INTRODUCTION TO ELLIPTIC CURVES

MATILDE LALÍN

ABSTRACT. These notes correspond to a mini-course taught by the author during the program "Two Weeks at Waterloo - A Summer School for Women in Math". Please send any comments or corrections to the author at mlalin@dms.umontreal.ca.

1. INTRODUCTION

One of the most popular questions that mathematicians ask is this: How do we find solutions to equations? For example, we can take $x^2 = 2$. In complex (or real) numbers, this equation has the solutions $\pm\sqrt{2}$ which we can find by making succesive approximations. If we look at the same equation for the rational numbers, we write $\left(\frac{a}{b}\right)^2 = 2$ with a and b integers. This equation does not have solutions. To see this, suppose that there is a solution. After simplification, we can assume that either a or b is (or both are) odd. The equation can be written as $a^2 = 2b^2$. Now this shows that 2 divides a^2, which means that a is even. But then $a = 2a_1$ and we can write $4a_1^2 = 2b^2$. This simplifies to $2a_1^2 = b^2$, which implies that b is even, which gives a contradiction.

Number theory is the study of the integers $\mathbb{Z}$, the rational $\mathbb{Q}$, and their properties. The above analysis shows an example of one of the big subjects of study in Number Theory, which is the search for integral or rational solutions to polynomial equations. This subject is known as *Diophantine arithmetic*, in honor of Diophantus, a Greek mathematician who lived in Alexandria in the third century A. D.

The simplest possible case of Diophantine equations are linear equations. For instance, $ax + by = c$ with $a, b, c \in \mathbb{Z}$ and we look for solutions $x, y \in \mathbb{Z}$. This case is well understood in the sense that one can tell a priori if there are solutions, and if there are solutions, one can give a list of them. See Appendix 6.1 for more information about linear equations.

Here is another example of a Diophantine equation. Fix an integer $n \geq 3$ and ask which integers X, Y, and Z satisfy the equation

$$X^n + Y^n = Z^n. \tag{1.1}$$

Equivalently, we could divide everything by Z^n and define new variables $x = X/Z$ and $y = Y/Z$ to rewrite the equation as

$$(X/Z)^n + (Y/Z)^n = 1 \text{ or } x^n + y^n = 1.$$

We see that searching for integer solutions (X, Y, Z) of the original equation is equivalent to searching for rational solutions (x, y) of the latter equation.

A French lawyer named Fermat conjectured in 1689 that equation (1.1) did not have integral solutions. This result was finally proved in 1995 by Sir Andrew Wiles, an English mathematician.

If we consider the case $n = 2$, the situation changes completely. The resulting equation has infinitely many solutions such as $(3, 4, 5)$ and $(5, 12, 13)$. These are called *Pythagorean triples*. See Appendix 6.2 for more information on quadratic equations.

***Exercise* 1.** Show that neither of the equations $x^2 + y^2 = -3$ and $x^2 + y^2 = 1003$ has a solution with integers x, y.

Key words and phrases. elliptic curves.

So far we have stated that linear equations and quadratic equations are well understood, while there is an equation of degree n that was very difficult to understand. The natural question is: When do equations start to get complicated in the sense that we do not have a general theory for understanding the solutions? The answer is degree 3.

Consider the following example: fix a rational number k. *Bachet's equation* is

$$y^2 - x^3 = k.$$

What are the *rational numbers* x and y that satisfy this equation? Suppose that (x, y) is a solution to the above equation with $y \neq 0$. Then one can verify by direct computation that

$$\left(\frac{x^4 - 8kx}{4y^2}, \frac{-x^6 - 20kx^3 + 8k^2}{8y^3}\right)$$

is also a solution of the Bachet equation. This is called the *duplication formula*. If x and y are rational numbers, the coordinates of the above expression will be rational numbers as well. We can start with one rational solution (x_0, y_0), substitute x_0 for x and y_0 for y into the above formula to get a new rational solution (x_1, y_1), then substitute x_1 for x and y_1 for y to get yet another rational solution (x_2, y_2) to the Bachet equation, and so on. In fact, if $k \neq 1, -432$, one can generate infinitely many rational solutions in this way.

***Exercise* 2.** What happens if we apply the duplication formula several times force
(a) $(2, 3)$ in the case of $k = 1$ and
(b) $(12, 36)$ in the case of $k = -432$?

Let us consider $k = -2$, i.e.,

$$y^2 - x^3 = -2.$$

It is easy to verify that $(3, 5)$ is a solution. Substituting into the duplication formula, we obtain

$$\left(\frac{129}{100}, \frac{-383}{1000}\right) \text{ and } \left(\frac{2340922881}{58675600}, \frac{113259286337279}{449455096000}\right),$$

which are also solutions.

What are the *integer numbers* x and y that satisfy this equation? This is actually a much harder problem; Axel Thue showed in 1908 that there are only *finitely many* integer solutions.

***Exercise* 3.** Show that

$$y^2 - x^3 = 7$$

has no solution with integers x, y. (Hint: Write $y^2 + 1 = x^3 + 8 = (x + 2)(x^2 - 2x + 4)$ and study the remainders of x and y under division by 4. Prove that there must be a prime number p dividing $y^2 + 1$ such that p is of the form $4k + 3$.)

How can one find a formula such as the duplication formula? Suppose $P = (x_0, y_0)$ is a rational solution to the equation which we may write as $y^2 = x^3 + k$. Now draw the tangent line to the curve through P and look at the point $Q = (x_1, y_1)$ at which it intersects the curve again. Since a line should, in principle, intersect a cubic in three points, and a tangency is counted as a "double intersection", such a point Q should exist (although it might be P again!) See Figure 1.

Explicitly, the tangent line L through P has equation $y = \lambda x + \nu$, where λ is the slope. To find the slope to the curve, we use implicit differentiation on the Bachet equation to get $2y\frac{dy}{dx} = 3x^2$ or $\frac{dy}{dx} = \frac{3x^2}{2y}$. So the slope λ at P is $\frac{3x_0^2}{2y_0}$, and $\nu = y_0 - \lambda x_0 = y_0 - \frac{3x_0^3}{2y_0} = \frac{2y_0^2 - 3x_0^3}{2y_0}$. To find the coordinates of Q, we intersect the line $y = \lambda x + \nu$ with the curve $y^2 = x^3 + k$. Substituting for y, we get:

$$(\lambda x + \nu)^2 = x^3 + k$$

or

$$x^3 - \lambda^2 x^2 - 2\nu\lambda x + (k - \nu^2) = 0.$$

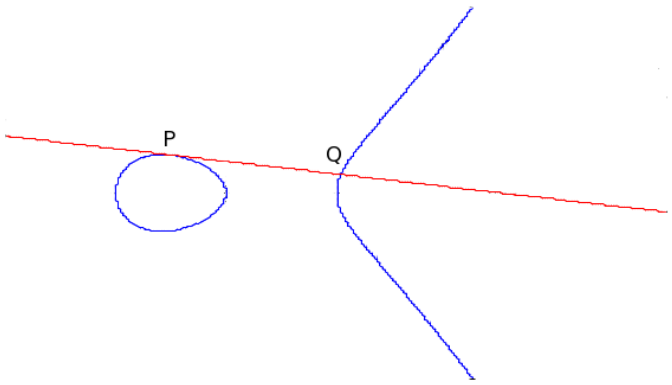

FIGURE 1. The tangent line at P intersects the curve in a "third" point Q.

Two of the solutions for x are given by x_0, since the line intersects the curve "twice" (tangency) at P. Now observe that the sum of the roots of the equation is the negative of the coefficient of x^2, so $x_0 + x_0 + x_1 = \lambda^2$, or

$$x_1 = \lambda^2 - 2x_0 = \frac{9x_0^4}{4y_0^2} - 2x_0 = \frac{9x_0^4 - 8x_0y_0^2}{4y_0^2}.$$

From the Bachet equation, $x_0^3 = y_0^2 - k$, or $x_0^4 = xy_0^2 - kx_0$, so the above expression becomes

$$\frac{8x_0^4 + x_0^4 - 8x_0y_0^2}{4y_0^2} = \frac{8x_0y_0^2 - 8kx_0 + x_0^4 - 8x_0y_0^2}{4y_0^2} = \frac{x_0^4 - 8kx_0}{4y_0^2},$$

which is the first coordinate of the duplication formula. To get the second coordinate, observe that $y_1 = \lambda x_1 + \nu$ and replace the values of λ, x_1, and ν to obtain y_1.

***Exercise* 4.** Complete the last steps to prove that

$$y_1 = \frac{-x^6 - 20kx^3 + 8k^2}{8y^3}.$$

In conclusion, given a solution to the Bachet equation, one can generate another solution by means of the simple geometric procedure described above.

Our goal is to show that we can do something even more general. Namely, given any plane curve $f(x, y) = 0$ defined by a cubic equation, we can make the rational points on the curve into a *group*.

2. Some facts of Group Theory

A group is a structure that appears often in number theory. It will be central to our study.

Definition 2.1. A *group* is a set G of objects together with an operation $* : G \times G \longrightarrow G$ such that:

(1) If $a, b, c \in G$, then $(a * b) * c = a * (b * c)$; that is, if we multiply three elements of G together, it does not matter which pair we multiply first. We say that the operation is *associative*.
(2) There exist an element $e \in G$ such that for every $a \in G$, we have $a * e = a = e * a$. This element is called the *identity* of G for $*$.
(3) For every $a \in G$ there exist a $b \in G$ such that $a * b = b * a = e$. We say that b is the *inverse* of a with respect to $*$. Sometimes we write $-a$ or a^{-1} for the inverse.

Examples.

- $\mathbb{Z}$ is a group with respect to addition (with 0 as the identity), but not with respect to multiplication (only 1 and -1 have inverses).

- Fix an integer n. The set of *residue classes mod n*, written $\mathbb{Z}_n$ is defined by $\{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}$. We can define a binary operation on $\mathbb{Z}_n$ by adding classes "modulo n". For example, if $n = 5$, then $\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{4}\}$ and we can define $\overline{1} + \overline{2} = \overline{3}$, $\overline{3} + \overline{4} = \overline{2}$ (because $3 + 4 = 7$, which is the same as 2 modulo 5). It is easy to verify that addition is associative, that $\overline{0}$ is the (only) identity element and that the inverse of $\overline{a}$ is $\overline{n-a}$.

  When there is no ambiguity, we drop the bar from the notation, writing $0, 1, \ldots$ instead of $\overline{0}, \overline{1}, \ldots$.
- The set $GL_2(\mathbb{R})$ of invertible 2 by 2 matrices with real entries forms a group under matrix multiplication. Indeed, if A and B are invertible matrices, then AB is also invertible with inverse $B^{-1}A^{-1}$. Once again, the matrix I_2 is the (only) identity element and the inverse of an element $A \in GL_2(\mathbb{R})$ is just the inverse matrix A^{-1}.
- If G_1, G_2 are two groups with operations $*_1$ and $*_2$, then we can form another group by considering $G_1 \times G_2$ with operation
$$(a, b) * (c, d) = (a *_1 c, b *_2 d).$$
In this case the identity is given by (e_1, e_2), and $(a, b)^{-1} = (a^{-1}, b^{-1})$.

  This construction can be extended to several groups $G_1, G_2, \ldots, G_n$.

  Examples of this construction are the real spaces $\mathbb{R}^n$ with the operation sum.

***Exercise* 5.** Prove that the identity element for G with operation $*$ is unique and that the inverse of any element is also unique.

Definition 2.2. A group G is called *abelian* or *commutative* if for every $a, b \in G$, we have $a * b = b * a$.

For example, $\mathbb{Z}$ with respect to addition is abelian. However, $GL_2(\mathbb{R})$ is not abelian. For example,
$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -1 & 2 \end{pmatrix} \text{ but } \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Remark.
For groups G, we often write ab instead of $a * b$ for the binary operation. If the group G is abelian, we usually write $a + b$ instead of $a * b$, label the identity element as 0 instead of e, and $-a$ for the inverse of a.

Definition 2.3. Let G be a finite group. The *cardinality* or *order* of G, denoted $|G|$, is the number of elements in G.

Definition 2.4. Let G be a group and $a \in G$ an element. The *order* of a, written $|a|$, is the smallest positive integer m (if it exists) such that $a^m = e$. If no such integer exists, we say that a has infinite order.

For example, the order of the identity element of any group is always 1; in fact, it is the only element of order 1. In $\mathbb{Z}$, every nonzero element has infinite order. In $\mathbb{Z}_n$, every element has order dividing n.

***Exercise* 6.** Show that in a group G, $|a| = |a^{-1}|$.

Proposition 2.5. *Let G be a finite group and $a \in G$ an element. Then $|a|$ divides $|G|$.*

Definition 2.6. Let G be a group. A *subgroup* of G is a subset $H \subseteq G$ which is a group with the same $*$ operation coming from G.

For example, $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$ has four subgroups: $\{0\}$, $\{0, 3\}$, $\{0, 2, 4\}$, $\{0, 1, 2, 3, 4, 5, \}$.

We also have
$$\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C},$$
with the sum.

Proposition 2.7. *Let G be a group. A subset $H \subseteq G$ is a subgroup if and only if H is nonempty, closed under the group operation, and closed under inverses.*

***Exercise* 7.** Let G be an *abelian group* and $n \geq 1$ an integer. Prove that the set

$$G_n = \{x \in G : |x| \text{ divides } n\}$$

is a subgroup of G.

Definition 2.8. A group G is said to be *generated* by a subset $S \subseteq G$ if every element of G can be obtained by composing some (finite) combination of elements of S and their inverses together.

For example, the group $\mathbb{Z}$ is generated by $\{1\}$ and so is $\mathbb{Z}_n$.

A group which has a generating subset consisting of one element is called a *cyclic* group. A group which has a finite generating subset is called a *finitely generated* group.

Consider the group $\mathbb{Z} \times \mathbb{Z}$. This group is abelian and not cyclic. It is finitely generated with a set of generators given by $\{(1,0),(0,1)\}$.

***Exercise* 8.** Find another set of generators for $\mathbb{Z} \times \mathbb{Z}$.

Definition 2.9. Let G and G' be groups. A *homomorphism* is a map $\phi : G \longrightarrow G'$ such that for all $x, y \in G$,

$$\phi(xy) = \phi(x)\phi(y).$$

An important example of a homomorphism is the *identity map* $id_G : G \longrightarrow G$ from a group to itself. It is defined by $id_G(g) = g$ for all $g \in G$.

For another example, consider the "doubling map" $D : \mathbb{Z} \longrightarrow \mathbb{Z}$ given by the formula $D(n) = 2n$. Then, for any $a, b \in \mathbb{Z}$, we have

$$D(a+b) = 2(a+b) = 2a + 2b = D(a) + D(b)$$

and so D is a homomorphism.

Definition 2.10. Let G and G' be groups. A homomorphism $\phi : G \longrightarrow G'$ is called an *isomorphism* if there exists a homomorphism $\psi : G' \longrightarrow G$ such that $\phi \circ \psi = id_{G'}$ and $\psi \circ \phi = id_G$. In this case we say that G and G' are *isomorphic* as groups and we write $G \cong G'$.

Groups which are isomorphic are considered to be essentially the same. For example, consider the set of rotations of the plane $G_4 = \{id, \rho_{\pi/2}, \rho_\pi, \rho_{3\pi/2}\}$, where id is the map which leaves everything fixed and ρ_θ is the counterclockwise rotation of the plane around the origin through an angle of θ radians. The set G_4 forms a group under composition of transformations; id is clearly the identity, and $\rho_{2\pi-\theta}$ is the inverse of ρ_θ. Furthermore, the group G_4 is isomorphic to $\mathbb{Z}_4$. One such isomorphism is obtained by sending id to 0, $\rho_{\pi/2}$ to 1, ρ_π to 2 and $\rho_{3\pi/2}$ to 3. The inverse map is defined by switching these definitions, and also defines a homomorphism.

***Exercise* 9.** Suppose G, H, and J are groups. Let $\phi : G \longrightarrow H$ and $\psi : H \longrightarrow J$ be homomorphisms. Prove that the composition $\psi \circ \phi : G \longrightarrow J$ is also a homomorphism.

***Exercise* 10.** (a) Let G be a group and $a, b \in G$ elements. Prove that

$$(ab)^{-1} = b^{-1}a^{-1}.$$

(b) Suppose G is a group. Prove that the function

$$\iota : G \longrightarrow G$$

defined by $\iota(g) = g^{-1}$ is a homomorphism if and only if G is an abelian group.

We close this section with the following result which will be very useful in the future.

Theorem 2.11. *[Fundamental Theorem of Finitely Generated Abelian Groups] Any finitely generated abelian group G can be written in the form*

$$G \cong \mathbb{Z}^r \times \mathbb{Z}_{q_1} \times \cdots \times \mathbb{Z}_{q_s},$$

where r is a non-negative integer and the q_i are (not necessarily distinct) powers of prime numbers. The value of r, and those of $q_1, \ldots, q_s$ are (up to reordering) uniquely determined by G.

3. Elliptic Curves and the Group Law

3.1. **The Projective Space.** The 2-*dimensional affine space over* $\mathbb{R}$ is given by

$$\mathbb{A}^2_{\mathbb{R}} = \{(x, y) : x, y \in \mathbb{R}\}.$$

In other words, this is just another notation for the Eucliden plane $\mathbb{R}^2$. This notation is useful because it allows us to replace $\mathbb{R}$ by $\mathbb{C}$ and define

$$\mathbb{A}^2_{\mathbb{C}} = \{(x, y) : x, y \in \mathbb{C}\}.$$

In fact, we could do the same thing for other sets such as $\mathbb{Q}$, $\mathbb{Z}$, etc.

Now consider the set $L = \{(X, Y, Z) \ : \ X, Y, Z \in \mathbb{R} \ \text{ and } \ X, Y, Z \text{ not all zero }\}$. (This set is essentially $\mathbb{A}^3_{\mathbb{R}} - \{(0, 0, 0)\}$.)

Define an equivalence relation on L by setting that $(A, B, C) \sim (\lambda A, \lambda B, \lambda C)$ for all nonzero real numbers λ. For instance, $\left(\frac{1}{4}, \frac{1}{2}, \frac{2}{3}\right)$ is equivalent to $(3, 6, 8)$ (use $\lambda = 12$).

The set $L/\sim$ of equivalence classes with respect to $\sim$ is called 2-*dimensional projective space over* $\mathbb{R}$ and is denoted $\mathbb{P}^2_{\mathbb{R}}$. This object is described as 2-dimensional because we started L, an essentially 3-dimensional object, and considered equivalence classes with respect to a linear (1-dimensional) relation. The equivalence class of (A, B, C) in $\mathbb{P}^2_{\mathbb{R}}$ is typically written $[A : B : C]$ to avoid confusion with the affine space $\mathbb{A}^3_{\mathbb{R}}$.

Thus, we have

$$\mathbb{P}^2_{\mathbb{R}} = \{[X : Y : Z] \ : \ X, Y, Z \in \mathbb{R} \ \text{ and } \ X, Y, Z \text{ not all zero }\}.$$

Consider the subsets

$$U_1 = \{[X : Y : Z] \ : \ Z \neq 0\} \text{ and } U_0 = \{[X : Y : 0]\}.$$

It is clear from the definition of $\sim$ that no element of U_1 is equivalent to any element of U_0.

A typical element $[X : Y : Z]$ of U_1 is equivalent to $\left[\frac{X}{Z} : \frac{Y}{Z} : 1\right]$. So the elements of U_1 are essentially of the form $[x : y : 1]$, where x and y are allowed to be any real numbers. In other words, U_1 is somewhat like a copy of $\mathbb{A}^2_{\mathbb{R}}$ sitting inside $\mathbb{P}^2_{\mathbb{R}}$.

A typical element of U_0 looks like $[X : Y : 0]$. Note that either X or Y is nonzero (since we excluded the triple consisting of all zeros). If $X \neq 0$, then $[X : Y : 0]$ is equivalent to $\left[1 : \frac{Y}{X} : 0\right]$, which is essentially a copy of $\mathbb{A}^1_{\mathbb{R}}$. If $X = 0$, then the typical point has the form $[0 : Y : 0]$, which is equivalent to $[0 : 1 : 0]$ since Y is nonzero.

Thus, U_0 is a union of $\mathbb{A}^1_{\mathbb{R}}$ and the point $[0 : 1 : 0]$. We can also think of U_0 as a copy of the *projective line over* $\mathbb{R}$ given by

$$\mathbb{P}^1_{\mathbb{R}} = \{[X : Y] \ : \ X, Y \in \mathbb{R} \ \text{ and } \ X, Y \text{ not both zero }\},$$

where $[X : Y]$ corresponds again to the equivalence class given by $(X, Y) \sim (\lambda X, \lambda Y)$.

In conclusion: the set $\mathbb{P}^2_{\mathbb{R}}$ is a disjoint union of $\mathbb{A}^2_{\mathbb{R}}$ and $\mathbb{P}^1_{\mathbb{R}}$. The former consists of all (equivalence classes of) points of the form $[X : Y : Z]$ with $Z \neq 0$ and the latter consists of (equivalence classes of) points of the form $[X : Y : 0]$. The latter set is often referred to as the "line at infinity", since it is one-dimensional and is not part of $\mathbb{A}^2_{\mathbb{R}}$.

3.2. **Weierstrass Equations.** Our goal is to study cubic equations in two variables. The most general cubic equation that one may consider is:

$$ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0. \tag{3.1}$$

Here $a, b, c, d, e, f, g, h, i, j$ are numbers (for instance, real numbers). The previous equation can be simplified. For instance, we can replace x by $x + k$, and obtain another similar equation. If we choose k wisely, we may make some of the coefficients equal zero. More generally, a birational transformation is a change of variables of the form $x_1 = \ell(x, y)$, $y_1 = m(x, y)$ where $\ell(x, y)$, $m(x, y)$ rational functions, and such that it is reversible with a transformation of the same shape.

One can prove that any cubic curve (3.1) is birationally equivalent an equation of the form:

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Such an equation is called Weierstrass equation. If we are working * in $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$ we can further assume that

$$y^2 = x^3 + ax^2 + bx + c. \tag{3.2}$$

*If the characteristic of the field is different from 2.

Moreover, by making the birational change [†] $x = x_1 - \frac{a}{3}$, we can eliminate the x^2 term and assume that the equation has the shape

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

We will work with equation (3.2) keeping in mind that we can assume $a = 0$ if necessary.

Now, instead of looking for solutions in $\mathbb{A}^2_{\mathbb{R}}$, consider solutions in the bigger set $\mathbb{P}^2_{\mathbb{R}}$. For this, recall that we can think of $\mathbb{A}^2_{\mathbb{R}}$ inside $\mathbb{P}^2_{\mathbb{R}}$ as the set of $\left[\frac{X}{Z} : \frac{Y}{Z} : 1\right]$. Setting $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$, and multiplying by Z^3, equation (3.2) becomes

$$Y^2Z = X^3 + aX^2Z + bXZ^2 + cZ^3. \tag{3.3}$$

Equation (3.3) has the same solutions (if any) as equation (3.2) plus possibly some solutions in the line at infinity. Let E be the set of solutions of (3.3)

$$E = \{[X : Y : Z] \ : \ Y^2Z = X^3 + aX^2Z + bXZ^2 + cZ^3\}.$$

To see the solutions at the line of infinity, set $Z = 0$. Substituting, we get $0 = X^3$, which forces $X = 0$; Y is allowed to be anything different from zero. All the points $[0 : Y : 0]$ are equivalent to $[0 : 1 : 0]$. Thus, E intersects the line at infinity at the point $[0 : 1 : 0]$.

Equation (3.3) is a *homogenization* of equation (3.2).

The moral: we can consider the curve E as the set of solutions to the so-called *Weierstrass equation* $y^2 = x^3 + ax^2 + bx + c$, together with a "point at infinity" $\mathcal{O} = [0 : 1 : 0]$, which sits outside the affine plane.

3.3. **When is a cubic an elliptic curve?** We will see that (under certains conditions) we can make the points on E into a group $E(\mathbb{R})$, and that the points with *rational* coordinates form a subgroup $E(\mathbb{Q})$ of this group.

Definition 3.1. Let E be a cubic curve with Weierstrass equation $y^2 = f(x) = x^3 + ax^2 + bx + c$. Let $F(x, y) = y^2 - f(x)$. The curve E is called *nonsingular* if there is no point on E at which the partial derivatives

$$\frac{\partial F}{\partial x} = -f'(x) \text{ and } \frac{\partial F}{\partial y} = 2y$$

vanish simultaneously.

A nonsingular curve of this form is called an *elliptic curve*.

What does this condition mean? From the equation $y^2 = f(x)$, we can try to find a formula for the slope $\frac{dy}{dx}$ of the tangent line to E at (x, y). Using implicit differentiation, we get $2ydy = f'(x)dx$ and

$$\frac{dy}{dx} = \frac{f'(x)}{2y} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}.$$

This expression makes sense if the denominator is nonzero, and can be interpreted as the "slope" of a vertical line if the numerator is nonzero and the denominator zero. However, if both quantities vanish, there is not a well-defined slope at that point.

Now we translate this condition in terms of algebra. Suppose that both derivatives vanish at a point $P = (x_0, y_0)$. We have $f'(x_0) = 0$ and $y_0 = 0$. However, since $y^2 = f(x)$, it follows that $f(x_0) = 0$. So we have both $f(x_0) = 0$ and $f'(x_0) = 0$. Now we use the following result.

Lemma 3.2. *A polynomial $p(x)$ has a repeated root α if and only if α is a common root of $p(x)$ and $p'(x)$.*

Proof. Suppose $p(x)$ and $p'(x)$ share a common root α. Since α is a root of $p(x)$, we can write

$$p(x) = (x - \alpha)g(x).$$

Differentiating, we have

$$p'(x) = g(x) + (x - \alpha)g'(x).$$

Evaluating in $x = \alpha$, we obtain $0 = g(\alpha)$. Therefore we can write $g(x) = (x - \alpha)h(x)$ and

$$p(x) = (x - \alpha)^2 h(x).$$

[†]This change can be made if the characteristic of the field is different from 3.

This means that α is a repeated root of $p(x)$.

Conversely, suppose some root α of $p(x)$ occurs with multiplicity ≥ 2. In other words,

$$p(x) = (x-\alpha)^2 h(x).$$

Using the product rule again, we have

$$p'(x) = 2(x-\alpha)h(x) + (x-\alpha)^2 h'(x) = (x-\alpha)(2h(x) + (x-\alpha)h'(x)),$$

and so α is also a root of $p'(x)$. ☺

It follows that $y^2 = f(x) = x^3 + ax^2 + bx + c$ is singular if and only if $f(x)$ has a repeated root α.

***Exercise* 11.** Prove that the conic given by the equation

$$x^2 - 3xy + 2y^2 - x + 1 = 0$$

is nonsingular.

There is a way to tell that a Weierstrass equation $y^2 = f(x) = x^3 + ax^2 + bx + c$ is nonsingular without having to find the roots of $f(x)$.

Definition 3.3. The *discriminant* of $f(x)$ is the quantity

$$\Delta = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2. \tag{3.4}$$

(Note that if $a = 0$, this reduces to $-4b^3 - 27c^2$, a formula that you may have seen before in the context of cubic equations.)

If we factor $f(x)$ over the complex numbers:

$$f(x) = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3),$$

it is not hard to check (by brute force computation) that

$$\Delta = (\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2(\alpha_2 - \alpha_3)^2.$$

It is then clear that the roots are not repeated if and only if $\Delta \neq 0$.

***Exercise* 12.** Verify the formula given above for the discriminant.

***Exercise* 13.** Let $f(x) = x^3 + ax^2 + bx + c$ have real coefficients. Prove that if $\Delta > 0$, then $f(x)$ has three real roots, while if $\Delta < 0$, then $f(x)$ has a real root and two complex conjugate roots.

To summarize, a Weierstrass equation $E : y^2 = f(x) = x^3 + ax^2 + bx + c$ is nonsingular (it corresponds to an elliptic curve) iff any of the following equivalent conditions is satisfied.

(1) At least one of the partial derivatives does not vanish at each point of E.
(2) The right hand side $f(x)$ of the Weierstrass equation for E has distinct roots in $\mathbb{C}$.
(3) The discriminant Δ given by equation (3.4) is different from zero.

3.4. **The Group Law.** From now on, we are going to assume that $E : y^2 = x^3 + ax^2 + bx + c$ is a nonsingular curve with Weierstrass equation in the above form.

Let $P, Q \in E$ be two points in the curve. For convenience, we suppose for the time being that $P \neq Q$ and that neither point is equal to $\mathcal{O}$, the point at infinity. We are going to "add" these points to get a new point. Let us take the line through P and Q and see where it intersects E. We call this third point of intersection $P * Q$. One could ask: does the operation $*$ make the points of E into a group? The answer is no, in fact, there is no identity element. However, if we define $P + Q$ to be the reflection of $P * Q$ in the x-axis (i.e. if $P * Q = (x, y)$, then $P + Q = (x, -y)$). If $P = Q$, then we should use the tangent line to P in place of the line between P and Q. Furthermore, if one of P and Q is $\mathcal{O}$, we declare $P + \mathcal{O} = P$ and $\mathcal{O} + Q = Q$. The resulting operation $+$ makes E a group.

It is easy to see that $+$ defines a binary operation on E and that $\mathcal{O}$ is the identity element. What is the inverse of a point $P = (x, y)$? Well, from the geometry one can see that $-P = (x, -y)$, is also a point of E, and that $P + (-P) = \mathcal{O} = (-P) + P$. The only property that is missing is associativity of $+$. Using

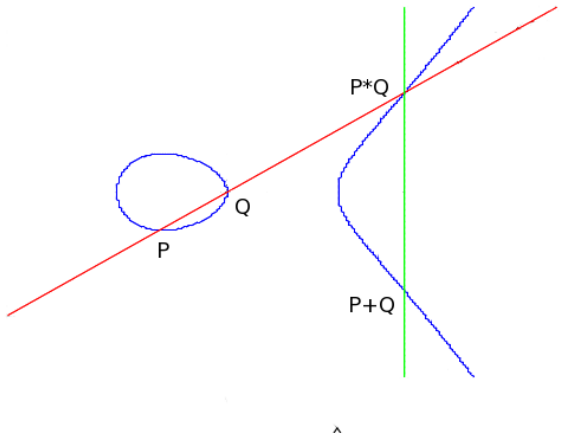

FIGURE 2. How to sum P and Q in the elliptic curve E.

the explicit formulas we are about to develop, one can check that $+$ is associative, but this proof is very tedious since there are many cases to consider. There are more elegant proofs that use algebraic geometry or complex analysis but we do not have the backgroup to explore them here.

Just as we did for the Bachet curve, we may also develop explicit formulas for the group law on E. Here we will sketch how to do it for distinct points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$, neither of which is the point $\mathcal{O}$ at infinity. If we label $P * Q = (x_3, y_3)$, then $P + Q = (x_3, -y_3)$ by our definition. Now, (x_3, y_3) is the third point of intersection of the line $y = \lambda x + \nu$ connecting P and Q to the curve E. We may calculate the slope of this line by using the slope formula

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1},$$

and then

$$\nu = y_1 - x_1\lambda = y_1 - \frac{x_1(y_2 - y_1)}{x_2 - x_1}.$$

Now, we substitute this expression for y into the Weierstrass equation for E to get:

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

or

$$x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x + (c - \nu^2) = 0.$$

As before, we note that the three roots of the equation are x_1, x_2, and x_3, and we obtain

$$\lambda^2 - a = x_1 + x_2 + x_3. \tag{3.5}$$

This allows us to calculate x_3. We may then find y_3 by substituting x_3 for x in the equation of the line connecting P and Q.

***Exercise* 14.** Find explicit formulas for x_3 and y_3 in terms of $x_1, x_2, y_1, y_2, a, b, c$.

In the case $P = Q = (x, y)$ with $y \neq 0$, the formula for the x-coordinate of $P + Q = 2P$ is

$$x(2P) = \frac{f'(x)^2}{4f(x)} - a - 2x = \frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4x^3 + 4ax^2 + 4bx + 4c}. \tag{3.6}$$

***Exercise* 15.** Verify the "duplication" formula above.

Now we will consider the case where $P = (x_1, y_1)$ and $Q = \mathcal{O}$. For this, consider the homogenization of the equation

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

In this system, we can think of P as $[x_1 : y_1 : 1]$ and Q as $[0 : 1 : 0]$. To simplify things, suppose that $y_1 \neq 0$. Thus, $P = \left[\frac{x_1}{y_1} : 1 : \frac{1}{y_1}\right]$. It then makes sense to divide by Y and consider

$$\bar{z} = \bar{x}^3 + a\bar{x}^2\bar{z} + b\bar{x}\bar{z}^2 + c\bar{z}^3 \tag{3.7}$$

where $\bar{x} = \frac{X}{Y}$ and $\bar{z} = \frac{Z}{Y}$. In the system of coordinates $(\bar{x}, \bar{z})$, we have that $P = \left(\frac{x_1}{y_1}, \frac{1}{y_1}\right)$ and $Q = (0, 0)$. We get $\lambda = \frac{1}{x_1}$ and $\nu = 0$, thus the line through P and $Q = \mathcal{O}$ is given by $\bar{z} = \frac{\bar{x}}{x_1}$. Going back to the $[X : Y : Z]$ and the (x, y) coordinates, we obtain $\frac{Z}{Y} = \frac{X}{x_1 Y}$ and $x = x_1$, which is the vertical line[‡]. It is now clear that $P * \mathcal{O} = (x_1, -y_1)$ and that $P + \mathcal{O} = P$.

Other cases are treated similarly.

***Exercise* 16.** (a) Find the tangent at $\mathcal{O}$ in the $(\bar{x}, \bar{z})$-coordinates.

(b) Find the third point of intersection of this tangent with E. What is $2\mathcal{O}$?

The final expressions for x_3 and y_3 are rational functions of $x_1, x_2, y_1,$ and y_2. In particular, if these latter four quantities are rational (i.e. P and Q are rational points), then x_3 and y_3 will be rational and $P + Q$ will be also a rational point. This shows that the set of rational points is closed under the group composition law.

It is also clear that if $P = (x, y)$ is a rational point, then $-P = (x, -y)$ must also be a rational point. By Proposition 2.7, the subset

$$E(\mathbb{Q}) = \{(x, y) : \ y^2 = x^3 + ax^2 + bx + c : \ x, y \text{ rational } \} \cup \{\mathcal{O}\}$$

forms a subgroup of E.

***Exercise* 17.** Consider the point $P = (3, 8)$ on the cubic curve $y^2 = x^3 - 43x + 166$. Compute P, $2P$, $3P$, $4P$, and $8P$. Comparing $8P$ with P, what can you conclude?

3.5. **Points of Order Two and Three.** A natural question which arises in this context is the following. Consider an elliptic curve E given by a Weierstrass equation $y^2 = f(x) = x^3 + ax^2 + bx + c$. We want to find the points $P = (x, y)$ of order 2. In other words, we want $P \neq \mathcal{O}$ but $2P = \mathcal{O}$. By adding $-P$ to both sides of the equation, we may write $P = -P$. Using our formula for the inverse of P, this now reads $(x, y) = (x, -y)$. Thus $y = -y$ and so $y = 0$. We look at the Weierstrass equation

$$0 = y^2 = x^3 + ax^2 + bx + c.$$

The number of points of order 2 depends on how many real roots the cubic on the right has. Certainly every cubic has at least one real root, and the other two roots are either two real numbers or complex conjugates (this can be determined by the sign of the discriminant). So, in terms of real coordinates, $E(\mathbb{R})$ has either one or three points of order 2. If we allow complex coordinates, $E(\mathbb{C})$ has three points of order 2. In sum allowing complex coordinates, E has *four* points of order *dividing* 2 (including $\mathcal{O}$).

As an example, let us look at E given by the Weierstrass equation $y^2 = x^3 + 8$. The cubic on the right, $x^3 + 8 = (x + 2)(x^2 - 2x + 4)$, has only one real root $x = -2$. The other two roots are given by the quadratic formula: $x = 1 \pm \sqrt{-3}$. So $E(\mathbb{R})$ has one point of order two and $E(\mathbb{C})$ has three points of order 2.

On the other hand, if we look at E given by $y^2 = (x - 1)(x - 2)(x - 3)$, then the right hand side has three real roots, so in this case $E(\mathbb{R})$ has 3 points of order 2.

[‡]This line also works when $y_1 = 0$. In this case the line is tangent to P.

Now let us look at points of order 3. The relation $3P = \mathcal{O}$ is equivalent to $2P = -P$. In particular, this means that $x(2P)$ and $x(-P)$ are the same, but since $x(-P) = x(P)$, this just means that $x(2P) = x(P)$. If we abbreviate $x(P)$ by x, by using formula (3.6), we get

$$x(2P) = \frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4x^3 + 4ax^2 + 4bx + 4c} = x,$$

and

$$\psi_3(x) = 3x^4 + 4ax^3 + 6bx^2 + 12cx + (4ac - b^2) = 0.$$

Thus, $P = (x, y)$ is a point of order three if and only if x satisfies the above equation.

How many points $P = (x, y)$ of order three are there? The polynomial $\psi_3(x)$ has at most four roots, so there are at most four possibilities for x. For each point $P = (x, y)$ of order 3, it is easy to check that $-P = (x, -y)$ also has order 3, so these points come in pairs. Now, P and $-P$ are distinct points as long as $y \neq 0$. However, we can safely assume $y \neq 0$, because all points with $y = 0$ have order 2, as we saw above.

Therefore, the number of points of order three is equal to twice the number of *distinct* roots of $\psi_3(x)$.

Proposition 3.4. *The polynomial $\psi_3(x)$ has four distinct roots.*

Proof. Recall how we obtained the polynomial $\psi_3(x)$. We set $x(2P) = x(P)$. By equation (3.6) this can be written as

$$x = \frac{f'(x)^2}{4f(x)} - a - 2x$$

and

$$\frac{f'(x)^2}{4f(x)} = 3x + a = \frac{f''(x)}{2}$$

so

$$\psi_3(x) = 2f(x)f''(x) - f'(x)^2.$$

To check that $\psi_3(x)$ has distinct roots, it suffices by Lemma 3.2 to check that $\psi_3(x)$ and $\psi_3'(x)$ have no common roots. However,

$$\psi_3'(x) = 2f(x)f'''(x) = 12f(x).$$

A common root of $\psi_3(x)$ and $\psi_3'(x)$ would be a common root of $2f(x)f''(x) - f'(x)^2$ and $12f(x)$, which would be a common root of $f(x)$ and $f'(x)$. However, this is not possible, because it would contradict the assumption of the nonsingularity of E.

***Exercise* 18.** Let E be an elliptic curve given by the usual Weierstrass equation

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

(a) Prove that

$$\frac{d^2y}{dx^2} = \frac{2f''(x)f(x) - f'(x)^2}{4yf(x)} = \frac{\psi_3(x)}{4yf(x)}.$$

(b) Use this to deduce that a point $P = (x, y) \in E$ is a point of order three if and only if $P \neq \mathcal{O}$ and P is a point of inflection on the curve E.

(c) Now suppose $a, b, c \in \mathbb{R}$. Prove that $\psi_3(x)$ has exactly two real roots, say α_1, α_2 with $\alpha_1 < \alpha_2$. Prove that $f(\alpha_1) < 0$ and $f(\alpha_2) > 0$. Use this to deduce that the points in $E(\mathbb{R})$ of order dividing 3 form a cyclic group of order three.

We may summarize our findings as follows:

Proposition 3.5. *Let E be an elliptic curve. Then E has exactly 8 points (allowing complex coordinates) of order 3 and 9 points of order dividing 3.*

Notice that, allowing complex coordinates, there is 1 point of order dividing 1, 4 points of order dividing 2, 9 points of order dividing 3. The following theorem (whose proof is beyond the scope of this discussion) should come as no surprise:

Theorem 3.6. *Let E be an elliptic curve and n a positive integer. Then the number of points on E of order dividing n is equal to n^2. In fact, this set of points forms a subgroup of E isomorphic to $\mathbb{Z}_n \times \mathbb{Z}_n$.*

4. The Lutz-Nagell Theorem

Let E be an elliptic curve with Weierstrass equation $y^2 = f(x) = x^3 + bx + c$ (recall that we can assume that $a = 0$) assume further that b, c are *rational* numbers. Write $b = b_1/b_2$, $c = c_1/c_2$ where the numerator and denominator of each expression is an integer, and let d be a large integer which is a multiple of each of the denominators b_2 and c_2. (For example, $d = b_2c_2$.) Making a change of variables, we set $\bar{x} = d^2x$ and $\bar{y} = d^3y$. Then our equation becomes:

$$\bar{y}^2 = \bar{x}^3 + d^4b\bar{x} + d^6c.$$

In particular, each of the coefficients is an *integer*. Therefore, we can assume from now on that our Weierstrass equation has integer coefficients.

Our goal is to state a theorem, due to Lutz and Nagell, that gives us a recipe for finding *all* the rational points of finite order. In particular, it also tells us that there are only finitely many such points.

Definition 4.1. The rational points of finite order are called *torsion* points. They make a group that is denoted $E(\mathbb{Q})_{\text{tors}}$.

Recall that the discriminat Δ of $f(x)$ given by equation (3.4). Since $a = 0$, we have a simpler expression given by

$$\Delta = -4b^3 - 27c^2.$$

***Exercise* 19.** Show that Δ is invariant by the change of variables $x = x_1 - \frac{a}{3}$ that makes the coefficient a equal 0.

A direct computation shows that we have the following explicit formula:

$$\Delta = ((3x^3 - 5bx - 27c)(x^3 + bx + c) - (3x^2 + 4b)(x^4 - 2bx^2 - 8cx + b^2). \tag{4.1}$$

We have the following lemma:

Lemma 4.2. *Let $P = (x, y)$ be a point on E such that both P and $2P$ have integer coordinates. Then either $y = 0$ or $y^2|\Delta$.*

Proof. We start by assuming that $y \neq 0$ and prove that $y^2|\Delta$. Because $y \neq 0$, we know that P does not have order 2 and therefore that $2P \neq \mathcal{O}$, so we may write $2P = (x_1, y_1)$. By assumption, x, y, x_1, y_1 are all integers. The duplication formula (3.6) asserts that

$$x_1 = \frac{x^4 - 2bx^2 - 8cx + b^2}{4y^2}.$$

Since x and x_1 are integers, it follows that y^2 divides $x^4 - 2bx^2 - 8cx + b^2$. We also have that $y^2 = f(x)$. Now we use relation (4.1) to conclude that y^2 divides Δ, too, as desired.
☺

The main statement is the following.

Theorem 4.3. *(Lutz–Nagell)*
Let

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

be an elliptic curve with integer coefficients a, b, c. Let Δ be the discriminant of the cubic polynomial. Let $P = (x, y)$ be a rational point of finite order. Then x and y are integers. If $y = 0$ then P has order 2; otherwise, y^2 divides Δ.

To prove the Lutz–Nagell theorem, it suffices to show that x and y are integers. The last statement will then follow from Lemma 4.2. The proof that x and y are integers is not hard but computationally complicated and we will not include it here.

Remark.
Note that the Lutz–Nagell Theorem is not an "if and only if" statement: a point may have integer coordinates without having finite order.

Now suppose we are trying to classify all the points $P = (x, y)$ of finite order on some elliptic curve E. We already know about the points of order 2. The Lutz–Nagell Theorem tells us that we after computing the value of the discriminant Δ (from the Weierstrass equation), the only y-values that we need to consider are *integers* such that their squares divide Δ. There are obviously only finitely many such values, so we just need to comb through the list and see which ones give rise to rational points lying on E.

Example.

Find all the torsion points on the curve $y^2 = x^3 + 4x$.

First, it is clear that the point $\mathcal{O}$ at infinity, which has projective coordinates $[0 : 1 : 0]$, is a rational point. The points of order 2 are determined by setting $x^3 + 4x = 0$, i.e. $x(x^2 + 4) = 0$. This yields one point: $(0, 0)$ (since $(2i, 0)$, $(-2i, 0)$, where $i = \sqrt{-1}$ are not rational points).

The discriminant yields $\Delta = -256$. Now Lutz–Nagell Theorem tells us that any rational point $P = (x, y)$ which does not have order 1 or 2 must have x and y both integers and y^2 dividing -256. Since $256 = 2^8$, the possibilities for y are

$$y = \pm 1, \pm 2, \pm 4, \pm 8, \pm 16.$$

We only have to test the positive values for y, since $P = (x, y)$ is a rational point if and only if $-P = (x, -y)$ is a rational point.

First consider the case $y = 1$, then we have $1 = x^3 + 4x$ for *integers* x. Factoring, we get $1 = x(x^2 + 4)$, and it is clear that no integer satisfies this equation. If $y = 2$, we get $4 = x(x^2 + 4)$; there are no solutions to this equation either. If $y = 4$, we get $16 = x(x^2 + 4)$. Since x has to be an integer, x has to divide 16. The possibilities are $x = \pm 1, \pm 2, \pm 4, \pm 8, \pm 16$. We can eliminate the negative values instantly, since they would make the right hand side of the equation negative. We find that the only value that satisfies this equation is $x = 2$. This gives the point $(2, 4)$, along with its inverse $(2, -4)$. If $y = 8$, we get $64 = x(x^2 + 4)$. Again, we can argue that x must divide 64 and that x must be positive. By repeated trial, we find that none of these values satisfy this equation. Finally, if $y = 16$, we get $256 = x(x^2 + 4)$. Once again, we find that x must be positive and divide 256, and that no such values satisfy this equation.

Thus, the integral points on E with $y^2|\Delta$ are:

$$G = \{\mathcal{O}, (0, 0), (2, 4), (2, -4)\}$$

To find the torsion points, we need to know which of these points have finite order. We know $\mathcal{O}$ has order 1 and $(0, 0)$ has order 2, and that the other two elements have order greater than 2. A simple application of the duplication formula (3.6) yields that $2(2, 4) = (0, 0)$. Thus, $(2, 4)$ and $(2, -4)$ are points of order 4. We conclude that the group of torsion points $E(\mathbb{Q})_{\text{tors}}$ is in fact isomorphic to $\mathbb{Z}_4$.

As a final note, Lutz–Nagell Theorem gives a bound (namely the square-root of the discriminant $\sqrt{\Delta}$) on the y-coordinate of a torsion point on the curve E. This bound depends on E since Δ does. One could also ask: exactly which orders are possible? This question was studied for a long time. The answer was finally provided by Barry Mazur in the following (very difficult) theorem:

Theorem 4.4. *(Mazur) Let E be an elliptic curve defined over $\mathbb{Q}$. (i.e. the coefficients in the Weierstrass equation for E may be chosen to lie in $\mathbb{Q}$.) Suppose that $E(\mathbb{Q})$ contains a point of finite order m. Then*[§]

$$1 \leq m \leq 10 \quad \textit{or} \quad m = 12.$$

***Exercise* 20.** For each of the following curves, determine all of the points of finite order, and determine the order of each such point. (a) $y^2 = x^3 - 2$.

(b) $y^2 = x^3 + 1$.

(c) $y^2 - y = x^3 - x^2$.

[§]Indeed, more is known, $E(\mathbb{Q})_{\text{tors}} \cong \mathbb{Z}_m$ with $1 \leq m \leq 10$ or $m = 12$, or $\mathbb{Z}_2 \times \mathbb{Z}_{2k}$ with $k = 1, 2, 3, 4$.

5. Mordell's Theorem

The result we are after is the following:

Theorem 5.1. *(Mordell) Let E be an elliptic curve whose Weierstrass equation has rational coefficients. Then the group $E(\mathbb{Q})$ is finitely generated.*

This theorem states that given just finitely many (rational) points on E, one can get *any* rational point on E by adding or subtracting some combination of these (finitely many) points.

The Fundamental Theorem of Finitely Generated Abelian Groups 2.11 implies the following.

Corollary 5.2. *Let E be an elliptic curve whose Weierstrass equation has rational coefficients. Then*

$$E(\mathbb{Q}) \cong E(\mathbb{Q})_{\text{tors}} \times \mathbb{Z}^r,$$

where r is a nonnegative integer (called rank).

For example, an elliptic curve will have an infinite number of points over $\mathbb{Q}$ (corresponding to solutions in the cubic equation) if and only if $r > 0$. We have seen that torsion points can be found relatively easy. The rank, on the other hand, is not so easy to find. ¶

We will prove a weak version of Mordell's theorem, for the case where the elliptic curve has all the four points of order 2 defined over $\mathbb{Q}$, i.e., the case where $f(x)$ has three integral roots.

5.1. Cosets: more abstract algebra.

Definition 5.3. Let G be a group and $H \subseteq G$ a subgroup. Let $a \in G$ be any element. The set

$$aH = \{ah : h \in H\}$$

is called a *coset* of H in G. The element a is called a *representative* for this coset.

As an example, let $G = \mathbb{Z}$ and $H = 2\mathbb{Z} = \{2n : n \in \mathbb{Z}\} = \{\text{all even integers}\}$. Then H is a subgroup of G. We could consider the coset $0 + H$ (we write $0 + H$ instead of $0H$ since we use the addition symbol '+' instead of multiplication to represent the group law). As a set $0 + H = \{0 + h : h \in H\}$ is simply H, the even integers. On the other hand, $1 + H = \{1 + h : h \in H\}$ is the set of all odd integers. What is $2 + H$? Actually $2 + H = H$, also. More generally, $n + H$ is the set of even numbers if n is even and the set of odd numbers if n is odd. Notice how there are only two distinct cosets of H in G.

Definition 5.4. Let G be a group and $H \subseteq G$ a subgroup. The *index* of H in G, written $[G : H]$ is the number of distinct cosets of H in G, if this quantity is finite. If not, we say that H has infinite index in G.

Proposition 5.5. *Let H be a subgroup of a group G and $a, b \in G$ two elements. Then:*

(1) Either $aH = bH$ or $aH \cap bH = \emptyset$.
(2) $aH = H$ if and only if $a \in H$.
(3) $aH = bH$ if and only if $b^{-1}a \in H$.

Proof. To prove the first assertion, suppose $aH \cap bH \neq \emptyset$, so let $x \in aH \cap bH$. This means that x may be written $x = ah$ for some $h \in H$ and as $x = bh'$ for some $h' \in H$. Thus $ah = bh'$, and multiplying both expressions on the right by h^{-1}, we have $ahh^{-1} = bh'h^{-1}$ or $a = bh'h^{-1}$. Since H is a subgroup and $h \in H$, it follows that $h^{-1} \in H$, too. Furthermore, since $h' \in H$ and $h^{-1} \in H$, the fact that H is a subgroup means that $h'h^{-1}$ is also a member of H. Thus, $a \in bH$ and hence $aH \subseteq bH$. By symmetric reasoning, $bH \subseteq aH$. Thus, $aH = bH$.

To prove the second assertion, first suppose that $aH = H$. Since $e \in H$, it follows that $a = ae \in aH$, which by hypothesis is just H. Thus $a \in H$. Conversely, suppose that $a \in H$. Then, since $e \in H$, we have that $a \in aH$ and thus $a \in H \cap aH$. By the first part, this implies that $H = aH$.

The third assertion follows almost immediately from the second:

$$aH = bH \text{ iff } b^{-1}aH = b^{-1}bH \text{ iff } b^{-1}aH = H.$$

¶The determination of the rank is tied to the Birch and Swinnerton-Dyer conjecture, one of the seven Millenium Prize Problems from the Clay Mathematics Institute with a prize of one million US dollars for the first correct proof.

By the second proposition this is true if and only if $b^{-1}a \in H$.
☺

By the first assertion of Proposition 5.5, we conclude the following.

Corollary 5.6. *With notation as above, the cosets of H in G partition G; that is, every element of G is contained in exactly one coset of H.*

***Exercise* 21.** Prove that if G finite and $H \subset G$ subgroup, then

$$[G : H] = \frac{|G|}{|H|}.$$

Conclude that for $g \in G$, $|g|$ divides $|G|$. (This proves Proposition 2.5.)

Definition 5.7. Let G be a group and $H \subseteq G$ a subgroup. The *quotient* of G by H, denoted G/H, is the set of different cosets.

Proposition 5.8. *If G is abelian, G/H is a group with the operation $aH * bH = abH$.*

Proof. Indeed, associativity is given by associativity in G, and one can verify that the identity is given by eH and the inverse of aH is $a^{-1}H$. The hard part of the proof is to see that the operation is well-defined. So if $aH = a_1H$, we want to see that $aH * bH = a_1H * bH$, but this is the same as proving that $abH = a_1bH$. By Proposition 5.5, this is equivalent to $b^{-1}a_1^{-1}ab \in H$, which is the same as $a_1^{-1}a \in H$ since the group is abelian. But this is true from the fact that $aH = a_1H$. ☺

For example, $\mathbb{Z}/2\mathbb{Z} \cong \mathbb{Z}_2$.

The proof of Mordell's theorem is done in two steps. The first and more difficult step is to prove that $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite. Given a point $P \in E(\mathbb{Q})$ we can write it as $Q + 2P_1$ with Q a representative of the finite set $E(\mathbb{Q})/2E(\mathbb{Q})$. We can continue this process with P_1 and so on. The second step consists of proving that P_1 is in a certain sense smaller than P and that this process can not continue forever because the points can not get arbitrarily smaller.

5.2. **Descent.** In this section, we prove that $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite. The name "descent theorem" was chosen because the proof is very much in the spirit of Fermat's method of infinite descent. Roughly speaking, one starts with an arbitrary point and tries to produce an infinite sequence of successively smaller points (size being measured by the height function); eventually, one is led to one of two conclusions: either the group is finitely generated or one reaches a contradiction in producing smaller points, since the height, being an integer, cannot be less than 1.

We are going to consider only the special case in which $f(x)$ has its three roots over $\mathbb{Q}$. Because of Lutz–Nagell, we can assume that the roots are indeed over $\mathbb{Z}$. In this section $\square$ will denote a square of a number in $\mathbb{Q}$, i.e., $\square = \frac{a^2}{b^2}$ with $a, b \in \mathbb{Z}$ and $b \neq 0$.

Proposition 5.9. *Let E be an elliptic curve over given by*

$$y^2 = (x - \alpha)(x - \beta)(x - \gamma),$$

with $\alpha, \beta,$ and $\gamma \in \mathbb{Z}$.

Let $P \neq \mathcal{O}$, $P = (x_1, y_1) \in E(\mathbb{Q})$. Then, there is a $Q = (x_2, y_2) \in E(\mathbb{Q})$ with $P = 2Q$ iff $x_1 - \alpha$, $x_1 - \beta$, and $x_1 - \gamma$ are squares in $\mathbb{Q}$.

Proof. First suppose that (x_2, y_2) exists. Let $y = \lambda x + \nu$ be the tangent line at Q. Then the roots of the equation

$$(x - \alpha)(x - \beta)(x - \gamma) - (\lambda x + \nu)^2 = 0$$

are x_1, x_2, x_2. Thus,

$$(x - \alpha)(x - \beta)(x - \gamma) - (\lambda x + \nu)^2 = (x - x_1)(x - x_2)^2.$$

Setting $x = \alpha$, we get $-\square = (\alpha - x_1)\square$. Since $x_2 \neq \alpha$, $\square$ in the right is not zero and we conclude $x_1 - \alpha = \square$. Similarly with β and γ.

Now assume that $x_1-\alpha$, $x_1-\beta$, and $x_1-\gamma$ are squares in $\mathbb{Q}$. After a change of variables, we can further assume that $x_1=0$. We write

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

Thus $-\alpha=\alpha_1^2$, $-\beta=\beta_1^2$, and $-\gamma=\gamma_1^2$. Further, we also have $y_1^2=-\alpha\beta\gamma=c$. After adjusting signs, we may assume

$$y_1=\alpha_1\beta_1\gamma_1.$$

Take the line $y=\lambda x+y_1$ going through P and tangent at an unknown point $Q=(x_2,y_2)$. The three roots of

$$(x-\alpha)(x-\beta)(x-\gamma)-(\lambda x+y_1)^2=0$$

are $0,x_2,x_2$. Therefore,

$$x^2+ax+b-\lambda^2x-2\lambda y_1$$

has the double root x_2, which means that its discriminat is zero,

$$(a-\lambda^2)^2=4(b-2\lambda y_1).$$

One can check that $\lambda_0=-\alpha_1-\beta_1-\gamma_1$ is a root of this equation. It yields a value $x_2=\frac{\lambda_0^2-a}{2}$ which gives $2(x_2,\lambda_0x_2+y_1)=(0,-y_1)$ and $2(x_2,-\lambda_0x_2-y_1)=(0,y_1)$, and thus $P=2Q$ as desired. ☺

In what follows, we consider the group $\mathbb{Q}^*/\mathbb{Q}^{*2}$ which is made of the quotient of the multiplicative group of nonzero rationals $\mathbb{Q}^*$ with its subgroup of nonzero square rationals $\mathbb{Q}^{*2}$. The elements of $\mathbb{Q}^*/\mathbb{Q}^{*2}$ can be described as

$$\mathbb{Q}^*/\mathbb{Q}^{*2}=\{\pm 2^{e_2}3^{e_3}5^{e_5}7^{e_7}\dots \mid e_2,e_3,e_5,e_7,\dots\in\{0,1\}\}. \tag{5.1}$$

Thus, an element of $\mathbb{Q}^*/\mathbb{Q}^{*2}$ can be thought of a sign and a string of 0's and 1's.

***Exercise* 22.** Describe $5/27$ and $24/13$ as elements in $\mathbb{Q}^*/\mathbb{Q}^{*2}$ with the above presentation.

Proposition 5.10. *Define* $\varphi_\alpha: E(\mathbb{Q})\to\mathbb{Q}^*/\mathbb{Q}^{*2}$ *by*

$$\varphi_\alpha(P)=\begin{cases}(x-\alpha)\mathbb{Q}^{*2} & \text{if } P=(x,y) \text{ with } P\neq\mathcal{O}, x\neq\alpha\\ (\alpha-\beta)(\alpha-\gamma)\mathbb{Q}^{*2} & \text{if } P=(\alpha,0)\\ \mathbb{Q}^{*2} & \text{if } P=\mathcal{O}\end{cases}$$

Then φ_α *is a homomorphism.*

Proof. Let $P_1+P_2=P_3$. We have to show that $\varphi_\alpha(P_1)\varphi_\alpha(P_2)\varphi_\alpha(P_3)^{-1}$ is a square in $\mathbb{Q}^*$, which is equivalent to prove that $\varphi_\alpha(P_1)\varphi_\alpha(P_2)\varphi_\alpha(P_3)$ is a square when $P_1+P_2+P_3=\mathcal{O}$. If any of P_i equals $\mathcal{O}$, the conclusion is trivial. Thus assume that $P_i=(x_i,y_i)$.

If no (x_i,y_i) is $(\alpha,0)$, let $y=\lambda x+\nu$ be the line through P_1,P_2,P_3. Then the roots of

$$(x-\alpha)(x-\beta)(x-\gamma)-(\lambda x+\nu)^2=0$$

are x_1,x_2,x_3. We can write

$$(x-\alpha)(x-\beta)(x-\gamma)-(\lambda x+\nu)^2=(x-x_1)(x-x_2)(x-x_3).$$

Setting $x=\alpha$ gives $(x_1-\alpha)(x_2-\alpha)(x_3-\alpha)=(\lambda\alpha+\nu)^2$ and thus $\varphi_\alpha(P_1)\varphi_\alpha(P_2)\varphi_\alpha(P_3)$ is a square.

If $(x_1,y_1)=(\alpha,0)$, then neither (x_2,y_2) nor (x_3,y_3) is $(\alpha,0)$, since otherwise the other point would be $\mathcal{O}$. Again, let $y=\lambda x+\nu$ be the line through P_1,P_2,P_3. Then the roots of

$$(x-\alpha)(x-\beta)(x-\gamma)-(\lambda x+\nu)^2=0$$

are α,x_2,x_3. We write

$$(x-\alpha)(x-\beta)(x-\gamma)-(\lambda x+\nu)^2=(x-\alpha)(x-x_2)(x-x_3).$$

Then $(x-\alpha)$ divides $(\lambda x+\nu)^2$, which means that $(\lambda x+\nu)=\lambda(x-\alpha)$. Dividing by $(x-\alpha)$, we obtain

$$(x-\beta)(x-\gamma)-\lambda^2(x-\alpha)=(x-x_2)(x-x_3).$$

Setting $x=\alpha$,

$$(\alpha-\beta)(\alpha-\gamma)=(\alpha-x_2)(\alpha-x_3),$$

which shows that $\varphi_\alpha(P_1) = \varphi_\alpha(P_2)\varphi_\alpha(P_3)$ as desired.
☺

Observe that we can think of φ_α as a function of $E(\mathbb{Q})/2E(\mathbb{Q})$ because $\varphi_\alpha(2P) = \varphi_\alpha(P)^2 \in \mathbb{Q}^{*2}$.

Corollary 5.11. *The homomorphism*

$$\varphi_\alpha \times \varphi_\beta : E(\mathbb{Q})/2E(\mathbb{Q}) \to \mathbb{Q}^*/\mathbb{Q}^{*2} \times \mathbb{Q}^*/\mathbb{Q}^{*2}$$

is one-to-one.

Proof. Suppose that $P = (x, y) \neq \mathcal{O}$ maps to $\mathbb{Q}^{*2}$ under both φ_α and φ_β.
If $P = (\alpha, 0)$, that means that

$$\varphi_\alpha(\alpha, 0) = (\alpha - \beta)(\alpha - \gamma)\mathbb{Q}^{*2} = \square$$

and

$$\varphi_\beta(\alpha, 0) = (\alpha - \beta)\mathbb{Q}^{*2} = \square.$$

Then, $(\alpha-\beta) = \square$ and $(\alpha-\gamma) = \square$. By Proposition 5.9, this implies that $(\alpha, 0) = 2Q$ and thus $(\alpha, 0) \in 2E(\mathbb{Q})$.
The case $P = (\beta, 0)$ is similar. Now suppose that $P = (x, y) \neq (\alpha, 0), (\beta, 0)$. Then we must have $(x - \alpha) = \square$ and $(x - \beta) = \square$. Since $(x - \alpha)(x - \beta)(x - \gamma) = \square$, we conclude that $(x - \gamma) = \square$. Again, Proposition 5.9 implies that $P = 2Q$ and thus $P \in 2E(\mathbb{Q})$. ☺

For a prime number p, an integer e, and a nonzero rational number n we will use the notation $p^e||n$ to express that $n = p^e \frac{a}{b}$ with a, b integers such that $p \nmid a, b$.

Proposition 5.12. *The image of φ_α satisfies that $e_p = 0$ if $p \nmid \Delta$.*

Proof. Recall that the discriminant is given by

$$\Delta = (\alpha - \beta)^2(\alpha - \gamma)^2(\beta - \gamma)^2.$$

Let $P = (x, y) \neq \mathcal{O}$. First assume that $x \neq \alpha, \beta, \gamma$. Fix a prime p and define integers a, b, c by

$$p^a||(x - \alpha), \quad p^b||(x - \beta), \quad p^c||(x - \gamma).$$

Since $(x - \alpha)(x - \beta)(x - \gamma)$ is a square, we have that $a + b + c$ is even.
Suppose that at least one of a, b, c is < 0. Say $a < 0$. Since α is an integer, $p^{|a|}||$(denominator of x). Then

$$p^a||(x - \alpha), \quad p^a||(x - \beta), \quad p^a||(x - \gamma),$$

i.e., $a = b = c$. The fact that $a + b + c$ is even implies that a, b, c are even. Thus, $e_p = 0$.
Suppose that at least one of a, b, c is > 0. Say $a > 0$. If $p \nmid \Delta$, then $p \nmid (\alpha - \beta)$ and it cannot occur in the numerator of

$$x - \beta = (x - \alpha) + (\alpha - \beta),$$

so that $b = 0$. Similarly, $c = 0$. Since $a + b + c$ is even, we conclude that a must be even. Thus we get again $e_p = 0$.
Now suppose that $x \in \{\alpha, \beta, \gamma\}$. The image of $\varphi_\alpha(x)$ will be some product of $\alpha - \beta, \alpha - \gamma$, and $\beta - \gamma$ up to sign. These numbers are prime to p if $p \nmid \Delta$ which implies that $e_p = 0$ in this case as well. ☺

Theorem 5.13. *$E(\mathbb{Q})/2E(\mathbb{Q})$ is finite.*

Proof. Corollary 5.11 implies that $E(\mathbb{Q})/2E(\mathbb{Q})$ is in one-to-one correspondence with its image in $\mathbb{Q}^*/\mathbb{Q}^{*2} \times \mathbb{Q}^*/\mathbb{Q}^{*2}$ while Proposition 5.12 implies that this image is finite. ☺

***Exercise* 23.** Find a bound for the size of $E(\mathbb{Q})/2E(\mathbb{Q})$ when $E : y^2 = x^3 - x$.

5.3. **Heights.** Heights are devices for measuring the arithmetic complexity of a rational number. For example, the numbers 1 and 57/58 are almost equal in absolute value, but in terms of prime factorizations, $57/58 = 3 \cdot 19/2 \cdot 29$ is much more complicated than 1. So, we want the height to be some measure of the complexity of the fractional representation of a rational number. To this end we define, for a rational number $x = \frac{m}{n}$ written in lowest terms,

$$H(x) = H\left(\frac{m}{n}\right) = \max\{|m|, |n|\}.$$

In our example above, $H(1) = 1$ and $H(57/58) = 58$. One of the most important properties of the height, however, is the following proposition, whose proof is quite basic.

Proposition 5.14. *Let $N \in \mathbb{Z}$ be a positive integer. Then $\{x \in \mathbb{Q} : H(x) \leq N\}$ is a finite set.*

***Exercise* 24.** Prove that the set of rational numbers x with height $H(x) \leq k$ contains at most $2k^2 + 1$ elements.

We can extend these ideas to elliptic curves. Given an elliptic curve E and a rational point $P = (x, y) \in E(\mathbb{Q})$, we define

$$H(P) = H(x),$$

and set also $H(\mathcal{O}) = 1$.

Notice that for any real number $M > 0$, we still have the property that

Lemma 5.15.

$$\{P \in E(\mathbb{Q}) : H(P) \leq M\}$$

is a finite set.

This is because $H(P) \leq M$ means that the x-coordinate of P must be a fraction $\frac{m}{n}$ with $|m|, |n| \leq M$. There are only finitely many such x, and to each such x there are at most two corresponding y-values.

In this section we are going to prove two lemmas about heights that are crutial for the proof of Mordell's theorem. We will work with the model $E : y^2 = x^3 + bx + c$.

Lemma 5.16. *There is a constant k_0 depending only on E such that for all $P \in E(\mathbb{Q})$,*

$$H(2P) \geq k_0 H(P)^4.$$

Proof. Let $P = \left(\frac{m}{n^2}, \frac{\ell}{n^3}\right)$. Then $H(P) = \max\{|m|, n^2\}$. Duplication formula (3.6) implies

$$x(2P) = \frac{(m^2 - bn^4)^2 - 8cmn^6}{4n^2(m^3 + bmn^4 + cn^6)} = \frac{m^4 - 2bm^2n^4 - 8cmn^6 + b^2n^8}{4n^2(m^3 + bmn^4 + cn^6)}.$$

Let $A = (m^2 - bn^4)^2 - 8cmn^6$ and $B = 4n^2(m^3 + bmn^4 + cn^6)$. We need to divide them both by $\gcd(A, B)$. First notice that $\gcd(n, A) = \gcd(n, m^4) = 1$. Now equation (4.1) gives

$$n^{12}\Delta = ((3m^3 - 5bmn^4 - 27cn^6)(m^3 + bmn^4 + cn^6) - (3m^2 + 4bn^4)(m^4 - 2bm^2n^4 - 8cmn^6 + b^2n^8). \tag{5.2}$$

From here we get that $\gcd(A, (m^3 + bmn^4 + cn^6))|n^{12}\Delta$. Since $\gcd(A, n) = 1$, we conclude that $\gcd(A, B)|4\Delta$. This means that

$$H(2P) \geq \frac{\max\{|A|, |B|\}}{4|\Delta|}.$$

Let $H(E) = \max\{|b|^3, c^2\}^{1/6}$. If $|m| \geq 3H(E)n^2$, then $H(P) = |m|$. We have that $m^2 \geq 9|b|n^4$ and that $|m|^3 \geq 27|c|n^6$. Therefore,

$$|A| = |m^4 - 2bm^2n^4 - 8cmn^6 + b^2n^8| \geq (m^2 - bn^2)^2 - 8|cmn^6| \geq \left(m^2 - \frac{m^2}{9}\right)^2 - \frac{8m^4}{27} = \frac{40m^4}{81} = \frac{40H(P)^4}{81}.$$

Thus,

$$\frac{|A|}{4|\Delta|} \geq kH(P)^4,$$

where k is a constant that only depends on the coefficients b, c.

If $|m| \leq 3H(E)n^2$ then equation (5.2) implies

$$\begin{aligned} 4n^{14}|\Delta| &\leq |(3m^3 - 5bmn^4 - 27cn^6)B| + 4n^2|(3m^2 + 4bn^4)A| \\ &\leq (81H(E)^3n^6 + 15|b|n^6H(E) + 27|c|n^6 + 108H(E)^2n^6 + 16|b|n^6)\max\{|A|, |B|\} \end{aligned}$$

This means that we can write

$$n^8 \leq k\frac{\max\{|A|, |B|\}}{4|\Delta|},$$

where k is a constant that only depends on the coefficients b, c. If $H(P) = n^2$ this proves the desired inequality. If not, then we multiply by $(3H(E))^4$ and obtain

$$m^4 \leq (3H(E))^4n^8 \leq k'\frac{\max\{|A|, |B|\}}{4|\Delta|}.$$

Thus, we have prove that $H(2P) \geq k_1H(P)^4$ in all the cases, for k_1 a constant that depends only on b, c. ☺

Lemma 5.17. *There is a constant k_1 depending on E such that for all $P, Q \in E(\mathbb{Q})$,*

$$H(P + Q) \leq k_1H(P)^3H(Q)^3.$$

Proof. Let $P = \left(\frac{m}{n^2}, \frac{\ell}{n^3}\right)$ and $Q = \left(\frac{M}{N^2}, \frac{L}{N^3}\right)$. Addition equation (3.5) implies

$$\begin{aligned} x(P + Q) &= \left(\frac{Ln^3 - \ell N^3}{nN(Mn^2 - mN^2)}\right)^2 - \frac{M}{N^2} - \frac{m}{n^2} \\ &= \frac{(Ln^3 - \ell N^3)^2 - (Mn^2 + mN^2)(Mn^2 - mN^2)^2}{n^2N^2(Mn^2 - mN^2)^2}. \end{aligned}$$

Notice that $|m|^3, n^6 \leq H(P)^3$, and $|M|^3, N^6 \leq H(Q)^3$. Also, $\ell^2 = m^3 + bmn^4 + cn^6$ implies that

$$\ell^2 \leq 3H(E)^3H(P)^3$$

and similarly with $L^2 \leq 3H(E)^3H(Q)^3$. Thus, both the numerator and denominator of $x(P+Q)$ are bounded by $k_1H(P)^3H(Q)^3$ for certain k_1 that depends only on b and c. ☺

As a final note, all these results are valid for $a = 0$. We have seen that to get $a = 0$ we need a change of variables $x \to x + \ell$.

***Exercise* 25.** Prove that Lemmas 5.16 and 5.17 remain valid (possibly with different constants k) under the change of variables $x \to x + \ell$.

5.4. **The end of the proof.** From here, the proof of the Mordell-Weil Theorem follows from the descent Theorem 5.13 and the three lemmas 5.15, 5.16 and 5.17.

Theorem 5.18 (Mordell)**.** *Then $E(\mathbb{Q})$ is finitely generated.*

Proof. Since we know that $[E(\mathbb{Q}) : 2E(\mathbb{Q})] = n$ is finite, choose representatives $Q_1, \ldots, Q_n$ for the (distinct) cosets of $2E(\mathbb{Q})$ in $E(\mathbb{Q})$. Now let $P \in E(\mathbb{Q})$ be an arbitrary element. Since $P + 2E(\mathbb{Q})$ is a coset, it must be one of $Q_1 + 2E(\mathbb{Q}), \ldots, Q_n + 2E(\mathbb{Q})$; say,

$$P + 2E(\mathbb{Q}) = Q_{i_1} + 2E(\mathbb{Q}).$$

By the third assertion of Proposition 5.5, this is equivalent to asserting that:

$$P - Q_{i_1} \in 2E(\mathbb{Q})$$

or that there exists $P_1 \in E(\mathbb{Q})$ such that

$$P - Q_{i_1} = 2P_1.$$

Now replace P with P_1 and keep repeating this procedure to get the following equations:

$$\begin{aligned} P_1 - Q_{i_2} &= 2P_2 \\ P_2 - Q_{i_3} &= 2P_3 \\ &\cdots \\ P_{m-1} - Q_{i_m} &= 2P_m. \end{aligned}$$

The essence of the proof is to show that since P_i is pretty much equal to $2P_{i+1}$, the height of P_{i+1} is smaller than the height of P_i. Thus the sequence P, P_1, P_2 produces points of successively decreasing height, and eventually we will get a set of points having bounded height, which by Lemma 5.15 is finite. We make this idea more precise.

From the first equation, we have

$$P = Q_{i_1} + 2P_1.$$

Now substitute the second equation $P_1 = Q_{i_2} + 2P_2$ into this to get

$$P = Q_{i_1} + 2Q_{i_2} + 4P_2.$$

Continuing in a similar manner, we obtain

$$P = Q_{i_1} + 2Q_{i_2} + 4Q_{i_3} + \ldots + 2^{m-1}Q_{i_m} + 2^m P_m.$$

In particular, this shows that P is in the subgroup generated by P_m and the Q_i, $1 \leq i \leq n$.

Now we apply Lemma 5.17 with any P and $-Q_i$. There is a k_i depending on Q_i such that

$$H(P - Q_i) \leq k_i H(P)^3 \text{ for all } P \in E(\mathbb{Q}).$$

If we do this for all $i = 1, \ldots, n$ and let $k = \max\{k_1, \ldots, k_n\}$, we clearly have

$$H(P - Q_i) \leq kH(P)^3 \text{ for all } P \in E(\mathbb{Q}) \text{ and all } 1 \leq i \leq n. \tag{5.3}$$

Now apply Lemma 5.16 we get a constant k_0 such that

$$H(2P) \geq k_0 H(P)^4 \text{ for all } P \in E(\mathbb{Q}).$$

Using the specific point $P = P_j$, we have $H(2P_j) \geq k_0 H(P_j)^4$. By the initial equations defining the P_j and Q_{i_j}, the expression on the left equals $H(P_{j-1} - Q_{i_j})$, which by equation (5.3) is $\leq kH(P_{j-1})^3$. Summarizing, we have a chain of inequalitites:

$$kH(P_{j-1})^3 \geq H(P_{j-1} - Q_{i_j}) \geq k_0 H(P_j)^4$$

Taking the 4th root, we get

$$\tilde{k}^4 H(P_{j-1})^{\frac{3}{4}} \geq H(P_j),$$

where $\tilde{k} = \frac{k}{k_0}$.

We claim that there exists some m such that $H(P_m) \leq \tilde{k}^2$. Let us examine the list:

$$P, P_1, P_2, \ldots$$

As long as $H(P_{j-1}) \geq \tilde{k}^2$, the above equation says that the next point will have height at most $H(P_j)^{\frac{7}{8}}$. However, repeated applications of the $\frac{7}{8}$-power will cause it to approach zero, so eventually we will find an index m such that $H(P_m) \leq \tilde{k}^2$.

Thus, we have shown that every element $P \in E(\mathbb{Q})$ may be written as

$$P = a_1 Q_1 + a_2 Q_2 + \ldots a_n Q_n + 2^m P_m$$

where $a_1, \ldots, a_n$ are integers (some of them could be zero), and P_m is a point satisfiying $H(P_m) \leq \tilde{k}^2$. Hence the set

$$\{Q_1, \ldots, Q_n\} \cup \{R \in E(\mathbb{Q}) : H(R) \leq \tilde{k}^2\}$$

generates $E(\mathbb{Q})$. The first set is clearly finite and the second is finite by Lemma 5.15. Thus, $E(\mathbb{Q})$ is finitely generated.

6. Appendix

6.1. **Linear equations.** The situation for linear equations is completely understood.

Theorem 6.1. *Let $a, b, c \in \mathbb{Z}$. The equation*

$$ax + by = c$$

has a solution $x_0, y_0 \in \mathbb{Z}$ iff $\gcd(a, b)$ divides c. In this case, all the solutions are given by

$$x = x_0 + \frac{bk}{(a,b)}, \qquad y = y_0 - \frac{ak}{(a,b)}, \qquad \text{with } k \in \mathbb{Z}.$$

It is not hard to see that the condition $\gcd(a, b)$ divides c is necessary. For if d divides both a and b, then it must divide $ax + by = c$. If, on the other hand, $\gcd(a, b)$ divides c, one can divide by $\gcd(a, b)$ all terms and assume that $\gcd(a, b) = 1$. To find such solution, one uses the Euclidean algorithm.

The Euclidean algorithm is a procedure to compute the greatest common divisor of two nonzero integers a, b that also provides a way to write $\gcd(a, b)$ as a linear combination of a and b with integral coefficients.

The base case is the division: given $a, b \in \mathbb{N}$, there exist unique q, $r \in \mathbb{Z}$ with $0 \leq r \leq b$ such that

$$a = qb + r.$$

q and r are called quotient and remainder respectively

If $r \neq 0$, this process is repeated with b and r. If the new remainder is different from zero, we do the same step again, and we continue until we get a zero remainder.

$$\begin{aligned} a &= q_0 b + r_0 \\ b &= q_1 r_0 + r_1 \\ r_0 &= q_2 r_1 + r_2 \\ &\vdots \\ r_{n-2} &= q_n r_{n-1} + r_n \\ r_{n-1} &= q_{n+1} r_n. \end{aligned}$$

It results that $r_n = \gcd(a, b)$. The equations can be reversed

$$\begin{aligned} r_n &= r_{n-2} - q_n r_{n-1} = x_n r_{n-2} + y_n r_{n-1} \\ &= r_{n-2} - q_n(r_{n-3} - q_{n-1} r_{n-2}) = x_{n-1} r_{n-3} + y_{n-1} r_{n-2} \\ &\vdots \\ &= x_2 r_0 + y_2 r_1 \\ &= x_1 b + y_1 r_0 \\ &= x_0 a + y_0 b. \end{aligned}$$

For example, suppose that we want to solve

$$48x + 74y = 6.$$

We first check that $\gcd(48, 74) = 2$ which divides 6, so this equation has integral solutions.

The Euclidean algorithm yields

$$\begin{aligned} 74 &= 1 \cdot 48 + 26 \\ 48 &= 1 \cdot 26 + 22 \\ 26 &= 1 \cdot 22 + 4 \\ 22 &= 5 \cdot 4 + 2 \\ 4 &= 2 \cdot 2. \end{aligned}$$

Thus,

$$\begin{aligned} 2 &= 22 - 5 \cdot 4 \\ &= 22 - 5 \cdot (26 - 1 \cdot 22) = -5 \cdot 26 + 6 \cdot 22 \\ &= -5 \cdot 26 + 6 \cdot (48 - 1 \cdot 26) = 6 \cdot 48 - 11 \cdot 26 \\ &= 6 \cdot 48 - 11 \cdot (74 - 1 \cdot 48) = -11 \cdot 74 + 17 \cdot 48. \end{aligned}$$

We have found, in particular, that

$$51 \cdot 48 - 33 \cdot 74 = 6.$$

Finally, the general solutions are given by

$$x = 51 + 37k, \qquad y = -33 - 24k, \qquad \text{with } k \in \mathbb{Z}.$$

6.2. **Quadratic equations.**

Theorem 6.2. *(Legendre) Suppose that a, b, c are square-free and pairwise coprime. Then the equation*

$$aX^2 + bY^2 + cZ^2 = 0$$

has solutions with X, Y, Z integers other than $X = Y = Z = 0$ if and only if $-bc$ is a square modulo a, $-ac$ is a square modulo b, and $-ab$ is a square modulo c, and a, b, c do not all have the same sign.

Once a solution is found, one can generate all the others by taking each line going through the solution and looking at the other point of intersection. As an example, let us look at the (dehomogenized) case of

$$x^2 + y^2 = 1.$$

We take the solution $(-1, 0)$. The lines through $(-1, 0)$ have equation

$$L_t : t(x + 1) = y, \quad t \in \mathbb{R}.$$

We want to look at the other point of intersection with the equation $x^2 + y^2 = 1$. (There is an extra line, $L_\infty : x = -1$, tangent to the circle at $(0, -1)$. We can interpret that this line intersects the circle at $(-1, 0)$ twice.)

It is easy to see that L_t intersects the circle again at $\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)$. If t is rational, then so is the new point.

This result can be translated in the following statement for the homogeneous equation.

Theorem 6.3. *Let X, Y, Z integers such that they are coprime and*

$$X^2 + Y^2 = Z^2.$$

Suppose that 2 divides X and that 2 does not *divide Y. Then there are $a, b \in \mathbb{Z}$ such that $\gcd(a, b) = 1$, one of them is even, and such that*

$$X = 2ab, \quad Y = a^2 - b^2, \quad Z = a^2 + b^2.$$

Observe that we necessarily have that one of X, Y is even and the other is odd, therefore, the assumption is natural.

7. Sources

These notes have extensively borrowed from several bibliographical sources. The general structure and many statements come from [Ak03a] and [ST92]. Part of section 2 and subsection 5.3 are taken from [Gr11]. Subsection 5.2 can be found in [Kn92].

References

[Ak03a] Reza Akhtar, An Introduction to Elliptic Curves. Notes from a summer course in Summer 2003 at Miami University. Available at `http://calico.mth.muohio.edu/reza/sumsri/2003/elliptic.pdf`

[Ak03b] Reza Akhtar, Elliptic Curve Cryptography. Notes from a summer course in Summer 2003 at Miami University. Available at `http://calico.mth.muohio.edu/reza/sumsri/2003/elliptic.pdf`

[Br01] Ezra Brown, Magic squares, finite planes, and points of inflection on elliptic curves. *College Math. J.* **32** (2001), no. 4, 260–267,

[Br00] A. E. Brouwer, Elliptic functions, integrals, and curves. Notes from a course given in 2000 at the Eindhoven University of Technology, Netherlands.

[Ca91] J. W. S. Cassels, Lectures on elliptic curves. London Mathematical Society Student Texts, 24. Cambridge University Press, Cambridge, 1991. vi+137 pp.

[Co99] Ian Connell, Elliptic Curve Handbook. February 1999. Available at `http://www.math.mcgill.ca/connell/public/ECH1/`

[Gr11] Andrew Granville, Rational and integral points on curves. Notes from a course given in Fall 2011 at the Université de Montréal.

[Kn92] Anthony W. Knapp, Elliptic curves. Mathematical Notes, 40. Princeton University Press, Princeton, NJ, 1992. xvi+427 pp.

[Ko93] Neal Koblitz, Introduction to elliptic curves and modular forms. Second edition. Graduate Texts in Mathematics, 97. Springer-Verlag, New York, 1993. x+248 pp.

[ST92] Joseph H. Silverman; John Tate, Rational points on elliptic curves. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1992. x+281 pp.

[Si09] Joseph H. Silverman, The arithmetic of elliptic curves. Second edition. Graduate Texts in Mathematics, 106. Springer, Dordrecht, 2009. xx+513 pp.

8. Suggested projects

Here is a list of possible projects.

8.1. **Congruent numbers.** Describe the congruent number problem and its relation to elliptic curves. References: [Kn92] pages 53-55 and 110-114, [Ko93].

8.2. **Division polynomials.** Explain what division polynomials are and what we can prove with them. References: [ST92] page 214, [Co99] pages 145-148.

8.3. **Construction of curves with prescribed torsion.** Show us how to construct elliptic curves with prescribed torsion. References: [Kn92] pages 145-148.

8.4. **Points of order 3.** Read the article [Br01] and tell us about it.

8.5. **The complex structure of elliptic curves.** Describe the structure of complex points on elliptic curves. References: [ST92] pages 41-46, [Kn92] chapter VI.

8.6. **Exploring the descent procedure.** Tells us about how to improve the bound on the rank that is given by the descent procedure. References: [Kn92] pages 107-114.

8.7. **Elliptic curve cryptography.** Explore the most basic idea of Elliptic curve cryptography. References: [Ak03b].

8.8. **Elliptic integrals.** Show how to compute the arc length of an ellipse and the relationship between this and elliptic curves. References: [ST92], pages 35-36, [Br00].

Département de mathématiques et de statistique, Université de Montréal. CP 6128, succ. Centre-ville. Montreal, QC H3C 3J7, Canada

E-mail address: mlalin@dms.umontreal.ca, http://www.dms.umontreal.ca/~mlalin