

# Finite Groups on Elliptic Curves

Michael Carter Woodbury

July 11, 2003

## Abstract

An explanation of cubic curves in the projective plane and the reduction modulo  $p$  map from the set of rational solutions to solutions mod  $p$  is given. The map is then proven to be a homomorphism in general. Also proven, is the fact that if the prime number  $p$  does not divide twice the discriminant of the curve the map injects the torsion subgroup into the group of the curve mod  $p$ . This fact together with theorems by Mazur and Nagell and Lutz are then exploited in a computer search to find specific curves with each of the possible groups. Finally, the problem which Mazur's theorem resolves is discussed—specifically, it is shown that there does not exist any elliptic curve with torsion subgroup of order eleven.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Elliptic Curves</b>	<b>2</b>
2.1	Projective Geometry and Elliptic Curves . . . . .	2
2.2	$C(\mathbb{Q})$ is a Group . . . . .	5
<b>3</b>	<b><math>C(\mathbb{F}_p)</math>: Elliptic Curves modulo <math>p</math></b>	<b>6</b>
3.1	Mapping $C(\mathbb{Q})$ to $C(\mathbb{F}_p)$ . . . . .	7
3.2	The Possible Torsion Groups on Elliptic Curves . . . . .	8
<b>4</b>	<b>Why Not Order 11?</b>	<b>10</b>
4.1	Torsion Subgroups of $C(\mathbb{R})$ and $C(\mathbb{C})$ . . . . .	11
4.2	A Rational Torsion Group of Order 11 is Not Possible . . . . .	13
<b>A</b>	<b>Search Results and Frequency of Groups</b>	<b>14</b>

## 1 Introduction

The study of elliptic curves is closely connected to many different topics within mathematics. The subject leads naturally into geometry and algebra, and since complex valued solutions can be considered, complex analysis is also important.

In proving some of the fundamental theorems involving the rational points on elliptic curves (a number theoretic question) one needs the tools of real analysis.

As the topic of elliptic curves is extremely broad any one article could only hope to scratch the surface of some particular aspect. Here, the group of rational solutions is the main point of interest. Specifically, its torsion subgroup and the finite group associated with the curve as a reduced equation mod  $p$  are discussed. As it turns out, there is a deep connection between these two groups.

For the most part, this exposition is self contained—enough is given to prove that the reduction mod  $p$  map is a homomorphism, and all of the background needed to understand it is presented. However, proofs of the associativity of the addition law on elliptic curves, the Nagell–Lutz theorem, and the Mazur theorem are not given. For more information, refer to the bibliography.

## 2 Elliptic Curves

A general cubic curve is of the form:

$$f(x, y) = ax^3 + bx^2y + cxy^2 + dy^2 + ex^2 + fxy + gy^2 + hx + iy + j = 0 \quad (1)$$

Such a cubic is called *rational* if the coefficients  $a, b, c, d, e, f, g, h, i$  and  $j$  are rational numbers. The notation  $C(F)$  refers to the set  $\{(x, y) : x, y \in F \text{ and } f(x, y) = 0\}$ . Of course,  $F$  could be  $\mathbb{C}, \mathbb{R}, \mathbb{Q}$  or  $\mathbb{Z}$ . As such,  $C(\mathbb{Q})$  will be called the rational solutions on the curve  $C$ .

### 2.1 Projective Geometry and Elliptic Curves

By considering the cubic curves in the projective plane ( $\mathbb{P}^2$ ) we gain a great deal of freedom. In particular, a *projective change of coordinates* will reduce the equation to a simpler form. Also, we have a more complete equation of the curve. Most importantly, it is within the context of the projective plane that our map from  $C(\mathbb{Q})$  to  $C(p)$  makes sense. Here,  $C(p)$  is the set of solutions to the curve taken modulo  $p$ . What is meant by these statements will be made clear in this section.

**Definition 1 (The Projective Plane).**

$$\mathbb{P}^2 = \left\{ \frac{[a, b, c] : a, b, c \text{ not all zero}}{\sim} \right\}$$

where  $\sim$  is the equivalence relation  $(x, y, z) \sim (a, b, c)$  if and only if  $(x, y, z) = (ta, tb, tc)$  for some  $t \neq 0$ .

This definition suggests a natural form for the general cubic. First, note that given a cubic curve  $f(x, y) = 0$  the relation  $\sim$  applied to pairs  $(x, y) \in C(F)$  does not make sense, because in general  $(tx, ty) \notin C(F)$ . Also, if the constant term  $j$  of equation (1) is zero then  $(0, 0) \in C(F)$ . This contradicts the definition. We want to make sense this by adding an additional variable to the equation

$f(x, y)$  so that if  $(a, b, c)$  is a solution then  $(ta, tb, tc)$  will always be a solution when  $t \neq 0$ . Changing  $f(x, y)$  in this way is called *homogenization*. This process is explained via the following:

**Definition 2.**  $\mathbb{Q}[x, y, z]$  refers to the ring of polynomials. If  $g(x, y, z) \in \mathbb{Q}[x, y, z]$  then  $g(x, y, z) = a_1x^{i_1}y^{j_1}z^{k_1} + a_2x^{i_2}y^{j_2}z^{k_2} + \dots + a_nx^{i_n}y^{j_n}z^{k_n}$  and each  $a_ix^{i_1}y^{j_1}z^{k_1}$  is called an element of  $f$ .

**Definition 3.** Given a constant  $a$  and variables  $x, y, z$  ( $i, j, k$  non-negative integers) the degree of  $ax^iy^jz^k$  is  $i + j + k$

**Definition 4.** If  $g \in \mathbb{Q}[x, y, z]$  then the degree of  $g$  is equal to the  $\max\{\text{degree of each element of } f\}$ .

**Definition 5.** If  $f \in \mathbb{Q}[x, y]$  then the homogenization of  $f(x, y)$  is given by  $g(x, y, z)$  where a power of the variable  $z$  is added to each element of  $f(x, y)$  such that the degree of each element of  $g(x, y, z)$  equals the degree of  $f$ .

These definitions could be extended to any polynomial ring  $F[x_1, x_2, \dots, x_n]$  but for our purposes they are sufficient. The result of homogenization is that  $g(x, y, z)$  does not vanish at  $x, y$  or  $z = 0$ , and if  $(x_0, y_0, z_0)$  is a solution to  $g(x, y, z) = 0$  then  $(x_1, y_1, z_1) \sim (x_0, y_0, z_0) \Rightarrow (x_1, y_1, z_1)$  is a solution too. Additionally,  $g(x, y, 1) = f(x, y)$ . Thus, we have a method of going back and forth between the general cubic in  $\mathbb{R}^2$  and a cubic in  $\mathbb{P}^2$ .

The next step is to show how to go from a point in  $C(\mathbb{R})$  to a point in  $C(\mathbb{P})$  and vice versa. The first way is clear from the definitions: If  $(x, y) \in C(\mathbb{R})$  then  $[x, y, 1]$  is in  $C(\mathbb{P})$ . Now for  $[x, y, z] \in C(\mathbb{P})$  and  $c \neq 0$

$$[a, b, c] \sim \left[ \frac{a}{c}, \frac{b}{c}, 1 \right] \Rightarrow g\left(\frac{a}{c}, \frac{b}{c}, 1\right) = f\left(\frac{a}{c}, \frac{b}{c}\right) = 0 \quad (2)$$

So, as long as  $c \neq 0$  this is well defined. In order to understand what we do when  $c$  is zero, we must understand two more concepts about  $\mathbb{P}^2$ . These are the so-called *projective change of coordinates*, and the *line at infinity*.

A line in the projective plane is given by an equation  $L : \alpha x + \beta y + \gamma z = 0$ . For  $\gamma \neq 0$  we can go back and forth from a line in  $\mathbb{R}^2$  to a line in  $\mathbb{P}^2$ :

$$(a, b, c) \in L \text{ (and } c \neq 0) \Rightarrow \alpha\left(\frac{a}{c}\right) + \beta\left(\frac{b}{c}\right) + \gamma = 0$$

i.e.,  $\left(\frac{a}{c}, \frac{b}{c}\right)$  is a solution of  $\alpha x + \beta y + \gamma = 0$ , a line in  $\mathbb{R}^2$ .

This is the same process as in equation (2) above. Again we see that if  $c = 0$  we encounter a problem.

The equation  $z = 0$  is a special line in  $\mathbb{P}^2$  that has no counterpart in  $\mathbb{R}^2$ . This is called the *line at infinity*. There is a geometric reason for naming it so. Definition 1 can be visualized as the collection of lines in  $\mathbb{R}^3$  that pass through the origin. Each line in  $\mathbb{P}^2$  is a plane in  $\mathbb{R}^3$  that passes through the origin. Thus, the line in  $\mathbb{R}^3$  that we obtained above is just the intersection of the planes  $\alpha x + \beta y + \gamma z = 0$  and  $z = 1$ . The only plane through the origin that does not intersect  $z = 1$  is, of course,  $z = 0$ .

In the following sense  $z = 0$  is the line obtained from the limiting value of  $L : ax + by = 0$  as  $L \rightarrow \infty$ . First, we must determine how *big* a given line in the affine plane is. This gives us a way of determining if a line is *close* to infinity.

**Definition 6.** *The magnitude of a line  $L : \alpha x + \beta y + \gamma = 0$ , (or distance from the origin) denoted  $\|L\|$  is the  $\inf\{|(x, y)| : (x, y) \in L\}$ . The absolute value  $\text{sign}|(x, y)|$  refers to the standard Euclidean norm.*

Now let  $\{\alpha_n\}$  and  $\{\beta_n\}$  be sequences that go to zero as  $n \rightarrow \infty$ . The intersection of the plane (i.e. the *line* in the projective plane)  $\alpha_n x + \beta_n y + \gamma z = 0$  and the plane  $z = 1$  gives a sequence of lines  $L_n : \alpha_n x + \beta_n y + \gamma = 0$  as in definition 6. It is now clear that:

$$\|L_n\| \geq \min \left\{ \frac{1}{2} \left| \frac{\gamma}{\alpha_n} \right|, \frac{1}{2} \left| \frac{\gamma}{\beta_n} \right| \right\} \rightarrow \infty \quad \text{as } n \rightarrow \infty$$

In this sense, the lines in  $\mathbb{R}^2$  go to infinity as the projective lines  $L_n$  go to  $z = 0$ .

We can now understand solutions of the type  $(a, b, 0) \in C(\mathbb{P}^2)$ . Since  $a$  and  $b$  are not both zero we can divide by one (say  $a \neq 0$ ) to obtain:

$$\left( 1, \frac{b}{a}, 0 \right) \in C(\mathbb{P}^2)$$

Meanwhile, the equation of the cubic curve goes from general cubic:

$$f(x, y) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j$$

to the projective cubic:

$$g(x, y, z) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2z + fxyz + gy^2z + hxz^2 + iyz^2 + jz^3$$

changing variables  $(x, y, z) \mapsto (\bar{z}, \bar{x}, \bar{y})$ :

$$\tilde{g}(x, y, z) = d\bar{x}^3 + g\bar{x}^2\bar{y} + i\bar{x}\bar{y}^2 + j\bar{y}^2\bar{z} + c\bar{x}^2\bar{z} + f\bar{x}\bar{y}\bar{z} + h\bar{y}^2\bar{z} + b\bar{x}z^2 + e\bar{y}z^2 + a\bar{z}^3$$

back to the affine plane:

$$\tilde{f}(x, y) = d\bar{x}^3 + g\bar{x}^2\bar{y} + i\bar{x}\bar{y}^2 + j\bar{y}^2 + c\bar{x}^2 + f\bar{x}\bar{y} + h\bar{y}^2 + b\bar{x} + e\bar{y} + a$$

True, the equation for the elliptic curve  $C$  changes to  $\tilde{C}$ , but projectively speaking we are actually studying the very same curve. The above transformation of  $f$  into  $\tilde{f}$  is an example of a *projective transformation* or *projective change of coordinates*.

**Definition 7.** *a projective transformation of a curve  $C : g(x, y, z)$  in  $\mathbb{P}^2$  to a curve  $C' : g'(x, y, z)$  is any change of coordinates that can be given by an invertible matrix  $M = [m_{ij}]$  such that:*

$$\begin{aligned} x &= m_{11}x' + m_{12}y' + m_{13}z' \\ y &= m_{21}x' + m_{22}y' + m_{23}z' \\ z &= m_{31}x' + m_{32}y' + m_{33}z' \end{aligned}$$

Since  $M$  is invertible, if we understand  $C'$  we understand  $C$ . We also know that the group structure is preserved because this transformation is linear. (i.e. it sends lines to lines.)

## 2.2 $C(\mathbb{Q})$ is a Group

For any cubic curve in the general form a birational change of coordinates can be made so that the curve is given in the form of the following equation:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \quad (3)$$

A simple change of coordinates (essentially *completing the square*) gives:

$$y^2 = f(x) = x^3 + ax^2 + bx + c \quad (4)$$

*Completing the cube* results in:

$$y^2 = f(x) = x^3 + bx + c \quad (5)$$

By way of note for working with reduction mod  $p$  completing the square is only permissible when division by two is allowed, and to complete the cube division by three is required. Thus, when working with  $C(\mathbb{F}_p)$  this form is only valid for  $p \geq 3$ . In general, these are two *normal* forms—i.e. forms in which any elliptic curve can be written. We call either of these two forms *Weierstrass normal form*.

In Weierstrass form the solutions  $C(\mathbb{R})$  take one of two shapes, topologically speaking. The real solutions to  $C : y^2 = f(x)$  make sense only for  $x$  such that  $f(x) > 0$ . This can be understood by looking at the *discriminant* of the function  $f$ .

**Definition 8.** *The discriminant of a curve in Weierstrass normal form is*

$$\Delta = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2.$$

If  $\Delta$  is positive there are three distinct roots to the function  $f(x)$ . Hence,  $f(x)$  is positive on two connected components. If  $\Delta < 0$   $f(x)$  has only one real root, and the region on which  $f(x)$  is greater than zero is connected giving only one component.

In both cases it appears that the curve is unbounded to the right. This is true in the affine plane. However, in the projective plane the curve closes back up at infinity. (This point at infinity is important because it will be the *zero element* in the group.)

By the way, if  $\Delta = 0$  there are repeated roots. A cubic curve is called *non-singular* if there aren't repeated roots. Much of the theory of elliptic curves *does not* apply to singular cubics. Therefore, this case is not treated here.

The addition law for points on an elliptic curve is a matter of intersections of lines with the curve. When dealing with the curves of equations (4)–(5) this reduces to solving a third degree polynomial equation.

**Remark 1.** *On curves in Weierstrass form the following are true:*

- (a) *The graph of such curves is symmetric. (i.e. if  $(x, y) \in C(F)$  then  $(x, -y) \in C(F)$ .)*

- (b) *If two of the intersection points of a line and the curve are rational (or real) then the third intersection is too.*
- (c) *Counting multiplicities, a tangent line to an elliptic curve intersects it three times at the point of intersection if it is an inflection point, or twice if it is not.*
- (d) *The line tangent to the curve at the point at infinity is the line at infinity. It is also an inflection point.*
- (e) *Any point that intersects the curve at the point at infinity and somewhere else must be vertical. (This is true because the point at infinity corresponds to the vertical direction. In other words, the tangent line to  $y^2 = g(x)$  goes to the vertical line as  $x \rightarrow \infty$ .)*

With these remarks we can now understand the addition law. Given the above, it makes sense to let the point at infinity be an element in the set  $C(\mathbb{Q})$ , or for that matter in  $C(F)$  for any  $F$ . In fact, this will be the *identity element* and will be referred to as  $\mathcal{O}$ . Take any two points  $P, Q \in C(\mathbb{Q})$ . The line  $\overline{PQ}$  will now intersect the curve in another point  $R^*$ . Of course, this is an *algebraic* intersection. In other words, the number of intersections of the curve is exactly three, counting multiplicities. So the *third* point could possibly be either of the first two or even the *identity*. It is also possible for  $P$  and  $Q$  to be the same, in which case the line would simply be the tangent at that point.

Now we take  $R$  to be the third point of intersection on  $C(\mathbb{Q})$  of the line connecting  $R^*$  and  $\mathcal{O}$ , and  $P \oplus Q$  is defined to be the point  $R$ . From the remarks it is clear that  $P \oplus Q = \mathcal{O}$  if and only if  $P = (x, y)$  and  $Q = (x, -y)$ , hence taking inverses is easy. Also, the rule is obviously commutative, and adding  $\mathcal{O}$  to any element gives that element back. To prove that  $\oplus$  and  $C(\mathbb{Q})$  define a commutative group, the only step left is to prove associativity. This turns out to be the most difficult step, and is not given here. (Silverman and Tate [5] outline the steps of the proof in their book.)

### 3 $C(\mathbb{F}_p)$ : Elliptic Curves modulo $p$

At this point the group law works fine when considering the rational (or real) solutions. Determining specific equations that give the addition law defined in the previous section is very simple. However, *a priori* if we consider the equation

$$y^2 = x^3 + bx + c \pmod{p} \tag{6}$$

it is not at all clear that solutions form a group and that we can add points by simply using those equations (mod  $p$ ). The question of whether or not there exists a map (let alone a homomorphism) from  $C(\mathbb{Q})$  to  $C(\mathbb{F}_p)$  is not obvious either.

### 3.1 Mapping $C(\mathbb{Q})$ to $C(\mathbb{F}_p)$

Let  $\varphi : \mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$  be the map that takes  $x \mapsto x \pmod{p}$ . The fact that  $\varphi$  is a homomorphism will be exploited throughout this section, and  $\varphi(x)$  will be denoted by  $\tilde{x}$ .

Looking at the curve in the projective plane will be essential. What does it mean for a point on  $C(\mathbb{P})$  to be rational? Well, if  $[a, b, c] \in \mathbb{P}^2$  and  $ta, tb, tc \in \mathbb{Q}$  for some  $t \neq 0$  then  $[a, b, c]$  is called rational. The set of all such points will be denoted by  $\mathbb{P}^2(\mathbb{Q})$ . For  $C$ , a general homogenized cubic curve with all of the coefficients rational, there is consistency between the points on  $C(\mathbb{Q})$  and the *projective* rational points  $\mathbb{P}^2(\mathbb{Q})$  on the curve.

Let  $P = \left[ \frac{m_x}{n_x}, \frac{m_y}{n_y}, \frac{m_z}{n_z} \right] \in \mathbb{P}^2(\mathbb{Q})$ . Multiplying by  $n_x n_y n_z$  and dividing by the  $\gcd(m_x n_y n_z, m_y n_x n_z, m_z n_x n_y)$  gives a unique (up to sign) representation of  $P$ . Written in this way  $P$  is called a *normalized coordinate triple*. Hence, the following well defined map:

**Definition 9.**

$$p^* : \mathbb{P}^2(\mathbb{Q}) \rightarrow \mathbb{P}^2(\mathbb{F}_p)$$

is the map such that for  $[a, b, c]$  in normalized form  $p^*(a, b, c) = [\tilde{a}, \tilde{b}, \tilde{c}]$ . As with the map  $\varphi$  we will denote  $p^*(P)$  as  $\tilde{P}$ .

For a projective cubic equation with rational points we can normalize the coefficients in the same way, and reduce a curve  $C$  to  $\tilde{C}$  just like we did in definition 9. The *reduction mod p* of points in  $C(\mathbb{Q})$  to  $C(\mathbb{F}_p)$  is obtained by first normalizing  $C$ , then sending points on the new curve to points on the reduced curve  $\tilde{C}$  via  $p^*$ . Since  $\varphi$  is a homomorphism, if  $(a, b, c) \in C(\mathbb{Q})$  then  $(\tilde{a}, \tilde{b}, \tilde{c}) \in \tilde{C}(\mathbb{F}_p)$ . Thus, the following is true:

$$(C(\mathbb{Q}) \widetilde{\cap} L(\mathbb{Q})) \subset \tilde{C}(\mathbb{F}_p) \cap \tilde{L}(\mathbb{F}_p) \quad (7)$$

The group law on  $C$  is defined by the intersections of  $C$  with a line  $L$ , so here we are interested in  $\tilde{C}$  and its intersections with the line  $\tilde{L}$ . To verify that the map of the above paragraph is a homomorphism we must determine that the reduced intersection points agree with the intersections of the reduced curve. We need to show that in the case that all of the intersection points are rational, the above containment is an equality.

**Theorem 1.** *If  $C$  is a non-singular rational cubic curve in  $\mathbb{P}^2$  and  $\tilde{C}$  is non-singular then the reduction mod p map is a homomorphism.*

We prove theorem 1 via two lemmas:

**Lemma 1.** *Suppose  $C$  is a rational cubic curve in  $\mathbb{P}^2$  and  $L$  the line at infinity. Suppose that all three of the complex intersection points are rational. ( $L$  is not a component of  $C$ .) Let  $C \cap L = \{P_1, P_2, P_3\}$  where each point is repeated in the list as many times as its multiplicity. Then  $\tilde{C} \cap \tilde{L} = \{\tilde{P}_1, \tilde{P}_2, \tilde{P}_3\}$  with the correct number of multiplicities.*

**Lemma 2.** *For any projective line given by  $L : ax + by + cz = 0$  with  $a, b, c$  integers there is an integral transformation which sends  $L$  to the line at infinity.*

*Proof of Lemma 1.* Let  $L$  be the line at infinity, and  $g(x, y, z)$  a normalized equation for the cubic curve. In the notation of the theorem  $L \cap C = \{P_1, P_2, P_3\}$  and must be given by the equation  $g(x, y, 0)$ . Let  $P_i = [l_i, m_i, 0]$  in normalized coordinates. Then

$$g(x, y, 0) = c(m_1x - l_1y)(m_2x - l_2y)(m_3x - l_3y)$$

for some constant  $c$ . Since  $\tilde{L}$  is not a component of  $\tilde{C}$ ,  $\tilde{g}(x, y, 0)$  does not vanish, and since  $P_i$  is in normalized form  $c$  is not divisible by  $p$ . Hence,

$$\tilde{g}(x, y, 0) = \tilde{c}(\tilde{m}_1x - \tilde{l}_1y)(\tilde{m}_2x - \tilde{l}_2y)(\tilde{m}_3x - \tilde{l}_3y)$$

□

*Proof of Lemma 2.* From Definition 7 we see that we need a matrix  $[m_{ij}]$  that sends  $L : ax + by + cz = 0$  to  $L' : z' = 0$ . In order for the reduction mod  $p$  to make sense, however, all of the entries must all be integers, and the inverse matrix should also have integer values. Thus, the determinant must equal one.

For  $L$  in normal form  $\gcd(a, b, c) = 1$ . Let  $d = \gcd(b, c)$ , and choose integers  $r$  and  $s$  such that  $rc - sb = d$ . Note that  $r$  and  $s$  are relatively prime. Now  $\gcd(a, d) = 1$ , so choose  $t$  and  $u$  so that  $td - ua = 1$ , and since  $\gcd(r, s) = 1$  choose  $v$  and  $w$  such that  $vs - wr = u$ . Now, the matrix

$$\begin{bmatrix} t & u & v \\ 0 & r & s \\ a & b & c \end{bmatrix}$$

meets all of the necessary requirements. □

Lemma 2 implies that lemma 1 is true even in the case that  $L$  is not the line at infinity. The proof of theorem 1 is now a simple corollary.

*Proof of Theorem 1.* Let  $P$  and  $Q \in C(\mathbb{Q})$ , and let  $R = P \oplus Q$ . The addition law on elliptic curves tells us the how to proceed:  $P$  and  $Q$  give a line  $L_1$  and  $C \cap L_1 = \{P, Q, S\}$ . Now,  $\mathcal{O}$  and  $S$  give a line  $L_2$  such that  $C \cap L_2 = \{S, \mathcal{O}, R\}$ . By the lemmas, taking reduction mod  $p$  is allowed. Thus,  $\tilde{P} \oplus \tilde{Q} = \tilde{R}$ . □

### 3.2 The Possible Torsion Groups on Elliptic Curves

As stated above, the set of rational points on an elliptic curve forms a group with the addition law. In this section some of the facts regarding the group of rational points of finite order are discussed. The first theorem that tells us about what this group looks like is:

**Theorem 2 (Nagell–Lutz).** *Let  $C$  be an elliptic curve given in Weierstrass form with integer coefficients, and let  $P = (x, y) \in C(\mathbb{Q})$ . Then  $x$  and  $y$  are integers, and either  $y = 0$  or  $y^2$  divides the discriminant  $(\Delta)$ .*

Also, Mazur proved:

**Theorem 3 (Mazur).** *Let  $C$  be a non-singular rational cubic curve, and suppose that  $C(\mathbb{Q})$  contains a point of finite order  $m$ . Then the torsion group has one of the following forms:*

- (i) *A cyclic group of order  $N$  with  $1 \leq N \leq 10$  or  $N = 12$ .*
- (ii) *The product of a cyclic group of order two and a cyclic group of order  $2N$  with  $1 \leq N \leq 4$ .*

The Nagell–Lutz theorem not only tells us that points of finite order are integer points, but it also tells us that there are only finitely many possibilities for the  $y$ -coordinate on any given curve. This suggests a way of finding curves with each of the torsion groups specified by Mazur’s theorem. A search could be made of elliptic curves in Weierstrass form in a certain range, then, since the list of possible values for  $y$  such that  $(x, y)$  is a point of finite order is itself finite, one could check to see if there is a corresponding  $x$  on the curve which gives a point of finite order. For each  $y$  there are finitely many  $x$  that could yield a solution. This process *can* be done, but the number of steps involved is extremely cumbersome.

Also, it turns out that for nearly all curves the torsion group is trivial, and when it is not trivial it is one of the *small* groups. (i.e. a group of order 2 through 8) A more efficient algorithm for finding curves with the larger groups on them uses the following theorem which identifies when the reduction mod  $p$  map is *injective* from  $T$  to  $C(\mathbb{F}_p)$  where  $T$  is the torsion group of  $C(\mathbb{Q})$ .

**Theorem 4.** *Let  $C$  be an elliptic curve given in Weierstrass form with integer coefficients and discriminant  $\Delta$ . Let  $T$  be the group of all points of finite order. For any odd prime  $p$  such that  $p$  does not divide  $\Delta$ , the reduction mod  $p$  map is an injective homomorphism of  $T$  into  $C(\mathbb{F}_p)$ .*

*Proof.* Since  $p$  does not divide  $\Delta$ , the reduced curve is non-singular, and provided that  $p \neq 2$  the reduced curve is a group. Thus, we can apply theorem 1 to see that the reduction mod  $p$  map is a homomorphism. Next, we note that any rational point on an elliptic curve in Weierstrass form is of the type:  $(x, y) = (\frac{m}{e^2}, \frac{n}{e^3})$  with  $\gcd(m, n, e) = 1$ . Thus, in normalized form, we can write  $(x, y)$  as  $[me, n, e^2]$ . This immediately implies that  $(\frac{m}{e^2}, \frac{n}{e^3}) \mapsto \tilde{O}$  if and only if  $p$  divides  $e$ . Finally, theorem 2 tells us that if the point is of finite order then it is an integer point. (i.e.  $e = 1$ ) Thus, the kernel of the map from  $T$  to  $C(\mathbb{F}_p)$  consists of the identity element only.  $\square$

Of course, since  $T$  is a subgroup of  $C(\mathbb{Q})$ , this map is always a homomorphism. But theorem 4 gives us the additional information that it is injective. In which case, we know that the image of  $T$  is a subgroup of  $C(\mathbb{F}_p)$ , and the order of  $T$  must divide the order of  $C(\mathbb{F}_p)$ . Now, these facts can all be used to find elliptic curves with some of the *large* torsion groups via the following steps:

1. Choose a list of primes:  $\{p_1, \dots, p_n\}$ .

2. Pick a curve in some normal form: (equations (3)–(5)) Convert (if necessary) to Weierstrass normal form.
3. Start with the list of possible orders of the finite group.

$$L = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 16\}$$

4. Reduce the curve by the first prime  $p_1$ .
5. If the reduced curve is non-singular calculate the order of the group on the curve mod  $p$ . Call the order  $o$ .
6. Throw out any numbers in  $L$  that don't divide  $o$ .
7. If 9, 10, 12 or 16 is still on the list repeat steps 4–6 for the next prime. Otherwise, go back to step 2.
8. If 9, 10, 12 or 16 is still on the list after going through the entire list of primes, use Nagell–Lutz to calculate the group.

The major reason why this algorithm is faster than simply using Nagell–Lutz for each curve is that reducing mod  $p$  and finding the order of the group on the reduced curve are *fast* calculations, whereas doing Nagell–Lutz is relatively slow. Hence, this method eliminates the vast majority of curves before using Nagell–Lutz.

It is also of interest to note that the most effective normal form for finding the large groups is the form of equation (3). Although there are five parameters in equation (3) versus the three or two free parameters of (4) and (5) respectively, from Cremona [2] we see that we have isomorphisms via the following change of coordinates that allows us to take  $a_1, a_3 \in \{0, 1\}$  and  $a_2 \in \{-1, 0, 1\}$ :

$$\begin{aligned}\bar{a}_1 &= a_1 + 2s \\ \bar{a}_2 &= a_2 + sa_1 + 3r - s^2 \\ \bar{a}_3 &= a_3 + ra_1 + 2t \\ \bar{a}_4 &= a_4 - sa_3 + 2ra_2 - (t + rs)a_1 + 2r^2 - 2st \\ \bar{a}_6 &= a_6 + ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1\end{aligned}$$

Since  $s$ ,  $r$  and  $t$  are arbitrary we can always choose them so that our new coordinates are of the form mentioned above. In this form, searches over a comparable number of different curves resulted in a higher frequency of the large groups. See appendix A for details.

## 4 Why Not Order 11?

In the late 1970s Mazur closed the book on a long standing open question in the study of elliptic curves with his proof of theorem 3. Many mathematicians had already proved that the groups in Mazur's list were all possibilities, however, proving that other groups cannot appear on *any* elliptic curve was a long process.

In this section we explore why it is that there cannot be an elliptic curve with a rational torsion group of order 11. This specific question was first answered by Billing and Mahler [1] in 1940.

#### 4.1 Torsion Subgroups of $C(\mathbb{R})$ and $C(\mathbb{C})$

Thus far, the only solutions on the elliptic curves with which we have been interested are the real-valued solutions. We have even limited our treatment of this set of solutions to the subset of rational solutions. As noted earlier, however, the construction of the group law, although done geometrically, can be dealt with on a strictly algebraic level: i.e. explicit rational equations can be used to “add” and “subtract” points on the curve. It is just as well to take the complex valued solutions or *any* subfield. The real and rational solutions are just two such subfields.

The topological shape of the real solutions was mentioned in section 2.2. Since the addition law is clearly continuous the real solutions give rise to a *lie group*. The theory of lie groups tells us that if  $C(\mathbb{R})$  is topologically equivalent to a circle ( $S^1$ ) then the group must be isomorphic to the group of rotations on a circle of radius 1. If it is two circles ( $S^1 \times S^1$ ) the group must be isomorphic to the direct product of the circle group and the cyclic group of order two.

In each of the two cases, we know exactly what the subgroups of order eleven are on  $S^1$  and  $S^1 \times S^1$ . In either situation it can be represented by the complex points:

$$\{1, e^{2\pi i/11}, e^{4\pi i/11}, \dots, e^{10\pi i/11}\}$$

We see immediately that there *is* a  $T_{11} \subset C(\mathbb{R})$  of order eleven for any elliptic curve with. Likewise, there exists  $T_n$  of order  $n$  for any  $n$ . One amazing consequence of Mazur’s theorem, however, is that it is impossible for all of the points in  $T_{11}$  to be rational on any given curve  $C$ .

Just as we have been able to better understand  $C(\mathbb{R})$  via isomorphism, there is a homomorphism from  $\mathbb{C}^2$  to any elliptic curve. Remarkably, there is an explicit map that accomplishes this goal. (For  $C(\mathbb{R})$  the existence of an isomorphism of groups is simply a consequence of lie group theory—there is no known explicit map.)

In section 2.2 it was noted that by completing the cube equation (4) becomes equation (5). Precisely, this is done by substituting  $x - \frac{1}{3}a$  for  $x$ , and then replacing both  $x$  and  $y$  by  $4x$  and  $4y$ . The resulting equation is:

$$y^2 = 4x^3 - g_2x - g_3 \tag{8}$$

This equation becomes equation (5) by making one more substitution. As with the change of coordinates used in obtaining the other equations, this is just another normal form, and gives a certain representation to a given elliptic curve. This representation is closely related to the present topic, because we will construct the map (actually a homomorphism) from the complex plane ( $\mathbb{C}$ ) to an elliptic curve in the form of equation (8). Although not shown, this map is invertible.

First, let  $\omega_1$  and  $\omega_2$  be two complex numbers with distinct arguments. Then a lattice  $\Lambda$  is defined to be the set  $\{n\omega_1, m\omega_2 : n, m \in \mathbb{Z}\}$ .  $\omega_1$  and  $\omega_2$  are called *periods* and the function

$$\wp(z) = \frac{1}{z^2} \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \left( \frac{1}{(z - \omega)^2} - \frac{1}{\omega} \right)$$

is the so-called *Weierstrass  $\wp$ -function*.

**Remark 2.** *Almost magically, the following are true:*

- *The Weierstrass  $\wp$ -function converges for all  $z \in \mathbb{C}^2$*
- *It is periodic over the periods:  $\wp(z + \omega_i) = \wp(z)$   $i = \{1, 2\}$*
- *The quantities  $g_2$  and  $g_3$  are well-defined.*

$$g_2 = 60 \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega^4}, \quad g_3 = 140 \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega^6}$$

- *The Weierstrass  $\wp$ -function is differentiable and satisfies the differential equation:*

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3$$

- *The coefficients  $g_2$  and  $g_3$  uniquely determine  $\Lambda$ . Hence, each lattice corresponds to an elliptic curve. The converse is also true—each elliptic curve in the form of equation (8) corresponds to a lattice.*
- *For every complex number  $z$  we have the map*

$$P(z) = (\wp(z), \wp'(z))$$

*which is a homomorphism of groups sending points in  $\mathbb{C}$  to an elliptic curve.*

$$\frac{P}{\Lambda} \rightarrow C(\mathbb{C}) \text{ is an isomorphism.}$$

The final item taken with the rest of the facts tells us that the complex solutions of a curve are topologically equivalent to a torus. Hence, the cyclic subgroups of  $C(\mathbb{C})$  are very simple. One of any order, say  $n$ , can be generated by taking any complex number  $z$  such that  $nz \in \Lambda$ . Then, via the Weierstrass  $\wp$ -function these points can be mapped to the curve. In particular, a finite group of order eleven is generated by any element  $n\omega_1/11 + m\omega_2/11$  such that  $n, m \in \{0, \dots, 10\}$  not both 0.

So finite groups of order eleven abound on elliptic curves, but theorem 3 says that none of these consist of *rational* points only. This fact is demonstrated in the next section.

## 4.2 A Rational Torsion Group of Order 11 is Not Possible

The level of difficulty of Mazur's theorem [3, 4] is evident in the fact that the paper by Billing and Mahler [1] came much earlier than Mazur's proof. Already, mathematicians were trying to determine what the possible torsion groups could be, but the task would not be completed until forty years after the results of this section were first discovered. Here, the proof will be demonstrated by a sequence of three steps.

**Step 1.** Suppose we have an elliptic curve  $C$  in Weierstrass form such that the torsion subgroup of  $C(\mathbb{Q})$  is the cyclic group of order eleven. Throughout,  $\mathcal{O}$  will refer to the identity element, and  $P$  will be a generator of the group. Hence,  $\{2P, 3P, \dots, 10P = -P\}$  will be distinct points.

**Remark 3.** *As is clear from the definition of the addition law, three points  $P$ ,  $Q$  and  $R$  on an elliptic curve are co-linear if and only if  $P \oplus Q \oplus R = \mathcal{O}$ .*

If we start with the five points  $\{\mathcal{O}, P, 2P, 3P, 4P\}$  the rest of the points can be determined by intersecting lines. If  $\overline{PQ}$  is the line through the points  $P$  and  $Q$ , and  $L_1 \cap L_2$  is the point of intersection of the lines  $L_1$  and  $L_2$  then

$$\begin{aligned} -3P &= \overline{\mathcal{O} 3P} \cap \overline{P 2P} & -4P &= \overline{\mathcal{O} 4P} \cap \overline{P 3P} \\ -5P &= \overline{P 4P} \cap \overline{2P 3P} & -P &= \overline{\mathcal{O} P} \cap \overline{-3P 4P} \\ -2P &= \overline{\mathcal{O} 2P} \cap \overline{-P 3P} \end{aligned}$$

**Step 2.** We can always make a projective change of coordinates that sends  $\mathcal{O}$  to  $[1, 1, 1]$ . Remark 3 shows that this transformation can't leave  $P$ ,  $2P$  and  $3P$  all on a straight line. We can assume, therefore, that  $P = [1, 0, 0]$ ,  $2P = [0, 1, 0]$ ,  $3P = [0, 0, 1]$  and  $4P = [x, y, z]$ . Using the facts from step 1 we can deduce the values of all of the other points in terms of  $x, y$  and  $z$ .

$$\begin{aligned} -P &= [x - y + z, z, z] \\ -2P &= [x - y + z, z, x - y + z] \\ -3P &= [1, 1, 0] \\ -4P &= [x - y, 0, z - y] \\ -5P &= [0, y, z] \\ 5P &= [xy + xz - y^2, xz, y(x - y + z)] \\ 6P &= [(x - y + z)(x^2y - xy^2 + y^2z), z(y - z)(xy + xz - y^2), xz(y - z)(x - y + z)] \end{aligned}$$

**Step 3.** Note that  $6P = -5P$  in our group. Since  $x - y + z = 0$  would force  $2P$  and  $-2P$  to coincide, the above group will only hold if

$$x^2y - xy^2 - xz^2 + y^2z = 0. \tag{9}$$

Of course, this equation can be put into Weierstrass form. Doing so, gives the curve:

$$\tilde{y}^2 = \tilde{x}^3 - 432\tilde{x} + 8208 \tag{10}$$

Using theorem 2 we immediately determine that this curve has a torsion subgroup of order five. The points are:

$$\tilde{\mathcal{O}}, (-12, 108), (-12, -108), (24, 108), (24, -108)$$

In equation (9) these five points correspond to

$$[1, 1, 1], [1, 0, 0], [0, 1, 0], [0, 0, 1] \text{ and } [1, 1, 0]$$

It is a fact that this curve has rank equal to zero. (see [1]) Hence, these are the *only* rational points on the curve. Now, proving the theorem is simple.

**Theorem 5.** *If  $C$  is a rational non-singular cubic curve then the order of the torsion subgroup of  $C(\mathbb{Q})$  is not eleven.*

*Proof.* If such a curve did exist it has been shown that the points  $\mathcal{O}$ ,  $P$ ,  $2P$  and  $3P$  can be taken as  $[1, 1, 1]$ ,  $[1, 0, 0]$ ,  $[0, 1, 0]$  and  $[0, 0, 1]$  respectively. This forces  $-3P = [1, 1, 0]$ .  $4P$  must be a unique point  $[x, y, z]$  satisfying the elliptic curve  $x^2y - xy^2 - xz^2 + y^2z = 0$ . However, the only rational points on this curve are the five that have already been mentioned.  $\square$

## References

- [1] Billing, G., Mahler, K., “On Exceptional Points on Cubic Curves”, *London Journal of Mathematics*, **15** (1940), 32–43.
- [2] Cremona, J.E., *Algorithms for Modular Elliptic Curves*, Cambridge University Press, Cambridge, 1992.
- [3] Mazur, B., “Modular curves and the Eisenstein ideal”, *Institut des Hautes tudes Scientifiques. Publications Mathematiques*, **47** (1977), 33–186.
- [4] Mazur, B., “Rational isogenies of prime degree”, *Inventiones Mathematicae*, **47** (1978), 129–162.
- [5] Silverman, J., Tate, J., *Rational Points on Elliptic Curves*, Springer-Verlag, New York, 1992.

## A Search Results and Frequency of Groups

The method for finding curves with each of the groups of order  $\geq 9$  used each of the normal forms: equations (3)–(5). Table 1 corresponds to

$$\begin{aligned} y^2 &= x^3 + bx + c \\ a &\in (-10000, 10000) \\ b &\in (-100000, 1000000) \end{aligned}$$

Table 2 corresponds to

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

$$a, b, c \in (-1000, 1000)$$

Table 3 corresponds to

$$y^2 + a_1xy + a_2y = x^3 + a_2x^2 + a_4x + a_6$$

$$a_1, a_3 \in \{0, 1\} \quad a_2 \in \{-1, 0, 1\}$$

$$a_4 \in (-2000, 2000) \quad a_6 \in (-10000, 10000)$$

group	occurences	frequency	example
$\mathbb{Z}_9$	1	$2.500 \cdot 10^{-10}$	$y^2 = x^3 - 219x + 1654$
$\mathbb{Z}_{10}$	0	0	n/a
$\mathbb{Z}_{12}$	0	0	n/a
$\mathbb{Z}_6 \times \mathbb{Z}_2$	0	0	n/a
$\mathbb{Z}_8 \times \mathbb{Z}_2$	0	0	n/a

Table 1: Search curves  $y^2 = x^3 + bx + c$

group	occurences	frequency	example
$\mathbb{Z}_9$	3	$3.756 \cdot 10^{-10}$	$y^2 = x^3 + 15x^2 - 144x + 684$
$\mathbb{Z}_{10}$	2	$2.504 \cdot 10^{-10}$	$y^2 = x^3 + 31x^2 - 400x - 916$
$\mathbb{Z}_{12}$	0	0	n/a
$\mathbb{Z}_6 \times \mathbb{Z}_2$	6	$7.511 \cdot 10^{-10}$	$y^2 = x^3 + 10x^2 - 263x + 828$
$\mathbb{Z}_8 \times \mathbb{Z}_2$	0	0	n/a

Table 2: Search curves:  $y^2 = x^3 + ax^2 + bx + c$

group	occurences	frequency	example (in Weierstrass form)
$\mathbb{Z}_9$	2	$2.084 \cdot 10^{-9}$	$y^2 = x^3 - 3x^2 - 216x + 1872$
$\mathbb{Z}_{10}$	5	$5.210 \cdot 10^{-9}$	$y^2 = x^3 + x^2 - 720x + 5184$
$\mathbb{Z}_{12}$	1	$1.042 \cdot 10^{-9}$	$y^2 = x^3 - 3x^2 - 1944x + 110160$
$\mathbb{Z}_6 \times \mathbb{Z}_2$	4	$4.168 \cdot 10^{-9}$	$y^2 = x^3 + x^2 - 296x + 1680$
$\mathbb{Z}_8 \times \mathbb{Z}_2$	1	$1.042 \cdot 10^{-9}$	$y^2 = x^3 + x^2 - 17120x + 499968$

Table 3: Search curves:  $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$