$ is a subset of the corresponding elliptic curve E, i.e. an element of $E(\mathbb{Q})$ is also an element of E. Therefore, the group structure of E asserts that the element $P * Q$ exists and is contained in E. It must be shown that this also implies that $P * Q \in E(\mathbb{Q})$. Recall that the elliptic curve E is defined by the following equation:

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

where $a, b, c \in \mathbb{Q}$.

Case I: P and Q are distinct points, where $Q = \mathcal{O}$. If $P = (x_1, y_1)$ and $Q = \mathcal{O}$, then the line $x = x_1$ intersects $E(\mathbb{Q})$ at both P and Q. If $y_1 \neq 0$ then the point $-P = (x_1, -y_1)$ is clearly a solution to $x = x_1$ as well as a solution to (1). As $x_1, y_1 \in \mathbb{Q}$ it is also clear that $-y_1 \in \mathbb{Q}$. Therefore $P * Q = -P \in E(\mathbb{Q})$. If $y_1 = 0$, then the line $x = x_1$ intersects E with multiplicity two at this point, hence the third point of intersection is P, which is an element of $E(\mathbb{Q})$. The case is symmetric when $P = \mathcal{O}$.

Case II: If $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ are distinct, $x_1 = x_2$ and $P, Q \neq \mathcal{O}$. By substitution in (1) one finds that $y_1^2 = y_2^2$, for the points to be distinct this implies that $y_1 = -y_2$, hence $Q = -P$. The line $x = x_1$ contains both P and Q as well as $\mathcal{O}$, which by Bezout's Theorem is the third and final point of intersection with E, but $\mathcal{O} \in E(\mathbb{Q})$ by definition.

Case III: If $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ are distinct, $x_1 \neq x_2$, and $P, Q \neq \mathcal{O}$. Consider the line $y = kx + m$, where $k = (y_1 - y_2)/(x_1 - x_2) \in \mathbb{Q}$ and $m = y_1 - kx_1 \in \mathbb{Q}$. By substituting the line equation into (1) it evaluates to

$$0 = x^3 + (a - k^2)x^2 + (b - 2km)x + (c - m^2). \tag{2}$$

As x_1 and x_2 are solutions to (2), it may be factorized into $0 = (x - x_1)(x - x_2)(x - x_3)$. Comparing the coefficients for x^2 in both (2) and the factorized expression, one gets the equality $-x_1 - x_2 - x_3 = a - k^2$, which shows that $x_3 = k^2 - a - x_1 - x_2$. As all terms in the right hand are rationals it follows that $x_3 \in \mathbb{Q}$. It also follows that $y_3 = kx_3 + m \in \mathbb{Q}$ as all terms in the right hand side are rationals in this equation as well. This shows that $P * Q = (x_3, y_3)$ and as both coordinates are rationals it follows that $P * Q \in E(\mathbb{Q})$.

Case IV: Finally consider the case when $P = Q$. If $P = \mathcal{O}$ then, as has been shown before, the tangent $Z = 0$ intersects E at $\mathcal{O}$ three times, so $P * P = \mathcal{O}$ in this case, which is an element of $E(\mathbb{Q})$. When $P \neq \mathcal{O}$, call $P = (x_1, y_1)$. If $y_1 = 0$, the tangent of P is the line $x = x_1$, which intersects $E(\mathbb{Q})$ at the point $\mathcal{O}$, and therefore $\mathcal{O}$ is the third point of intersection, where $\mathcal{O} \in E(\mathbb{Q})$ by construction. Finally, if $y_1 \neq 0$, one may then express the tangent of P by the equation $y = kx + m$ where k is derived using implicit differentiation of (1), that is:

$$2y\frac{dy}{dx} = 3x^2 + 2ax + b \iff \frac{dy}{dx} = \frac{3x^2 + 2ax + b}{2y}. \tag{3}$$

As all terms in the numerator and denominator of (3) are rational, when evaluated at (x_1, y_1), it implies that $k = \frac{dy}{dx}(x_1, y_1) \in \mathbb{Q}$. Let $m = y_1 - kx_1 \in \mathbb{Q}$. When substituting the line equation

into (1), the equation once more evaluates to (2). The difference this time is that the factorized expression becomes $0 = (x-x_1)^2(x-x_3)$. By similar reasoning it follows that $x_3 = k^2-a-2x_1 \in \mathbb{Q}$ and $y_3 = kx_3 + m \in \mathbb{Q}$. As both coordinates of $P * P = (x_3, y_3)$ are rational, it follows that $P * P \in E(\mathbb{Q})$. □

Lemma 4.3. *$\forall P, Q \in E(\mathbb{Q})$ where E is some elliptic curve over $\mathbb{Q}$, $P + (-Q) \in E(\mathbb{Q})$.*

Proof. If $Q = \mathcal{O}$, by lemma 3.9 it follows that $-Q = \mathcal{O}$ and thus $P + (-Q) = P + \mathcal{O} = P \in E(\mathbb{Q})$.

When $Q \neq \mathcal{O}$, let $Q = (x_1, y_1)$ where $x_1, y_1 \in \mathbb{Q}$. By the discussion following lemma 3.9 it follows that $-Q = (x_1, -y_1)$, which is an element of $E(\mathbb{Q})$ as both coordinates are rationals. Since both P and $-Q$ are elements of $E(\mathbb{Q})$, lemma 4.2 implies that $P * (-Q) \in E(\mathbb{Q})$. By applying the lemma again it follows that $(P * (-Q)) * \mathcal{O} \in E(\mathbb{Q})$. But $P + (-Q) = (P * (-Q)) * \mathcal{O}$ by definition, so $P + (-Q) \in E(\mathbb{Q})$. □

Theorem 4.4. *Let E be an elliptic curve over $\mathbb{Q}$, then $E(\mathbb{Q})$ is a subgroup of E.*

Proof. As $\mathcal{O}$ is always contained in the set of rational points of E by definition, $E(\mathbb{Q})$ is non-empty. In addition, by lemma 4.3, $\forall P, Q \in E(\mathbb{Q})$, $P + (-Q) \in E(\mathbb{Q})$. By the subgroup criterion [2, p. 47] it follows that $E(\mathbb{Q})$ is a subgroup of E. □

The theorem implies that the set of rational points on an elliptic curve over $\mathbb{Q}$ is also an abelian group in and of itself. The main result of this thesis, the Nagell-Lutz Theorem, describes a property of any such group when all coefficients are integers.

5 Points of order two and the duplication formula

The purpose of this chapter is to introduce two theorems that are used in proving the Nagell-Lutz theorem. Both theorems are also applied in two examples in order to illustrate their usefulness.

Definition 5.1. For some integer $n \geq 1$ and some point P on an elliptic curve over $\mathbb{C}$, let nP denote the operation of adding the point P to itself $n-1$-times. That is:

$$nP = \underbrace{P + P + ... + P}_{n \text{ summands}}.$$

For $n = 0$, let $nP = \mathcal{O}$. △

Definition 5.2. [1, p. 38] An element of any group is said to have *order* m if $mP = \mathcal{O}$ and $m'P \neq \mathcal{O}$ for all integers $1 \leq m' < m$. If such an m exists, then P is said to be of *finite order*; otherwise it is said to be of *infinite order*. △

Theorem 5.3 (Points of order two)**.** [1, p. 40] *Let $E(\mathbb{C})$ be an elliptic curve over $\mathbb{C}$. Denote the corresponding curve equation by C where*

$$C : y^2 = f(x) = x^3 + ax^2 + bx + c.$$

Then a point $P = (x', y') \neq \mathcal{O}$ on $E(\mathbb{C})$ has order two if and only if $y' = 0$.

Proof. Let $\hat{C}$ be the homogenization of C, such that $\hat{C} : F(X, Y, Z) = ZY^2 - X^3 - aX^2Z - bXZ^2 - cZ^3 = 0$. Recall that, in the projective plane, the point P has the homogeneous coordinates $[x', y', 1]$.

Assume that P has order two, i.e. $2P = \mathcal{O}$. In order to derive the point $2P$, the group action asserts that one must first calculate the tangent to the point P, that is

$$\frac{\partial F}{\partial X}(P)X + \frac{\partial F}{\partial Y}(P)Y + \frac{\partial F}{\partial Z}(P)Z = 0,$$

which simplifies to

$$(-3x'^2 - 2ax' - b)X + (2y')Y + (y'^2 - ax'^2 - 2bx' - 3c)Z = 0. \tag{4}$$

As $2P = \mathcal{O}$ it follows that $P * P$ must be the point $\mathcal{O}$. The reason that it follows is because if $P * P = Q \neq \mathcal{O}$, then $2P = \mathcal{O}$ implies that $Q * \mathcal{O} = \mathcal{O}$. This means that the line, L, between Q and $\mathcal{O}$ would intersect $E(\mathbb{C})$ twice at $\mathcal{O}$. It then follows that L would have to be the tangent to $\mathcal{O}$, i.e.

$$L : \frac{\partial F}{\partial X}(\mathcal{O})X + \frac{\partial F}{\partial Y}(\mathcal{O})Y + \frac{\partial F}{\partial Z}(\mathcal{O})Z = 0 \iff Z = 0.$$

Hence, L would only contain points where $Z = 0$. But the only point on $\hat{C}$ where $Z = 0$ is $\mathcal{O}$, which contradicts the existence of the point Q. Therefore it follows that $P * P = \mathcal{O}$. By evaluating $\mathcal{O}$ in the equation (4) one sees that $2y' = 0$. This shows the first implication that $2P = \mathcal{O} \Rightarrow y' = 0$.

Conversely, assume that $y' = 0$. Consider once more the tangent to the point P

$$\frac{\partial F}{\partial X}(P)X + \frac{\partial F}{\partial Y}(P)(Y) + \frac{\partial F}{\partial Z}(P)Z = 0,$$

which simplifies to

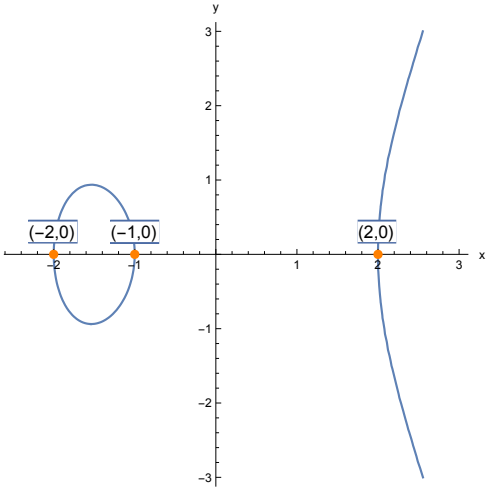
$$(-3x'^2 - 2ax' - b)X + (-ax'^2 - 2bx' - 3c)Z = 0. \tag{5}$$

It is clear that $\mathcal{O}$ is a solution to (5). Bezout's Theorem therefore implies that this must be the third and final intersection with $\hat{C}$. As was shown in the proof for the other implication, the tangent at $\mathcal{O}$ intersects $E(\mathbb{C})$ at $\mathcal{O}$ with multiplicity three. Hence, we may conclude that $y' = 0 \Rightarrow 2P = \mathcal{O}$. □

It can be noted that the theorem clearly also holds for elliptic curves over $\mathbb{Q}$ seeing as they are a particular type of elliptic curve over $\mathbb{C}$. Whenever there is a point $P = (x, 0)$, $x \in \mathbb{Q}$, on an elliptic curve over $\mathbb{Q}$, call it E, P would be an element of $E(\mathbb{Q})$. As $2P$ remains the same point regardless of whether one looks at E or $E(\mathbb{Q})$, it follows that the theorem is also applicable to the set of rational points on an elliptic curve over $\mathbb{Q}$.

Example 5.4 (Rational points of order two on an elliptic curve over $\mathbb{Q}$)**.** Consider the elliptic curve over $\mathbb{Q}$: $E = \{(x, y) \in \mathbb{C}^2 | y^2 = x^3 + x^2 - 4x - 4\} \cup \{\mathcal{O}\}$. Any point $P = (x', y') \in E$ of order two should have $y' = 0$ by theorem 5.3. Thus, finding any such point amounts to finding the roots of $0 = x^3 + x^2 - 4x - 4$. Any rational roots will then correspond to a point of order two on $E(\mathbb{Q})$. As the particular equation can be factorized as $0 = (x + 1)(x + 2)(x - 2)$ one finds the following rational points $(-2, 0), (-1, 0)$ and $(2, 0)$. Figure 4 illustrates a segment of E with the points of order two highlighted.

Figure 4



Another way of interpreting the result is that we have found that the line $y = 0$ intersects $E(\mathbb{Q})$ at these three distinct points. This implies that for any two of these points P_1 and P_2, $P_1 * P_2$ is the third point P_3. As the line from P_3 to $\mathcal{O}$ is the vertical line that intersects $E(\mathbb{Q})$ twice at P_3, it is also the case that $P_3 * \mathcal{O} = P_3$, and thus $P_1 + P_2 = P_2 + P_1 = P_3$. It follows that the subset $\{(-2, 0), (-1, 0), (2, 0), \mathcal{O}\}$ forms a subgroup that is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. It can be shown that any elliptic curve over $\mathbb{C}$ contains such a subgroup [1, p. 40]. △

With theorem 5.3 proven the focus of this chapter shifts to the derivation of a so called duplication formula. That is, for a point P on an elliptic curve over $\mathbb{C}$, the duplication formula gives the coordinates of the point $2P$. Before deriving the formula an additional definition is introduced.

Definition 5.5. Let $E(\mathbb{C})$ be an elliptic curve over $\mathbb{C}$. For any point $P = (x, y) \in E(\mathbb{C})\backslash\{\mathcal{O}\}$ define the functions $x(P) = x$ and $y(P) = y$. △

Theorem 5.6 (Duplication formula)**.** *Let $E(\mathbb{C})$ be an elliptic curve over $\mathbb{C}$ and C be the equation describing the elliptic curve, i.e. $C : y^2 = x^3 + ax^2 + bx + c$. If $P = (x', y') \in E(\mathbb{C})$ such that $2P \neq \mathcal{O}$, then*

$$x(2P) = k^2 - a - 2x' \quad \textit{and} \quad y(2P) = -k^3 + ka + 2kx' - m$$

where k is the slope and m is the intersection of the tangent of P.

Proof. The group operation asserts that in order to compute $2P$ one needs to find the tangent of P. Let the tangent be described by the line $y = kx + m$, the slope k is then determined by implicitly differentiating C. It follows that

$$k = \frac{dy}{dx}(P) = \frac{3x'^2 + 2ax' + b}{2y'}$$

and

$$m = y' - kx'.$$

Similarly to what was discussed in the proof of lemma 4.2, by substituting the linear expression for y into C, the equation becomes

$$0 = x^3 + x^2(a - k^2) + x(b - 2km) + (c - m^2). \tag{6}$$

However, as $y = kx + m$ intersects C twice at x', (6) can also be expressed as $0 = (x - x')^2(x - x_0)$, where x_0 is the third intersection of the line with C and thus $x_0 = x(2P)$. By collecting the terms for x^2 in both (6) and the factorization of the same equation, it follows that

$$x(2P) = x_0 = k^2 - a - 2x'.$$

By evaluating the tangent line at $x(2P)$ one gets

$$y_0 = kx_0 + m = k^3 - ka - 2kx' + m.$$

From the discussion following lemma 3.9 it follows that

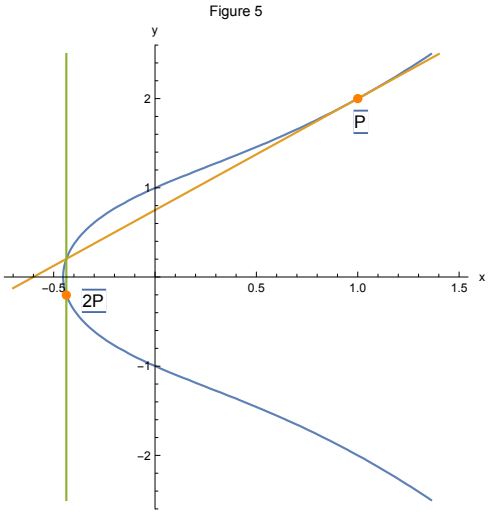
$$y(2P) = -y_0 = -k^3 + ka + 2kx' - m.$$

□

Example 5.7 (Application of the duplication formula)**.** Consider the curve $y^2 = f(x) = x^3 + 2x + 1$. The point $P = (1, 2)$ is a rational point on the curve where $y \neq 0$. Calculating the coordinates of the point $2P$ can therefore be done using the duplication formula. The slope of the tangent to the point P is derived by implicit differentiation such that $k = 5/4$. The intersection m can then easily be computed, yielding $m = 3/4$. By applying the duplication formula it follows that $2P = (-7/16, -13/64)$, which, as expected by the group structure of the rational points on an elliptic curve over $\mathbb{Q}$, is another rational point on the curve. The points are illustrated in figure 5. △

6 The Nagell-Lutz Theorem

In this final chapter, the Nagell-Lutz Theorem is finally stated and proven. The proof closely follows that of Silverman and Tate [1, Ch 2.3-2.4] where this thesis includes a lemma, lemma 6.12, that corrects an error in the original text noted by the authors themselves [4, p. 4].



Theorem 6.1 (The Nagell-Lutz Theorem)**.** [1, p. 56] *Let E be an elliptic curve over $\mathbb{Q}$ defined by the curve*

$$y^2 = f(x) = x^3 + ax^2 + bx + c, \tag{7}$$

where a, b and c are integers. Let D be the discriminant of the polynomial $f(x)$,

$$D = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^3. \tag{8}$$

Let $P \in E(\mathbb{Q})$ be a point of finite order. Then $x(P)$ and $y(P)$ are integers; and either $y(P) = 0$, in which case P has order 2, or else $y(P)$ divides D.

Remark: This theorem is useful in the sense that calculating the discriminant is computationally a very simple operation. Then, by using the information contained in the discriminant, the theorem reduces the search space for possible points of finite order to a finite number of candidates on the curve. Finding all rational points of finite order on the curve can therefore always be done in a finite number of steps. Proving this theorem will, however, require the introduction of additional lemmas and definitions.

Strategy: The overall strategy of the proof is to show that whenever either coordinate of a point $P \in E(\mathbb{Q})$ are rationals, P cannot be of finite order. This is done by showing that for any prime p, whenever the denominator of $x(P)$ and $y(P)$ are divisible by p, P cannot be of finite order. As this will be seen to be true for any prime p, it follows that the denominator for $x(P)$ and $y(P)$ can only be 1, which means that they must both be integers.

Lemma 6.2. [1, p. 48] *Let $P = (x, y)$ be a point on the cubic curve (7) such that both P and $2P$ have integer coordinates. Then either $y = 0$ or y divides D.*

Proof. If $y \neq 0$, P cannot be a point of order two by theorem 5.3, i.e. $2P \neq \mathcal{O}$. Therefore it follows that $2P = (X, Y)$ where, by assumption, $X, Y \in \mathbb{Z}$. Using the duplication formula 5.6 one finds that

$$X = k^2 - a - 2x, \tag{9}$$

where

$$k = \frac{3x^2 + 2ax + b}{2y} = \frac{f'(x)}{2y} \tag{10}$$

for $f(x)$ as in (7).

By assumption $a, x, X \in \mathbb{Z}$, which, when considering (9), implies that $k^2 \in \mathbb{Z}$. Similarly in (10) it follows that $k \in \mathbb{Q}$, i.e. $k = i/j$ where $i, j \in \mathbb{Z}$ and $\gcd(i, j) = 1$. If $j \neq 1$ it would imply that $i^2/j^2 \notin \mathbb{Z}$, which contradicts that $k^2 \in \mathbb{Z}$, hence $j = 1$ and thus $k \in \mathbb{Z}$. In turn, this means that $2y | f'(x)$, and in particular $y | f'(x)$. As $y^2 = f(x)$ it clearly also follows that $y | f(x)$.

It can be shown that the discriminant D, as defined in equation (8), can be rewritten as

$$D = r(x)f(x) + s(x)f'(x)$$

where

$$r(x) = (18b - 6a^2)x - (4a^3 - 15ab + 27c)$$

and

$$s(x) = (2a^2 - 6b)x^2 + (2a^3 - 7ab + 9c)x + (a^2b + 3ac - 4b^2).$$

As $a, b, c \in \mathbb{Z}$, both $r(x)$ and $s(x)$ will attain integer values whenever x is an integer. For the particular point P, $x, y \in \mathbb{Z}$, and therefore it follows that $y|D$ as $y|f(x)$ and $y|f'(x)$. □

The preceding lemma will be used to prove the final statement of the Nagell-Lutz Theorem. The following definitions and lemmas will instead focus on showing that the coordinates of a point must be integers.

Definition 6.3. For any prime p, any non-zero rational number x can be represented on the reduced form

$$x = \frac{m}{n}p^v,$$

where $m, n, v \in \mathbb{Z}$, $\gcd(m, n) = 1$, $\gcd(m, p) = 1$, $\gcd(n, p) = 1$ and $n > 0$. For any such x, call this form its *reduced form with respect to p*. △

Definition 6.4. For any prime p, let x be any non-zero rational number written on its reduced form with respect to p,

$$x = \frac{m}{n}p^v.$$

Then define the *order of x with respect to p*, denoted $\mathrm{ord}_p(x)$, to be v. For $x = 0$, let $\mathrm{ord}_p(x) = \infty$. △

Lemma 6.5. *Let $P = (x, y)$ be a rational point on (7). If for some prime p the order with respect to p is less than 0 for either x or y, then there exists some integer $v > 0$ such that $ord_p(x) = -2v$ and $ord_p(y) = -3v$.*

Proof. Assume that $\mathrm{ord}_p(x) = -\mu$ where μ is an integer and $\mu > 0$. Let $\mathrm{ord}_p(y) = -\sigma$ for some integer σ. Using definition 6.3 the coordinates may be expressed as

$$x = \frac{m}{np^\mu} \qquad \text{and} \qquad y = \frac{u}{wp^\sigma}. \tag{11}$$

Substituting (11) into (7) yields the equation

$$\frac{u^2}{w^2p^{2\sigma}} = \frac{m^3 + am^2np^\mu + bmn^2p^{2\mu} + cn^3p^{3\mu}}{n^3p^{3\mu}}. \tag{12}$$

By definition 6.3, $p \not| m, p \not| n, p \not| u$ and $p \not| w$. For the left hand side of (12) this means that

$$\mathrm{ord}_p\left(\frac{u^2}{w^2p^{2\sigma}}\right) = -2\sigma.$$

Similarly for the right hand side of (12) it follows that p does not divide the numerator and thus

$$\mathrm{ord}_p\left(\frac{m^3 + am^2np^\mu + bmn^2p^{2\mu} + cn^3p^{3\mu}}{n^3p^{3\mu}}\right) = -3\mu.$$

The equality between the expressions then implies that $2\sigma = 3\mu$. As $\mu > 0$ by assumption it follows that $\sigma > 0$ and therefore $\mathrm{ord}_p(y)$ is negative. In addition, this implies that $2|\mu$ and $3|\sigma$. Therefore, let v be an integer such that $\mu = 2v$ and $\sigma = 3v$, clearly it follows that $v > 0$.

When instead assuming that $\mathrm{ord}_p(y) = -\sigma$ where σ is an integer and $\sigma > 0$ and letting $\mathrm{ord}_p(x) = -\mu$ the same result follows by a symmetric argument. □

Example 6.6. Consider the elliptic curve over $\mathbb{Q}$ defined by the equation $y^2 = x^3 + 2x^2 + 13$. Then the point $(1/4, 29/8)$ is a solution to the equation and hence a rational point on the curve. By the same notation as in equation (11), it follows that $p = 2$, $\mu = 2$ and $\sigma = 3$. Thus, $v = 1$ is an integer satisfying the statement of the preceding lemma. △

The lemma says that whenever there exists a rational point on the curve (7) where some prime p divides the denominator of either coordinate of that point, it follows that p also divides the denominator of the other coordinate as well. In addition, there is a relationship between the order, with respect to p, of both coordinates. These results are the foundation of the following discussion.

Definition 6.7. Let $E(\mathbb{Q})$ be as in theorem 6.1. Then for any prime p and positive integer v, let $C(p^v)$ be defined as follows

$$C(p^v) = \{(x, y) \in E(\mathbb{Q}) | \text{ord}_p(x) \leq -2v \wedge \text{ord}_p(y) \leq -3v\} \cup \{\mathcal{O}\}.$$

△

It is clear that $C(p^v) \subset E(\mathbb{Q})$, and by the way the set is constructed, $C(p^{(1+i)v}) \subset C(p^{iv})$ for every positive integer i. In particular, when $v = 1$ the following chain of inclusions holds for any prime p:

$$E(\mathbb{Q}) \supset C(p) \supset C(p^2) \supset C(p^3)... \tag{13}$$

As part of the proof for Nagell-Lutz Theorem, it must be shown that $C(p^v)$ is a subgroup of $E(\mathbb{Q})$. In order to do this, the following change of variable is introduced for points in $C(p^v)$:

$$t = \frac{X}{Y} \qquad \text{and} \qquad s = \frac{Z}{Y}. \tag{14}$$

It follows that $\mathcal{O}$ is mapped to the point $(0, 0)$ in the ts-plane. All points on (7), except for the points of order two where $Y = 0$, are mapped to the ts-plane by this mapping.

Lemma 6.8. *Points of order two on $E(\mathbb{Q})$ as defined in theorem 6.1 are not elements of any subset $C(p^v)$.*

Proof. This is a direct consequence of the rational root theorem for monic polynomials. As a point of order two, P, on $E(\mathbb{Q})$ will have $y(P) = 0$ by theorem 5.3, the problem is reduced to finding roots of a monic polynomial with integer coefficients. The rational root theorem for monic polynomials then states that any rational roots to such a polynomial must be integers. Hence, $x(P)$ must be an integer and therefore $\text{ord}_p(x(P)) \geq 0$, it follows that $P \notin C(p^v)$. □

Example 6.9 (Rational root theorem for monic polynomials)**.** Consider the equation $0 = x^3 + 2x^2 - x - 2$. The right hand side, call it $p(x)$, is clearly a monic polynomial with integer coefficients. The rational root theorem for monic polynomials states that if there exists a rational root, r, to $p(x)$, then it can be written as $r = \pm a/b$ where $\gcd(a, b) = 1$, $a | -2$ (i.e. the constant term) and $b|1$ (i.e. the leading coefficient). The important part here is that the leading coefficient is 1, which in this particular case means that if r exists it must be one of the following integer values: $\pm 1, \pm 2$. As $p(-2) = p(-1) = p(1) = 0$ it follows that the polynomial can be factorized as $p(x) = (x + 2)(x + 1)(x - 1)$. △

The consequence of this lemma is that when mapping some subset $C(p^v)$ by the mapping (14), there is no need to consider any points where $Y = 0$, and thus where $y = 0$.

Definition 6.10. For any point $P = (x, y)$ on (7), let P^{ts} denote the corresponding point in the ts-plane when mapped by (14). In addition, let $t(P) = x/y$ and $s(P) = 1/y$. △

Applying the mapping (14) to (7) one gets the cubic

$$s = t^3 + at^2 s + bts^2 + cs^3. \tag{15}$$

Importantly, it also follows that any line in the xy-plane is mapped to a line in the st-plane. For instance, the line $y = kx + m$ corresponds to the line $s = -k/mt + 1/m$. Consider the three points P_1, P_2 and $P_1 * P_2$ on (7) and some line L_{xy} in the xy-plane. By the mapping (14) they all lie on (15) as well as some line L_{ts}, i.e. the line L_{xy} mapped by (14), in the ts plane. Hence, drawing a line between two points in the xy-plane and deriving the third point of intersection and applying the mapping (14) to that point will result in the same point as if one first mapped the two initial points to the ts-plane and derived the third point of intersection of the line between the points and the curve (15). Thus, it is possible to add points in the ts-plane by the same method as in the xy-plane, with the difference that $\mathcal{O}$ is located at $(0, 0)$ in the ts-plane.

Definition 6.11. For any prime p, let R_p be defined as

$$R_p = \{x \in \mathbb{Q} | \text{ord}_p(x) \geq 0\}.$$

△

In fact, R_p is a ring called the localization of $\mathbb{Z}$ at p [2, p. 708]. Consider a rational point $P = (x, y)$, $2P \neq \mathcal{O}$, in the subset $C(p^v)$ as defined in definition 6.7. By this definition, the coordinates of P may be written on its reduced form with respect to p in the following way:

$$x = \frac{m}{np^{2(v+i)}} \qquad \text{and} \qquad y = \frac{u}{wp^{3(v+i)}}$$

for some $i \geq 0$. Applying the mapping (14) then gives the coordinates for P^{ts}:

$$t = \frac{x}{y} = \frac{mw}{nu}p^{v+i} \qquad \text{and} \qquad s = \frac{1}{y} = \frac{w}{u}p^{3(v+i)}.$$

This means that for any rational point $P \in C(p^v)$ it follows that $t(P) \in p^v R_p$ and $s(P) \in p^{3v} R_p$. The same holds when applying the mapping to the point $\mathcal{O}$ as $t(\mathcal{O}) = 0 \in p^v R_p$ and $s(\mathcal{O}) = 0 \in p^{3v} R_p$. Thus, (14) induces a one-to-one mapping between elements of $C(p^v)$ and a subset of $p^v R_p \times p^{3v} R_p$. Therefore, an indirect way of showing that $C(p^v)$ is closed under addition is to show that for $P_1, P_2 \in C(p^v)$, then $t(P_1 + P_2) \in p^v R_p$ and $s(P_1 + P_2) \in p^{3v} R_p$, where $(P_1 + P_2)^{ts}$ is on the curve (15). Then, if it can be shown that for any $P \in C(p^v)$ there exists an element $-P \in C(p^v)$, it follows that $C(p^v)$ is a subgroup. The proof of the Nagell-Lutz theorem relies on some of the properties of the subgroup $C(p^v)$.

Lemma 6.12. *Let $C(p^v)$ be a subset of $E(\mathbb{Q})$ as defined in definition 6.7. Then there cannot exist two points $P_1, P_2 \in C(p^v)$ such that $P_1 \neq P_2$, but $t(P_1) = t(P_2)$.*

Proof. Recall that in the assumptions of the Nagell-Lutz Theorem, the coefficients in the curve (7), a, b and c, are all integers. Assume for contradiction that there exists two points $P_1, P_2 \in C(p^v)$ such that $P_1^{ts} \neq P_2^{ts}$ but $t(P_1) = t(P_2)$. Let $P_1 = (t_1, s_1)$ and $P_2 = (t_2, s_2)$, then from the assumptions it follows that $s_1 \neq s_2$. By assumption, both P_1^{ts} and P_2^{ts} are points on the curve (15). Using this equation, subtract s_2 from s_1:

$$s_1 - s_2 = (t_1^3 - t_2^3) + a(t_1^2 s_1 - t_2^2 s_2) + b(t_1 s_1^2 - t_2 s_2^2) + c(s_1^3 - s_2^3).$$

As $s_1 - s_2 \neq 0$ by assumption and $t_1 = t_2$ by assumption, it is possible to divide by $s_1 - s_2$ and exchange t_2 for t_1. This yields the equivalent expression

$$1 = at_1^2 + bt_1(s_1 + s_2) + c(s_1^2 + s_1 s_2 + s_2^2).$$

Observe that $s_1^2 + s_1 s_2 + s_2^2 = 0$ only when $s_1 = s_2 = 0$, which is impossible by the assumption that $s_1 \neq s_2$. Hence, divide by this expression and collect terms such that c is on the left hand side of the equation:

$$c = \frac{1 - at_1^2 - bt_1(s_1 + s_2)}{(s_1^2 + s_1 s_2 + s_2^2)}. \tag{16}$$

Consider the case when $c = 0$, then $1 - at_1^2 - bt_1(s_1 + s_2) = 0$. It follows that $t_1 \neq 0$, as $t_1 = 0$ would give the absurd result $1 = 0$. The numerator may therefore be rewritten as:

$$a = \frac{1 - bt_1(s_1 + s_2)}{t_1^2}. \tag{17}$$

When $a = 0$ it follows that $1 - bt_1(s_1 + s_2) = 0$. By similar reasoning as in the previous paragraph, it follows that $t_1(s_1 + s_2) \neq 0$. Dividing the equation by $t_1(s_1 + s_2)$ then gives:

$$b = \frac{1}{t_1(s_1 + s_2)}. \tag{18}$$

As $P_1, P_2 \in C(p^v)$ it was shown that $t_1, t_2 \in p^v R_p$ and $s_1, s_2 \in p^{3v} R_p$ and hence p divides the denominator of (18). This means that b cannot be an integer, which contradicts the fact that b is an integer. Hence $a = 0$ leads to a contradiction.

Consider the case when $a \neq 0$. As p divides the denominator of (17) but not the numerator, it also follows that a cannot be an integer, which is a contradiction. Hence $c = 0$ leads to a contradiction.

The only remaining possibility is when $c \neq 0$. By similar reasoning for the fraction (16), p does not divide the numerator, but does divide the denominator. Therefore c cannot be an integer, which is also a contradiction. It follows that there cannot exist two such points in $C(p^v)$. □

Lemma 6.13. *Let $C(p^v)$ be a subset of $E(\mathbb{Q})$ as defined in definition 6.7. Then for two points $P_1, P_2 \in C(p^v)$ it follows that $P_1 + P_2 \in C(p^v)$.*

Proof. Case I: Consider the case when $P_1 \neq P_2$. By lemma 6.12 it follows that $t(P_1) \neq t(P_2)$. Thus, the line between any two points in the st-plane may be defined as

$$s = \alpha t + \beta \tag{19}$$

where α is the slope defined as

$$\alpha = \frac{s_2 - s_1}{t_2 - t_1}. \tag{20}$$

However, a more useful expression for the slope will be derived that gives more information about α. Consider the difference $s_2 - s_1$ when expressed in terms of the cubic (15):

$$s_2 - s_1 = (t_2^3 - t_1^3) + a(t_2^2 s_2 - t_1^2 s_1) + b(t_2 s_2^2 - t_1 s_1^2) + c(s_2^3 - s_1^3),$$

which is equivalent to

$$s_2 - s_1 = (t_2^3 - t_1^3) + a((t_2^2 - t_1^2)s_2 - t_1^2(s_2 - s_1)) + b((t_2 - t_1)s_2^2 + t_1(s_2^2 - s_1^2)) + c(s_2^3 - s_1^3).$$

Moving all terms with $s_2 - s_1$ as a factor to the left hand side and factoring $s_2 - s_1$ and $t_2 - t_1$ yields:

$$(s_2 - s_1)(1 - at_1^2 - bt_1(s_1 + s_2) - c(s_1^2 + s_1 s_2 + s_2^2)) = (t_2 - t_1)((t_2^2 + t_2 t_1 + t_1^2) + a(t_1 + t_2)s_2 + bs_2^2)$$

As $t_1 \neq t_2$ by assumption it is possible to divide by $(t_2 - t_1)$. For the expression $1 - at_1^2 - bt_1(s_1 + s_2) - c(s_1^2 + s_1 s_2 + s_2^2)$, a little more analysis is required. It was shown previously that if $P \in C(p^v)$, then $t(P) \in p^v R_p$ and $s(P) \in p^{3v} R_p$, thus, $p|s_1, p|s_2$ and $p|t_1$. As a, b, c are all integers, this means that $p \nmid 1 - at_1^2 - bt_1(s_1 + s_2) - c(s_1^2 + s_1 s_2 + s_2^2)$ and that $1 - at_1^2 - bt_1(s_1 + s_2) - c(s_1^2 + s_1 s_2 + s_2^2) \neq 0$. Hence, the expression can be divided by $1 - at_1^2 - bt_1(s_1 + s_2) - c(s_1^2 + s_1 s_2 + s_2^2)$ as well. This leads to the following expression for α:

$$\alpha = \frac{s_2 - s_1}{t_2 - t_1} = \frac{t_2^2 + t_2 t_1 + t_1^2 + a(t_1 + t_2)s_2 + bs_2^2}{1 - at_1^2 - bt_1(s_1 + s_2) - c(s_1^2 + s_1 s_2 + s_2^2)}. \tag{21}$$

Case II: If instead $P_1 = P_2$, the slope is calculated by implicit differentiation:

$$\frac{ds}{dt} = 3t^2 + 2ast + \frac{ds}{dt}at^2 + bs^2 + 2bts\frac{ds}{dt} + 3cs^2\frac{ds}{dt},$$

which is equivalent to

$$\frac{ds}{dt}(1 - at^2 - 2bts - 3cs^2) = 3t^2 + 2ast + bs^2.$$

For the point P_1, it follows that $1 - at_1^2 - 2bt_1 s_1 - 3cs_1^2$ is non-zero and that p is not a divisor by similar reasoning as when deriving (21). This gives the slope α when $P_1 = P_2$:

$$\alpha = \frac{ds}{dt}(P_1) = \frac{3t_1^2 + 2as_1 t_1 + bs_1^2}{1 - at_1^2 - 2bt_1 s_1 - 3cs_1^2} \tag{22}$$

Combining cases I and II: It turns out that the right hand expression in (21) evaluates to exactly the right hand expression in (22) when $P_1 = P_2$, therefore it is enough to consider (21) in either case.

Finalizing the proof: Note that p^{2v} divides each term in the numerator, but p does not divide the denominator of (21). As a and b are just integers, it follows that $p^{2v}|\alpha$. Finally, let $\beta = s_1 - \alpha t_1$. Since $p^{3v}|s_1$ and $p^{3v}|t_1\alpha$ it follows that $p^{3v}|\beta$.

Let $P_3^{ts} = (t_3, s_3)$ be the third point of intersection of the line $s = \alpha t + \beta$ with the cubic (15). By lemma 4.2 the point $P_3 \in E(\mathbb{Q})$, it remains to show that it is also in $C(p^v)$. To do this, substitute s with $\alpha t + \beta$ in (15):

$$\alpha t + \beta = t^3 + at^2(\alpha t + \beta) + bt(\alpha t + \beta)^2 + c(\alpha t + \beta)^3 \iff$$

$$0 = t^3(1 + a\alpha + b\alpha^2 + c\alpha^3) + t^2(a\beta + 2b\alpha\beta + 3c\alpha^2\beta) + t(b\beta^2 + 3c\alpha\beta^2 - \alpha) + (c\beta^3 - \beta). \quad (23)$$

As $p|\alpha$ and a, b, c are integers it follows that the leading coefficient is non-zero and that p is not a divisor. Therefore the whole expression may be divided by this coefficient such that it becomes a monic polynomial. Following this division, also note that the polynomial can be factored as

$$0 = (t - t_1)(t - t_2)(t - t_3). \quad (24)$$

Equating the coefficients of t^2 in both (23) and (24) gives the following equality:

$$t_1 + t_2 + t_3 = -\frac{a\beta + 2b\alpha\beta + 3c\alpha^2\beta}{1 + a\alpha + b\alpha^2 + c\alpha^3}. \quad (25)$$

As $p^{3v}|\beta$ and p is not a divisor of the denominator, it can be seen that $t_3 \in p^v R_p$. In turn, this means that, since $s_3 = \alpha t_3 + \beta$, $p^{3v}|s_3$, such that $s_3 \in p^{3v}R_p$. By the one-to-one mapping (14) and lemma 6.5 this means that $P_3 \in C(p^v)$. Now, $P_3 = P_1 * P_2$, so it is still needed to verify that $P_1 + P_2 \in C(p^v)$. As $\mathcal{O}$ maps to $(0, 0)$ in the ts-plane under the mapping (14), this means that the final line needed to be considered goes through the origin in the ts-plane. By analyzing the curve (15), it follows that if (t_3, s_3) is a solution to the curve, then $(-t_3, -s_3)$ is also a solution. Hence, the line through (t_3, s_3) and $(0, 0)$ must also pass through $(-t_3, -s_3)$, which is the final point of intersection by Bezout's Theorem and thus corresponds to $P_1 + P_2$. As $t_3 \in p^v R_p$ and $s_3 \in p^{3v}R_p$, it clearly follows that $-t_3 \in p^v R_p$ and $-s_3 \in p^{3v}R_p$ and thus $P_1 + P_2 \in C(p^v)$. □

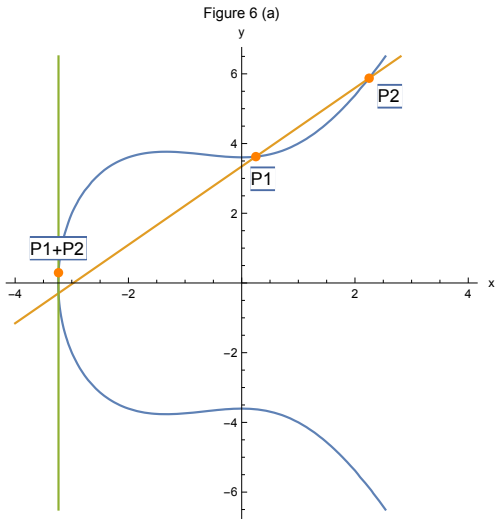
Example 6.14. Consider the elliptic curve over $\mathbb{Q}$ defined by the curve $y^2 = x^3 + 2x^2 + 13$. Both $P_1 = (1/4, 29/8)$ and $P_2 = (9/4, 47/8)$ are rational points on the curve such that they belong to the subset $C(2^1)$. By the additive operation on elliptic curves it follows that $(1/4, 29/8) + (9/4, 47/8) = (-207/64, 151/512)$. Where clearly, $(-207/64, 151/512) \in C(2^1)$ as was expected by lemma 6.13. In fact, the point actually belongs to the subset $C(2^3)$, which is not a coincidence and a result that will be utilized in the proof of lemma 6.16. The addition is illustrated in figure 6(a). In addition, figure 6(b) illustrates the corresponding addition in the ts-plane, where $P_1^{ts} = (2/29, 8/29)$ and $P_2^{ts} = (18/47, 8/47)$. It is clear that $2/29, 18/47 \in 2R_2$ and $8/29, 8/47 \in 2^3R_2$ as expected. By adding the points in the ts-plane one derives the point $P_1^{ts} + P_2^{ts} = (-207 \cdot 8/151, 512/151)$ where $-207 \cdot 8/151 \in 2^3R_2 \subset 2R_2$ and $512/151 \in 2^9R_2 \subset 2^3R_2$. By applying the mapping (14) on $P_1 + P_2$ it is also clear that $(P_1 + P_2)^{ts} = P_1^{ts} + P_2^{ts}$. △

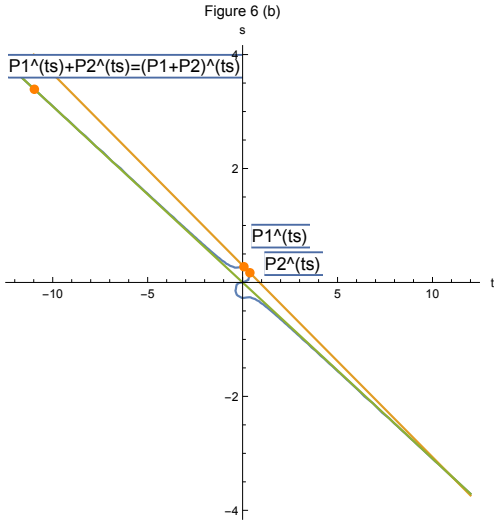
Lemma 6.15. *Let $C(p^v)$ be a subset of $E(\mathbb{Q})$ as defined in definition 6.7. Then $C(p^v)$ is a subgroup of $E(\mathbb{Q})$.*

Proof. For any point $P = (x', y') \in C(p^v)$, by definition P is also an element of $E(\mathbb{Q})$. This means that P has an additive inverse, $-P = (x', -y')$, in $E(\mathbb{Q})$. By definition of $C(p^v)$, $\text{ord}_p(y') \leq -3v$. But as the reduced form with respect to p of y' and $-y'$ only differs by sign, $\text{ord}_p(y') = \text{ord}_p(-y')$. Thus $-P \in C(p^v)$ also and therefore it follows that each point in $C(p^v)$ has an additive inverse in $C(p^v)$. By lemma 6.13 this means that for any points $P, Q \in C(p^v)$, $P + (-Q) \in C(p^v)$ as $-Q \in C(p^v)$. $C(p^v)$ is non-empty as $\mathcal{O} \in C(p^v)$ for any p and v. Finally, as $C(p^v) \subset E(\mathbb{Q})$ it follows by the subgroup criterion [2, p. 47] that $C(p^v)$ is a subgroup of $E(\mathbb{Q})$. □

Note that the prime p and integer $v > 0$ were chosen arbitrarily. Hence, the lemma defines a large set of groups that are all subgroups of $E(\mathbb{Q})$. In addition, by the chain of inclusions (13) this means that for each prime p there exists a chain of subgroups.

Lemma 6.16. *For any prime p and integer $v > 0$, the quotient group $C(p^v)/C(p^{3v})$ is isomorphic to a subgroup of the quotient group $p^v R_p/p^{3v}R_p$.*





Proof. As has been shown in this chapter, for a point $P \in C(p^v)$ it follows that $t(P) \in p^v R_p$. In addition, in lemma 6.13 it was also seen that for $P, Q \in C(p^v)$, $t(P+Q) \in p^v R_p$. Consider equation (25). It was shown that $p^{3v} | \beta$ and hence, as $-t_3$ is the t-coordinate for $P+Q$, it follows that $t(P) + t(Q) - t(P+Q) \in p^{3v} R_p$. This leads to the congruence relation

$$t(P+Q) \equiv t(P) + t(Q) \pmod{p^{3v} R_p}. \tag{26}$$

That is, the mapping from $C(p^v)$ to the quotient group $p^v R_p / p^{3v} R_p$ where $P = (x, y) \mapsto t(P) = x/y$ is a group homomorphism. The kernel of this homomorphism are all the elements of $P \in C(p^v)$ such that $t(P) \in p^{3v} R_p$. By definition 6.7 this means that $P \in C(p^{3v})$. Thus, the mapping from the quotient group $C(p^v)/C(p^{3v})$ to $p^v R_p / p^{3v} R_p$, defined as above, is a one-to-one homomorphism. By the first isomorphism theorem [2, p. 97], $C(p^v)/C(p^{3v})$ is isomorphic to a subgroup of $p^v R_p / p^{3v} R_p$. □

Lemma 6.17. *For any prime p and integer $v > 0$, the quotient groups $p^v R_p / p^{3v} R_p$ and $R_p / p^{2v} R_p$ are isomorphic.*

Proof. Let $\varphi : R_p \to p^v R_p / p^{3v} R_p$ be the group homomorphism $x \mapsto xp^v + p^{3v} R_p$. Consider an element $r \in p^v R_p / p^{3v} R_p$, then $r = a/b + p^{3v} R_p$ where $a, b \in \mathbb{Z}$, $\gcd(a, b) = 1$, $\gcd(b, p) = 1$. If $a = 0$, then $x = 0$ is mapped onto r by φ. For $a \neq 0$ it follows that $\operatorname{ord}_p(a) \geq p^v$. To show that φ is surjective, there has to exist an element $x \in R_p$ such that $xp^v + p^{3v} R_p = a/b + p^{3v} R_p$, or equivalently $xp^v - a/b = (xbp^v - a)/b \in p^{3v} R_p$. As $p \nmid b$, it is sufficient that $p^{3v} | xbp^v - a$, i.e. $xbp^v \equiv a \pmod{p^{3v}}$. As $p^v | a$, it is equivalent to $xb \equiv a' \pmod{p^{2v}}$, where $a = a' p^v$. As $\gcd(b, p) = 1$, b has a multiplicative inverse $b^{-1} \in \mathbb{Z}$ modulo p^{2v}. Hence, all $x \in R_p$ such that $x \equiv a' b^{-1} \pmod{p^{2v}}$ are mapped onto r. As $a', b^{-1} \in \mathbb{Z}$, it follows that such an x exists in $\mathbb{Z}$. This shows that φ is a surjective mapping.

The kernel of φ are all elements $x \in R_p$ such that $xp^v + p^{3v} R_p \equiv 0 + p^{3v} R_p$, that is $p^{3v} | xp^v$, or $p^{2v} | x$. This is true whenever $\operatorname{ord}_p(x) \geq 2v$, note that $\operatorname{ord}_p(0) = \infty$ by definition. But this corresponds exactly to $p^{2v} R_p$. This proves the isomorphism in the lemma. □

Lemma 6.18. *For any prime p and integer $v > 0$, the quotient groups $R_p / p^{2v} R_p$ and $\mathbb{Z} / p^{2v} \mathbb{Z}$ are isomorphic.*

Proof. Let $\varphi : \mathbb{Z} \to R_p / p^{2v} R_p$ be the group homomorphism $x \mapsto x + p^{2v} R_p$. Consider any element $r \in R_p / p^{2v} R_p$ such that $r = a/b + p^{2v} R_p$, where $a, b \in \mathbb{Z}, \gcd(a, b) = 1$ and $\gcd(b, p) = 1$. For $a = 0$ it follows that $x = 0$ maps onto r. When $a \neq 0$, then showing that φ is surjective amounts to showing that $x + p^{2v} R_p \equiv a/b + p^{2v} R_p$, i.e. $x - a/b = (xb - a)/b \in p^{2v} R_p$, similarly as in the proof of lemma 6.17 this is equivalent to $xb \equiv a \pmod{p^{2v}}$, and since p does not divide b there exists a multiplicative inverse $b^{-1} \in \mathbb{Z}$ in this case also. Thus, all $x \in \mathbb{Z}$ such that $x \equiv ab^{-1} \pmod{p^{2v}}$ are mapped onto r. Hence φ is surjective.

The kernel of φ are all elements $x \in \mathbb{Z}$ such that $x + p^{2v}R_p \equiv 0 + p^{2v}R_p$, i.e. all x such that $p^{2v}|x$. This corresponds to the elements of the subgroup $p^{2v}\mathbb{Z}$. This proves the isomorphism in the lemma. □

Lemma 6.19. *For any prime p and integer $v > 0$, the quotient group $C(p^v)/C(p^{3v})$ is a cyclic group of order p^σ for some $0 \leq \sigma \leq 2v$.*

Proof. By lemma 6.17 and lemma 6.18 it follows that $p^vR_p/p^{3v}R_p$ is isomorphic to the cyclic group of order p^{2v}. By lemma 6.16, the group $C(p^v)/C(p^{3v})$ is isomorphic to a subgroup of $p^vR_p/p^{3v}R_p$. As all subgroups of a cyclic group are cyclic [2, p. 58], it follows that $C(p^v)/C(p^{3v})$ is a cyclic group and by Lagrange's theorem [2, p. 89] its order is p^σ for some σ such that $0 \leq \sigma \leq 2v$. □

Lemma 6.20. [1, p. 55]
(a) For every prime p, the subgroup $C(p)$ contains no points of finite order (other than $\mathcal{O}$)
(b) Let $P = (x, y) \neq \mathcal{O}$ be a rational point of finite order on $E(\mathbb{Q})$ as defined in theorem 6.1. Then x and y are integers.

Proof. (a) Let $P = (x, y) \neq \mathcal{O}$ be an element of order m. Assume for contradiction that $P \in C(p)$. This means that x and y are rational numbers where $\text{ord}_p(x) \leq -2$ and $\text{ord}_p(y) \leq -3$. Consider the case where $x = 0$, it follows by substitution in (7) that $y^2 = c$ where c is an integer. As y is a rational number by definition, $y^2 \in \mathbb{Z}$ implies that $y \in \mathbb{Z}$, and hence $P \notin C(p)$. Thus, $x \neq 0$. This means that $\text{ord}_p(x)$ must be a finite number. Also, by lemma 6.8 it follows that $y \neq 0$. Hence, it is possible to find some subgroups, $C(p^v)$ and $C(p^{v+1})$, of $C(p)$ such that $P \in C(p^v)$ but $P \notin C(p^{v+1})$ for some integer $v \geq 1$.

Using this observation, consider the case when p does not divide the order m. By repeated application of the congruence relation (26), one finds that $t(mP) \equiv mt(P) \pmod{p^{3v}R_p}$. By assumption $mP = \mathcal{O}$, so $t(mP) = t(\mathcal{O}) = 0$. This implies that $0 \equiv mt(P) \pmod{p^{3v}R_p}$. But since m does not contain any factor p, it must be the case that $0 \equiv t(P) \pmod{p^{3v}R_p}$. This in turn implies that $t(P) \in p^{3v}R_p$, which means that $P \in C(p^{3v})$. But by assumption $P \notin C(p^{v+1}) \supset C(p^{3v})$. Hence, $p \not| m$ leads to a contradiction.

Consider the case when $p|m$, so that $m = np$. Then consider $nP = P'$. By assumption, P has order $m = np$. It follows that $pP' = pnP = mP = \mathcal{O}$. If there is some positive integer $i < p$ such that i is the order of P', then $\mathcal{O} = iP' = inP$, which contradicts that $m = np$ is the order of P, hence P' has order p. As $P \in C(p)$, it follows that $nP = P' \in C(p)$. Once more it is possible to find an integer $v \geq 1$ such that $P' \in C(p^v)$ but $P' \notin C(p^{v+1})$. Using the same congruence relation it follows that $0 = t(\mathcal{O}) = t(pP') \equiv pt(P') \pmod{p^{3v}R_p}$. As multiplication by p increases the order of $t(P')$ by 1, it means that $0 \equiv t(P') \pmod{p^{3v-1}R_p}$, i.e. $t(P') \in p^{3v-1}R_p$ and hence $P' \in C(p^{3v-1})$. As $3v - 1 \geq v + 1$ (for $v \geq 1$), it follows that $P' \in C(p^{v+1}) \supset C(p^{3v-1})$, which contradicts $P' \notin C(p^{v+1})$. That is, both cases leads to a contradiction. It can therefore be concluded that when $P \in C(p)$, P cannot be of finite order. This proves (a).

(b) If $P = (x, y)$ has finite order m, by a) it was shown that $P \notin C(p)$ for any prime p. By lemma 6.5 it follows whenever $(x, y) \in E(\mathbb{Q})$ and p is a factor in the denominator of x or y, then p is a factor in the denominator of both, i.e. $(x, y) \in C(p)$. As this holds for all primes p, the denominator of x and y must be exactly 1 and hence both x and y must be integers. □

Finally, all the necessary lemmas for proving the Nagell-Lutz Theorem have been introduced. All this preparatory work allows for the following concise proof of the theorem.

Proof. By lemma 6.20 b) it follows that all points of finite order on $E(\mathbb{Q})$ must have integer coordinates, hence $x(P), y(P) \in \mathbb{Z}$. Whenever $y(P) = 0$ it follows by theorem 5.3 that P has order 2. When $y(P) \neq 0$, P generates a a group of finite order that contains $2P$. This means that the point $2P$ must also be a point of finite order, and hence $x(2P), y(2P) \in \mathbb{Z}$ by lemma 6.20 b). By lemma 6.2 this means that $y(P)|D$. This proves the Nagell-Lutz Theorem. □

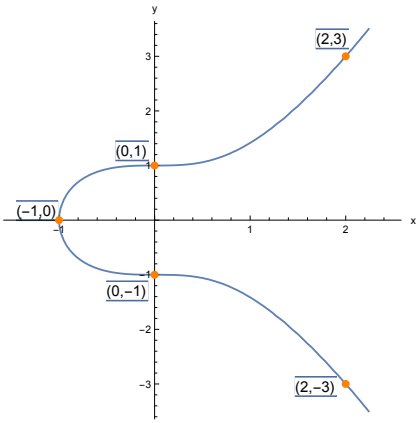
Example 6.21 (An application of Nagell-Lutz Theorem)**.** Consider the rational points on the elliptic curve $y^2 = x^3 + 1$. By equation (8) the discriminant of the polynomial is -27. By the Nagell-Lutz Theorem, it follows that any possible points of finite order should have y-coordinate $0, \pm 1, \pm 3, \pm 9$ or ± 27. The only point where $y = 0$, i.e. $(-1, 0)$, has order 2 by the theorem. For

the point $(x, 27)$, by substitution it follows that $x^3 = 27^2 - 1 = 728$, which does not have an integer cubic root and therefore it cannot be a point of finite order by the Nagell-Lutz Theorem. Similarly, the point $(x, 9)$ does not have integer coefficients, as $x^3 = 9^2 - 1 = 80$ which does not have an integer cubic root either. The arguments are symmetric for $(x, -27)$ and $(x, -9)$.

The potentially remaining points that can have finite order are $(0, \pm 1)$ and $(2, \pm 3)$. Consider the point $(0, 1)$, it follows that $2(0, 1) = (0, -1)$. It is easy to see that the point $(0, -1) + (0, 1) = \mathcal{O}$, hence the point $(0, 1)$ has order 3. Symmetrically, it can be seen that the point $(0, -1)$ also has order 3. Finally, consider $(2, 3)$, one finds that $2(2, 3) = (0, 1)$, which was shown to be a point of order 3. Hence, $6(2, 3) = 3(0, 1) = \mathcal{O}$, which means that $(2, 3)$ has order 6. Symmetrically, $(2, -3)$ also has order 6.

The Nagell-Lutz Theorem lets us conclude that the only points of finite order on $y^2 = x^3 + 1$ are the points $\mathcal{O}, (-1, 0), (0, \pm 1)$ and $(2, \pm 3)$, which are illustrated in figure 7.

Figure 7



By repeatedly applying the group addition, one can see that the element $(2, 3)$ generates all elements:

$$1(2,3) = (2,3) \qquad 2(2,3) = (0,1) \qquad 3(2,3) = (-1,0)$$

$$4(2,3) = (0,-1) \qquad 5(2,3) = (2,-3) \qquad 6(2,3) = \mathcal{O}$$

Hence, the group generated by $(2, 3)$ is isomorphic to the cyclic group of order 6. △

An interesting result about the rational points on an elliptic curves over $\mathbb{Q}$ is Mordell's Theorem, which can be formulated as:

Theorem 6.22 (Mordell's Theorem)**.** [1, p. 22] *Let E be an elliptic curve over $\mathbb{Q}$. If $E(\mathbb{Q})$ is non-empty, then $E(\mathbb{Q})$ is finitely generated.* □

The theorem essentially states that all points of $E(\mathbb{Q})$, as defined in the Nagell-Lutz Theorem, can be generated by a finite set of elements in $E(\mathbb{Q})$. As was shown in the preceding example, the element $(2, 3)$ generated a particular subgroup of the elliptic curve. The theorem of Mordell therefore states that there should be some additional finite set of points that generates the remaining points on the curve, i.e. $E(\mathbb{Q})$ is a finitely generated abelian group. By the fundamental theorem of finitely generated abelian groups [2, p. 158] it follows that

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \times Z_{n_1} \times Z_{n_2} \times ... \times Z_{n_s},$$

where $r, s, n_1, ...n_s \in \mathbb{Z}$ and $r \geq 0$, $n_j \geq 2$ for all $j \in [1, s]$ and $n_{i+1} | n_i$ for $i \in [1, s-1]$. The Nagell-Lutz Theorem allows one to find all elements of the *torsion subgroup*, i.e. the part $Z_{n_1} \times Z_{n_2} \times ... \times Z_{n_s}$. The integer r is called the *rank* of $E(\mathbb{Q})$. In the particular example, the torsion subgroup was generated by $(2, 3)$ and had *torsion order* 6. The following important theorem of Mazur gives the possible structures of the torsion subgroups of the rational points on an elliptic curve over $\mathbb{Q}$.

Theorem 6.23 (Mazur's Theorem)**.** [1, p. 58] *Let E be an elliptic curve over $\mathbb{Q}$ and suppose that $E(\mathbb{Q})$ contains a point of finite order m. Then either*

$$1 \leq m \leq 10 \qquad \text{or} \qquad m = 12.$$

More precisely, the set of all points of finite order in $E(\mathbb{Q})$ forms a subgroup which has one of the following two forms:
(i) A cyclic group of order N with $1 \leq N < 10$ or $N = 12$.
(ii) The product of a cyclic group of order two and a cyclic group of order $2N$ with $1 \leq N \leq 4$. □

Although the theorems of Mordell and Mazur gives a more general understanding of the torsion subgroup of the rational points on an elliptic curve over $\mathbb{Q}$, the Nagell-Lutz Theorem is a very effective tool in analyzing specific elliptic curves.

References

[1] Joseph H. Silverman and John Tate. *Rational points on elliptic curves*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1992.

[2] David S. Dummit and Richard M. Foote. *Abstract algebra*. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2004.

[3] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition, 2009.

[4] Joseph H. Silverman and John Tate. Errata and corrections to Rational points on elliptic curves. https://www.math.brown.edu/~jhs/RPEC/RPECErrata.pdf, 2011. [Online; accessed 19-Apr-2018].