# Diophantine Equations

Author: Salman Ahmad Faris
Supervisor: Prof. Payman L. Kassaei

Module Code: 6CCM345A

2020/2021

### Abstract

Diophantine analysis concerns the understanding of integer and rational solutions to multivariate polynomial equations. Our aim in this project is to study them, giving an introduction to arithmetic geometry along the way. In particular, we shall see how viewing Diophantine equations as geometrical objects have powerful consequences. Moreover, we shall look at some modern number-theoretic tools — such as the $p$-adic numbers, $L$-functions and modular forms — that have aided number theorists in solving some of the hardest problems in mathematics.

We will start our analysis with a demonstration of how geometry comes into the picture of number theory, motivating via a concrete example. We then look at the conics, the curve defined by quadratic polynomials. Here we show that there is a nice criterion of checking the existence of rational points on a conic, and that if one has a rational point, then this point can be used to find infinitely many other rational points on that conic. Then we look at the cubics, the curve defined by cubic polynomials. It turns out that the problem of finding rational points on general cubics reduces to the problem of finding rational points on elliptic curves, which, unfortunately, we have yet to completely understand. However, we shall see that the study of elliptic curves motivates some of the most captivating advancements in mathematics, such as the proof of Fermat's last theorem.

# Contents

# Standard notation

The following symbols

$$\mathbb{Z},\ \mathbb{Q},\ \mathbb{R},\ \mathbb{C},\ \mathbb{F}_p$$

denote the integers, rational numbers, real numbers, complex numbers and the finite field of $p$ elements, respectively. We will not use the notation $\mathbb{N}$ for the natural numbers due to its ambiguity. Furthermore, if $A$ is any set, then $A^n$ denotes the $n$-fold Cartesian product of $S$. If $R$ is a ring, then $R[x_1, \ldots, x_n]$ denotes the polynomial ring over $R$ in $n$ indeterminates. Furthermore, the notation $R^\times$ denotes the group of units in $R$. If $K$ is a field, then $K(x_1, \ldots, x_n)$ denotes the field of rational functions. Moreover, $\bar{K}$ denotes an algebraic closure of $K$.

# 1 What are Diophantine Equations?

Our central object of discussion are *Diophantine equations.*

**Definition 1.1.** Let $f \in \mathbb{Q}[x_1, \ldots, x_n]$ be a polynomial. The polynomial equation

$$f(x_1, \ldots, x_n) = 0$$

is said to be a **Diophantine equation** if we seek solutions in $\mathbb{Z}$ or $\mathbb{Q}$.

**Remark 1.** We could have equally require $f$ to have coefficients in $\mathbb{Z}$ instead of $\mathbb{Q}$. This is because if we have rational coefficients, we can always clear denominators by multiplication by an appropriate integer.

Finding solutions in $\mathbb{Z}$ or $\mathbb{Q}$ is far harder compared to finding solutions in, say, $\mathbb{R}$. This leads to asking the following reasonable questions whenever solving Diophantine equations.

(1). Are there any integer and rational solutions?

(2). If there are, can we deduce finitely (or infinitely) many other solutions?

(3). If there are only finitely many solutions, can we prove that they are the *only* solutions?

(4). Are the solutions dependent on the coefficients of $f$?

With computational aid, the first one is not always difficult because sometimes we can do it by inspection. However, the next few questions are usually quite difficult, hard enough to be a problem unsolved for almost four centuries (to be precise, it was unsolved for 358 years). This problem is the celebrated Fermat's last theorem which is related to perhaps the most famous Diophantine equation in existence, the Fermat equation

$$X^n + Y^n - Z^n = 0.$$

It was proven by Taylor-Wiles (see [30], [31]) that this equation does not have any solutions $(x, y, z) \in \mathbb{Z}^3$ with $xyz \neq 0$ for $n \geqslant 3$. However, the methods they used are modern and highbrow. In particular, it would be impossible that Fermat had actually found a proof as he once claimed. We shall see what these methods are and discuss (very) briefly how they were used to prove Fermat's last theorem at the end of the paper (see Section 5.2).

Before we move on to more theory, let us give some definitions that we will use freely later.

**Definition 1.2.** Let $f \in \mathbb{Q}[x_1, \ldots, x_n]$ be a polynomial and let $d$ be a positive integer. We say $f$ is **homogeneous** of degree $d$ if all the monomials in $f$ occured with the same degree $d$.

**Remark 2.** A homogeneous polynomial is also known as a **form**. We will look at a special case of forms later on as they will be very useful. We will use the convention of writing indeterminates in forms with capital letter $X, Y, Z$ whereas writing indeterminates in non-homogeneous polynomials with lowercased letter $x, y, z$.

**Definition 1.3.** We say a Diophantine equation is **homogeneous** of degree $d$ if the polynomial defining it is homogeneous of degree $d$.

As usual in mathematics, the examples are more important than the definitions.

**Example.** *Examples of homogeneous Diophantine equations.*

(i). $X + Y + Z + W = 0$ is a *linear* (i.e. degree 1) homogeneous Diophantine equation.

(ii). $X^2 + Y^2 = 3Z^2$ is a homogeneous Diophantine equation of degree 2; the polynomial defining it is known as a *quadratic form*. We will look more at this later in Chapter 3.

(iii). $X^3 + 2XY^2 - Y^2Z + XYZ = 0$ is a homogeneous Diophantine equation of degree 3; the polynomial defining it is known as a *cubic form*.

(iv). The Fermat equation $X^n + Y^n = Z^n$ is a degree $n$ homogeneous Diophantine equation.

**Example.** *Examples of non-homogeneous Diophantine equations.*

(i). The equation $y^2 = x^3 + 7$ is not a homogeneous Diophantine equation. This equation is an example of a *Mordell's equation*.

(ii). The equation $x^2 - 22y^2 = 1$ is not a homogeneous Diophantine equation. This equation is an example of a *Pell's equation*.

Observe that in the case of a homogeneous Diophantine equation $f(x_1, \ldots, x_n) = 0$, we would always have $(0, \ldots, 0)$ as a solution. This is what we call a *trivial* solution of the Diophantine equation, and we will ignore it. We only care about nontrivial solutions, and sometimes, we demand an even stronger condition. Let $(x_1, \ldots, x_n) \in \mathbb{Q}^n$ be a solution to some Diophantine equation. By demanding the condition $\prod_i x_i \neq 0$, we have that $x_i$ is not all zero and so $x$ is nontrivial by definition. This condition allows us to disregard solutions that look like $(1, 0)$ or $(0, 0, 0, 2/3)$ which in some sense, are also trivial depending on the Diophantine equation. For example, $(1, 0, 1)$ can be seen as a trivial solution to the Fermat equation for any $n \geqslant 3$.

We note that we are only interested in nonlinear Diophantine equations in this paper. A treatment of the linear case can be found in virtually any elementary number theory textbook, for example, in Rosen [19] and Silverman [25]. In the next chapter, Chapter 2, we begin our analysis on Diophantine equations by demonstrating how to view a Diophantine equation as a *plane curve*. We then proceed with our first attempt in finding *rational points* on these curves, and we do this concretely by looking at the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 3$. We then look at the *conics*, the curve defined by quadratic polynomials, in Chapter 3. Here we show that there is a nice criterion due to Legendre (later generalized by Hasse and Minkowski) of checking, in a finite number of steps, the existence of rational points on a conic. Moreover, by generalizing the method we used on the unit circle, we will show that if one has a rational point on a conic, then this point can be used to find infinitely many other rational points on the conic. In other words, the rational points on a conic with at least one rational point is parametrizable.

Chapter 4 considers the *cubics* which are curves defined by cubic polynomials. We will first see that all the rational points on *singular* cubics are parametrizable and showing this will be the first goal of the chapter. For nonsingular cubics, the methods we know so far fail and we require new insights. It turns out that the problem of finding rational points on the general cubics reduces to the problem of finding such points on so-called elliptic curves, which unfortunately, we have yet to completely understand. Here, the true power of arithmetic geometry comes into sight as it turns out that the rational points on an elliptic curve $E$ can be made into an abelian group, denoted by $E(\mathbb{Q})$. Poincaré conjectured in 1901 that $E(\mathbb{Q})$ is a finitely generated abelian group. This conjecture was eventually proven by Mordell in 1922, and later generalized by Weil in 1928. Due to the structure theorem of finitely generated abelian groups, understanding $E(\mathbb{Q})$ now amounts to understanding its torsion part and free part. The torsion part is well-understood thanks to the work of Mazur in 1977; and so what is left is the free part which our hope right now is a \$1,000,000 problem called the Birch and Swinnerton-Dyer conjecture.

In the concluding chapter, we define an important invariance in the setting of algebraic geometry called the *genus* which is a non-negative integer. As it turns out, the case of conics (resp. cubics) corresponds to curves of genus 0 (resp. genus 1). This means that when the genus equals 0, the existence of one rational point implies the existence of infinitely many other rational points. On the other hand, for curves of genus 1, the existence of one rational point may or may not guarantee the existence of infinitely many other such points. It is then natural to ask whether it is possible or not that nonsingular curves defined over $\mathbb{Q}$ with genus $> 1$ have infinitely many rational points. Mordell conjectured in 1922 that this is impossible, and this was proven by Faltings in 1983. Finally, we shall look at the progress in proving Fermat's last theorem throughout the centuries. In particular, we will see how Taylor and Wiles, building on the work of many other mathematicians such as Taniyama, Shimura, Weil, Frey, Ribet and Serre, used the theory of elliptic curves in proving Fermat's last theorem.

# 2 Motivation: Geometry and Number Theory

We will first look at two special Diophantine equations and discuss their solutions.

## 2.1 The equation $x^2 + y^2 = 1$

Let us start with a discussion of solutions to the equation

$$x^2 + y^2 = 1. \tag{2.1}$$

What would the solutions $(x, y)$ to (2.1) in $\mathbb{R}^2$ be? It is obvious that $(\pm 1, 0)$ and $(0, \pm 1)$ are solutions. Some trial and error also give $(\frac{\sqrt{3}}{2}, \frac{1}{2})$ and $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ as solutions. These are immediate ones but at least we know there are solutions. If we are seeking solutions in $\mathbb{Z}^2$, it is easy to verify that in fact $(\pm 1, 0)$ and $(0, \pm 1)$ are the only solutions. As a consequence, these are also the only solutions in $(\mathbb{Z}/n\mathbb{Z})^2$ for any integer $n \geqslant 2$.

How about solutions in $\mathbb{Q}^2$? Aside from the integer ones that we have mentioned earlier, $(\frac{3}{5}, \frac{4}{5})$ is one solution in $\mathbb{Q}^2$ since $(3/5)^2 + (4/5)^2 = 1$. But now observe that by multiplying $5^2$ on both sides of the equation, we have $3^2 + 4^2 = 5^2$. That is, the triple $(3, 4, 5)$ is a Pythagorean triple. This suggests that there is a correspondence between solutions to (2.1) and solutions to the Pythagorean equation $X^2 + Y^2 = Z^2$. In fact, it is not too difficult to see this correspondence as follows. If $(x_0, y_0)$ is a solution in $\mathbb{Q}^2$ to (2.1), then we can choose a nonzero common denominator $c \in \mathbb{Z}$ to write $x_0 = a/c$ and $y_0 = b/c$ where $a, b \in \mathbb{Z}$, so that $(a, b, c)$ satisfies $X^2 + Y^2 = Z^2$. Conversely, if $(a, b, c) \in \mathbb{Z}^3$ with $c \neq 0$ is a Pythagorean triple, then the pair $(a/c, b/c)$ solves (2.1). Thus, if we can describe all the solutions in $\mathbb{Q}^2$ to (2.1), we get to describe all the solutions in $\mathbb{Z}^3$ to $X^2 + Y^2 = Z^2$. We will do so using methods of geometry.

From basic analytic geometry, we know that equation (2.1) defines a circle of radius 1 centred at the origin which, from now on, we will denote as $C$. So finding rational solutions to this equation is equivalent to finding *rational points* on the circle $C$. The question is how? Before we proceed, let us make this definition precise.

**Definition 2.1.** Let $K$ be a field. An **algebraic plane curve** defined over $K$ is the set of points $(x, y) \in \bar{K}^2$ such that $f(x, y) = 0$ for some $f \in K[x, y]$. In this case, we say that it is **associate to** $f$ and we denote it as $C_f$. An element of the plane curve is said to be a **point** on the curve. If the context is clear, we will just write $C$, and call it a plane curve, or just curve.

**Remark 3.** The word *algebraic* in "algebraic plane curve" is for technical reasons due to algebraic geometry. Since we will be only dealing with algebraic plane curves in this paper, we will simply refer to it as plane curves. A handy piece of notation that we will use when defining plane curves is the following: we will usually write a variant of "let $C\colon f(x, y) = 0$ be the plane curve" to mean one of "let $C$ be the plane curve associated to $f(x, y)$" or " let $C$ be the plane curve $f(x, y) = 0$". This makes things more readable.

Using our new definition, the unit circle is then the plane curve $C : x^2 + y^2 = 1$ defined over the field $\mathbb{Q}$. As we have seen, $(1, 0)$ and $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ are example of points on $f$ as they satisfy $f(x, y) = 0$.

**Definition 2.2.** Let $C$ be a plane curve defined over $\mathbb{Q}$. We say $(x, y)$ is a **rational point** on $C$ if $(x, y) \in C \cap \mathbb{Q}^2$. That is, it is a point on $C$ with coordinates in $\mathbb{Q}$. The set of rational points on $C$ will be denoted $C(\mathbb{Q})$.

**Example.** Since $(-1, 0)$ and $(\frac{3}{5}, \frac{4}{5})$ have rational coordinates and satisfy the equation $f(x, y) = 0$, they are rational points on $C$. However, the point $(\frac{\sqrt{3}}{2}, \frac{1}{2})$ is not a rational point despite satisfying $f(x, y) = 0$ as $\frac{\sqrt{3}}{2} \notin \mathbb{Q}$.

Let us fix a rational point on $C$, say, $\boldsymbol{P} = (-1, 0)$. If $(x, y)$ is any other arbitrary rational point on $C$, then a straight line through the two points would have a rational slope as $y/(x + 1) \in \mathbb{Q}$. This is quite straightforward. There is an even more interesting observation which is less obvious.

**Proposition 2.1.** *Let $C$ be defined by (2.1) and let $\boldsymbol{P} = (-1, 0)$, which is a point on $C$. If $\lambda \in \mathbb{Q}$ and $L_\lambda$ is a line through $\boldsymbol{P}$ with slope $\lambda$, then it intersects another point $\boldsymbol{Q}_\lambda$ in $C$ which is necessarily rational.*

Note that by restricting $\lambda \in \mathbb{Q}$, we do not allow $L_\lambda$ to be a vertical line. This is an important observation to note when we try to generalize this idea to general conics in Chapter 3.

*Proof.* Consider the line $L_\lambda : y = \lambda(x + 1)$ with $\lambda \in \mathbb{Q}$. A point $(x, y)$ that is in the intersection $C \cap L_\lambda$ must satisfy the simultaneous equation

$$\begin{cases} y = \lambda(x+1), & (2.2) \\ x^2 + y^2 = 1. & (2.3) \end{cases}$$

If we substitute (2.2) into (2.3), we would end up with an equation of the form

$$x^2 + ax + b = 0, \tag{2.4}$$

for some $a, b \in \mathbb{Q}$. Since we know that the $x$-coordinate of $\boldsymbol{P}$, which is $x_0 = -1$, must be a root of (2.4), Vieta's sum of roots formula gives us that $x_0 + x_1 = -a$ where $x_1$ is our other root. Consequently, $x_1 \in \mathbb{Q}$ as $a$ and $x_0$ are both rational. From (2.2), we then know $y_1 = \lambda(x_1 + 1)$ is rational. So $(x_1, y_1)$ gives a rational solution to the original simultaneous equation. The proof is done by taking $\boldsymbol{Q}_\lambda = (x_1, y_1)$. ■

In fact, it is not too hard to explicitly solve the system of equations defined by (2.2) and (2.3). Plugging (2.2) into (2.3), we have

$$x^2 + \lambda^2(1 + x)^2 = 1 \implies \lambda^2(1 + x)^2 = 1 - x^2 = (1 - x)(1 + x).$$

Assume that $x \neq -1$ because otherwise we get our obvious rational point $\boldsymbol{P}$. Then we can divide both side by $1 + x$ to have

$$\lambda^2(1 + x) = 1 - x.$$

Solving for $x$ in terms of $\lambda$, we have $x = (1 - \lambda^2)/(1 + \lambda^2)$. Consequently, we have $y = 2\lambda/(1 + \lambda^2)$. This implies that $\boldsymbol{Q}_\lambda$ in our preceding proposition is given by the explicit formula

$$\boldsymbol{Q}_\lambda = \left(\frac{1 - \lambda^2}{1 + \lambda^2}, \frac{2\lambda}{1 + \lambda^2}\right).$$

Notice that we have actually established a bijection $\mathbb{Q} \to C(\mathbb{Q}) \setminus \{\boldsymbol{P}\}$ via Proposition 2.1. Explicitly, this bijection is given by the map $\lambda \mapsto \boldsymbol{Q}_\lambda$. Such a correspondence is called a *birational equivalence* which is a kind of isomorphism in the setting of algebraic geometry. This definition will be made precise later on.

As promised earlier, we can now get a description for all solutions in $\mathbb{Z}^3$ to the Pythagorean equation. To see this, write $\lambda = s/t$ for some $s \in \mathbb{Z}$, $t \in \mathbb{Z}^+$ with $\gcd(s, t) = 1$. Then using our explicit formula for $\boldsymbol{Q}_\lambda$, we get

$$\boldsymbol{Q}_\lambda = \left(\frac{1 - \lambda^2}{1 + \lambda^2}, \frac{2\lambda}{1 + \lambda^2}\right) = \left(\frac{s^2 - t^2}{s^2 + t^2}, \frac{2st}{s^2 + t^2}\right).$$

By using the relation $x^2 + y^2 = 1$, we see that

$$\left(\frac{s^2 - t^2}{s^2 + t^2}\right)^2 + \left(\frac{2st}{s^2 + t^2}\right)^2 = 1,$$

and further multiplying through by $s^2 + t^2$, we have

$$(2st)^2 + (s^2 - t^2)^2 = (s^2 + t^2)^2.$$

From here, we immediately see that $(2st, s^2 - t^2, s^2 + t^2)$ is a Pythagorean triple. In fact, we can easily show that this triple is a *primitive* Pythagorean triple. That is, the greatest common divisor of all three entries of the triple is 1. This is important as since the Pythagorean equation is homogeneous of degree 2, then any integral multiple of this triple is also a solution. In other words, the solution given above parametrized by $s, t$ describes all possible Pythagorean triples.

## 2.2 The equation $x^2 + y^2 = 3$

Consider the equation $x^2 + y^2 = 3$. This has solutions in $\mathbb{R}^2$, and also defines a circle on the plane, but now of radius $\sqrt{3}$. Unlike the case when the radius is 1, there are no rational points on this curve. This can be seen by using a pure number-theoretic argument. Let us first recall the definition of the Legendre symbol.

**Definition 2.3.** Let $p$ be an odd prime and $a \in \mathbb{Z}$. We define the Legendre symbol $\left(\frac{a}{p}\right)$ by

$$\left(\frac{a}{p}\right) = \begin{cases} +1, & \text{if } a \text{ is a quadratic residue mod } p, \\ 0, & \text{if } a \text{ is divisible by } p, \\ -1, & \text{if } a \text{ is not a quadratic residue mod } p. \end{cases}$$

**Lemma 2.1.** *Let $x, y \in \mathbb{Z}$ and let $p$ be a prime such that $p \equiv 3 \pmod 4$. If $p$ divides $x^2 + y^2$, then $p$ divides both $x$ and $y$.*

*Proof.* Suppose for contradiction that $p$ does not divide $x$. Since $p$ divides $x^2 + y^2$, we have $y^2 \equiv -x^2 \pmod p$. This implies that

$$1 = \left(\frac{-x^2}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{x^2}{p}\right) = \left(\frac{-1}{p}\right).$$

But this is a contradiction as $-1$ is not a quadratic residue modulo a prime $p \equiv 3 \pmod 4$. The same argument holds if we assume that $p$ does not divide $y$. ■

**Proposition 2.2.** *The equation $X^2 + Y^2 = 3Z^2$ has no nontrivial integer solutions.*

*Proof.* Suppose $(x, y, z)$ is a nontrivial integer solution. Without loss of generality, we may assume that $\gcd(x, y, z) = 1$, for otherwise we can just divide by their common divisor. Since $x^2 + y^2 = 3z^2$, 3 clearly divides $x^2 + y^2$. By our preceding lemma, this implies that 3 divides both $x$ and $y$. This further implies that 9 divides both $x^2$ and $y^2$. Consequently, 9 divides $3z^2$ and so 3 divides $z^2$. But since 3 is prime, 3 divides $z$ which is a contradiction to our hypothesis that $\gcd(x, y, z) = 1$. ■

By dividing the equation $X^2 + Y^2 = 3Z^2$ by $Z^2$ on both sides and labelling $x = X/Z$ and $y = Y/Z$, the preceding proposition implies that the equation $x^2 + y^2 = 3$ has no solutions in $\mathbb{Q}^2$.

**Corollary 2.1.** *There are no rational points on the circle $x^2 + y^2 = 3$.*

Observe the argument we used more closely. The idea in showing that the circle $x^2 + y^2 = 3$ has no rational points is by checking the related equation $X^2 + Y^2 = 3Z^2$ has no solutions modulo the prime $p = 3$. We will see in the next chapter that this idea of looking at a *related equation* has a nice consequence which enables us to generalize the above argument to arbitrary plane curves defined by quadratic polynomials.

# 3 Rational Points on Conics

Let us now consider the rational solutions to a general quadratic equation in two variables with coefficients in $\mathbb{Q}$. That is, we want to look at rational points on the curve $C$ associated to the polynomial

$$g(x,y) = ax^2 + bxy + cy^2 + dx + ey + f,$$

where $a, b, c, d, e, f \in \mathbb{Q}$. The curve $C$ is called a *conic* because it defines a conic section — the *nondegenerate* case being an ellipse, hyperbola or parabola. There are also the *degenerate* (sometimes called singular) ones. A classification given by [17] tells us that a conic is degenerate if it is a union of two lines — a single point, the empty set, and a line (single, double or parallel).

## 3.1 From one rational point to infinitely many

We have seen that there is a possibility that there are no rational points at all on a conic. This was shown via the circle $x^2 + y^2 = 3$ which has no solutions in $\mathbb{Q}^2$. A natural question to ask is then the following.

**Question 3.1.** *When do conics have a rational point?*

It turns out that this seemingly elementary question requires deeper insights and the goal of the upcoming sections is to develop a theory to answer this. For now, let us take a step back and try to find rational points on the general conic by using the same method we used to find rational points on the unit circle $C : x^2 + y^2 = 1$. To do this, we take a closer look at the steps of our argument:

(1). We found one rational point $\boldsymbol{P} = (-1, 0)$ on $C$.

(2). We *project* lines with rational slope through $\boldsymbol{P}$.

(3). All such lines intersect with another distinct rational point of $C$.

Clearly, the most important part of the argument is to first find one rational point for otherwise, the succeeding steps become meaningless. Once we find such a point, we can project lines through this point. But one question that is still lingering is the following.

**Question 3.2.** *Does every projected line, except those tangent to the conic at this point, intersects the conic at another distinct point?*

The answer is no. Here are some immediate counterexamples.

**Example.**

(i). The conic $C : x^2 + y^2 = -3$ tautologically gives a no to our question as the set of real (not even rational) points on $C$ is the empty set.

(ii). The circle with zero radius $C : x^2 + y^2 = 0$ is a conic. However, $C$ contains only the single point $(0, 0)$. So, a line through this point with any slope does not hit the conic at another point.

(iii). The single line $L : y - 2x = 0$ through the origin is itself a conic. But choose any real point $\boldsymbol{P}$ and no non-tangential lines projected through $\boldsymbol{P}$ meets $L$ at some other point.

(iv). The double line $L : xy = 0$ is also a conic. However, this case is no different than the preceding example (iii).

What do these examples have in common? They are degenerate conics — (i) is the empty set, (ii) is a single point, (iii) and (iv) are lines. There is also a slight problem for the nondegenerate ones and we demonstrate this via an example. Consider running the argument on the parabola $C : y = x^2$ by starting with the minima $\mathbf{0} = (0, 0)$ which is a rational point on $C$. When projecting lines through $\mathbf{0}$, we may encounter two possible problematic lines that were not present in our unit circle example. These are the tangent line $y = 0$ and the vertical line $x = 0$. The tangent line $y = 0$ is not truly a problem as it still have rational slope and we can just say that the line $y = 0$ intersects $C$ at $\mathbf{0}$ twice. We then say that the *intersection multiplicity of $C$ and the line $y = 0$ at $\mathbf{0}$* is 2. When the curve $C$ that is intersecting is clear, we simply say that $\mathbf{0}$ *occurs with multiplicity* 2. The vertical line $x = 0$ is much more problematic. However, we can solve this issue by allowing the line to have a slope of *infinity.* To do this, we say that the line through a point $(x_0, y_0)$ on the conic has slope *infinity*, denoted $\infty$, if it is the line $x = x_0$. For the parabola $C$ and the line $x = 0$, we see that $\mathbf{0}$ again occurs with multiplicity 2. By taking care of all these cases, we have a quite satisfying answer whenever we already have one rational point on a conic.

**Theorem 3.1.** *Let $C$ be a nondegenerate conic defined by $f \in \mathbb{Q}[x, y]$, and let $\boldsymbol{P} \in C(\mathbb{Q})$. If $L_\lambda$ is a line through $\boldsymbol{P}$ with slope $\lambda \in \mathbb{Q} \cup \{\infty\}$, then it intersects at a point $\boldsymbol{Q}_\lambda \in C(\mathbb{Q})$, dependent on $\lambda$, which is distinct from $\boldsymbol{P}$ unless $L_\lambda$ is the tangent line to $C$ at $\boldsymbol{P}$.*

We give a proof that builds on one that was given by Lozano-Robledo [12].

*Proof.* Let $C$ be the conic $f(x, y) = 0$, and let $\boldsymbol{P} \in C(\mathbb{Q})$, say $\boldsymbol{P} = (x_0, y_0)$. Let $L_\lambda$ be the line

$$L_\lambda : \begin{cases} y - y_0 = \lambda(x - x_0), & \text{if } \lambda \in \mathbb{Q}, \\ x = x_0, & \text{if } \lambda = \infty. \end{cases}$$

If $\lambda = \infty$, the intersection $C \cap L_\lambda$ is defined by the polynomial

$$p_\infty(y) = f(x_0, y).$$

If otherwise $\lambda \in \mathbb{Q}$, the intersection $C \cap L_\lambda$ is defined by the polynomial

$$p_\lambda(x) = f(x,\ \lambda(x - x_0) + y_0).$$

Observe that both these polynomials are defined over $\mathbb{Q}$ and have degree at most 2 (because $f$ defines a conic). For brevity, we can define a single polynomial of intersection

$$p_\lambda(\xi) = \begin{cases} f(\xi,\ \lambda(\xi - x_0) + y_0), & \text{if } \lambda \in \mathbb{Q}, \\ f(x_0,\ \xi), & \text{if } \lambda = \infty, \end{cases}$$

but keeping in mind that the variables admit different values. By applying the Fundamental Theorem of Algebra, we know that either $p_\lambda(\xi)$ has at most 2 real roots (counting multiplicities) or $p_\lambda(\xi)$ is identically zero (i.e. zero for all $\xi$). But if $p_\lambda(\xi)$ is identically zero, then every point on $L_\lambda$ is a point on $C$. That is, $C$ contains the line $L_\lambda$ and so must be a degenerate conic. This contradicts our hypothesis, so we must have that $p_\lambda(\xi)$ is not the zero polynomial, and thus has at most 2 real roots counting multiplicities. If there is only one unique root, we are done by taking $\boldsymbol{Q}_\lambda = \boldsymbol{P}$. Otherwise, assume we have exactly two distinct roots so that $p_\lambda(\xi)$ has degree exactly 2. We now want to prove two things: that the other root is in $\mathbb{Q}$, and that it defines a rational point on $C$. To do this, we look at the two cases of $p_\lambda(\xi)$ separately. If $\lambda \in \mathbb{Q}$, then clearly $x_0$ is the first root of $p_\lambda(\xi)$. Accordingly, we may write $p_\lambda(\xi) = (x - x_0)\, q(\xi)$ for some polynomial $q \in \mathbb{Q}[\xi]$ with $q(x_0) \neq 0$. Clearly, $\deg q = 1$ (because $\deg p_\lambda = 2$) and so we can write $q(\xi) = a\,\xi + b$ for some $a, b \in \mathbb{Q}$ where $a \neq 0$. Our second root is thus given by $x_1 = -b/a$, the root of $q$, which is rational since $q$ is defined over $\mathbb{Q}$. To get the second rational point on $C$, we use the equation of the line

$L_\lambda$. We see that $y_0 = \lambda(x_1 - x_0) + y_0$ which is rational and so $\boldsymbol{Q}_\lambda = (x_1, y_1)$ gives another rational point on $C$. On the other hand, if $\lambda = \infty$, the argument is identical. Instead of $x_0$, we see that $y_0$ is the first root of $p_\lambda(\xi)$. By running the same argument, we get a second rational root $y_1$ and so $\boldsymbol{Q}_\lambda = (x_0, y_1)$ gives another rational point on $C$. ∎

## 3.2 Non-triviality condition 1: Legendre's theorem

So we now know that if we can find one rational point on a nondegenerate conic, then we can find infinitely many other rational points. But how do we guarantee the existence of one rational point? This was the question posed in Question 3.1. Now, we ask a more precise question.

**Question 3.3.** *How do we guarantee the existence of a single rational point on a conic? Is there a characterization of the conic which says something about the existence of rational points on it?*

For example, is it possible to say with absolute certainty that the conic

$$C : 66564x^2 + 8084y^2 - 1548x - 91 = 0 \tag{3.1}$$

has a rational point on it? To tackle this question, we need to first talk about quadratic forms.

### 3.2.1 Quadratic forms

Recall that under a change of variables, there is a map to go from the polynomial $x^2 + y^2 - 3$ to the homogeneous polynomial $X^2 + Y^2 - 3Z^2$. As we have seen in Chapter 1, we call the latter polynomial a quadratic form. Here we redefine this carefully.

**Definition 3.1.** A **quadratic form** $q(X_1, \ldots, X_n)$ is a degree 2 homogeneous polynomial defined over $\mathbb{Q}$. We can extend this definition to define quadratic forms defined over $\mathbb{Z}$ which we will call an **integral quadratic form**. We will usually write the variables in a tuple $\boldsymbol{X} = (X_1, \ldots, X_n)$ and so write $q(\boldsymbol{X}) = q(X_1, \ldots, X_n)$.

**Example.** *We look at one example and one non-example.*

(i). A particular case of interest for us is the case $n = 3$, called a *ternary quadratic form.* For example, $aX^2 + bY^2 = cZ^2$ is an equation defined by a ternary quadratic form. By taking $a = b = c = 1$, we recover Fermat's equation with exponent 2.

(ii). $X^3 + 9XYZ - 9Z^3 + 8YW^2$ is not a quadratic form since it is a homogeneous polynomial of degree 3.

Our first result regarding quadratic forms is a very useful one.

**Lemma 3.1.** *Let $q(\boldsymbol{X})$ be a quadratic form. Then for any $\lambda \in \mathbb{C}$, we have*

$$q(\lambda\boldsymbol{X}) = \lambda^2 q(\boldsymbol{X}).$$

*Proof.* Since $q(\boldsymbol{X})$ is a quadratic form, it is a homogeneous polynomial by definition so we can write

$$q(\boldsymbol{X}) = q(X_1, \ldots, X_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}X_iX_j,$$

where $a_{ij} \in \mathbb{Q}$. Take any $\lambda \in \mathbb{C}$ and observe that

$$\begin{aligned} q(\lambda \boldsymbol{X}) = q(\lambda X_1, \ldots, \lambda X_n) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}(\lambda X_i)(\lambda X_j) \\ &= \lambda^2 \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} X_i X_j \\ &= \lambda^2 q(X_1, \ldots, X_n) \\ &= \lambda^2 q(\boldsymbol{X}), \end{aligned}$$

as desired. ■

There's nothing special about quadratic forms here. If $f(\boldsymbol{X})$ is a homogeneous polynomial of degree $d$ over $\mathbb{Q}$, then the identity $f(\lambda\boldsymbol{X}) = \lambda^d f(\boldsymbol{X})$ also holds for all $\lambda \in \mathbb{C}$. The proof is similar as one can use the same construction as above.

**Lemma 3.2.** *If $q(\boldsymbol{X})$ is an integral quadratic form, then the equation $q(\boldsymbol{X}) = 0$ has a solution in $\mathbb{Z}^n$ if and only if it has a solution in $\mathbb{Q}^n$.*

*Proof.* The direction ($\Rightarrow$) is clear since $\mathbb{Z} \subseteq \mathbb{Q}$, and so $\mathbb{Z}^n \subseteq \mathbb{Q}^n$. We prove the converse. Suppose $(x_1, \ldots, x_n) \in \mathbb{Q}^n$ is a solution of $q(\boldsymbol{X}) = 0$. By choosing $\lambda$ to be the least common multiple of the denominators of $x_i$, we have that $\lambda x_i \in \mathbb{Z}$ for all $1 \leqslant i \leqslant n$. Now applying Lemma 3.1, we see that

$$q(\lambda x_1, \ldots, \lambda x_n) = \lambda^2 q(x_1, \ldots, x_n) = 0.$$

This implies that $(\lambda x_1, \ldots, \lambda x_n) \in \mathbb{Z}^n$ is a solution to $q(\boldsymbol{X}) = 0$. ■

Recall the Pythagorean equation $X^2 + Y^2 = Z^2$ which is defined by a ternary quadratic form. We call a triple $(x, y, z)$ a primitive Pythagorean triple if it is a Pythagorean triple with $\gcd(x, y, z) = 1$. Motivated by this definition, we define an analogous definition for solution of general quadratic forms.

**Definition 3.2.** Let $q(\boldsymbol{X})$ be an integral quadratic form. The tuple $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$ is called a **primitive solution** of $q(\boldsymbol{X}) = 0$ if

(1). $q(\boldsymbol{x}) = 0$,

(2). $\gcd(x_1, \ldots, x_n) = 1$.

The set of primitive solutions of $q(\boldsymbol{X}) = 0$ is denoted as $\mathrm{Prim}(q)$.

**Example.** Primitive solutions of $X^2 + Y^2 = Z^2$ is the primitive Pythagorean triples.

We now show that we can always construct a primitive solution once we have a nontrivial integer solution.

**Lemma 3.3.** *Let $q(\boldsymbol{X})$ be an integral quadratic form. If $q(\boldsymbol{X}) = 0$ has a nontrivial solution in $\mathbb{Z}^n$, then it has a primitive solution.*

*Proof.* Suppose $(y_1, \ldots, y_n)$ is a nontrivial solution in $\mathbb{Z}^n$ to the equation $q(\boldsymbol{X}) = 0$, and let $d = \gcd(y_1, \ldots, y_n)$. By hypothesis, at least one $y_i$ is nonzero, and so $d$ is well-defined. By definition, $d \mid y_i$ for each $i$ and thus $y_i/d = x_i$ for some $x_i \in \mathbb{Z}$. Then observe that

$$q(x_1, \ldots, x_n) = q\left(\frac{y_1}{d}, \ldots, \frac{y_n}{d}\right) = \frac{1}{d^2} q(y_1, \ldots, y_n) = 0,$$

so $\boldsymbol{x} = (x_1, \ldots, x_n)$ is a solution to $q(\boldsymbol{X}) = 0$. Since at least one $y_i$ is nonzero, we have at least one $x_i$ being nonzero so $\boldsymbol{x}$ is a nontrivial solution. Moreover, we have that $\gcd(x_1, \ldots, x_n) = 1$ by construction. Together, this gives $\boldsymbol{x}$ as a primitive solution of $q(\boldsymbol{X}) = 0$, as desired. ■

The reason why quadratic forms are useful is the following. It turns out that we can always *homogenize* a quadratic equation. Let $g(x, y)$ be a generic non-homogeneous quadratic polynomial

$$g(x, y) = ax^2 + bxy + cy^2 + dx + ey + f,$$

defined over $\mathbb{Z}$, and consider the integral quadratic form

$$q(X, Y, Z) = aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2,$$

with the same coefficients. Then we can *homogenize* $g(x, y)$ by the change of variables $x = X/Z$ and $y = Y/Z$ so that we have

$$q(X, Y, Z) = Z^2 g\left(\frac{X}{Z}, \frac{Y}{Z}\right).$$

Of course, to do this we need to impose the condition that $Z \neq 0$ which we will always implicitly assume when homogenizing a polynomial. We thus should expect a natural correspondence: if $q(X, Y, Z) = 0$ has a nontrivial integral solution in $\mathbb{Z}^3$, then $g(x, y) = 0$ has a rational point and vice-versa. Due to this correspondence, we give it a name.

**Definition 3.3.** Let $g \in \mathbb{Z}[x, y]$ be a quadratic polynomial. The ternary quadratic form $q(X, Y, Z)$ is said to be **associated to** $g$ if $q = Z^2 g(X/Z, Y/Z)$.

**Remark 4.** Since $g \in \mathbb{Z}[x, y]$, it should be well understood that the homogenization $Z^2 g(X/Z, Y/Z)$ is also defined over $\mathbb{Z}$. So whenever we say a (ternary) quadratic form associated to $g$, we mean that it is an integral quadratic form.

**Example.** *Some immediate examples of associated quadratic forms.*

(i). The quadratic form associated to $x^2 + y^2 - 1$ is $X^2 + Y^2 - Z^2$.

(ii). The quadratic form associated to the polynomial defining (3.1) is

$$q(X, Y, Z) = 66564X^2 + 8084Y^2 - 91Z^2 - 1548XZ.$$

Consider $q(X, Y, Z) = X^2 + Y^2 - Z^2$, the quadratic form associated to $f(x, y) = x^2 + y^2 - 1$. Then we know from Chapter 2 that by imposing certain conditions on $s, t \in \mathbb{Z}$, we get that $(2st, s^2 - t^2, s^2 + t^2) \in \text{Prim}(q)$. That is, the triple parametized by $s, t$ are primitive Pythagorean triples. We have also seen that the rational numbers

$$\left(\frac{s^2 - t^2}{s^2 + t^2}, \frac{2st}{s^2 + t^2}\right),$$

gives a rational point on $C_f$. So we can see that there is a correspondence here between primitive solutions of $q = 0$ and rational points on $C_f$ given by

$$(2st, s^2 - t^2, s^2 + t^2) \longleftrightarrow \left(\frac{s^2 - t^2}{s^2 + t^2}, \frac{2st}{s^2 + t^2}\right). \tag{3.2}$$

It is then natural to ask whether this holds for a general quadratic polynomial and its associated quadratic form — the answer is positive.

**Theorem 3.2.** *Let $g \in \mathbb{Z}[x, y]$ be a quadratic polynomial and let $q(X, Y, Z)$ be its associated quadratic form. Then there is a bijection* $\text{Prim}(q) \longrightarrow C_g(\mathbb{Q})$.

Of course, now, we do not have the luxury of having an explicit form for elements in $\mathrm{Prim}(q)$, but we can still mimic (3.2). The map $\alpha : \mathrm{Prim}(q) \to C_g(\mathbb{Q})$ defined by

$$(a, b, c) \longmapsto \left(\frac{a}{c}, \frac{b}{c}\right),$$

should be a good candidate for our bijection. Instead of showing this directly, it is easier to show that $\alpha$ is invertible by finding an explicit inverse map. This is both a sufficient and necessary condition for $\alpha$ to be bijective. The question now is, what is the inverse map of $\alpha$? We need to find a map that sends rational points to primitive solutions; with emphasis on the word *primitive* (this is the tricky part). Here is our construction.

*Proof of Theorem 3.2.* Consider the map $\beta : C_g(\mathbb{Q}) \to \mathrm{Prim}(q)$ defined by

$$(x, y) \longmapsto \left(\frac{\lambda x}{d}, \frac{\lambda y}{d}, \frac{\lambda}{d}\right),$$

where $\lambda$ is any integer such that $\lambda x, \lambda y \in \mathbb{Z}$ and $d = \gcd(\lambda x, \lambda y, \lambda)$. We claim that $\beta$ is the inverse map of $\alpha$. Let $(x, y) \in C_g(\mathbb{Q})$. Then its image under $\alpha \circ \beta$ is

$$(x, y) \xrightarrow{\beta} \left(\frac{\lambda x}{d}, \frac{\lambda y}{d}, \frac{\lambda}{d}\right) \xrightarrow{\alpha} (x, y),$$

and so $\alpha \circ \beta$ is the identity map $C_g(\mathbb{Q}) \to C_g(\mathbb{Q})$. If we consider $(a, b, c) \in \mathrm{Prim}(q)$, then its image under $\beta \circ \alpha$ is

$$(a, b, c) \xrightarrow{\alpha} \left(\frac{a}{c}, \frac{b}{c}\right) \xrightarrow{\beta} \left(\frac{a\lambda}{cd}, \frac{b\lambda}{cd}, \frac{\lambda}{d}\right) = (a, b, c),$$

where $\lambda = c$ and $d = \gcd(a, b, c) = 1$. So $\beta \circ \alpha$ is the identity map $\mathrm{Prim}(q) \to \mathrm{Prim}(q)$. Therefore, we conclude that $\beta$ is the inverse map of $\alpha$ and so, $\alpha$ is invertible as desired. ■

Our bijection due to Theorem 3.2 now says that to talk about a generic conic, it suffices to talk about its associated quadratic form. So we will do exactly that from now, but first, let us look at one example.

**Example.** There is a bijection between rational points on the circle $ax^2 + by^2 = r^2$ and primitive solutions of the equation $aX^2 + bY^2 = r^2Z^2$, where $a, b, r \in \mathbb{Z}$.

*Proof.* The bijection is immediate by considering the quadratic $g(x, y) = ax^2 + by^2 - c$, where $a, b, c \in \mathbb{Z}$, in Theorem 3.2. Since $c \in \mathbb{Z}$ is arbitrary, we can take it to be a perfect square $c = r^2$ for some $r \in \mathbb{Z}$. ■

This example is essentially a generalization (for circles) of the bijection we had between rational points on the unit circle $x^2 + y^2 = 1$ and the primitive Pythagorean triples (which are solutions to the Pythagorean equation $X^2 + Y^2 = Z^2$).

### 3.2.2 Legendre's theorem

We now discuss the main theorem of this section. Observe that the ternary quadratic form

$$q(X, Y, Z) = aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2, \tag{3.3}$$

where $a, b, c, d, e, f \in \mathbb{Z}$ can be written as the product $q = \boldsymbol{X}^\top A\boldsymbol{X}$ where $\boldsymbol{X} = (X, Y, Z)$ and $A$ is the real (in fact, rational) symmetric matrix

$$A = \begin{pmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{pmatrix}.$$

It should be clear that we can actually do this in general. That is, we can always write $q(\boldsymbol{X}) = \boldsymbol{X}^\top A\boldsymbol{X}$ where $\boldsymbol{X} = (X_1, \ldots, X_n)$ and $A$ is some real symmetric matrix. Now recall this beautiful theorem from linear algebra which we will state without proof.

**Theorem 3.3.** *Every real matrix is symmetric if and only if it is orthogonally diagonalizable.*

This gives the following powerful theorem which will be very useful for us.

**Theorem 3.4** (Principal axis theorem)**.** *Let $q(\boldsymbol{X}) = \boldsymbol{X}^\top A\boldsymbol{X}$ be a quadratic form where $\boldsymbol{X} = (X_1, \ldots, X_n)$ and $A$ is a real symmetric matrix. Then there exists an orthogonal matrix $\mathcal{O}$ so that with $\boldsymbol{Y} = (Y_1, \ldots, Y_n) = \mathcal{O}\boldsymbol{X}$, we have*

$$q(\boldsymbol{Y}) = \sum_{i=1}^{n} d_i Y_i^2,$$

*where $d_1, \ldots, d_n$ is the diagonal entries of the diagonal matrix $\mathcal{O}A\mathcal{O}^\top$.*

In other words, there is an (invertible) change of variables so that the product $Y_iY_j$ vanishes for all $i \neq j$. In the case of ternary quadratic forms, this means that there is a change of variables so that (3.3) becomes

$$q(U, V, W) = \alpha U^2 + \beta V^2 + \gamma W^2.$$

for some $\alpha, \beta, \gamma \in \mathbb{Z}$. We call this process the *diagonalization* of the quadratic form $q$. The proof of Theorem 3.4 is just a simple application of Theorem 3.3.

*Proof.* Since $A$ is real symmetric, Theorem 3.3 implies the existence of an orthogonal matrix $\mathcal{O}$ such that $D = \mathcal{O}A\mathcal{O}^\top$ is diagonal. We can then rewrite this relation as $A = \mathcal{O}^\top D\mathcal{O}$, so that we have $\boldsymbol{X}^\top A\boldsymbol{X} = (\mathcal{O}\boldsymbol{X})^\top D(\mathcal{O}\boldsymbol{X}) = \boldsymbol{Y}^\top D\boldsymbol{Y}$ where we have put $\boldsymbol{Y} = (Y_1, \ldots, Y_n) = \mathcal{O}\boldsymbol{X}$. Since $D$ is diagonal, we get precisely $q(\boldsymbol{Y})$ as desired. ■

The principal axis theorem tells us the following: from now on, to talk about the solutions of a ternary quadratic form, it suffices to talk about the solutions of the (much simpler) equation

$$q(X, Y, Z) = aX^2 + bY^2 + cZ^2 = 0. \tag{3.4}$$

In fact, we can make this already simple equation much simpler. We observe that there is no loss in generality here if we assume $\gcd(a, b, c) = 1$; for otherwise, just divide by their greatest common divisor. Now, suppose one of the coefficients in (3.4) is zero, say, $b = 0$. Then

$$aX^2 + cZ^2 = 0 \iff -\frac{a}{c} = \left(\frac{Z}{X}\right)^2,$$

and so (3.4) has a nontrivial solution if and only if $-a/c$ is the square of some rational number. So we may assume that none of $a, b, c$ are zero for otherwise we are essentially done with the problem of searching for solutions. We may also assume that $a, b, c$ are squarefree. To see this, we have to look at factorizations. For example, factorize $b$ into a squarefree part and a nonsquarefree part by writing $b = b's^2$ where $b'$ is squarefree. Then we can write $bY^2 = b'(sY)^2 = b'Y'^2$. So by applying a transformation similar to $sY \mapsto Y'$, we can always make the coefficients squarefree and this idea applies to $a$ and $c$ as well. Finally, we may also assume that $a, b, c$ are pairwise coprime. Let $p$ be a prime such that $p \mid a$ and $p \mid b$ so that $p \mid (aX^2 + bY^2) = -cZ^2$. Since $\gcd(a, b, c) = 1$, $p$ does not divide $c$ and so $p$ divides $Z$. This implies that we can write $Z = pU$ for some $U \in \mathbb{Z}$ and so (3.4) becomes

$$aX^2 + bY^2 + c(pU)^2 = 0.$$

Dividing both sides by $p$, we further have

$$\frac{a}{p}X^2 + \frac{b}{p}Y^2 + cpU^2 = 0.$$

Observe that initially we had $p$ as a common divisor of $a$ and $b$, but now $p$ is a common divisor of only $cp$ in this set of coefficients $a/p$, $b/p$, $cp$. We can keep repeating this procedure until the coefficients defining the quadratic form are eventually pairwise coprime. Under these assumptions, there is an elegant theorem due to Legendre that gives the existence of solutions to (3.4).

**Theorem 3.5** (Legendre, 1785)**.** *Let $a, b, c \in \mathbb{Z}$ be nonzero, pairwise coprime and squarefree. Then the equation*

$$aX^2 + bY^2 + cZ^2 = 0 \tag{3.5}$$

*has a nontrivial solution in $\mathbb{Z}^3$ if and only if*

(1). *$a, b, c$ do not all have the same sign.*

(2). *$-bc$, $-ac$ and $-ab$ are quadratic residues modulo $a$, $b$ and $c$ respectively.*

We will give a proof due to Niven et al. [15]. For this, we will need three lemmas.

**Lemma 3.4.** *Let $n \in \mathbb{Z}$ and let $\alpha, \beta, \gamma$ be positive real numbers such that $\alpha\beta\gamma = n$. Then for any $a, b, c \in \mathbb{Z}$, there is a solution $(x, y, z) \in \mathbb{Z}^3$ not all zero to the congruence equation $ax + by + cz \equiv 0 \pmod{n}$ which satisfies $|x| \leqslant \alpha$, $|y| \leqslant \beta$ and $|z| \leqslant \gamma$.*

If $x \in \mathbb{R}$, we will write $\lfloor x \rfloor$ to denote the greatest integer less than or equal to $x$. The operator $\lfloor \cdot \rfloor$ is called the *floor function* and should be familiar. By definition of the floor function, we know that $\lfloor x \rfloor \leqslant x < \lfloor x \rfloor + 1$. This is an important property which we will use in the proof of this lemma.

*Proof.* Fix $a, b, c \in \mathbb{Z}$. Consider the following set of triples

$$S = \left\{(x, y, z) \in \mathbb{Z}^3 : 0 \leqslant x \leqslant \lfloor \alpha \rfloor,\ 0 \leqslant y \leqslant \lfloor \beta \rfloor,\ 0 \leqslant z \leqslant \lfloor \gamma \rfloor \right\}.$$

We easily see that $|S| = (1 + \lfloor \alpha \rfloor)(1 + \lfloor \beta \rfloor)(1 + \lfloor \gamma \rfloor)$, and so by definition of the floor function, we deduce that $|S| > \alpha\beta\gamma = n$. By the pigeonhole principle, there must then exist two triples $(x_1, y_1, z_1)$, $(x_2, y_2, z_2) \in S$ such that

$$ax_1 + by_1 + cz_1 \equiv ax_2 + by_2 + cz_2 \pmod{n},$$

or equivalently

$$a(x_1 - x_2) + b(y_1 - y_2) + c(z_1 - z_2) \equiv 0 \pmod{n}.$$

Surely, $|x_1 - x_2| \leqslant \lfloor \alpha \rfloor \leqslant \alpha$, $|y_1 - y_2| \leqslant \lfloor \beta \rfloor \leqslant \beta$ and $|z_1 - z_2| \leqslant \lfloor \gamma \rfloor \leqslant \gamma$, and so the desired triple is given by $(x, y, z)$ with $x = x_1 - x_2$, $y = y_1 - y_2$ and $z = z_1 - z_2$. Since $a, b, c$ were arbitrary, the claim follows. ■

**Lemma 3.5.** *Let $m, n$ be coprime positive integers and let $q(X, Y, Z) = aX^2 + bY^2 + cZ^2$ be an integral quadratic form. If $q(X, Y, Z)$ can be written as a product of linear factors modulo $m$ and*

*$n$, then it can be written as a product of linear factors modulo $mn$.*

*Proof.* Suppose that we can write

$$\begin{aligned} q(X,Y,Z) &\equiv (a_mX + b_mY + c_mZ)(d_mX + e_mY + f_mZ) \pmod{m}, \\ q(X,Y,Z) &\equiv (a_nX + b_nY + c_nZ)(d_nX + e_nY + f_nZ) \pmod{n}. \end{aligned}$$

Since $m$ and $n$ are coprime, the Chinese remainder theorem implies that there exists an integer $\alpha$ such that $\alpha \equiv a_m \bmod m$ and $\alpha \equiv a_n \bmod n$. Similarly, there exist integers $\beta, \gamma, \delta, \varepsilon, \varphi$ with $\beta \equiv b_m \bmod m$ and $\beta \equiv b_n \bmod n$, $\gamma \equiv c_m \bmod m$ and $\gamma \equiv c_n \bmod n$, so on and so forth. The existence of these integers then implies that we can write

$$q(X,Y,Z) \equiv (\alpha X + \beta Y + \gamma Z)(\delta X + \varepsilon Y + \varphi Z) \pmod{mn},$$

which is a product of linear factors modulo $mn$ as desired. ■

**Lemma 3.6.** *Let $n$ be a positive integer such that $-1$ is a quadratic residue modulo $n$. Then $n$ can be written as a sum of two squares.*

We shall prove this lemma using Fermat's *method of infinite descent*. For the readers unfamiliar with this concept, this is a method that is very useful in both proving and disproving existence of solutions to a certain Diophantine equation. In our case, we will see that the assumptions of the lemma allow us to find integers $u, v$ such that $u^2 + v^2 = \alpha n$ for some positive integer $\alpha < n$. We then proceed with the *descent step*, where we do some mathematical magic to get a new pair of integers $s, t$ such that $s^2 + t^2 = \beta n$ with $\beta < \alpha < n$. The fact that $\beta < \alpha$ is the main point of the process, and this justifies the terminology *descent*. It turns out that there is nothing stopping us to do this indefinitely until we get a pair of integers $x, y$ such that $x^2 + y^2 = 1n$, for which $n$ is now a sum of two squares. That is the idea of the proof, so let us begin.

*Proof.* By assumption, there exists $u \in \mathbb{Z}$ such that $-1 \equiv u^2 \bmod n$. So by definition, we can find an integer $0 < \alpha < n$ such that $u^2 + 1 = \alpha n$. If we put $v = 1$, then $(u, v)$ gives a solution to the equation $x^2 + y^2 = \alpha n$. That is, we have

$$u^2 + v^2 = \alpha n.$$

If $\alpha = 1$, we are done. So assume that $1 < \alpha < n$. Now, choose $a, b \in \mathbb{Z}$ so that the integers

$$s = u - a\alpha, \quad t = v - b\alpha$$

satisfy the constraints $|s|, |t| \leqslant \alpha/2$. We then see that the sum $s^2 + t^2$ is bounded

$$0 < s^2 + t^2 \leqslant 2\left(\frac{\alpha}{2}\right)^2 = \frac{\alpha^2}{2} < \alpha^2. \tag{3.6}$$

We now make three observations:

(i). Firstly, observe that

$$s^2 + t^2 \equiv (u^2 + \alpha \text{ terms}) + (v^2 + \alpha \text{ terms}) \equiv u^2 + v^2 \equiv 0 \pmod{\alpha}.$$

By definition, this means that that there exists $\beta \in \mathbb{Z}$ such that $s^2 + t^2 = \beta\alpha$. Moreover the bound (3.6) implies that $0 < \beta < \alpha$.

(ii). Secondly, $su + tv \equiv u(u - a\alpha) + v(v - b\alpha) \equiv u^2 + v^2 \equiv 0 \pmod{\alpha}$.

(iii). Thirdly, $sv - tu \equiv v(u - a\alpha) - u(v - b\alpha) \equiv 0 \pmod{\alpha}$.

Then, we consider the product $\beta\alpha^2 n$ to see that

$$\begin{aligned}\beta\alpha^2 n = (\beta\alpha)(\alpha n) &= (s^2 + t^2)(u^2 + v^2) \\ &= s^2u^2 + s^2v^2 + t^2u^2 + t^2v^2 + 2suvt - 2suvt \\ &= (su + tv)^2 + (sv - tu)^2.\end{aligned}$$

Dividing through by $\alpha^2$, we thus have

$$\beta n = \left(\frac{su + tv}{\alpha}\right)^2 + \left(\frac{sv - tu}{\alpha}\right)^2.$$

From observation (ii) and (iii), it follows that the pair $\left(\frac{su+tv}{\alpha}, \frac{sv-tu}{\alpha}\right)$ defines an integer solution to the equation $x^2 + y^2 = \beta n$. We emphasize observation (i) that $0 < \beta < \alpha$, and so this completes the descent step as we have made the appearing factor smaller. Now if $\beta = 1$, we are done. If otherwise $\beta > 1$, we can repeat the descent step indefinitely until the factor is 1, for which then $n$ is a sum of two squares. ■

We are now in a position to prove Legendre's theorem.

*Proof of Legendre's Theorem 3.5.* We first prove the ($\Rightarrow$) direction. Suppose $(x, y, z) \in \mathbb{Z}^3$ is a nontrivial integer solution to (3.5) so that we get

$$ax^2 + by^2 + cz^2 = 0. \tag{3.7}$$

Then surely $a, b, c$ must not all have the same sign for otherwise the equation does not even have real solutions; so we get statement (1). We now want to show (2). Without loss of generality, we can assume that the triple $(x, y, z)$ is primitive for if not, we can just divide by their greatest common divisor.

We first claim that $\gcd(a, z) = 1$. Suppose not. Then there is a prime number $p$ such that $p$ divides both $a$ and $z$. From here, we can make two observations:

(i). We have $p^2 \mid z^2$.

(ii). We have $p \mid z^2$ and so $p \mid by^2$ since we see from (3.7) that $by^2 = -(ax^2 + cz^2)$.

From observation (ii), we deduce that $p \mid y^2$ since $p$ cannot possibly divide $b$ (as $a, b, c$ are pairwise coprime). But since $p$ is prime, then this implies that $p \mid y$ and so $p^2 \mid y^2$. Using observation (i), we then further deduce that $p^2$ divides $by^2 + cz^2$ which we see from (3.7) is equal to $-ax^2$. So $p^2 \mid ax^2$. But since $a$ is squarefree by assumption, it follows that $p^2 \nmid a$ and thus we conclude that $p \mid x$. However, this is a contradiction since we have essentially demonstrated that $p$ divides all three of $x, y, z$ whereas we assumed initially that $(x, y, z)$ is a primitive solution. So we must have $\gcd(a, z) = 1$.

Since $\gcd(a, z) = 1$, there exists an element $u \in \mathbb{Z}$ such that $uz \equiv 1 \pmod{a}$ i.e. $u$ is the inverse of $z$ modulo $a$. We can reduce equation (3.7) modulo $a$ to get

$$by^2 + cz^2 \equiv 0 \pmod{a}.$$

Multiplying this congruence equation by $u^2 b$ on both sides, we have

$$u^2b^2y^2 + bc\underbrace{(uz)^2}_{\equiv 1} \equiv 0 \pmod{a},$$

which implies that

$$-bc \equiv (uby)^2 \pmod{a}.$$

That is, $-bc$ is a quadratic residue modulo $a$. By symmetry, we can apply the same reasoning to conclude that $-ac$ and $-ab$ are quadratic residues modulo $b$ and $c$, respectively. So we get statement (2).

($\Leftarrow$). Assume that $a, b, c$ do not all have the same sign, and that $-bc$, $-ac$ and $-ab$ are quadratic residues modulo $a$, $b$ and $c$ respectively. We now want to show that there is a nontrivial integer solution $(x, y, z) \in \mathbb{Z}^3$ to the equation $q(X, Y, Z) = 0$ where $q$ is the quadratic form $q(X, Y, Z) = aX^2 + bY^2 + cZ^2$.

Firstly, we make the observation that if any of $a, b, c$ changes sign, then the latter assumption above still hold true. So without loss of generality, we may assume that $a > 0$ and $b, c < 0$. We now consider two integers that will make our computation neat:

(i). Since $-bc$ is a quadratic residue modulo $a$, then there exists an integer $s$ such that $-bc \equiv s^2 \bmod a$.

(ii). Since $\gcd(a, b) = 1$, then there exists an integer $u$ such that $ub \equiv 1 \bmod a$.

Then, reduce $q(X, Y, Z)$ modulo $a$ to get

$$\begin{aligned} q(X, Y, Z) &\equiv bY^2 + cZ^2 \pmod{a}, \\ &\equiv ub(bY^2 + cZ^2) \pmod{a}, \\ &\equiv u(b^2Y^2 + bcZ^2) \pmod{a}, \\ &\equiv u((bY)^2 - (sZ)^2) \pmod{a}, \\ &\equiv u(bY - sZ)(bY + sZ) \pmod{a}, \\ &\equiv (Y - usZ)(bY + sZ) \pmod{a}. \end{aligned}$$

So we see that $q(X, Y, Z)$ is a product of linear factors modulo $a$. By symmetry, one can further show that $q(X, Y, Z)$ can be written as a product of linear factors modulo $b$ and $c$ as well. Since $a, b, c$ are pairwise coprime, we can then apply Lemma 3.5 with integers $a$ and $b$, and then apply Lemma 3.5 again with integers $ab$ and $c$ to conclude that $q(X, Y, Z)$ can be written as a product of linear factors modulo $abc$. That is, there exist integers $\alpha, \beta, \gamma, \delta, \varepsilon, \varphi$ such that

$$q(X, Y, Z) \equiv (\alpha X + \beta Y + \gamma Z)(\delta X + \varepsilon Y + \varphi Z) \pmod{abc}. \tag{3.8}$$

Now, we make the fundamental observation that the real numbers $\sqrt{bc}$, $\sqrt{|ac|}$, $\sqrt{|ab|}$ are positive and satisfy $\sqrt{bc}\sqrt{|ac|}\sqrt{|ab|} = \sqrt{a^2b^2c^2} = abc \in \mathbb{Z}$. It follows that we can apply Lemma 3.4 to find a solution $(x, y, z) \in \mathbb{Z}^3$ not all zero to the congruence equation

$$\alpha X + \beta Y + \gamma Z \equiv 0 \pmod{abc},$$

satisfying the constraints $|x| \leqslant \sqrt{bc}$, $|y| \leqslant \sqrt{|ac|}$, $|z| \leqslant \sqrt{|ab|}$. But $\alpha X + \beta Y + \gamma Z$ is just a linear factor modulo $abc$ in (3.8). So $(x, y, z)$ passes to a solution of the congruence equation

$$q(X, Y, Z) \equiv 0 \pmod{abc}. \tag{3.9}$$

In other words, $abc$ divides $q(x, y, z)$.

Next, we look closely at the constraints on $|x|, |y|, |z|$. Since $a, b, c$ are squarefree and pairwise coprime, it follows that any two product of them is squarefree. This implies that $\sqrt{bc} \in \mathbb{Z}$ if and only if $\sqrt{bc} = 1$. In other words, the inequality $|x| \leqslant \sqrt{bc}$ attains equality if and only $bc = 1$ i.e. when $b = c = -1$. Using the same line of reasoning, we can conclude that the inequality $|y| \leqslant \sqrt{|ac|}$ attains equality if and only if $a = 1, c = -1$, and that the inequality $|z| \leqslant \sqrt{ab}$ attains equality if and only if $a = 1, b = -1$. In these last two cases, we can get a nontrivial solution quite easily. If $a = 1, c = -1$, the problem becomes solving $X^2 + bY^2 - Z^2 = 0$ which has the nontrivial solution $(1, 0, 1)$. If on the other hand $a = 1, b = -1$, the problem becomes solving $X^2 - Y^2 + cZ^2 = 0$ which has the nontrivial solution $(1, 1, 0)$. In the case $b = c = -1$, the problem becomes solving $aX^2 - Y^2 - Z^2 = 0$, which unfortunately has no obvious nontrivial solution. However, the assumption $b = c = -1$ implies that $-1$ is a quadratic residue modulo $a$ (because $-bc = -1$ and our hypothesis is that $-bc$ is a quadratic residue modulo $a$). So since $a > 0$, we can thus invoke Lemma 3.6 to deduce the existence of integers $y', z'$ satisfying $y'^2 + z'^2 = a$. It then follows that $(1, y', z')$ is a nontrivial solution as $a - (y'^2 + z'^2) = a - a = 0$.

We now assume that $a, b, c$ are not of the special cases above so that all the constraints cannot attain equality. Since we assumed that $b, c < 0$, it follows that $-abc < by^2, cz^2 \leqslant 0$. Since we also assumed $a > 0$, we thus have the bound

$$-2abc < ax^2 + by^2 + cz^2 \leqslant ax^2 < abc,$$

or in other words,

$$-2abc < q(x, y, z) < abc.$$

But from (3.9), we know that $q(x, y, z)$ is an integer divisible by $abc$. So we are forced to conclude that either $q(x, y, z) = 0$ or $q(x, y, z) = -abc$. If $q(x, y, z) = 0$, we are done. So assume that $q(x, y, z) = -abc$. In this case, we can construct a solution $(\tilde{x}, \tilde{y}, \tilde{z})$ defined in the following way:

$$\tilde{x} = -by + xz, \quad \tilde{y} = ax + yz, \quad \tilde{z} = z^2 + ab.$$

To verify that $(\tilde{x}, \tilde{y}, \tilde{z})$ is indeed a solution, we compute

$$\begin{aligned}
q(\tilde{x}, \tilde{y}, \tilde{z}) &= a\tilde{x}^2 + b\tilde{y}^2 + c\tilde{z}^2 \\
&= a(-by + xz)^2 + b(ax + yz)^2 + c(z^2 + ab)^2 \\
&= ab^2y^2 + ax^2z^2 - \cancel{2abxyz} + a^2bx^2 + by^2z^2 + \cancel{2abxyz} + cz^4 + a^2b^2c + 2abcz^2 \\
&= ab\underbrace{(ax^2 + by^2 + cz^2)}_{=-abc} + z^2\underbrace{(ax^2 + by^2 + cz^2)}_{=-abc} + abcz^2 + a^2b^2c \\
&= -a^2b^2c + -abcz^2 + abcz^2 + a^2b^2c \\
&= 0.
\end{aligned}$$

What happens if $(\tilde{x}, \tilde{y}, \tilde{z})$ is trivial? Then necessarily, we have $\tilde{z} = 0$ and so $z^2 = -ab$. As established previously, the product $ab$ is squarefree and so this implies that $z = \pm 1$, which in turn implies that $-ab = 1$. This is equivalent to $a = 1, b = -1$ which is one of the special cases that we have already settled. ■

With Legendre's theorem, we have a satisfying answer to Question 3.3 thanks to Lemma 3.3 and Theorem 3.2 combined. Given a conic $C$ with associated quadratic form $q(X, Y, Z)$, the link between these results is captured in the following diagram:

$$\begin{tikzcd} q(X,Y,Z) = 0 \arrow[r, "\text{Legendre (existence)}"] & \{ \text{ nontrivial solutions } \} \arrow[r, leftrightarrow, "\text{Lem. 3.3}"] & \mathrm{Prim}(q) \arrow[r, leftrightarrow, "\text{Thm. 3.2}"] & C(\mathbb{Q}). \end{tikzcd}$$

In particular, we see that there is a direct connection

$$\begin{tikzcd} q(X,Y,Z) = 0 \arrow[r, "\text{Legendre (existence)}"] & C(\mathbb{Q}). \end{tikzcd}$$

Thus, in a sense, Legendre's theorem gives us an algorithm of checking rational points on a conic:

(1). A conic $C : f(x, y) = 0$ is given.

(2). We look at its associated quadratic form $q(X, Y, Z) = Z^2 f(X/Z, Y/Z)$.

(3). Diagonalize $q$ to get $q(U, V, W) = aU^2 + bV^2 + cW^2$ for some $a, b, c \in \mathbb{Z}$ that are nonzero, pairwise coprime, squarefree and not all having the same sign.

(4). Compute $\alpha = -bc$, $\beta = -ac$ and $\gamma = -ab$.

(5). If $\alpha$, $\beta$ and $\gamma$ are **all** quadratic residues modulo $a$, $b$ and $c$ respectively, then return: there exists a rational point $C$.

(6). Otherwise, return: there does not exist any rational point on $C$.

Let us now look at a concrete original example.

**Example.** We revisit the conic (3.1) that we mentioned at the start of this section. We claim that there exists a rational point on the conic

$$C : 66564x^2 + 8084y^2 - 1548x - 91 = 0.$$

*Proof.* By Theorem 3.2 and Lemma 3.3, it is enough to prove that the associated quadratic form

$$q(X, Y, Z) = 66564X^2 + 8084Y^2 - 91Z^2 - 1548XZ,$$

has a nontrivial solution in $\mathbb{Z}^3$. We shall first diagonalize $q$. To do this, we need to complete the square in $X$ with a suitable $Z$ coefficient so that the $XZ$ term vanishes. By inspection, we see that this is achieved if we write

$$q(X, Y, Z) = 36\left(43X - \frac{Z}{2}\right)^2 + 8084Y^2 - 100Z^2.$$

Then, we consider the change of variables $R = 43X - Z/2$, $S = Y$, $T = Z$ to get the quadratic form

$$q'(R, S, T) = 36R^2 + 8084S^2 - 100T^2.$$

Now, we divide by $\gcd(36, 8084, 100) = 4$ to get

$$\frac{q'(R, S, T)}{4} = 9R^2 + 2021S^2 - 25T^2.$$

Finally, we apply the change of variables $U = 3R$, $V = S$, $W = 5T$ which make the coefficients squarefree (note that $2021 = 43 \times 47$ and so is already squarefree) to get the integral quadratic form

$$q''(U, V, W) = U^2 + 2021V^2 - W^2.$$

We are now in the setting of Legendre's theorem with $a = 1$, $b = 2021$ and $c = -1$. Since any integer modulo 1 is a quadratic residue, and that 1 is a quadratic residue modulo any integer, it follows by Legendre's theorem that $q'' = 0$ has a nontrivial solution. ■

Using the computer algebra system SageMath [28], we verify that indeed $C$ has a rational point via the following code:

```
> P.<X, Y, Z> = ProjectiveSpace(QQ, 2)
> C = Conic(66564*X^2 + 8084*Y^2 - 91*Z^2 - 1548*X*Z)
> C.has_rational_point(point=True)
(True, (13/258 : 0 : 1))
```

The code returned $\boldsymbol{P} = (13/258, 0)$ to be a rational point on $C$, and so we can parametrize all other rational points on $C$ using $\boldsymbol{P}$.

It turns out that Legendre's theorem is a special case of the celebrated Hasse-Minkowski theorem. It is a shame if we do not discuss this a little as this is a good chance to introduce the theory of $p$-adic analysis, a very powerful tool in arithmetic geometry and number theory in general.

## 3.3 Non-triviality condition 2: Hasse-Minkowski theorem

We begin with a discussion on *absolute values* on $\mathbb{Q}$, finishing with a theorem of Ostrowski which classify all possible nontrivial absolute values on the rational numbers. Then we move on to constructing the $p$-adic numbers and finally, discuss the Hasse-Minkowski theorem.

### 3.3.1 Absolute values

There are many ways to formulate the Hasse-Minkowski theorem albeit they all eventually lead to the same idea. Our preferred way is the formulation using $p$-adic fields $\mathbb{Q}_p$. For this, we first need to construct the $p$-adic numbers and deduce basic facts about it.

**Definition 3.4.** An **absolute value** on a field $K$ is a map $|\cdot| : K \to [0, \infty)$ satisfying the following conditions:

(1). (Positive). $|x| = 0$ if and only if $x = 0$,

(2). (Multiplicative). $|xy| = |x||y|$ for all $x, y \in K$,

(3). (Triangle inequality). $|x + y| \leqslant |x| + |y|$ for all $x, y \in K$.

Recall that in the field of real numbers $\mathbb{R}$, we have the usual absolute value, denoted $|\cdot|_\infty$, which gives a notion of size and distance. This passes to an absolute value on the field of rational numbers $\mathbb{Q}$. Of course, this is not the only absolute value on $\mathbb{Q}$ as, for example, we have the *trivial absolute value* given by $|x| = 1$ for any $x \neq 0$ and $|0| = 0$. Absolute values different from the trivial absolute value is said to be *nontrivial*, with $|\cdot|_\infty$ being an immediate example. Our next goal is to define a new family of nontrivial absolute values on $\mathbb{Q}$ using something called a *$p$-adic valuation.*

**Definition 3.5.** Let $p$ be a prime number. For any nonzero $a \in \mathbb{Z}$, we define the **$p$-adic valuation** of $a$, denoted $\nu_p(a)$, to be the greatest positive integer $k$ such that $p^k \mid a$. Furthermore, we can extend the definition to any $a/b \in \mathbb{Q}^\times$ by defining $\nu_p(a/b) = \nu_p(a) - \nu_p(b)$.

In simpler words, the $p$-adic valuation of any nonzero $a \in \mathbb{Z}$ is the power $k$ appearing in the factorization $a = bp^k$, for some $b \in \mathbb{Z}$ such that $\gcd(b, p) = 1$. For $a = 0$, we shall define $\nu_p(0) = \infty$. Gouvêa [8] reasons this in the following way: a natural way to compute, say, $\nu_2(80)$ is to keep dividing by 2 to get a sequence

$$80 \to 40 \to 20 \to 10 \to 5,$$

where we stopped at 5 because it is not divisible by 2 anymore. The number of steps is then $\nu_2(80)$ which in this case is 4. If we apply the same logic to $\nu_p(0)$, we would have to keep dividing 0 by $p$ infinitely many times

$$0 \to 0 \to 0 \to 0 \to \cdots,$$

which justifies our definition. For the case of $a/b \in \mathbb{Q}^\times$, there is an even simpler way to formulate its $p$-adic valuation: $\nu_p(a/b)$ is the power $k$ appearing in the factorization

$$\frac{a}{b} = \left(\frac{m}{n}\right) p^k,$$

such that $\gcd(m,p) = \gcd(n,p) = 1$. This can be seen by unpacking definitions.

**Example.** *Some immediate examples.*

(i). $\nu_{13}(21) = 0$. It is trivial that for any nonzero $a \in \mathbb{Z}$, $\nu_p(a) = 0$ if $p \nmid a$.

(ii). $\nu_7(1715) = 3$. This is because $1715 = 343 \cdot 5 = 7^3 \cdot 5$.

(iii). $\nu_5(2020) = 1$. This is because $2020 = 2^2 \cdot 5 \cdot 101$ and observe that $\gcd(5, 2^2 \cdot 101) = 1$.

(iv). $\nu_{43}\left(\frac{2020}{2021}\right) = -1$. To see this, we use our latest formulation to have

$$\frac{2020}{2021} = \left(\frac{2020}{47}\right)\left(\frac{1}{43}\right) = \left(\frac{2020}{47}\right) 43^{-1}.$$

We can see that the $p$-adic valuation on $\mathbb{Z}$ have a shared property with the logarithm function: $\nu_p(mn) = \nu_p(m) + \nu_p(n)$. Morover, we have that

$$\nu_p(m+n) \geqslant \min\{\nu_p(m), \nu_p(n)\}, \tag{3.10}$$

for any $m, n \in \mathbb{Z}$. To see both of this facts, just consider the factorization of $m$ and $n$, and look at the highest power of $p$. The first one should be immediate. However, one might falsely conclude that we have an equality instead in (3.10). Here's an immediate counterexample to this false conclusion: $\nu_5(2020) = 1$ and $\nu_5(5) = 1$ but $\nu_5(2020+5) = \nu_5(2025) = 2 > 1 = \min\{\nu_5(2020), \nu_5(5)\}$.

**Lemma 3.7.** *Let $p$ be a prime number. Then for any $x, y \in \mathbb{Q}^\times$, we have*

(1). $\nu_p(xy) = \nu_p(x) + \nu_p(y)$,

(2). $\nu_p(x+y) \geqslant \min\{\nu_p(x), \nu_p(y)\}$.

*Proof.* Let $x, y \in \mathbb{Q}$. From our previous discussion, we know that we can write

$$x = \left(\frac{a}{b}\right) p^m, \quad y = \left(\frac{c}{d}\right) p^n,$$

for some $a, b, c, d \in \mathbb{Z}$ and for some positive integers $m, n$ such that $a, b, c, d$ are all coprime to $p$. We then have $\nu_p(x) = m$ and $\nu_p(y) = n$. Now consider their product,

$$xy = \left(\frac{ac}{bd}\right) p^{m+n}.$$

By coprimality of $a, b, c, d$ and $p$, we have that $\gcd(ac, p) = \gcd(bd, p) = 1$. So it follows that $\nu_p(xy) = m + n = \nu_p(x) + \nu_p(y)$. Now, assume without loss of generality that $m \leqslant n$. If we consider their sum, we can take out all the common $p$ powers:

$$x + y = p^m \left(\frac{a}{b} p^{n-m} + \frac{c}{d}\right) = p^m \left(\frac{adp^{n-m} + bc}{bd}\right).$$

Since $\gcd(bd, p) = 1$, it follows that $\nu_p(x+y) \geqslant m = \min\{m, n\} = \min\{\nu_p(x), \nu_p(y)\}$. ■

Using the notion of $p$-adic valuation, we can define an absolute value on $\mathbb{Q}$.

**Definition 3.6.** The map $|\cdot|_p : \mathbb{Q} \to [0, \infty)$ defined by

$$|x|_p = \begin{cases} p^{-\nu_p(x)}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

is called the ***p*-adic absolute value** on $\mathbb{Q}$.

We will prove that this is actually an absolute value in Proposition 3.1 below. In fact, we will prove a stronger condition.

**Definition 3.7.** Let $K$ be a field. We say that an absolute value $|\cdot| : K \to [0, \infty)$ is **non-Archimedean** if $|x+y| \leqslant \max\{|x|, |y|\}$ for all $x, y \in K$. Otherwise, we say that it is **Archimedean**.

Observe that the non-Archimedean property is stronger than the triangle inequality. It gives a sharper bound as $\max\{|x|, |y|\} \leqslant |x| + |y|$ is true for all $x, y \in K$. Our usual absolute value $|\cdot|_\infty$ on $\mathbb{R}$ is Archimedean. It turns out, our absolute value of interest $|\cdot|_p$ is a non-Archimedean one.

**Proposition 3.1.** *The p-adic absolute value $|\cdot|_p$ is a non-Archimedean absolute value on $\mathbb{Q}$.*

*Proof.* Fix $p$ a prime number. We have to check three things.

(Positive). Clearly, if $x = 0$, then $|x|_p = 0$ by definition. The converse is true as well as $p^{-\nu_p(x)} \neq 0$ for any $x \neq 0$. So $|x|_p = 0$ must imply $x = 0$.

(Multiplicative). Let $x, y \in \mathbb{Q}$. Then

$$|xy|_p = p^{-\nu_p(xy)} = p^{-\nu_p(x)-\nu_p(y)} = p^{-\nu_p(x)}p^{-\nu_p(y)} = |x|_p\,|y|_p,$$

where we have used (1) of Lemma 3.7 in the second equality.

(Triangle inequality). We will prove that $|\cdot|_p$ is non-Archimedean as this implies the triangle inequality. Let $x, y \in \mathbb{Q}$. If either $x$ or $y$ is zero, or even if their sum is zero, then we are done. So suppose not. Now, $\nu_p(x+y) \geqslant \min\{\nu_p(x), \nu_p(y)\}$ by (2) of Lemma 3.7. Consequently,

$$|x+y|_p = p^{-\nu_p(x+y)} \leqslant \max\left\{p^{-\nu_p(x)}, p^{-\nu_p(y)}\right\} = \max\{|x|_p, |y|_p\} \leqslant |x|_p + |y|_p.$$

The first inequality gives non-Archimedean whereas the second inequality gives the triangle inequality. ■

Before we move on to more theory, we prove some basic facts about absolute values that will be useful for us later on. It is easy to confuse the multiplicative identity in a general field $K$ and the one in $\mathbb{R}$ when we put them side by side. To emphasize distinction, we shall write $1_K \in K$ and $1_\mathbb{R} \in \mathbb{R}$ for their respective multiplicative identities when needed.

**Lemma 3.8.** *Let $|\cdot|$ be an absolute value on a field $K$. Then $|1_K| = 1_\mathbb{R}$ and $|x| = 1_\mathbb{R}$ for any $x \in K$ such that $x^n = 1_K$ for some $n \in \mathbb{Z}^+$.*

A useful reminder before we prove this lemma is to note that the image of the absolute value is a subset of $[0, \infty) \subseteq \mathbb{R}$. That is, we are only dealing with real numbers in the image, so we can apply known results from calculus.

*Proof.* To prove the first claim, we first observe that $1_K = 1_K^2$. By the multiplicative property of absolute values, we thus have $|1_K| = |1_K|^2$. Due to the positive property, the only way this is true is if $|1_K| = 1_\mathbb{R}$. We can use the same argument to prove the second claim. We have

$$|x|^n = |x^n| = |1_K| = 1_\mathbb{R},$$

where the last equality is due to the first part of the lemma. There are two cases to consider when solving the equation $z^n = 1_{\mathbb{R}}$ for real solutions. If $n$ is odd, the equation has $1_{\mathbb{R}}$ as its only solution. On the other hand, if $n$ is even, we have $\pm 1_{\mathbb{R}}$ as solutions. Now replace $z$ with $|x|$ to deduce that we must have $1_{\mathbb{R}}$ as a solution in both cases since absolute values must be positive by definition. ■

Recall what a *metric* on a set is. It is a map $d : X \times X \to [0, \infty)$ where $X$ is any non-empty set, such that for all $x, y \in X$ we have $d(x, y) = 0$ if and only if $x = y$; $d(x, y) = d(y, x)$; and $d(x, y) \leqslant d(x, z) + d(z, y)$ for all $z \in X$. The pair $(X, d)$ is then called a metric space. We know that any norm $\| \cdot \|$ defined on a field induces a metric by taking $d(x, y) = \|x - y\|$. How about absolute values? The answer is positive as well by using the same construction.

**Lemma 3.9.** *Let $| \cdot |$ be an absolute value on a field $K$. The metric $d(x, y) = |x - y|$ is a metric on $K$. We call $d$ the* ***metric induced by*** $| \cdot |$.

*Proof.* The first metric axiom is straightforward from the absolute value axioms: $d(x, y) = 0$ if and only if $|x - y| = 0$ which is true if and only if $x = y$. The second metric axiom is also immediate due to $K$ being a field: $d(x, y) = |x - y| = |y - x| = d(y, x)$. The only thing left to prove is the triangle inequality which is also straightforward:

$$d(x, y) = |x - y| = |(x - z) + (z - y)| \leqslant |x - z| + |z - y| = d(x, z) + d(z, y).$$

Since the $x, y, z$ appearing above are all arbitrary, we are done. ■

**Definition 3.8.** Let $X$ be a non-empty set. We say that a metric $d : X \times X \to [0, \infty)$ is **non-Archimedean** if $d(x, y) \leqslant \max \{d(x, z), d(z, y)\}$ for any $x, y, z \in K$. Otherwise, we say that it is Archimedean.

We have two things named non-Archimedean, so they better be related.

**Lemma 3.10.** *Let $| \cdot |$ be an absolute value on a field $K$, and let $d$ be the metric induced by $| \cdot |$. Then $| \cdot |$ is a non-Archimedean absolute value if and only if $d$ is a non-Archimedean metric.*

*Proof.* Suppose $| \cdot |$ is a non-Archimedean absolute value. Then

$$d(x, y) = |x - y| \leqslant |(x - z) + (z - y)| \leqslant \max \{|x - z|, |z - y|\} = \max \{d(x, z), d(z, y)\},$$

which implies that $d$ is a non-Archimedean metric. Conversely, suppose $d$ is a non-Archimedean metric. Then

$$|x + y| = |x - (-y)| = d(x, -y) \leqslant \max \{d(x, 0), d(0, -y)\} = \max \{|x|, |-y|\}.$$

Since $(-1)^2 = 1$, it follows that $|-1| = 1$ by Lemma 3.8. This implies that $|-y| = |-1||y| = |y|$, so we are done. ■

So for us, the $p$-adic absolute value on $\mathbb{Q}$ induces a non-Archimedean metric $d(x, y) = |x - y|_p$. This allows us to talk about Cauchy sequences in $\mathbb{Q}$ with respect to this metric. We recall what it means for a sequence to be Cauchy.

**Definition 3.9.** A sequence $\{x_n\}$ in a metric space $(X, d)$ is said to be **Cauchy** if for every $\varepsilon > 0$, there is a positive integer $N$ such that $d(x_m, x_n) < \varepsilon$ whenever $m, n \geqslant N$.

So if $K$ is a field equipped with an absolute value $|\cdot|$, it makes sense to talk about Cauchy sequences in $K$ by considering the metric $d$ induced by $|\cdot|$. More explicitly, we say that a sequence $\{x_n\}$ in $K$ is *Cauchy with respect to* $|\cdot|$ if it is Cauchy in the metric space $(K, d)$.

**Lemma 3.11.** *Let $|\cdot|$ be a non-Archimedean absolute value on a field $K$. Any sequence $\{x_n\}$ in $K$ is Cauchy with respect to $|\cdot|$ if and only if $|x_{n+1} - x_n| \to 0$.*

*Proof.* The ($\Rightarrow$) direction should be obvious. For the converse ($\Leftarrow$), suppose that $|x_{n+1} - x_n| \to 0$. Then for every $\varepsilon > 0$, there is a positive integer $N$ such that $|x_{n+1} - x_n| < \varepsilon$ for all $n \geqslant N$. Consider the index $m = n + k > n$. Then observe that

$$|x_m - x_n| = |x_{n+k} - x_n| = |x_{n+k} - \underbrace{x_{n+k-1} + x_{n+k-1}}_{=0} - \underbrace{x_{n+k-2} + x_{n+k-2}}_{=0} - \underbrace{\cdots + x_{n-1}}_{=0} - x_n|.$$

By the non-Archimedean hypothesis on $|\cdot|$, we can apply the non-Archimedean property repeatedly so that the right hand side is $\leqslant \max\{|x_{n+k} - x_{n+k-1}|, |x_{n+k-1} - x_{n+k-2}|, \ldots, |x_{n+1} - x_n|\}$. Since we assumed $n + k > n$ and we have convergence to zero for $n \geqslant N$, it follows that this bound goes to 0 as well. But $m, n$ was arbitrary, so the claim follows. ■

This claim is not true for Archimedean absolute values. For example, consider the well-known harmonic series from real analysis

$$x_n = \sum_{k=1}^{n} \frac{1}{k}.$$

We have $|x_{n+1} - x_n| = 1/(n+1) \to 0$ where $|\cdot|$ is the usual absolute value in $\mathbb{R}$. But it is a standard real analysis fact that the harmonic series diverges, and so cannot be Cauchy.

Let $(X, d)$ be metric space. Just as a norm induces a metric, the metric $d$ itself induces a topology on $X$. Such a topology is the one generated by the collection of all open balls defined by $d$. Write

$$B(x, \varepsilon) = \{y \in X \mid d(x, y) < \varepsilon\},$$

to mean the *$d$-open ball with center $x$ and radius $\varepsilon$*. We then say any subset $U \subseteq X$ is *$d$-open* if for any $x \in U$, there exists $\varepsilon_x > 0$ such that $B(x, \varepsilon_x) \subseteq U$.

**Definition 3.10.** Let $d_1, d_2$ be two metrics on a set $X$, and let $x \in X$. We say $d_1$ and $d_2$ are **equivalent** if they induce the same topology on $X$. In other words, a subset $U \subseteq X$ is $d_1$-open if and only if it is $d_2$-open. If $X$ is a field, then we say that two absolute values on $X$ are **equivalent** if they induce equivalent metrics.

The punchline of our discussion about Cauchy sequences, equivalent metrics and absolute values is the theorem of Ostrowski, which says that the usual absolute value together with the $p$-adic absolute value defines all the possible nontrivial absolute values on $\mathbb{Q}$.

**Theorem 3.6** (Ostrowski). *Every nontrivial absolute value $|\cdot|$ on $\mathbb{Q}$ is equivalent to either the usual Archimedean absolute value $|\cdot|_\infty$ or $|\cdot|_p$ for some prime number $p$.*

*Proof.* See Chapter 1, Theorem 1 of Koblitz [10] or Chapter 3, Theorem 3.1.4 of Gouvêa [8]. ■

### 3.3.2 $p$-adic field, Hasse's principle, and the main theorem

Suppose $|\cdot|_p$ is a non-Archimedean absolute value on $\mathbb{Q}$. By Ostrowski's theorem, we know that this is one of the $p$-adic absolute values (which explains why we wrote $|\cdot|_p$). Now consider the set

$$\mathcal{C}_p = \{\{x_n\} \subseteq \mathbb{Q} : \{x_n\} \text{ is Cauchy with respect to } |\cdot|_p\}.$$

It turns out that we can give $\mathcal{C}_p$ a commutative ring structure by defining addition + and multiplication × in the following way: for any $\{x_n\}, \{y_n\} \in \mathcal{C}_p$, define

$$\{x_n\} + \{y_n\} = \{x_n + y_n\}, \quad \{x_n\} \times \{y_n\} = \{x_n y_n\}. \tag{3.11}$$

Furthermore, the additive and multiplicative identities are given by the sequences

$$0_{\mathcal{C}_p} = \{0, 0, \ldots\}, \quad 1_{\mathcal{C}_p} = \{1, 1, \ldots\}.$$

**Remark 5.** We will drop the multiplication symbol × and instead write $\{x_n\}\{y_n\} = \{x_n y_n\}$; this is standard practice in abstract algebra.

We shall omit the proof of (3.11) as it is identical to the case of Cauchy sequences in $\mathbb{R}$ with respect to the usual Archimedean absolute value. The point now is that we have the following.

**Proposition 3.2.** *Let $|\cdot|_p$ be any non-Archimedean absolute value on $\mathbb{Q}$. Then $\mathcal{C}_p$ is a unital commutative ring.*

**Remark 6.** Note that $\mathcal{C}_p$ is not a field as it is not even an integral domain. For example, the sequence $\{1, 0, 0, \ldots\}$ and $\{0, 0, \pi, \ldots\}$ are zero divisors as

$$\{1, 0, 0, \ldots\}\{0, 0, \pi, \ldots\} = \{0, 0, 0, \ldots\} = 0_{\mathcal{C}_p}.$$

We now claim that $\mathbb{Q}$ lives inside this monstrous ring $\mathcal{C}_p$.

**Lemma 3.12.** *There is an injective ring homomorphism $\mathbb{Q} \hookrightarrow \mathcal{C}_p$.*

To see this, we first observe that for any $x \in \mathbb{Q}$, the constant sequence $\bar{x} := \{x, x, x, \ldots\}$ is an element of $\mathcal{C}_p$. This is true as constant sequences are trivially Cauchy. We can then construct an injective homomorphism by taking each element of $\mathbb{Q}$ to its constant sequence representation.

*Proof.* The map $\mathbb{Q} \to \mathcal{C}_p$ defined by $x \mapsto \bar{x}$ gives the desired injective ring homomorphism. ■

Now consider the set $\mathfrak{m}_p = \{\{x_n\} \in \mathcal{C}_p \mid x_n \to 0\}$, which is a subset of $\mathcal{C}_p$. This set consists of Cauchy sequences which converges to 0 with respect to the non-Archimedean absolute value *inherited* from $\mathcal{C}_p$. We claim that this set is an ideal in $\mathcal{C}_p$.

**Lemma 3.13.** *$\mathfrak{m}_p$ is an ideal in $\mathcal{C}_p$.*

While this seems like an algebraic claim, the proof is really a routine of real analysis. We shall use the standard fact that Cauchy sequences in any metric space is bounded.

*Proof.* It should be clear that $\mathfrak{m}_p$ is an additive subgroup of $\mathcal{C}_p$. Let $\{x_n\} \in \mathcal{C}_p$ and let $\{z_n\} \in \mathfrak{m}_p$. We want to show that $x_n z_n \to 0$ with respect to $|\cdot|_p$. Since $\{x_n\}$ is Cauchy, then it is bounded, say, by $M > 0$. Since $z_n \to 0$, there exists a positive integer $N_\varepsilon$ such that for any $n \geqslant N_\varepsilon$, we have $|z_n|_p = |z_n - 0|_p < \varepsilon/M$ for any $\varepsilon > 0$. It then follows that for sufficiently large $n$,

$$|z_n x_n - 0|_p = |x_n z_n|_p \leqslant M|z|_p < M\left(\frac{\varepsilon}{M}\right) = \varepsilon$$

for any $\varepsilon > 0$. So $\{x_n z_n\} \in \mathfrak{m}_p$, and $\mathfrak{m}_p$ is an ideal in $\mathcal{C}_p$ as desired. ■

Recall what it means for an ideal in a ring to be *maximal.* If $R$ is a ring and $\mathfrak{m} \subseteq R$ is an ideal, we say that $\mathfrak{m}$ is *maximal* in $R$ if for any ideal $J \subseteq R$ such that $\mathfrak{m} \subseteq J$, we either have $J = \mathfrak{m}$ or $J = R$. We now make the following claim about $\mathfrak{m}_p$.

**Lemma 3.14.** *$\mathfrak{m}_p$ is a maximal ideal in $\mathcal{C}_p$.*

*Proof.* We will follow closely a proof given by Gouvêa [8]. Let $\{x_n\} \in \mathcal{C}_p$ be a Cauchy sequence such that $\{x_n\} \notin \mathfrak{m}_p$. Consider the ideal $I$ generated by two elements $\{x_n\}$ and $\mathfrak{m}_p$. We will show that $\bar{1} \in I$ as this would imply $I = \mathcal{C}_p$.

By definition of $\{x_n\} \notin \mathfrak{m}_p$, there exists a real number $c > 0$ and a positive integer $N$ such that for any $n \geqslant N$, we have $|x_n|_p \geqslant c > 0$. So, $x_n \neq 0$ for $n \geqslant N$. Now consider the sequence $\{y_n\}$ defined by

$$y_n = \begin{cases} 1/x_n, & \text{if } n \geqslant N, \\ 0, & \text{if } n < N. \end{cases}$$

We claim that $\{y_n\} \in \mathcal{C}_p$ i.e. $\{y_n\}$ is Cauchy with respect to $|\cdot|_p$. To prove this, we first observe that

$$|y_{n+1} - y_n|_p = \left| \frac{1}{x_{n+1}} - \frac{1}{x_n} \right|_p = \frac{|x_{n+1} - x_n|_p}{|x_{n+1} x_n|_p} \leqslant \frac{|x_{n+1} - x_n|_p}{c^2} \longrightarrow 0.$$

Since $|\cdot|_p$ is non-Archimedean, Lemma 3.11 implies that $\{y_n\}$ is Cauchy. Now put $\{z_n\} = \{x_n\}\{y_n\}$. Since $\{x_n\}$ and $\{y_n\}$ are both Cauchy, it follows that $\{z_n\} \in \mathcal{C}_p$. But what is $\{z_n\}$? It is the sequence

$$z_n = \begin{cases} 1, & \text{if } n \geqslant N \\ 0, & \text{if } n < N \end{cases} = \{0, 0, \ldots, 0, 1, 1, \ldots\}.$$

If we further consider the sequence $\bar{1} - \{z_n\}$, we see that

$$\bar{1} - \{z_n\} = \{1, 1, \ldots\} - \{0, 0, \ldots, 0, 1, 1, \ldots\} = \{1, 1, \ldots, 1, 0, 0, \ldots\} \longrightarrow 0,$$

That is, we have $\bar{1} - \{z_n\} \in \mathfrak{m}_p$. But this is equivalent to saying that

$$\bar{1} = \{z_n\} + \{w_n\} = \{x_n\}\{y_n\} + \{w_n\},$$

for some $\{w_n\} \in \mathfrak{m}_p$. That is, $\bar{1}$ is a $\mathcal{C}_p$-linear combination of $\{x_n\}$ and $\mathfrak{m}_p$, and so belongs to $I$. ■

The motivation for us in proving the preceding lemma is the following: let $R$ be a ring and let $\mathfrak{m} \subseteq R$ be an ideal. Then the correspondence theorem of ideals (for those unfamiliar with this result, see Chapter 10, Section 4 of Artin [2]) gives us that $\mathfrak{m}$ is a maximal ideal if and only if $R/\mathfrak{m}$ is a field. So with the knowledge that $\mathfrak{m}_p$ is maximal in $\mathcal{C}_p$, we can construct a new field which is of most interest to us.

**Definition 3.11.** The quotient ring $\mathbb{Q}_p = \mathcal{C}_p/\mathfrak{m}_p$ is called the **field of $p$-adic numbers**.

Recall the fact that $\mathbb{Q}$ is embedded in the ring $\mathcal{C}_p$ by taking $x \mapsto \bar{x}$; this was the statement of Lemma 3.12. In fact, the same ring homomorphism passes to an inclusion $\mathbb{Q} \hookrightarrow \mathbb{Q}_p$. The crucial fact is that, for any two distinct constant sequences $\bar{x}, \bar{y} \in \mathcal{C}_p$, we have that

$$\bar{x} - \bar{y} = \{x, x, \ldots\} - \{y, y, \ldots\} = \{x - y, x - y, \ldots\} \not\longrightarrow 0.$$

This implies that they belong to different equivalence classes modulo $\mathfrak{m}_p$. So we can take any $x \in \mathbb{Q}$ to the equivalence class $(\bar{x} + \mathfrak{m}_p) \in \mathbb{Q}_p$ in a one-to-one fashion.

Now consider the Diophantine equation $f = 0$ where $f \in \mathbb{Q}[x_1, \ldots, x_n]$ is some polynomial. Our discussion above tells us that it should be expected that a solution in $\mathbb{Q}$ to $f = 0$ gives a solution in $\mathbb{R}$ and in $\mathbb{Q}_p$ for all prime number $p$. What about the converse? This yields the following question.

**Question 3.4.** *Does the existence of solutions in $\mathbb{R}$ and $\mathbb{Q}_p$, for all prime number $p$, to the Diophantine equation $f = 0$ implies the existence of a solution in $\mathbb{Q}$?*

Hasse's *local-global principle* suggests that in this case, it may be good to study solutions in $\mathbb{R}$ and $\mathbb{Q}_p$ to deduce the existence of solutions in $\mathbb{Q}$. Of course, this is merely a principle and in fact, it does not always work as we shall see. But the general idea is the important part: look at things *locally*, and then combine them to see if they give anything *globally*. This strategy works really well for puzzles although its usage in number theory might feel a bit unusual or counterintuitive at first. In our case, finding solutions in $\mathbb{Q}$ is the global problem, whereas finding solutions in $\mathbb{R}$ and $\mathbb{Q}_p$ is the local problem. To our delight, the Hasse principle works extremely well for quadratic forms.

**Theorem 3.7** (Hasse-Minkowski theorem)**.** *Let $q(\boldsymbol{X})$ be a quadratic form. Then the equation $q(\boldsymbol{X}) = 0$ has a nontrivial solution in $\mathbb{Q}^n$ if and only if it has a nontrivial solution in $\mathbb{R}^n$ and $\mathbb{Q}_p^n$ for all prime numbers $p$.*

We will not prove this theorem as this would simply digress too much from our goal, but the motivated reader can find a proof in either Serre [22] or Gerstein [7].

Clearly, we can restrict our attention to integral quadratic forms and the Hasse-Minkowski theorem still holds. The useful thing for us (in the ternary case) is the necessary condition of the theorem. Why? We know by Lemma 3.2 that $q(\boldsymbol{X}) = 0$ having a nontrivial solution in $\mathbb{Q}^n$ implies the existence of a nontrivial solution in $\mathbb{Z}^n$; which by Lemma 3.3, further implies the existence of a primitive solution. By applying Theorem 3.2, we thus get our desired rational points. This tells us that it suffices to look at *local solutions* of $q(\boldsymbol{X}) = 0$ to deduce the existence of *rational points* on the conic $C$ whose associated quadratic form is $q$. Note that the Hasse-Minkowski theorem is much more powerful as it is true for quadratic forms in any number of indeterminates. Legendre's Theorem 3.5 is simply the special case for when the quadratic form is ternary.

Since the Hasse principle works so well for quadratics, it is a natural question to ask if it works for cubics, quartics or even higher degree polynomials. Unfortunately, the answer is negative.

**Example.** *Counterexamples to the Hasse principle.*

(i). It can be shown (for example, using Hensel's lemma) that the equation

$$(X^2 - 2)(X^2 - 17)(X^2 - 34) = 0$$

has a solution in $\mathbb{R}$ and $\mathbb{Q}_p$ for every prime number $p$, but has no roots in $\mathbb{Q}$.

(ii). Selmer proved in [21] that the equation $3X^3 + 4Y^3 + 5Z^3 = 0$ has a nonzero solution in $\mathbb{R}$ and $\mathbb{Q}_p$ for every prime number $p$, but that its only solution in $\mathbb{Q}$ is the trivial one.

(iii). Aitken and Lemmermeyer gave a whole class of counterexamples to the Hasse principle in [1] building on the work of Lind and Reichardt. For example, they proved that the equation $U^2 - 17W^2 - 19Z^2 = 0$, where $UW = V^2$, does not satisfy the Hasse principle.

# 4 Rational Points on Cubics

We now turn our attention to finding rational points on the curve $C$ associated to the cubic polynomial

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

which is defined over $\mathbb{Q}$. The curve is called a *cubic* because it is defined by a cubic polynomial. Looking back at how we did it for the conics, some questions that should be natural to ask are the following.

**Question 4.1.** *A first attack on the cubic curve problem:*

(i). *Given one rational point on the cubic, can we find other rational points?*

(ii). *If the answer to* (i) *is positive, can we use the same method as we did for conics?*

(iii). *If the answer to* (i) *is negative, then what methods can we use?*

As we shall see, the first question has a positive answer for special cubics called *singular* cubics. However, the way we find other rational points for singular cubics uses a simpler argument compared to what we did for the conics. In fact, it is not even a geometric argument and we do not know whether the same method for conics work for singular cubics. What we do know is that the method we used for conics breaks for general cubics due to a special case of Bézout's theorem (see Theorem 4.1). This gives a negative answer to our second question. The third question is something that we will explore further on throughout the chapter. In particular, it will be discussed in Section 4.3 and beyond.

## 4.1 Tools that will be helpful

This section gives the vocabulary needed to talk about plane curves in a more general setting. A particular important topic that we shall discuss is that of projective geometry which gives us a completely new way of viewing plane curves.

### 4.1.1 Curves and lines

**Definition 4.1.** Let $K$ be a field. A nonzero polynomial $f \in K[x,y]$ is said to be **irreducible** if $f = gh$ for some $g, h \in K[x,y]$ implies that either $g$ or $h$ is a constant polynomial.

**Definition 4.2.** Let $K$ be a field and let $f \in K[x,y]$. The **degree** of the plane curve $C : f(x,y) = 0$ is the degree of the polynomial $f$. Morever, we say that $C$ is **irreducible** if $f$ is an irreducible polynomial.

So degree 1 curves are lines, degree 2 curves are conics, and degree 3 curves are cubics. Before we even start to talk about cubics, we prove a special case of the well-known Bézout's theorem. For this, we first need a lemma.

**Lemma 4.1.** *Let $K$ be a field and let $f \in K[x]$ be a nonzero polynomial with* $\deg f = d$. *Then $f$ has at most $d$ roots in $K$, counting multiplicities.*

*Proof.* We will prove by induction on $d \in \mathbb{Z}^+$.

(Base case). If $d = 0$, then there is nothing to prove as $f$ is nonzero and is constant, so it has no roots in $K$.

(Inductive hypothesis). Suppose the result holds for all $f \in K[x]$ with $\deg f < d$.

(Inductive step). Let $f \in K[x]$ be a polynomial of degree $d$. If $f$ has no roots, then we are done. So suppose there exists $\alpha \in K$ such that $f(\alpha) = 0$. Then there is a factorization

$$f(x) = (x - \alpha)g(x),$$

for some $g(x) \in K[x]$. But since $K$ is a field, and so is an integral domain, we have

$$d = \deg f = \deg(x - \alpha) + \deg g = 1 + \deg g.$$

This implies that $\deg g = d - 1$ and so by the inductive hypothesis, $g$ has at most $d - 1$ roots. Obviously, these $d-1$ roots are roots of $f$ as well. Moreover, if $\beta \in K$ is another root of $f$ distinct from $\alpha$, then it must be one of the $d - 1$ roots of $g$. So together with $\alpha$ earlier, $f$ has at most $d$ roots. This completes the inductive step. ■

**Theorem 4.1** (Bézout's theorem, easy case)**.** *Let $K$ be a field and let $d$ be a positive integer. Let $C$ be a degree $d$ curve and let $L$ be a line, both defined over $K^2$. If $L \nsubseteq C$, then $\#(C \cap L) \leqslant d$.*

This theorem says that there is at most $d$ intersection points between a degree $d$ curve and a line. For example, when $d = 2$, this theorem says that a line meets a conic at 0, 1 or 2 points. For a concrete example, consider the unit circle $C : x^2 + y^2 = 1$. Then observe that the line $x = 2$ meets $C$ at 0 points, the line $x = 1$ meets $C$ at 1 point and the line $x = 0$ meets $C$ at 2 points. The proof of this theorem uses a similar argument to the one we used to prove Theorem 3.1.

*Proof.* Let $C$ be the curve associated to $f$. Since $C$ has degree $d$, so does $f$ by definition. We now have two cases to take care of.

(Non-vertical line case). Suppose $g$ defines the line $L$ such that it is not a vertical line. Then $\deg g = 1$ and we can write

$$g(x, y) = ay - mx - b, \qquad (*)$$

for some $a, b, m \in K$ with $a \neq 0_K$. Without loss of generality, we may assume that $a = 1_K$ (for otherwise, just divide by $a$ when we solve $g(x, y) = 0$). Now suppose $(x, y) \in K^2$ satisfies the system of equations

$$\begin{cases} f(x, y) = 0, \\ y = mx + b. \end{cases}$$

Then $x$ satisfies $p_m(x) = f(x, mx + b) = 0$ which is a polynomial in $x$, defined over $K$, of degree at most $d$. By Lemma 4.1, we know that $p_m$ has at most $d$ roots (counting multiplicities) for otherwise it is identically zero and so $L \subseteq C$. Consequently, there is at most $d$ such $x$ that satisfies $p_m(x) = 0$. This implies that there is at most $d$ such $(x, y)$ that satisfies the system of polynomial equations.

(Vertical line case). If $L$ is a vertical line, then $g$ is still given by $(*)$ but now with $a = 0_K$. Without loss of generality, we may also assume that $m = 1_K$ now. Suppose $(x, y) \in K^2$ satisfies the system of equations

$$\begin{cases} f(x, y) = 0, \\ x + b = 0. \end{cases}$$

Then $y$ satisfies $f(-b, y) = 0$ which is a polynomial in $y$ of degree at most $d$. By the same argument using Lemma 4.1, we see that there are most $d$ such $(x, y)$ for otherwise we have $L \subseteq C$ and get a contradiction. ■

Now consider the theorem with $K = \mathbb{Q}$. The main point of the theorem for us is that it tells us that the idea we used to find infinitely many rational points on the conic from a single rational point via projecting onto a line cannot be extended to the cubics. Why? Again, think about our prototypical example, the unit circle. We found one rational point $\boldsymbol{P} = (-1, 0)$ and fixed it. Then, we project lines through this point. Accordingly, only *one point* is identified with *one line*. In cubics, there is a possibility that *two points* are identified with *one line*; so, we lose information. For this reason, we need to be more clever when dealing with cubics.

### 4.1.2 Projective geometry

For this subsection, we fix $K$ to be a field and let $\bar{K}$ be its algebraic closure. As a reminder, we write $R^\times$ to mean the group of units in $R$ where $R$ is any ring.

Recall that we dealt with conics by talking about their associated quadratic forms. This strategy of homogenizing the quadratic polynomial defining the conic is nice, so we want to try to extend it to cubics. To do this, we need to establish some new ideas first.

**Definition 4.3.** The **affine $n$-space** over $K$ is the set

$$\mathbb{A}^n = \mathbb{A}^n(\bar{K}) = \left\{(x_1, \ldots, x_n) \mid x_i \in \bar{K}\right\}.$$

An element $\boldsymbol{P} = (x, y) \in \mathbb{A}^n$ is called a **point**, and the $x, y$ are called the **coordinates** of $\boldsymbol{P}$. The following similar set

$$\mathbb{A}^n(K) = \{(x_1, \ldots, x_n) \mid x_i \in K\}$$

is called the **set of $K$-rational points in $\mathbb{A}^n$**. We will be particularly interested in the affine 2-space also known as the **affine plane**.

Now consider the set $\mathbb{A}^{n+1} \setminus \{0\}$ and define a relation $\sim$ on it by

$$(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n) \iff x_i = \lambda y_i \text{ for some } \lambda \in \bar{K}^\times.$$

We claim that this relation is an equivalence relation.

**Lemma 4.2.** *$\sim$ on $\mathbb{A}^{n+1} \setminus \{0\}$ as defined above is an equivalence relation.*

*Proof.* Let $S = \mathbb{A}^{n+1} \setminus \{0\}$ for simplicity and assume that every $(n+1)$-tuple written below are elements of $S$. We have to prove three things.

(Reflexive). If $(x_0, \ldots, x_n) \in S$, then clearly, we have $x_i = 1x_i$ where 1 is the unital identity in $\bar{K}$. So $(x_0, \ldots, x_n) \sim (x_0, \ldots, x_n)$.

(Symmetric). Suppose $(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$. Then there exists $\lambda \in \bar{K}^\times$ such that $x_i = \lambda y_i$. Since $\lambda$ is a unit in $\bar{K}$, there exists $\mu \in \bar{K}^\times$ such that $\mu\lambda = 1$. This implies that $\mu x_i = y_i$. That is, $(y_0, \ldots, y_n) \sim (x_0, \ldots, x_n)$.

(Transitive). Suppose $(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$ and $(y_0, \ldots, y_n) \sim (z_0, \ldots, z_n)$. Then there exists $\lambda, \mu \in \bar{K}^\times$ such that $x_i = \lambda y_i$ and $y_i = \mu z_i$ for all $i$. Then simply observe that

$$x_i = \lambda(\mu z_i) = (\lambda\mu) z_i,$$

and so $(x_0, \ldots, x_n) \sim (z_0, \ldots, z_n)$ since the product $\lambda\mu$ is itself a unit in $\bar{K}$. ■

Now write $[x_0, \ldots, x_n]$ for the equivalence class that contains $(x_0, \ldots, x_n)$. Then we can define the set of equivalence classes under $\sim$. This set is the so-called *projective $n$-space.*

**Definition 4.4.** The **projective $n$-space** over $K$ is the set

$$\mathbb{P}^n = \mathbb{P}^n(\bar{K}) = \left\{[x_0, \ldots, x_n] : x_i \in \bar{K} \text{ not all zero}\right\}.$$

The $x_i$ are called the **homogeneous coordinates** of the point $[x_0, \ldots, x_n]$. The following similar set

$$\mathbb{P}^n(K) = \{[x_0, \ldots, x_n] : x_i \in K \text{ not all zero}\}$$

is called the **set of $K$-rational points in $\mathbb{P}^n$**.

We will be primarily interested in the projective 1- and 2-space called the *projective line* and *projective plane* respectively.

To visualize the projective plane geometrically, it is easier to set $K = \mathbb{R}$ and focus on the $\mathbb{R}$-rational points in $\mathbb{P}^2$. In this setting, $\mathbb{P}^2$ consists of all the lines through the origin in $\mathbb{R}^3$. Now, observe that if $[x, y, z] \in \mathbb{P}^2$ and $z \neq 0$, we have an equivalence $(x, y, z) \sim (\frac{x}{z}, \frac{y}{z}, 1)$. This implies that there is a decomposition of $\mathbb{P}^2$ into two parts:

$$\mathbb{P}^2 = \{[x, y, 1] : x, y \in \mathbb{R}\} \cup \{[a, b, 0] : a, b \in \mathbb{R}\}. \tag{4.1}$$

This decomposition gives us a nice picture of $\mathbb{P}^2$. First, observe that there is a bijection $\mathbb{A}^2(\mathbb{R}) \to \{[x, y, 1] : x, y \in \mathbb{R}\}$ simply by taking $(x, y)$ to $[x, y, 1]$. This implies that $\mathbb{P}^2$ contains a copy of the affine plane $\mathbb{A}^2$ together with some extra points, so we have at most an inclusion $\mathbb{A}^2 \hookrightarrow \mathbb{P}^2$. These extra points are what we call *points at infinity* and they form a projective line $\mathbb{P}^1$. But note that the decomposition (4.1) is not unique. We could have equally decompose $\mathbb{P}^2$ in a way such that the first set in the union is not $\{[x, y, 1]\}$ but instead $\{[x, 1, z]\}$ or $\{[1, y, z]\}$, or even replace the 1 with, say, $e^\pi$, all of which are still in bijection with $\mathbb{A}^2$. Such a subset of $\mathbb{P}^2$ is called an *affine chart.*

Recall how we defined plane curves in $K^2$ in Chapter 2. They are independent of how the plane sits in space, so we really defined them on $K^2$ viewed as an affine space. Due to this, we shall call them *affine curves* from now on. The reasoning behind this emphasis is because we want to extend the notion of affine curves to the projective plane.

**Definition 4.5.** Let $F \in K[X, Y, Z]$ be a homogeneous polynomial. We call the set of solutions

$$C_F = \left\{[x, y, z] \in \mathbb{P}^2(\bar{K}) : F(x, y, z) = 0\right\},$$

the **projective plane curve** (associated to $F$). We say that $C_F$ is defined over $K$ since $F$ is defined over $K$. An element of $C_F$ is called a **projective point**. When the context is clear, we will simply call $C_F$ a projective curve or simply a curve, and write $C$. Similarly, we may instead just say a point instead of a projective point.

Just like plane curves in the affine plane, we will occasionally write $C : F(X, Y, Z) = 0$ to mean the projective curve $C$ associated to $F$. And just like affine curves, we have a notion of degree and irreducibility.

**Definition 4.6.** Let $C$ be a projective plane curve associated to $F$. The **degree** of $C$ is the degree of $F$; and we say that $C$ is **irreducible** if $F$ is an irreducible polynomial.

As before with affine curves, a degree 1 projective curve is called a projective line, a degree 2 projective curve is called a projective conic, and a degree 3 projective curve is called a projective cubic.

When dealing with conics, we homogenize a quadratic polynomial $f(x, y)$ to get its associated quadratic form $q$ via the transformation $q = Z^2 f(X/Z, Y/Z)$, where we then get a correspondence between rational points on $C_f$ and nontrivial integer solutions to $q = 0$. But observe that there is nothing special about the word "quadratic" here. We could have equally homogenize any polynomial $f \in K[x, y]$ of degree $d$ by the same affine transformation $x = X/Z$ and $y = Y/Z$ to get its *associated form $F$* via

$$F(X, Y, Z) = Z^d f\left(\frac{X}{Z}, \frac{Y}{Z}\right). \tag{4.2}$$

Accordingly, if $K = \mathbb{Q}$, we should expect the same correspondence between rational points on $f$ and nontrivial integer solutions to $F = 0$. We call (4.2) the *homogenization* of $f$. But since $F$ is homogeneous, $F = 0$ defines what we now call a projective curve, and so contains points that are in correspondence with points on $C_f$ together with some extra points at infinity. From a topological point of view (although we never defined a topology on curves), this homogenization process can be thought of taking the *closure* of the curve $C_f$.

**Definition 4.7.** Let $C$ be the affine curve associated to $f \in K[x, y]$ of degree $d \geqslant 1$. The projective curve associated to the homogeneous polynomial

$$F(X, Y, Z) = Z^d f\left(\frac{X}{Z}, \frac{Y}{Z}\right),$$

is called the **projective closure** of $C$, and is denoted $\widehat{C}$.

**Example.** Let $f(x, y)$ be a quadratic polynomial with associated quadratic form $q$. The curve $q = 0$ is the projective closure of $C_f$.

We have a converse of this notion as well. That is, if we have a projective curve, then there is a natural related affine curve to it.

**Definition 4.8.** Let $\widehat{C}$ be the projective curve associated to $F \in K[X, Y, Z]$. The affine curve $C$ associated to the polynomial $f(x, y) = F(x, y, 1)$ is called the **projection** of $\widehat{C}$ onto $\mathbb{A}^2$.

Let $f \in K[x_1, \dots, x_n]$ be a polynomial. We define the *formal derivative* (or *partial derivative*) of $F$ with respect to $x_i$ in the usual analysis sense, and write this as $\partial f/\partial x_i$. We then define the *tangent vector* of $f$ to be the $n$-tuple

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right)$$

of polynomials. Note that if $K = \mathbb{R}$ (or any of its subfields), then this coincide with what we are used to in analysis. If $K = \mathbb{C}$, we let the reader decide to use real analysis on $\mathbb{R}^2$ or the rich theory of complex analysis (they are inherently the same anyways). But for general $K$, just consider this as symbols as we never defined what does it mean to "differentiate" in arbitrary fields.

**Definition 4.9.** Let $C$ be a projective plane curve associated to $F(X, Y, Z)$ and let $\boldsymbol{P} \in C$. We say that $\boldsymbol{P}$ is a **singular point** of $C$ (or $C$ is singular at $\boldsymbol{P}$) if

$$\nabla F(\boldsymbol{P}) = 0.$$

Otherwise, we say that $\boldsymbol{P}$ is a **nonsingular** point of $C$. If $C$ contains a singular point, we say that $C$ is singular. If otherwise $C$ does not contain any singular points, we say that $C$ is **nonsingular** (or **smooth**). Finally, we say that an affine curve $C$ is nonsingular if its projective closure $\widehat{C}$ is nonsingular, and vice-versa.

**Example.** Nonsingular degree 2 affine curves are just nondegenerate conics.

## 4.2 Case 1: Singular cubics

We are now finally in a position to tackle the cubic curve

$$C : \gamma(x,y) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0, \tag{4.3}$$

where $a, b, c, d, e, f, g, h, i, j \in \mathbb{Q}$. In this section, we shall first deal with the case where the cubic has a singular point (i.e. the projective closure of (4.3) has a singular point). It turns out that in this case, all the rational points are well-understood.

First, we show that there is a reduction of (4.3) into a simpler form when $C$ is singular. Suppose $\boldsymbol{S} = (x_0, y_0)$ is a singular point of $C$ such that $\boldsymbol{S} \in C(\mathbb{Q})$. Without loss of generality, we may assume that $x_0 = y_0 = 0$ since we can always apply the affine transformation $x \mapsto x + x_0$ and $y \mapsto y + y_0$. Geometrically, this is just a shifting of the affine plane. We now claim that the constant term $j$, and linear terms $hx$ and $iy$ vanish:

(i). Since $\boldsymbol{S} = (0,0) \in C(\mathbb{Q})$, it follows that $j = \gamma(\boldsymbol{S}) = 0$.

(ii). Since we assumed that $\boldsymbol{S}$ is a singular point of $C$, this implies that $\nabla\gamma(\boldsymbol{S})$ vanishes or equivalently,

$$h = \frac{\partial\gamma}{\partial x}(\boldsymbol{S}) = 0, \quad i = \frac{\partial\gamma}{\partial y}(\boldsymbol{S}) = 0.$$

So we conclude that the singular point assumption allows us to rewrite (4.3) as

$$C : \gamma(x,y) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 = 0.$$

From here, we can concretely describe rational points on singular cubic curves.

**Theorem 4.2.** *Let $C : \gamma(x,y) = 0$ be the cubic curve defined over $\mathbb{Q}$ by*

$$\gamma(x,y) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 \tag{4.4}$$

*with singularity at $(0,0)$. Define the maps $\nu, \delta : \mathbb{Q} \to \mathbb{Q}$ by*

$$\nu(\lambda) = e\lambda^2 + f\lambda + g, \quad \delta(\lambda) = a\lambda^3 + b\lambda^2 + c\lambda + d.$$

*Then the rational points on $C$ is given by the union of three sets*

$$C(\mathbb{Q}) = \left\{(0,0), \left(-\frac{e}{a}, 0\right)\right\} \cup \left\{\left(\frac{s}{t}, \frac{1}{t}\right) : t \in \mathbb{Q}^\times, \nu(s) = 0\right\} \cup \left\{\left(-s\frac{\nu(s)}{\delta(s)}, -\frac{\nu(s)}{\delta(s)}\right) : \delta(s) \neq 0\right\},$$

*for all possible $s \in \mathbb{Q}$.*

We will give a proof based on one given by Lozano-Robledo [12].

*Proof.* Let $(x,y) \in C(\mathbb{Q})$. Let us look at this with a case by case approach.

**The case $y = 0$.** If $y = 0$, then $\gamma(x,0) = ax^3 + ex^2 = x^2(ax + e) = 0$. This implies that $x = 0$ or $x = -e/a$. This gives the first set $\{(0,0), (-e/a, 0)\}$ in the union.

**The case $y \neq 0$.** If $y \neq 0$, then we can divide (4.4) by $y^3$ to get

$$\frac{\gamma(x,y)}{y^3} = a\left(\frac{x}{y}\right)^3 + b\left(\frac{x}{y}\right)^2 + c\left(\frac{x}{y}\right) + d + e\left(\frac{x}{y}\right)^2\left(\frac{1}{y}\right) + f\left(\frac{x}{y}\right)\left(\frac{1}{y}\right) + g\left(\frac{1}{y}\right).$$

Note that this equation still define the same affine curve $C$. Next, we apply a change of variables $s = x/y$ and $t = 1/y$ to get a new affine curve

$$C' : \gamma'(s,t) = (as^3 + bs^2 + cs + d) + (es^2t + fst + gt) = \delta(s) + t\,\nu(s) = 0,$$

where $\nu, \delta$ are maps as defined in the theorem's statement. We also get a map between affine curves $\phi : C \setminus \{(0,0), (-e/a, 0)\} \hookrightarrow C'$ defined by

$$(x, y) \longmapsto \left(\frac{x}{y}, \frac{1}{y}\right).$$

This map $\phi$ has an inverse given by

$$\phi^{-1}(s,t) = \left(\frac{s}{t}, \frac{1}{t}\right) = \phi(s,t),$$

and so is a bijection. Consequently, it is enough to look at rational points on $C'$ because the map $\phi^{-1}$ passes the rational points on $C'$ to rational points on $C$. There are two cases to look at when considering rational points on $C'$.

**The case $\nu(s) = 0$.** If $(s,t) \in \mathbb{Q}^2$ is such that $\nu(s) = 0$, then it follows immediately that $\delta(s) = 0$ and so $(s,t) \in C'(\mathbb{Q})$. Via the inverse map $\phi^{-1}$, we thus get $(s/t, 1/t) \in C(\mathbb{Q})$ for all $t \in \mathbb{Q}^\times$ in this particular case. This gives the second set in the union.

**The case $\nu(s) \neq 0$.** If on the other hand $(s,t) \in \mathbb{Q}^2$ is such that $\nu(s) \neq 0$, then we can solve for $t$ easily to get a rational point on $C'$. For any $s \in \mathbb{Q}$, the point $(s,t)$ with

$$t = -\frac{\nu(s)}{\delta(s)},$$

defines a rational point on $C'$. Via the inverse map $\phi^{-1}$, it follows that

$$\left(-s\frac{\nu(s)}{\delta(s)}, -\frac{\nu(s)}{\delta(s)}\right) \in C(\mathbb{Q}),$$

and this gives our third and final set in the union. ■

This theorem summarizes all we need to know about rational points on singular cubic curves. So, let us move on to the nonsingular case.

## 4.3 Case 2: Nonsingular cubics

It turns out that characterizing all the rational points on a nonsingular cubic is far from a trivial problem and requires some very clever insights. So, this section has a completely different aim compared to the preceding Section 4.2 on singular cubics. The goal here is to prove that every nonsingular cubic curve can be put into a so-called *Weierstrass form*.

**Definition 4.10.** A **Weierstrass equation** (or form) over a field $K$ is an equation of the form

$$y^2 + axy + by = x^3 + cx^2 + dx + e, \tag{4.5}$$

where $a, b, c, d, e \in K$ are constants. The special case with $a = b = c = 0$ reduces the equation to

$$y^2 = x^3 + dx + e, \tag{4.6}$$

which is called a **Weierstrass normal form**.

Over the rational numbers $\mathbb{Q}$, we have the following lemma.

**Lemma 4.3.** *There exists an invertible change of variables so that any Weierstrass equation over $\mathbb{Q}$ can be transformed into a Weierstrass normal form.*

*Proof.* Given the Weierstrass equation (4.5), we can complete the square (with respect to $y$) on the left-hand side to get

$$\begin{aligned}y^2 + axy + by &= y^2 + y(ax+b) + \left(\frac{ax+b}{2}\right)^2 - \left(\frac{ax+b}{2}\right)^2 \\ &= \left(y + \frac{ax+b}{2}\right)^2 - \left(\frac{ax+b}{2}\right)^2.\end{aligned}$$

So we can rewrite (4.5) as

$$\left(y + \frac{ax+b}{2}\right)^2 = x^3 + cx^2 + dx + e + \left(\frac{ax+b}{2}\right)^2. \tag{4.7}$$

Now consider the change of variables

$$u = x, \quad v = y + \frac{ax+b}{2}.$$

Then equation (4.7) under this transformation becomes

$$v^2 = u^3 + Au^2 + Bu + C, \tag{4.8}$$

where

$$A = c + \frac{a^2}{4}, \quad B = d + \frac{ab}{2}, \quad C = e + \frac{b^2}{4}.$$

Next, we seek a change of variables so that the square term vanishes. This can be achieved by considering the transformation

$$s = u + \frac{A}{3}, \quad t = v.$$

This transforms equation (4.8) into

$$t^2 = s^3 + Ds + E,$$

where

$$D = \frac{A^2}{3} + B, \quad E = A\left(\frac{2A^2}{27} - \frac{B}{3}\right) + C.$$

We are done now as this equation is in Weierstrass normal form. ■

In the future, we shall also be concerned with Weierstrass equations which have integer coefficients so we discuss this now. Consider the Weierstrass equation

$$y^2 + axy + by = x^3 + cx^2 + dx + e$$

where $a, b, c, d, e \in \mathbb{Q}$. To make the coefficients integers, we need to remove the denominators occuring in the coefficients. The obvious way to do it is thus to multiply both sides by the lowest common multiple of the denominators of the coefficients. But this is not enough as the resulting equation is not in Weierstrass form. The transformation that we need is the following: $u = \delta^2 x$, $v = \delta^3 y$, where $\delta \in \mathbb{Z}$ is the lowest common multiple that we required earlier. This puts the equation into

$$v^2 + (a\delta)uv + (b\delta^3)v = u^3 + (c\delta^2)u^2 + (d\delta^4)u + e\delta^6,$$

which has integer coefficients and is in Weierstrass form.

To give the main theorem of this section, we need the notion of *birational equivalence* (which we actually have been using implicitly up until now). We shall give a simple definition of birational equivalence due to Mordell [13]. The more complete definition of this can be found in almost all algebraic geometry textbooks; one that we recommend is by Shafarevich which first gives an undergraduate-friendly definition (see Section 1.4 of [23]) and then gives the standard definition (see Section 3.3 of [23]). For us, Mordell's is good enough.

**Definition 4.11.** Let $C : \gamma(x, y) = 0$ and $C' : \Gamma(u, v) = 0$ be two affine curves defined over $\mathbb{Q}$. We say that $C$ and $C'$ are **birationally equivalent** if there is a relation

$$x = f(u, v) \quad y = g(u, v); \quad u = h(x, y), \quad v = \iota(x, y) \tag{4.9}$$

except for a finite set of values, where $f, g \in \mathbb{Q}(x, y)$ and $h, \iota \in \mathbb{Q}(u, v)$ are rational functions defined over $\mathbb{Q}$.

**Remark 7.** It is worth emphasizing (since our definition does not make it too obvious) that the relation (4.9) are inverses to each other. That is, the map that sends $(x, y) \mapsto (f(u, v), g(u, v))$ is the inverse to the map that sends $(u, v) \mapsto (h(x, y), \iota(x, y))$, and vice-versa. This justifies the terminology *birational* as these are maps of rational functions (the right terminology is that these are called *rational maps*, see [23] for more details).

We have seen that the unit circle is birationally equivalent to a line; and every bijection we have between affine curves so far are also instances of birational equivalence. For example, the map that allows us to talk *only* about rational points on $C'$ in the proof of Theorem 4.2. In this proof, we have seen just how powerful this idea is. Essentially, knowing the rational points on one curve gives knowledge of the rational points on the other, and vice-versa. In fact, as we have mentioned before, birational equivalence is actually one notion of isomorphism in algebraic geometry.

We now claim that every cubic curve with at least one nonsingular rational point is birationally equivalent to a cubic defined by a Weierstrass normal form. This is our main result of this section.

**Theorem 4.3.** *Any cubic curve $C$ defined over $\mathbb{Q}$, with at least one nonsingular rational point $\boldsymbol{P}$ is birationally equivalent to a cubic curve defined by a Weierstrass normal form.*

Note that we do not impose the condition that $C$ is nonsingular everywhere. Our proof only demands that $C$ has one known nonsingular rational point $\boldsymbol{P}$ and allows $C$ to be singular at any other points different from $\boldsymbol{P}$. We give a proof following closely one given by Lozano-Robledo [12], and filling in some gaps based on idea from Silverman and Tate [26].

*Proof.* Let $\boldsymbol{P}$ be a nonsingular rational point on the cubic curve

$$C : \gamma(x, y) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0,$$

which is defined over $\mathbb{Q}$. The projective closure $\widehat{C}$ of $C$ is defined by the equation

$$\Gamma(X, Y, Z) = aX^3 + bX^2Y + cXY^2 + dY^3 + eX^2Z + fXYZ + gY^2Z + hXZ^2 + iYZ^2 + jZ^3 = 0.$$

Suppose $\boldsymbol{P}$ in projective coordinates is given by $[x_0, y_0, z_0] \in \mathbb{P}^2$; and consider the tangent line to $\widehat{C}$ at $\boldsymbol{P}$ given by

$$L : w(X, Y, Z) = X\frac{\partial\Gamma}{\partial X}(\boldsymbol{P}) + Y\frac{\partial\Gamma}{\partial Y}(\boldsymbol{P}) + Z\frac{\partial\Gamma}{\partial Z}(\boldsymbol{P}) = a_{31}X + a_{32}Y + a_{33}Z = 0,$$

where we have put

$$a_{31} = \frac{\partial\Gamma}{\partial X}(\boldsymbol{P}), \quad a_{32} = \frac{\partial\Gamma}{\partial Y}(\boldsymbol{P}), \quad a_{33} = \frac{\partial\Gamma}{\partial Z}(\boldsymbol{P}).$$

By definition (of $L$ being tangent to $\widehat{C}$ at $\boldsymbol{P}$), we know that $L$ intersects $C$ at $\boldsymbol{P}$ with multiplicity of at least two. There is, however, a possibility that this intersection occurs with multiplicity three but proving this is easier and similar to proving the case that the intersection occurs with multiplicity exactly two. So we focus only on the case of multiplicity exactly two.

In this case, we have $L \cap \widehat{C} = \{\boldsymbol{P}, \boldsymbol{Q}\}$ where $\boldsymbol{Q}$ is a point different from $\boldsymbol{P}$ such that $L$ intersects $C$ at $\boldsymbol{Q}$ with multiplicity one. It should be easy to see that $\boldsymbol{Q}$ has rational coordinates and so belongs to $\widehat{C}(\mathbb{Q})$. We now consider two more projective lines:

(i). Let $M : u(X, Y, Z) = a_{11}X + a_{12}Y + a_{13}Z = 0$ be a line which goes through $\boldsymbol{Q}$ but is different from $L$.

(ii). Let $N : v(X, Y, Z) = a_{21}X + a_{22}Y + a_{23}Z = 0$ be any other line which goes through $\boldsymbol{P}$, where the coefficients $a_{21}, a_{22}, a_{23}$ are picked such that the following matrix

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

has nonzero determinant, and so is invertible.

Using these lines, we can apply a projective transformation so that $L$ acts as the new $Z$-axis, $M$ acts as the new $X$-axis and $N$ acts as the new $Y$-axis. Such a projective transformation is defined by the matrix $\mathcal{A}$, or more explicitly by the change of variables

$$U = u(X, Y, Z), \quad V = v(X, Y, Z), \quad W = w(X, Y, Z).$$

Since $N$ is chosen so that $\mathcal{A}$ is invertible, there is a clear passage to go from the $(X, Y, Z)$ coordinates to the $(U, V, W)$ coordinates and vice-versa. Most importantly, note that in this new coordinate system, we necessarily have that

$$\boldsymbol{P} = [1, 0, 0], \quad \boldsymbol{Q} = [0, 1, 0]. \tag{4.10}$$

This is true because, $\boldsymbol{P} \in N$ and $\boldsymbol{P} \in L$ which implies that $V = v(\boldsymbol{P}) = 0$ and $W = w(\boldsymbol{P}) = 0$. Similarly, $\boldsymbol{Q} \in M$ and $\boldsymbol{Q} \in L$ which implies that $U = u(\boldsymbol{Q}) = 0$ and $W = w(\boldsymbol{Q}) = 0$.

Now, suppose that $\widehat{C}$ is given by the equation $\Delta(U, V, W) = 0$ in this new coordinate system with $\Delta$ being the polynomial

$$\Delta(U, V, W) = aU^3 + bU^2V + cUV^2 + dV^3 + eU^2W + fUVW + gV^2W + hUW^2 + iVW^2 + jW^3,$$

where the constants $a, b, c, d, e, f, g, h, i, j \in \mathbb{Q}$ and are *not necessarily the same* as the one in $\Gamma$.

**Lemma.** We claim that the terms $aU^3$, $bU^2V$ and $dV^3$ vanish. For this, we require (4.10).

(i). By hypothesis, $\boldsymbol{P} \in \widehat{C}$. So, by definition $a = \Delta([1, 0, 0]) = \Delta(\boldsymbol{P}) = 0$.

(ii). By our choice, $\boldsymbol{Q} \in \widehat{C}$. So, by definition $d = \Delta([0, 1, 0]) = \Delta(\boldsymbol{Q}) = 0$.

(iii). Consider the intersection $\widehat{C} \cap L$ in $(U, V, W)$ coordinates. In this intersection $W$ vanishes and so together with our discovery in (i) and (ii) that $a = d = 0$, we have the reduced form

$$\Delta(U, V, W) = bU^2V + cUV^2 = UV(bU + cV).$$

Now recall that $\widehat{C} \cap L = \{\boldsymbol{P}, \boldsymbol{Q}\}$ where $\boldsymbol{P}$ occurs with multiplicity two and $\boldsymbol{Q}$ occurs with multiplicity one. If we view the right-hand side of the reduced $\Delta$ as a product of three linear factors, we then see that $\boldsymbol{Q}$ solves the equation $U = 0$ and $\boldsymbol{P}$ solves the equation $V = 0$; but most importantly, $\boldsymbol{P}$ solves the equation $bU + cV = 0$ which implies that $b = 0$, as desired.

So we have shown that $\Delta$ can be further reduced to

$$\Delta(U,V,W) = cUV^2 + eU^2W + fUVW + gV^2W + hUW^2 + iVW^2 + jW^3.$$

Next, we consider the projection of $\widehat{C}$ that is currently defined by $\Delta(U,V,W) = 0$ onto the affine plane $\mathbb{A}^2$. This can be done by taking $\delta(s,t) = \Delta(s,t,1)$ and so we get the affine curve

$$C : \delta(s,t) = cst^2 + es^2 + fst + gt^2 + hs + it + j = 0.$$

We can group the $t^2$ terms and write $\delta(s,t) = (cs+g)t^2 + es^2 + fst + hs + it + j$. We then apply our first change of variables via the affine transformation $(s,t) \mapsto (cs+g,t)$ so that our curve is defined by

$$C' : \delta'(s,t) = st^2 + es^2 + fst + hs + it + j = 0.$$

Next, we multiply $\delta'(s,t)$ by $s$ to get $s\delta'(s,t) = (st)^2 + es^3 + fs(st) + hs^2 + i(st) + js$. Then, we consider the invertible (for $t \neq 0$) change of variables

$$\tilde{x} = s, \quad \tilde{y} = st,$$

to get a new affine curve

$$C'' : \delta''(\tilde{x},\tilde{y}) = \tilde{y}^2 + f\tilde{x}\tilde{y} + i\tilde{y} = k\tilde{x}^3 + \ell\tilde{x}^2 + m\tilde{x}.$$

Finally, if we consider the change of variables

$$x = k\tilde{x}, \quad y = k^2\tilde{y},$$

we get a new affine curve

$$C''' : y^2 + \mathfrak{a}xy + \mathfrak{b}y = x^3 + \mathfrak{c}x^2 + \mathfrak{d}x + \mathfrak{e},$$

for some $\mathfrak{a},\mathfrak{b},\mathfrak{c},\mathfrak{d},\mathfrak{e} \in \mathbb{Q}$. The equation defining $C'''$ is a Weierstrass equation, and so by Lemma 4.3, this equation can be put into Weierstrass normal form. All the change of variables we have used are invertible except for a finite set of points, and so their composition defines a bijection $C \to C'''$. That is, $C$ and $C'''$ are birationally equivalent, as desired. ■

So to understand rational points on nonsingular cubic curves, the birational equivalence we have just established above says that it suffices to study rational points on cubic curves given by a Weierstrass normal form. Such cubic curves are captured by a class of curves called *elliptic curves*.

## 4.4 Elliptic curves

**Definition 4.12.** An **elliptic curve** over $\mathbb{Q}$ is a nonsingular projective cubic curve $E$ defined over $\mathbb{Q}$ with at least one rational point $\mathcal{O}$, called the **origin**. We will denote such an elliptic curve as $E/\mathbb{Q}$, and denote its set of rational points as $E(\mathbb{Q})$.

**Remark 8.** This definition extends naturally to general fields — just replace $\mathbb{Q}$ with a field $K$ everywhere. In particular, we shall write $E/K$ for an elliptic curve over $K$.

Despite an elliptic curve $E$ being defined over projective planes, we would usually only consider affine charts of $E$. That is, we talk about the projection of $E$ onto the affine plane, but always being conscious that we would miss some points by doing it this way. The points that we missed are of course, the points at infinity (cf. discussion (4.1) about $\mathbb{P}^2$ decomposition).

**Definition 4.13.** Let $E$ be an elliptic curve with origin $\mathcal{O}$, and let $\tilde{E}$ be an elliptic curve with origin $\tilde{\mathcal{O}}$. We say that $E$ and $\tilde{E}$ are **isomorphic over** $\mathbb{Q}$ if there exists an invertible change of variables $\phi : E \to \tilde{E}$ (called an **isomorphism**) defined by

$$[x, y, z] \mapsto [f(x, y, z), g(x, y, z), h(x, y, z)] ,$$

where $f, g, h$ are rational functions with coefficients in $\mathbb{Q}$, such that $\phi(\mathcal{O}) = \tilde{\mathcal{O}}$.

**Proposition 4.1** ([11], Chapter 2.2, Proposition 2.2.2)**.** *Let $K$ be a field of characteristic not 2 or 3, and let $E/K$ be an elliptic curve with origin $\mathcal{O}$. Then there exists an elliptic curve $\tilde{E}$, whose origin is $[0, 1, 0]$, defined by the equation*

$$ZY^2 = X^3 + aXZ^2 + bZ^3, \tag{4.11}$$

*where $a, b \in K$ is such that $4a^3 + 27b^2 \neq 0$. Moreover, $E$ and $\tilde{E}$ are isomorphic over $\mathbb{Q}$, and so $\mathcal{O}$ is mapped to $[0, 1, 0]$ under this isomorphism.*

The proof of this proposition is beyond the scope of this paper as it uses the Riemann-Roch theorem. For the adventurous readers, we invite you to read Chapter III, Proposition 3.1 of Silverman [24] for a proof.

Before we talk about what Proposition 4.1 means for us, let us look at the subtle hypothesis and consequences. Firstly, notice that the origin $[0, 1, 0]$ is the *unique* point at infinity on $\tilde{E}$. To see this, observe what happens when we set $Z = 0$. The equation reduces to $X^3 = 0$ which implies that $X = 0$. Since any element $[x, y, z] \in \mathbb{P}^2$ cannot have $x, y, z$ all simultaneously zero, it follows that $y \neq 0$ in our case. So after scaling, $[0, 1, 0]$ gives the only point at infinity on $\tilde{E}$. Next, notice that the projection of $\tilde{E}$ onto $\mathbb{A}^2$ is given by

$$y^2 = x^3 + ax + b, \tag{4.12}$$

which is in Weierstrass normal form. From here, the hypothesis that $4a^3 + 27b^2 \neq 0$ in (i) of the proposition should make more sense. This condition ensures that (4.12) has distinct roots and so is nonsingular. This condition of nonsingularity will be important for us, especially when we talk about elliptic curves over finite fields in Section 4.5.

**Definition 4.14.** Let $E/\mathbb{Q}$ be an elliptic curve with Weierstrass equation $y^2 = x^3 + ax + b$ where $a, b \in \mathbb{Q}$. The **discriminant** of $E$ is defined to be $\Delta_E = -16(4a^3 + 27b^2)$.

Proposition 4.1 tells us that it suffices to talk about the projective curve (4.11) if we want to talk about elliptic curves over any field of characteristic different from 2 or 3. This is good as we are mostly interested when $K = \mathbb{Q}$ which has characteristic 0.

### 4.4.1 A group structure on $E(\mathbb{Q})$

Perhaps, one of the most amazing feat about the study of elliptic curves is the idea that we can give the set $E(\mathbb{Q})$ a group structure by purely using a geometric argument. As we shall see, this has powerful consequences. The question now is how do we do this?

First, let us make sure that we are on the same page, and thinking about the same things. We will focus only on the elliptic curve $E : y^2 = x^3 + ax + b$ with $a, b \in \mathbb{Q}$. Next, recall that we define the notation $C(\mathbb{Q}) = C \cap \mathbb{Q}$ only for affine curves $C$. We could think of $E(\mathbb{Q})$ as the set of rational points on the affine curve $y^2 = x^3 + ax + b$. But as discussed before, we still need to keep track of the points at infinity. We solve this subtle issue by defining

$$E(\mathbb{Q}) := (E \cap \mathbb{Q}) \cup \{\mathcal{O}\} ,$$

with $\mathcal{O} = [0, 1, 0]$ being the (unique) point at infinity of $E$.

To put a group structure on $E(\mathbb{Q})$, we first need to define a binary operation on it. Let $\boldsymbol{P} = (x_1, y_1)$ and $\boldsymbol{Q} = (x_2, y_2)$ be rational points on $E$, and let $L$ be the line through the points $\boldsymbol{P}$ and $\boldsymbol{Q}$. We first claim the following.

**Lemma 4.4** (Partial closure)**.** *The line $L$ that goes through the points $\boldsymbol{P}, \boldsymbol{Q} \in E(\mathbb{Q})$ meets a third point in $E(\mathbb{Q})$, which we will denote as $\boldsymbol{P} * \boldsymbol{Q}$.*

This is called *partial closure* because as we shall see, this is *not* how we would define a binary operation on $E(\mathbb{Q})$, and so is not the full closure that we want for a binary operation. However, the partial closure implies the closure of the desired operation. The proof of this lemma should feel routine by now.

*Proof.* Let $\boldsymbol{P} = (x_1, y_1)$ and $\boldsymbol{Q} = (x_2, y_2)$, and let $L$ be the line through $\boldsymbol{P}$ and $\boldsymbol{Q}$ . We have two cases to take care of.

**Case 1: $x_1 \neq x_2$.** It should be immediate to see that $L$ has rational slope. To see that $L$ meets at a third point $\boldsymbol{P} * \boldsymbol{Q} \in E(\mathbb{Q})$, we solve the system given by the intersection $L \cap E$. Suppose $L$ is defined by the equation $y = mx + c$, where

$$m = \frac{y_2 - y_1}{x_2 - x_1}, \quad c = y_1 - mx_1 = y_2 - mx_2,$$

are rational numbers. Then, solving the system

$$\begin{cases} y = mx + c, \\ y^2 = x^3 + ax + b, \end{cases}$$

we get the equation $p(x) = x^3 - m^2x^2 + Ax + B$, for some $A, B \in \mathbb{Q}$. Since the sum of roots of $p(x)$ add up to $m^2 \in \mathbb{Q}$ and $x_1, x_2$ are roots, it follows that the third root $x_3 = m^2 - x_1 - x_2$ is rational. Consequently, $y_3 = mx_3 + c$ is rational and thus

$$\boldsymbol{P} * \boldsymbol{Q} = (x_3, y_3) = (x_3,\, mx_3 + c) \tag{4.13}$$

gives a third rational point on $E$.

**Case 2: $x_1 = x_2$.** In this setting, we are forced to either have $y_1 = -y_2$ or $y_1 = y_2$. If $y_1 = -y_2$, then $L$ is a vertical line so we take $\boldsymbol{P} * \boldsymbol{Q} = \mathcal{O}$ (this also hints towards a possible candidate for the *inverse* of the group law). On the other hand, if $y_1 = y_2$, then we have $\boldsymbol{P} = \boldsymbol{Q}$. The approach earlier does not fail completely in this setting. The only problem is that the slope $m$ is no longer defined if we compute it in the same way as before. Otherwise, once we find a suitable line, the remaining steps should be identical. A reasonable line to use in this setting is the tangent line to $E$ at $\boldsymbol{P}$, so let $L$ be this line. The slope of $L$ can be computed by implicit differentiation of the equation $y^2 = x^3 + ax + b$ and substituting $\boldsymbol{P}$. This gives that

$$m = \frac{3x_1^2 + a}{2y_1}$$

is the slope of $L$, and use this $m$ in (4.13) to get the desired third rational point. ■

**Remark 9.** Notice that equation (4.13) effectively gives us an explicit algorithm to compute the third point $\boldsymbol{P} * \boldsymbol{Q}$ for both cases that $\boldsymbol{P} = \boldsymbol{Q}$ and $\boldsymbol{P} \neq \boldsymbol{Q}$.

Knowing this fact, we now attempt to define *addition*, the desired binary operation, of two points $\boldsymbol{P}, \boldsymbol{Q} \in E(\mathbb{Q})$. Unsurprisingly, we will denote addition by +. Firstly, if $\boldsymbol{P} = \boldsymbol{Q} = \mathcal{O}$, then we shall define $\mathcal{O}+\mathcal{O}=\mathcal{O}$. This can be taken as a pure convention, but was actually justified when we discussed the subtle assumptions in Proposition 4.1. Now, suppose $\boldsymbol{P},\boldsymbol{Q}$ are not both $\mathcal{O}$. It is tempting to define $\boldsymbol{P}+\boldsymbol{Q}=\boldsymbol{P}*\boldsymbol{Q}$, but then what is the identity element of the resulting group $(E(\mathbb{Q}),+)$? Playing around with points using equation (4.13), we see that there is none so it is not a group! Instead, after getting $\boldsymbol{P}*\boldsymbol{Q}$, we draw a vertical line $L'$ through $\boldsymbol{P}*\boldsymbol{Q}$ and $\mathcal{O}$, and define the third intersection of $L'\cap E$ to be $\boldsymbol{P}+\boldsymbol{Q}$. In other words, we shall define

$$\boldsymbol{P}+\boldsymbol{Q}=(\boldsymbol{P}*\boldsymbol{Q})*\mathcal{O}.$$

There is an easy geometrical interpretation of this. Observe that the projective closure of the affine (vertical) line $\ell : x=c$ is given by the projective line $\widehat{\ell} : X=cZ$. Clearly, this line passes through $[0,1,0]$. Since we have seen that our choice of elliptic curve $E$ has the unique point at infinity $\mathcal{O}=[0,1,0]$, this guarantees that any line that passes through $\mathcal{O}$ is a vertical line. Accordingly, for any point $\boldsymbol{R}\in E(\mathbb{Q})$, the point $\mathcal{O}*\boldsymbol{R}$ is the reflection of $\boldsymbol{R}$ in the $x$-axis. Note however that this is not true if the equation defining the elliptic curve is not given by a Weierstrass normal form, so we have to be careful about generalizing this.

**Remark 10** (Addition algorithm)**.** Since we now know that $\boldsymbol{P}+\boldsymbol{Q}=(\boldsymbol{P}*\boldsymbol{Q})*\mathcal{O}$ is just the reflection of $\boldsymbol{P}*\boldsymbol{Q}$ in the $x$-axis, it follows that equation (4.13) defines an algorithm to compute $\boldsymbol{P}+\boldsymbol{Q}$ as well simply by putting a minus sign in the $y$-coordinate. Let us summarize the algorithm: let $\boldsymbol{P}=(x_1,y_1)$ and $\boldsymbol{Q}=(x_2,y_2)$ be rational points on the curve $E/\mathbb{Q} : y^2=x^3+ax+b$, then

(i). If $\boldsymbol{P}\neq\boldsymbol{Q}$ but $x_1=x_2$, then $\boldsymbol{P}+\boldsymbol{Q}=\mathcal{O}$.

(ii). If $\boldsymbol{P}=\boldsymbol{Q}$ but $y_1=y_2=0$, then $\boldsymbol{P}+\boldsymbol{Q}=\mathcal{O}$.

(iii). Otherwise, $\boldsymbol{P}+\boldsymbol{Q}=(x_3,-mx_3-c)$, where $x_3=m^2-x_1-x_2$ and

$$m=\begin{cases}\dfrac{y_2-y_1}{x_2-x_1}, & \text{if } \boldsymbol{P}\neq\boldsymbol{Q} \text{ and } x_1\neq x_2,\\[2ex] \dfrac{3x_1^2+a}{2y_1}, & \text{if } \boldsymbol{P}=\boldsymbol{Q} \text{ and } y_1\neq 0,\end{cases}$$

and $c=y_1-mx_1=y_2-mx_2$.

From (iii), we further get an explicit formula for the $x$-coordinate of points of the form $2\boldsymbol{P}=\boldsymbol{P}+\boldsymbol{P}$. Denote this as $x(2\boldsymbol{P})$. Then

$$x(2\boldsymbol{P})=m^2-2x_1=\left(\frac{3x_1^2+a}{2y_1}\right)^2-2x_1=\frac{9x_1^3+6ax_1^2+a^2}{4y_1^2}-2x_1.$$

If we replace $y^2$ by $x^3+ax+b$, then we end up with

$$x(2\boldsymbol{P})=\frac{x_1^4-2ax_1^2-8bx_1+a^2}{4x_1^3+4ax_1+4b}.$$

This formula for computing $x(2\boldsymbol{P})$ using only information about the $x$-coordinate of the point $\boldsymbol{P}$ is known as the *duplication formula* and is important both as a theoretical tool (i.e. to prove theorems) and as a computational tool.

By the partial closure lemma (applied twice), $(\boldsymbol{P}*\boldsymbol{Q})*\mathcal{O}$ is an element of $E(\mathbb{Q})$. Accordingly, $\boldsymbol{P}+\boldsymbol{Q}\in E(\mathbb{Q})$, and so $+$ defines a binary operation on $E(\mathbb{Q})$.

**Theorem 4.4.** *The pair $(E(\mathbb{Q}),+)$ is an abelian group with identity $\mathcal{O}$.*

*Proof.* We first prove that it is a group, and then prove commutativity of $+$. The only convention that we shall consider is that $\mathcal{O}+\mathcal{O}=\mathcal{O}$.

**$\mathcal{O}$ is the identity element.** Some books actually take this as a convention as well. To prove this, we have to prove that $\boldsymbol{P} + \mathcal{O} = \mathcal{O} + \boldsymbol{P} = \boldsymbol{P}$ for all $\boldsymbol{P} \in E(\mathbb{Q})$. If $\boldsymbol{P} = \mathcal{O}$, we are done so suppose not. We have that $\boldsymbol{P} + \mathcal{O} = (\boldsymbol{P} * \mathcal{O}) * \mathcal{O}$. The line $L$ that goes through $\boldsymbol{P}$ and $\mathcal{O}$ is a vertical line which meets $\boldsymbol{P} * \mathcal{O}$. Consequently, the line that goes through $\mathcal{O}$ and $\boldsymbol{P} * \mathcal{O}$ is the same vertical line which meets $\boldsymbol{P}$ as the third point. Therefore, $\boldsymbol{P} + \mathcal{O} = (\boldsymbol{P} * \mathcal{O}) * \mathcal{O} = \boldsymbol{P}$. By a similar argument, $\mathcal{O} + \boldsymbol{P} = \boldsymbol{P}$.

**Additive inverse of a point $\boldsymbol{P}$.** For our model of the elliptic curve, we claim that if $\boldsymbol{P} = (x, y)$, then its additive inverse is given by $-\boldsymbol{P} = (x, -y)$, the reflection of $\boldsymbol{P}$ in the $x$-axis. For this, we need to prove that $\boldsymbol{P} + (-\boldsymbol{P}) = \mathcal{O}$. Again, we unpack definitions: $\boldsymbol{P} + (-\boldsymbol{P}) = (\boldsymbol{P} * (-\boldsymbol{P})) * \mathcal{O}$. The line that goes through $\boldsymbol{P}$ and $-\boldsymbol{P}$ is the vertical line that meets $\mathcal{O}$ as the third intersection point. The line that goes through $\mathcal{O}$ and $\mathcal{O}$ is the line at infinity and so meets $\mathcal{O}$ as the third intersection point, as desired. Note that this argument works for $\boldsymbol{P} \neq \mathcal{O}$, but it should be easy to see that $-\mathcal{O} = \mathcal{O}$. Also note that this argument would not be as effective if our model of elliptic curve is different as we have mentioned before.

**Associativity of $+$.** Associativity is the hardest part which we shall not prove. For a proof, see Section 2.4 of Washington [29].

**Commutativity of $+$.** The line that goes through $\boldsymbol{P}$ and $\boldsymbol{Q}$ is the same line that goes through $\boldsymbol{Q}$ and $\boldsymbol{P}$. Accordingly, $\boldsymbol{P} * \boldsymbol{Q} = \boldsymbol{Q} * \boldsymbol{P}$ and so $\boldsymbol{P} + \boldsymbol{Q} = \boldsymbol{Q} + \boldsymbol{P}$, as desired. ■

Since $E(\mathbb{Q})$ is now a group, we can talk about order of elements. A point $\boldsymbol{P} \in E(\mathbb{Q})$ is said to have *finite order* $n \in \mathbb{Z}^+$ if $n$ is the least integer such that $n\boldsymbol{P} = \boldsymbol{P} + \cdots + \boldsymbol{P} = \mathcal{O}$. Any point of finite order $n$ is also called a *torsion point* of order $n$. If no such $n$ exists, we say $\boldsymbol{P}$ has *infinite order.*

### 4.4.2 Mordell-Weil theorem

From the structure theorem of abelian groups, we know that any finitely generated abelian group can be decomposed into a torsion part and a free part (for readers that are not familiar with this theorem, see Chapter 12, Section 6 of Artin [2]). A remarkable result proved by Mordell tells us that $E(\mathbb{Q})$ is in fact a finitely generated abelian group, and so has such a decomposition.

**Theorem 4.5** (Mordell-Weil theorem)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Then $E(\mathbb{Q})$ is a finitely generated abelian group.*

More explicitly, the Mordell-Weil theorem tells us that there is a finite set of rational points

$$\{\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_n\} \subseteq E(\mathbb{Q}),$$

which generates all the rational points on $E$. In other words, any other point $\boldsymbol{P} \in E(\mathbb{Q})$ is just a $\mathbb{Z}$-linear combination of the $\boldsymbol{Q}_i$. This theorem was initially proven by Mordell in 1922 and later generalized by Weil in 1928 to *abelian varities* over *number fields.* We note that these are deep results of number theory, especially the generalization. For those unfamiliar with number fields, they are fields containing $\mathbb{Q}$, such that as $\mathbb{Q}$-vector spaces, have finite dimension. The simplest example of a number field is $\mathbb{Q}$ itself — it has dimension 1 as a $\mathbb{Q}$-vector space. The definition of an abelian variety is a bit too technical to describe in only a few sentences. But one can think of it as simply a generalization of elliptic curves. Abelian varieties form one of the most important objects in algebraic geometry.

In honor of Mordell and Weil, the group $E(\mathbb{Q})$ is often called the *Mordell-Weil group* of $E$. The proof of the Mordell-Weil theorem relies on three main ingredients:

(1). The *weak Mordell-Weil theorem* (see Theorem 4.6 below).

(2). The idea of a *height function* on an abelian group $A$.

(3). The *descent theorem* which says that an abelian group $A$ with a height function, such that $A/2A$ is finite, is finitely generated.

We will give a brief but precise overview of what these statements mean. First, the weak Mordell-Weil theorem.

**Theorem 4.6** (Weak Mordell-Weil theorem)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Then the quotient group $E(\mathbb{Q})/mE(\mathbb{Q})$ is a finite group for all $m \geqslant 2$.*

For the rest of this discussion, we fix $E/\mathbb{Q}$ to be the elliptic curve defined by the Weierstrass equation $y^2 = x^3 + ax + b$ where $a, b \in \mathbb{Z}$.

An elementary proof of the weak Mordell-Weil theorem for $m = 2$ can be found in Lozano-Robledo [11], Section 2.8 and Washington [29], Section 8.2. However, here they make the assumption that $E(\mathbb{Q})$ has four distinct torsion points of order 2. So their proof does not work for elliptic curves in its full generality, but that is the usual price for elementary proofs. The reason behind this quite restrictive assumption is so that we can write the Weierstrass equation of $E$ as a product

$$x^3 + ax + b = (x - e_1)(x - e_2)(x - e_3),$$

where the $e_i$ are all distinct integers. The proof is done by ingeniously considering a group homomorphism

$$E(\mathbb{Q}) \longrightarrow (\mathbb{Q}^\times/{\mathbb{Q}^\times}^2) \times (\mathbb{Q}^\times/{\mathbb{Q}^\times}^2) \times (\mathbb{Q}^\times/{\mathbb{Q}^\times}^2) = (\mathbb{Q}^\times/{\mathbb{Q}^\times}^2)^3$$

whose kernel is $2E(\mathbb{Q})$. It turns out that this homomorphism induces an injection of $E(\mathbb{Q})/2E(\mathbb{Q})$ into a finite subgroup $\Gamma$ of $(\mathbb{Q}^\times/{\mathbb{Q}^\times}^2)^3$. Since $\Gamma$ is finite, the claim follows. The proof that takes care of the general case as in Theorem 4.6 (in fact, more general as $\mathbb{Q}$ is replaced by a number field) is done in Silverman [24], Section VIII.1. Here, they use a quite advanced tool from cohomology theory called *group cohomology* (for a basic exposition of this, see Appendix B of the same book [24]).

We now move to the second statement about height functions.

**Definition 4.15.** Let $x \in \mathbb{Q}$, and suppose $x = m/n$ is written in lowest terms. The **height** of $x$ is the positive integer defined by $H(x) = \max\{|m|, |n|\}$.

Using the notion of height on rational numbers, we can define a height function on $E(\mathbb{Q})$. For any point $\boldsymbol{P} \in E(\mathbb{Q})$, write $\boldsymbol{P} = (x_{\boldsymbol{P}}, y_{\boldsymbol{P}})$.

**Definition 4.16.** The (logarithmic) **height** on $E(\mathbb{Q})$ is a map $h : E(\mathbb{Q}) \to \mathbb{R}$ defined by

$$h(\boldsymbol{P}) = \begin{cases} \log H(x_{\boldsymbol{P}}), & \text{if } \boldsymbol{P} \neq \mathcal{O}, \\ 0, & \text{if } \boldsymbol{P} = \mathcal{O}. \end{cases}$$

**Remark 11.** Note that the definition of height on $E(\mathbb{Q})$ is relative to the Weierstrass equation defining $E$. Also, note that $h(\boldsymbol{P}) \geqslant 0$ for all $\boldsymbol{P}$.

The purpose of a height function on $E(\mathbb{Q})$ is that we want a device to measure the complexity of a rational point, at least in a number-theoretic sense. In particular, if given two points, we want to be able to tell which point is "smaller" relative to the other. This allows us to run Fermat's descent argument, just like we did when proving Lemma 3.6.

We now state and prove the descent theorem mentioned in the third statement.

**Theorem 4.7** (Descent theorem)**.** *Let $A$ be an abelian group, and suppose there exists a map $h : A \to [0, \infty)$ satisfying the following three properties:*

(i) *For every constant $C_1 \in \mathbb{R}$, the set $\{\boldsymbol{P} \in A : h(\boldsymbol{P}) \leqslant C_1\}$ is finite.*

(ii) *For every $\boldsymbol{P}_0 \in A$, there exists a constant $C_2$ such that*

$$h(\boldsymbol{P} + \boldsymbol{P}_0) \leqslant 2h(\boldsymbol{P}) + C_2, \text{ for all } \boldsymbol{P} \in A.$$

(iii) *There exists a constant $C_3$ such that*

$$h(2\boldsymbol{P}) \geqslant 4h(\boldsymbol{P}) - C_3, \text{ for all } \boldsymbol{P} \in A.$$

*Suppose further that the quotient group $A/2A$ is finite. Then $A$ is finitely generated.*

We give a proof that is adapted from Silverman and Tate [26].

*Proof.* Pick an element $\boldsymbol{Q}_i$ for each coset of $2A$ in $A$ to get a finite list $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_n$ of coset representatives. We can do this since we assumed that $A/2A$ is a finite group. Now let $\boldsymbol{P} \in A$. Since the cosets of $2A$ partitions $A$, we can find an index $1 \leqslant i_1 \leqslant n$ such that

$$\boldsymbol{P} = 2\boldsymbol{P}_1 + \boldsymbol{Q}_{i_1}, \tag{4.14}$$

for some $\boldsymbol{P}_1 \in A$. We repeat the same idea with $\boldsymbol{P}_1$, and keep repeating to get a list of points

$$\begin{aligned} \boldsymbol{P}_1 &= 2\boldsymbol{P}_2 + \boldsymbol{Q}_{i_2}, \\ \boldsymbol{P}_2 &= 2\boldsymbol{P}_3 + \boldsymbol{Q}_{i_3}, \\ &\vdots \\ \boldsymbol{P}_{m-1} &= 2\boldsymbol{P}_m + \boldsymbol{Q}_{i_m}, \end{aligned}$$

where $1 \leqslant i_j \leqslant n$. From here, we can do back substitution. We start by substituting the equation $\boldsymbol{P}_1 = 2\boldsymbol{P}_2 + \boldsymbol{Q}_{i_2}$ into its preceding equation (4.14). We then substitute $\boldsymbol{P}_2 = 2\boldsymbol{P}_3 + \boldsymbol{Q}_{i_3}$ into $\boldsymbol{P}_1 = 2\boldsymbol{P}_2 + \boldsymbol{Q}_{i_2}$, so on and so forth. This results in $\boldsymbol{P}$ being written as a linear combination of $\boldsymbol{P}_m$ and the $\boldsymbol{Q}_i$,

$$\boldsymbol{P} = 2^m \boldsymbol{P}_m + \sum_{j=1}^{m} 2^{j-1} \boldsymbol{Q}_{i_j}.$$

Equivalently, we have that the set $\{\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_n, \boldsymbol{P}_m\}$ forms a (not necessarily finite) generating set for $A$. Our aim now is to show that we can choose $m$ sufficiently large so that $h(\boldsymbol{P}_m)$ is bounded by a constant $\kappa$ *independent* of $\boldsymbol{P}$. This would imply that the set

$$\mathcal{G}_\kappa = \{\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_n\} \cup \{\boldsymbol{R} \in A : h(\boldsymbol{R}) \leqslant \kappa\}$$

forms a *finite* generating set for $A$, where the finiteness is due to property (i).

To bound $h(\boldsymbol{P}_m)$, we need some important constants which we can get by playing around with properties (ii) and (iii). We apply property (ii) on each $-\boldsymbol{Q}_i$ to get constants $c_i$ such that

$$h(\boldsymbol{P} - \boldsymbol{Q}_i) \leqslant 2h(\boldsymbol{P}) + c_i,$$

for all $\boldsymbol{P} \in A$. If we set $C_2 = \max_{1 \leqslant i \leqslant n} \{c_i\}$, then we in fact have

$$h(\boldsymbol{P} - \boldsymbol{Q}_i) \leqslant 2h(\boldsymbol{P}) + C_2, \tag{4.15}$$

for all $\boldsymbol{P} \in A$ and all $1 \leqslant i \leqslant n$. Next, for any index $j$, we apply property (iii) to get a constant $C_3$ such that

$$\begin{aligned} h(\boldsymbol{P}_j) \leqslant \frac{1}{4}(h(2\boldsymbol{P}_j) + C_3) &= \frac{1}{4}(h(\boldsymbol{P}_{j-1} - \boldsymbol{Q}_{i_j}) + C_3) \\ &\leqslant \frac{1}{4}(2h(\boldsymbol{P}_{j-1}) + C_2 + C_3) = \frac{1}{2}h(\boldsymbol{P}_{j-1}) + \frac{C_2 + C_3}{4}, \end{aligned}$$

where we have used (4.15) in the second inequality. We can then cleverly manipulate the right-hand side to get

$$h(\boldsymbol{P}_j) \leqslant \frac{3}{4}h(\boldsymbol{P}_{j-1}) - \frac{1}{4}\left(h(\boldsymbol{P}_{j-1}) - (C_2 + C_3)\right).$$

From here, we can clearly see that if $h(\boldsymbol{P}_{j-1}) \geqslant C_2 + C_3$, then

$$h(\boldsymbol{P}_j) \leqslant \frac{3}{4}h(\boldsymbol{P}_{j-1}).$$

Moreover, this is true for any index $j$. So if we still have $h(\boldsymbol{P}_j) \geqslant C_2 + C_3$, then the next point $\boldsymbol{P}_{j+1}$ must have a smaller height

$$h(\boldsymbol{P}_{j+1}) \leqslant \frac{3}{4}h(\boldsymbol{P}_j) \leqslant \left(\frac{3}{4}\right)^2 h(\boldsymbol{P}_{j-1}).$$

If we do this repeatedly, we accumulate a factor of 3/4 at each step; and so the height decreases for each point until some point $\boldsymbol{P}_m$, whose height satisfies $h(\boldsymbol{P}_m) \leqslant C_2 + C_3$. It then follows that the set

$$\mathcal{G}_{C_2+C_3} = \{\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_n\} \cup \{\boldsymbol{R} \in A : h(\boldsymbol{R}) \leqslant C_2 + C_3\}$$

is a finite generating set for $A$, and hence, $A$ is finitely generated. ■

So for us, if we can prove that the height function on $E(\mathbb{Q})$ satisfy the three properties (i)-(iii) above, we get the Mordell-Weil theorem by taking $A = E(\mathbb{Q})$. Property (i) is quite trivial to prove.

**Lemma 4.5.** *For every constant $C_1 \in \mathbb{R}$, the set $\{\boldsymbol{P} \in E(\mathbb{Q}) : h(\boldsymbol{P}) \leqslant C_1\}$ is finite.*

*Proof.* If $x = m/n$ is written in lowest terms, and that its height $H(x) \leqslant \kappa$ for some constant $\kappa \in \mathbb{R}$, it follows that $|m|, |n| \leqslant \kappa$. Since $m, n$ are integers, then there are only finitely many possibilities for $m, n$. So we have proved that for any constant $\kappa \in \mathbb{R}$, the set

$$\{x \in \mathbb{Q} : H(x) \leqslant \kappa\} \subseteq \mathbb{Q}$$

is finite. Now given any $x \in \mathbb{Q}$, there are only two possibilities that $y$ can take if $(x, y) \in E(\mathbb{Q})$. So if we restrict to finitely many possibilities for $x$, we have only finitely many $(x, y) \in E(\mathbb{Q})$. This proves that for any constant $\kappa'$, the set

$$\{\boldsymbol{P} \in E(\mathbb{Q}) : H(\boldsymbol{P}) \leqslant \kappa'\}$$

is finite, and the same holds true if replace $H$ with $h$. ■

Property (ii) and (iii) are a bit more involved and so we only state them here as a lemma. For a proof, see Silverman-Tate [26], Section 3.1 or Silverman [24], Section VIII.4.

**Lemma 4.6.** *Let $E/\mathbb{Q}$ be an elliptic curve defined by $y^2 = x^3 + ax + b$ where $a, b \in \mathbb{Z}$.*

(a). *Let $\boldsymbol{P}_0 \in E(\mathbb{Q})$. Then there exists a constant $C_2$, depending on $\boldsymbol{P}_0$, $a$, and $b$, such that*

$$h(\boldsymbol{P}+\boldsymbol{P}_0) \leqslant 2h(\boldsymbol{P}) + C_2, \textit{ for all } \boldsymbol{P} \in E(\mathbb{Q}).$$

(b). *There exists a constant $C_3$, that depends on $a$ and $b$, such that*

$$h(2\boldsymbol{P}) \geqslant 4h(\boldsymbol{P}) - C_3, \textit{ for all } \boldsymbol{P} \in E(\mathbb{Q}).$$

For readers who are interested in proving the Mordell-Weil theorem in more generality (but still not to the extent of Weil's result), see Silverman [24], Chapter VIII. Here, they prove the theorem for elliptic curves over an arbitrary number field $K$ instead of just $\mathbb{Q}$. The idea is to define a general height function $H_K$ on the projective $n$-space $\mathbb{P}^n$. Since elliptic curves are really just subsets of $\mathbb{P}^n$, we can then look at the properties of $H_K$ when the points on $\mathbb{P}^n$ are restricted to the points on the elliptic curve.

### 4.4.3 Mazur's torsion theorem and the Nagell-Lutz theorem

As mentioned before, the structure theorem implies that $E(\mathbb{Q})$ decomposes into abelian groups

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

for some positive integer $r$. The integer $r$ is called the *rank* of $E(\mathbb{Q})$, and $E(\mathbb{Q})_{\text{tors}}$ is called the *torsion subgroup of* $E(\mathbb{Q})$ which contains the torsion (rational) points on $E$. Note that if $r = 0$, then $E(\mathbb{Q})$ is torsion and so is a finite abelian group. It turns out that the torsion part $E(\mathbb{Q})_{\text{tors}}$ is well-understood thanks to the work of Barry Mazur in 1977. Consequently, we have understood all the rational points on elliptic curves over $\mathbb{Q}$ with rank $r = 0$.

**Theorem 4.8** (Mazur's torsion theorem). *Let $E/\mathbb{Q}$ be an elliptic curve. Then $E(\mathbb{Q})_{\text{tors}}$ can be identified with exactly one of the following 15 groups:*

(1). *$\mathbb{Z}/m\mathbb{Z}$ for $1 \leqslant m \leqslant 10$ or $m = 12$.*

(2). *$(\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2m\mathbb{Z})$ for $1 \leqslant m \leqslant 4$.*

*In particular, the cardinality of $E(\mathbb{Q})_{\text{tors}}$ is bounded: $|E(\mathbb{Q})_{\text{tors}}| \leqslant 16$.*

The proof of this theorem is quite deep and we will not discuss it. Mazur's theorem is nice but it does not give us a method to compute the torsion subgroup of $E(\mathbb{Q})$. An effective method that allows us to compute $E(\mathbb{Q})_{\text{tors}}$ is given by the following theorem of Nagell and Lutz.

**Theorem 4.9** (Nagell-Lutz theorem). *Let $E/\mathbb{Q}$ be an elliptic curve defined by the Weierstrass equation $y^2 = x^3 + ax + b$, where $a, b \in \mathbb{Z}$. If $(x_0, y_0) \in E(\mathbb{Q})$ is a nonzero torsion point, then*

(1). *$(x_0, y_0) \in \mathbb{Z}^2$.*

(2). *Either $y_0 = 0$ and so $(x_0, y_0)$ has order 2, or $y_0^2$ divides $4a^3 + 27b^2$.*

*Proof.* See Silverman and Tate [26]. ■

Let us use the Nagell-Lutz theorem to compute torsion points.

**Example.** Consider the elliptic curve $E : y^2 = x^3 + 1849$. Note that $1849 = 43^2$ and 43 is a prime number. Observe that $E$ follows the setup of the Nagell-Lutz theorem with $a = 0$ and $b = 1849$ so we can apply it.

*Is there a point of order 2*? This is easily verified to be no. By the rational root theorem, the only possible rational solutions to $x^3 + 1849 = 0$ are $\{\pm 43, \pm 1849\}$ but none of these satisfy our equation. It follows that $y^2 \neq 0$ and so $y \neq 0$. So any torsion point on $E$ cannot possibly have order 2.

*What are the torsion points on $E$?* Let $(x_0, y_0) \in E(\mathbb{Q})$ be a torsion point. It cannot have order 2, so $y_0^2$ divides $4a^3 + 27b^2 = 27 \cdot 1849^2 = 3^3 \cdot 43^4$. The possible candidates for $y_0$ are then

$$S = \{\pm 1,\ \pm 3,\ \pm 43,\ \pm 129,\ \pm 1849,\ \pm 5547\}.$$

By inspection, we see that $(0, \pm 43) \in E(\mathbb{Q})$. We claim that $\boldsymbol{Q} = (0, 43)$ is a torsion point of order 3. We see that the tangent line to $E$ at $\boldsymbol{Q}$ is defined by the line $\ell : y = 43$. Using the explicit algorithm of adding points (see Remark 10) in $E(\mathbb{Q})$, we have

$$2\boldsymbol{Q} = \boldsymbol{Q} + \boldsymbol{Q} = (0, -43) = -\boldsymbol{Q}.$$

It thus follows that $3\boldsymbol{Q} = \boldsymbol{Q} + 2\boldsymbol{Q} = \boldsymbol{Q} - \boldsymbol{Q} = \mathcal{O}$ as desired. Note that consequently, $-\boldsymbol{Q}$ is also a torsion point of order 3. We now make another claim that in fact $E(\mathbb{Q})_{\text{tors}} = \{\mathcal{O}, \boldsymbol{Q}, 2\boldsymbol{Q}\}$ and so is isomorphic (as groups) to $\mathbb{Z}/3\mathbb{Z}$. To prove this, we simply have to rule out all the other possibilities we had earlier. But this is easy as for any $y_1 \in S$, the real solution $x_1$ to the equation $x^3 = y_1^2 - 1849$ is not an integer. By the Nagell-Lutz theorem, it follows that $(x_1, y_1) \notin E(\mathbb{Q})_{\text{tors}}$. We thus conclude that $E(\mathbb{Q})_{\text{tors}} \cong \mathbb{Z}/3\mathbb{Z}$ which agrees with Mazur's torsion theorem.

## 4.5 Elliptic curve over finite fields

Recall that $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ is the finite field of $p$ elements where $p$ is a prime number. We shall now look at elliptic curves defined over $\mathbb{F}_p$.

Let $p$ be a prime number and suppose $E/\mathbb{Q}$ is an elliptic curve defined by the Weierstrass equation $y^2 = x^3 + ax + b$ where $a, b \in \mathbb{Z}$. We can reduce the coefficients modulo $p$ to get a new cubic curve $\tilde{E}$, called the *reduction curve modulo $p$*, which is now defined over $\mathbb{F}_p$. However, $\tilde{E}/\mathbb{F}_p$ is not necessarily an elliptic curve as it may be singular!

**Definition 4.17.** Let $E/\mathbb{Q}$ be an elliptic curve defined by the Weierstrass equation $y^2 = x^3 + ax + b$ where $a, b \in \mathbb{Q}$. Define $S$ to be the set of elliptic curves

$$S = \{E' : E' \text{ is isomorphic to } E \text{ over } \mathbb{Q}, \text{ and } \Delta_{E'} \in \mathbb{Z}\}.$$

The **minimal discriminant** of $E$ is defined to be the integer

$$\Delta_{\min,E} = \min_{E' \in S} |\Delta_{E'}|.$$

The elliptic curve model $E'$ with minimal discriminant of $E$ is then said to be a **minimal model** for $E$.

**Remark 12.** To summarize, a minimal model for $E$ is an elliptic curve isomorphic to $E$ such that its discriminant is an integer and (in absolute value) is as small as possible.

Due to the possibilities of cubic curves being singular once again, we need to talk about special type of singularities that will hold some interest for us. Let $\boldsymbol{P} = (x_0, y_0)$ be a singular point of the cubic curve

$$C : f(x, y) = y^2 + axy + by - x^3 - cx^2 - dx - e = 0,$$

which is defined over $K$ and given by a Weierstrass equation. By definition of a singular point, we have that $\nabla(\boldsymbol{P}) = 0$. So, there exist $\alpha, \beta \in \bar{K}$ such that we can write the Taylor series expansion of $f$ around $\boldsymbol{P}$ in the following way:

$$f(x,y) - f(x_0,y_0) = ((y-y_0) - \alpha(x-x_0))\,((y-y_0) - \beta(x-x_0)) - (x-x_0)^3.$$

**Definition 4.18.** With notation as used above, we say that the singular point $\boldsymbol{P} = (x_0, y_0)$ is

(i). A **node** if $\alpha \neq \beta$, in which case there are two different tangent lines to $C$ at $\boldsymbol{P}$ given by:

$$y - y_0 = \alpha(x-x_0), \quad y - y_0 = \beta(x-x_0).$$

(ii). A **cusp** if $\alpha = \beta$, in which case there is a unique tangent line to $C$ at $\boldsymbol{P}$ given by

$$y - y_0 = \alpha(x-x_0).$$

**Definition 4.19.** Let $E/\mathbb{Q}$ be an elliptic curve, and let $\tilde{E}$ be the reduction curve modulo $p$ of a minimal model for $E$. We say that

(1). $E$ has **good** (or stable) reduction if $\tilde{E}$ is nonsingular.

(2). $E$ has **multiplicative** (or semistable) reduction if $\tilde{E}$ has a node.

(3). $E$ has **additive** (or unstable) reduction if $\tilde{E}$ has a cusp.

We say that $E$ has **bad reduction** if it has either multiplicative or additive reduction. If $E$ has multiplicative reduction at a singular point $\boldsymbol{P}$, we further say that the reduction is **split** multiplicative if the slopes of the tangent line to $E$ at $\boldsymbol{P}$ are in $\mathbb{F}_p$. Otherwise, we say that it is **non-split**.

We now claim that if $E/\mathbb{Q}$ is an elliptic curve given by a minimal model, then $E$ has bad reduction at a prime $p$ if and only if $p \mid \Delta_E$. For this, we first need some lemmas.

**Lemma 4.7.** *Let $f \in K[x]$ be a polynomial over a field $K$, and let its derivative be $f'$. Then $\lambda$ is a double root of $f$ if and only if $f(\lambda) = f'(\lambda) = 0$.*

*Proof.* Suppose that $\lambda$ is a double root of $f$. Then there exists a polynomial $g \in K[x]$ such that $f(x) = (x-\lambda)^2 g(x)$. Taking the derivative of $f$, we get

$$f'(x) = 2(x-\lambda)g(x) + (x-\lambda)^2 g'(x).$$

It is then immediate to see that $f'(x)$ also vanishes at $\lambda$.

The converse is a bit tricky. Suppose now that $f(\lambda) = f'(\lambda) = 0$. By the division algorithm (applied to $(x-\lambda)^2$ and $f$), there exist polynomials $q, r \in K[x]$ such that

$$f(x) = (x-\lambda)^2 q(x) + r(x),$$

where either $\deg r < \deg(x-\lambda)^2 = 2$ or $r(x) = 0$. If $r(x) = 0$, we are done so suppose not. Then $r(x) = ax + b$ for some $a, b \in K$ with derivative $r'(x) = a$. Taking the derivative of $f$, we get

$$f'(x) = 2(x-\lambda)q(x) + (x-\lambda)^2 q'(x) + r'(x).$$

Applying our hypothesis that $f(\lambda) = f'(\lambda) = 0$, we get the relation $r(\lambda) = r'(\lambda) = 0$. But this implies that $a = b = 0$ and so $r(x)$ is identically zero. So we conclude that $\lambda$ is a double root of $f$ as desired. ■

**Lemma 4.8.** *Let $f \in K[x]$ be a monic cubic polynomial over a field $K$ and consider the cubic curve $C : y^2 = f(x)$. Then the singular points on $C$, if they exist, are of the form $(\lambda, 0)$ where $\lambda$ is a double root of $f$.*

*Proof.* Suppose that $\boldsymbol{P} = (x_0, y_0)$ is a singular point on $C$. We want to prove that $y_0 = 0$ and $x_0 = \lambda$ for some root $\lambda \in K$ of $f$. Suppose $f(x) = x^3 + ax^2 + bx + c$ and consider the polynomial $F(x, y) = y^2 - f(x)$. For $\boldsymbol{P}$ to be a singular point, it follows that we must have $\nabla F(\boldsymbol{P}) = 0$. So, we require that

$$\frac{\partial F}{\partial y}(\boldsymbol{P}) = 2y_0 = 2\sqrt{f(x_0)} = 0,$$

which implies that we require $y_0 = 0$ and $f(x_0) = 0$. And, we require

$$\frac{\partial F}{\partial x}(\boldsymbol{P}) = -3x_0^2 - 2ax_0 - b = -f'(x_0) = 0,$$

which implies that $f'(x_0) = 0$. By Lemma 4.7, it follows that $x_0$ is a double root of $f$. ■

**Proposition 4.2** ([11], Chapter 2, Proposition 2.5.8)**.** *Let $C : y^2 = f(x)$ be a cubic curve where $f(x) \in K[x]$ is a monic cubic polynomial. Suppose that $f$ is given by $f(x) = (x - \alpha)(x - \beta)(x - \gamma)$ where $\alpha, \beta, \gamma \in \bar{K}$ and define $D = (\alpha - \beta)^2(\alpha - \gamma)^2(\beta - \gamma)^2$. Then $C$ is nonsingular if and only if $D \neq 0$.*

Note that $\alpha, \beta, \gamma$ are elements of the algebraic closure of $K$ so that the factorization of $f$ into linear factors make sense in general.

*Proof.* We shall prove that $C$ is singular if and only if $D = 0$.

Suppose that $C$ is singular. Then by Lemma 4.8, the singular points on $C$ are of the form $(\lambda, 0)$ where $\lambda$ is a double root of $f$. Accordingly, at least two of $\alpha, \beta, \gamma$ are equal and so $D = 0$.

The proof of the converse is similar to the proof of Lemma 4.8. Assume that $D = 0$. Then at least two of $\alpha, \beta, \gamma$ are equal, so suppose $\alpha = \beta$. This implies that $f(x) = (x - \alpha)^2(x - \gamma)$ and so $\alpha$ is a double root of $f$. By Lemma 4.7, it follows that $f'(\alpha) = 0$. Now let $F(x, y) = y^2 - f(x)$ and consider its tangent vector

$$\nabla F = (-f'(x), 2y) = \left(-f'(x), 2\sqrt{f(x)}\right).$$

We then immediately see that $(\alpha, 0)$ defines a singular point on $C$, and so $C$ is singular. ■

Notice that the number $D$ is the usual discriminant of a cubic. If we consider the cubic polynomial $f(x) = x^3 + ax + b$, then one can compute that this polynomial has discriminant $D = -4a^3 - 27b^2$. This quantity should look strikingly familiar. This is because we have defined the discriminant of the elliptic curve $E : y^2 = f(x)$ to be $\Delta_E = -16(4a^3 + 27b^2) = 16D$. This observation combined with Proposition 4.2 gives rise to the following corollary.

**Corollary 4.1.** *Let $p$ be a prime number, and let $E/\mathbb{Q}$ be an elliptic curve given by a minimal model. Then $E$ has bad reduction at $p$ if and only if $p \mid \Delta_E$.*

Now consider the elliptic curve $E/\mathbb{F}_q$ defined over a finite field where $q = p^r$ is a prime power. A question that we can then ask is how many points are there in $E(\mathbb{F}_q)$? We know that there are only a finite number of points as $E(\mathbb{F}_q) \subseteq \mathbb{P}^2(\mathbb{F}_q)$, but is there an estimate for this number $N_q := \#E(\mathbb{F}_q)$? This is answered by the following result, conjectured by Emil Artin in his thesis and finally proved by Hasse in 1933.

**Theorem 4.10** (Hasse, 1933). *Let $E/\mathbb{F}_q$ be an elliptic curve with coefficients in $\mathbb{F}_q$. Then there is a bound*

$$|q + 1 - N_q| \leqslant 2\sqrt{q},$$

*where $N_q = \#E(\mathbb{F}_q)$.*

*Proof.* See Chapter V, Theorem 1.1 of Silverman [24]. ■

By using Hasse's bound, Schoof [20] established an algorithmic approach to get the exact number of $N_q$ which is definitely worth a read.

### 4.5.1 Introduction to $L$-functions and modular forms

Let $E/\mathbb{Q}$ be an elliptic curve defined by the Weierstrass equation

$$y^2 + axy + by = x^3 + cx^2 + dx + e,$$

where $a, b, c, d, e \in \mathbb{Z}$. For a prime number $p$, we define a new quantity

$$a_p = \begin{cases} p + 1 - N_p, & \text{if } E \text{ has good reduction at } p. \\ 1, & \text{if } E \text{ has split multiplicative reduction at } p. \\ -1, & \text{if } E \text{ has non-split multiplicative reduction at } p. \\ 0, & \text{if } E \text{ has additive reduction at } p. \end{cases}$$

where $N_p = \#E(\mathbb{F}_p)$ is as defined before. One can think of $a_p$ as a piece of data which describes the behaviour of the reduction curve modulo $p$. Next, we define the *local part at* $p$ to be:

$$L_p(X) = \begin{cases} 1 - a_pX + pX^2, & \text{if } E \text{ has good reduction at } p; \\ 1 - X, & \text{if } E \text{ has split multiplicative reduction at } p; \\ 1 + X, & \text{if } E \text{ has non-split multiplicative reduction at } p; \\ 1, & \text{if } E \text{ has additive reduction at } p. \end{cases}$$

Using the local part at $p$, we can now associate any elliptic curve $E/\mathbb{Q}$ with a new object called the *$L$-function*.

**Definition 4.20.** Let $E/\mathbb{Q}$ be an elliptic curve. The **$L$-function** of $E$ is defined to be

$$L(E, s) = \prod_{p \text{ prime}} \frac{1}{L_p(p^{-s})}.$$

More explicitly, the $L$-function of $E$ is given by the Euler product

$$L(E, s) = \prod_{p \text{ bad}} (1 - a_pp^{-s})^{-1} \prod_{p \text{ good}} (1 - a_pp^{-s} + p^{1-2s})^{-1},$$

where the first product is over primes where $E$ has bad reduction (called *bad primes*) and the second product is over primes where $E$ has good reduction (called *good primes*). We previously mentioned that $a_p$ contains modulo $p$ data for the elliptic curve $E$. So looking at how the $L$-function is defined above, we see that $L(E, s)$ contains information for reduction curves $\tilde{E}$ modulo $p$ for all prime numbers $p$. While this seems like nothing unusual at the moment, there is an important conjecture which claims that it is enough to understand $L(E, s)$ if we want to understand the rank of the Mordell-Weil group $E(\mathbb{Q})$ (which is still currently a mystery).

The nice thing about the $L$-function is that we can expand it into a Dirichlet series. Using the well-known geometric series identity, we can write

$$\frac{1}{1-a_pp^{-s}} = \sum_{k=0}^{\infty}(a_pp^{-s})^k = a_{p^k}p^{-sk}$$

for the primes of bad reduction, where we have put $a_{p^k} = (a_p)^k$. For the primes of good reduction, it is a bit more complicated. We expand the good prime factor to get

$$\begin{aligned}\frac{1}{1-a_pp^{-s}+p^{1-2s}} &= \sum_{k=0}^{\infty}(a_pp^{-s} - p^{1-2s})^k \\ &= 1 + a_pp^{-s} + (a_p^2 - p)p^{-2s} + (a_p^3 - 2pa_p)p^{-3s} + \cdots \\ &= 1 + a_pp^{-s} + a_{p^2}p^{-2s} + a_{p^3}p^{-3s} \cdots ,\end{aligned}$$

where we have put $a_{p^2} = a_p^2 - p$ and $a_{p^3} = a_p^3 - 2pa_p$. In general, we define this recursively

$$a_{p^{k+1}} = a_{p^k}a_p - pa_{p^{k-1}},$$

for all positive integer $k$. Then, we put $a_{mn} = a_m a_n$ if $\gcd(m, n) = 1$ for the other coefficients. This is a consequence of the observation that $a_n$ defines a multiplicative function. Finally, we define $a_1 = 1$. We then make the fundamental observation that in both good and bad prime factors, we get a series that looks like

$$\sum_{k=0}^{\infty}\frac{a_{p^k}}{(p^k)^s}.$$

Since the $L$-function is the Euler product of these factors, and since $a_n$ is multiplicative, it follows by unique factorization in $\mathbb{Z}$ that indeed the $L$-function can be written as a Dirichlet series

$$L(E, s) = \sum_{n=1}^{\infty}\frac{a_n}{n^s}.$$

Using Hasse's estimate (Theorem 4.10), one can show that the Dirichlet series of $L(E, s)$ converges for all $s \in \mathbb{C}$ in the half-plane $\text{Re}(s) > \frac{3}{2}$. What is really not obvious is the fact that $L(E, s)$ has an analytic continuation to all of $\mathbb{C}$. This is a consequence of the Taniyama-Shimura-Weil conjecture (now a theorem called the modularity theorem) which is a very deep result. To talk about this, we have to first define modular and cusp forms.

**Definition 4.21.** Let $N$ be a positive integer. The **principal congruence subgroup of level** $N$ is defined to be the set

$$\Gamma(N) = \{\mathcal{M} \in \text{SL}_2(\mathbb{Z}) : \mathcal{M} \equiv \mathbb{I} \bmod N\},$$

where $\mathbb{I}$ is the $2 \times 2$ identity matrix.

**Definition 4.22.** A subgroup $\Gamma \subseteq \text{SL}_2(\mathbb{Z})$ is called a **congruence subgroup (of level $N$)** if $\Gamma(N) \subseteq \Gamma$ for some positive integer $N$.

It can be easily proven that these are really subgroups of $\text{SL}_2(\mathbb{Z})$ for all positive integer $N$.

**Example.** *Typical examples.*

(i). $\Gamma_0(N) = \left\{\mathcal{M} \in \text{SL}_2(\mathbb{Z}) : \mathcal{M} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod N\right\}.$

(ii). $\Gamma_1(N) = \left\{\mathcal{M} \in \text{SL}_2(\mathbb{Z}) : \mathcal{M} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \bmod N\right\}.$

We shall need some more definitions. From now on, denote $\mathcal{H} = \{z \in \mathbb{C} : \text{Im}(\tau) > 0\}$ to be the *upper half plane.*

**Definition 4.23.** For any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, we define the following:

(1). The **factor of automorphy** for $z \in \mathcal{H}$ is defined to be $j(\gamma, z) = cz + d$.

(2). Let $k \in \mathbb{Z}$. The **weight-$k$ operator** $[\gamma]_k$ is a map $\mathrm{Hom}(\mathcal{H}, \mathbb{C}) \to \mathrm{Hom}(\mathcal{H}, \mathbb{C})$ defined by

$$(f[\gamma]_k)(z) = j(\gamma, z)^{-k} f(\gamma(z)), \qquad \text{for all } z \in \mathcal{H},$$

where we have written $[\gamma]_k(f) = f[\gamma]_k$.

Note that the definition of the weight-$k$ operator does not require the function $f$ to be either holomorphic or meromorphic on $\mathcal{H}$.

**Definition 4.24.** Let $k \in \mathbb{Z}$ and let $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$. A function $f : \mathcal{H} \to \mathbb{C}$ is said to be **weakly modular of weight $k$ with respect to** $\Gamma$ if it is meromorphic on $\mathcal{H}$ and $f[\gamma]_k = f$ for all $\gamma \in \Gamma$.

By definition of the congruence subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$, we can always find a translation matrix of the form

$$\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}, \tag{4.16}$$

which maps $z \mapsto z + h$ for some *minimal* $h \in \mathbb{Z}^+$. This is due to $\Gamma$ containing the principal subgroup $\Gamma(N)$ which contains matrices of the form

$$\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix}.$$

For example, we see that the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ is an element of the congruence subgroups $\Gamma_0(N)$ and $\Gamma_1(N)$. Now, define $\mathcal{D} = \{z \in \mathbb{C} : |z| < 1\} \subseteq \mathbb{C}$ to be the open unit disk centred at 0, and define its punctured disk by $\mathcal{D}^* = \mathcal{D} \backslash \{0\}$. With notation of $h$ and $\Gamma$ as defined above (so $\Gamma$ is a congruence subgroup), we give the following proposition.

**Proposition 4.3.** *Let $f : \mathcal{H} \to \mathbb{C}$ be a weakly modular function of weight $k$ with respect to $\Gamma$. Then $f$ is periodic of period $h$ and there is a function $g : \mathcal{D}^* \to \mathbb{C}$ such that $f(z) = g(q_h)$, where $q_h = q_h(z) = e^{2\pi i z/h}$.*

*Proof.* Let $\gamma$ be the matrix (4.16). We compute the factor of automorphy to be $j(\gamma, z) = 1$ and so by definition of weakly modular, we observe that

$$f(z) = f[\gamma]_k = f(\gamma(z)) = f(z + h).$$

This implies that $f$ is periodic of period $h$. From complex analysis, we know that the function $q_h(z) = e^{2\pi i z/h}$, which is periodic of period $h$, is a holomorphic function $\mathcal{H} \to \mathcal{D}^*$. Now consider the function $g : \mathcal{D}^* \to \mathbb{C}$ defined by

$$g(q_h) = f\left(\frac{h \log q_h}{2\pi i}\right)$$

so that $f(z) = g(q_h)$. Since $f$ is periodic of period $h$, we are allowed to choose any branch of $\log q_h$ in $\mathcal{H}$ and so $g$ is well-defined. ■

Using the notation in the proof above, notice that $f$ being holomorphic on $\mathcal{H}$ implies that $g$ is holomorphic on $\mathcal{D}^*$ and so has a Laurent series expansion on $\mathcal{D}^*$,

$$f(z) = \sum_{n \in \mathbb{Z}} a_n(f)\, q_h^n,$$

where $q_h = e^{2\pi i z/h}$ and $a_n \in \mathbb{C}$. This series expansion is also called the *Fourier expansion of $f$*. If $g$ is in fact bounded on the punctured neighbourhood $\mathcal{D}^*$, then $q = 0$ is a removable singularity of $g$ or equivalently, $g$ is holomorphically extendable to $q = 0$. This implies that the negative power terms in the Fourier expansion of $f$ vanishes. That is, the Fourier expansion of $f$ is given by

$$f(z) = \sum_{n=0}^{\infty} a_n(f)\, q_h^n.$$

In this special case, we say that $f$ is *holomorphic at* $\infty$.

To make our discussion of cusp forms complete, we have to define the so-called *cusp* of a congruence subgroup $\Gamma$.

**Definition 4.25.** Let $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. The set of $\Gamma$-equivalence classes of points in $\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}$ is called the **cusp of** $\Gamma$.

Let $\alpha \in \mathrm{SL}_2(\mathbb{Z})$. It can be shown that if $f$ is a weakly modular form of weight $k$ with respect to $\Gamma$, then $f[\alpha]_k$ is a weakly modular form of weight $k$ with respect to $\alpha^{-1}\Gamma\alpha$ (for this, one need to show that $\alpha^{-1}\Gamma\alpha$ is a congruence subgroup). Consequently, it makes sense to say that $f[\alpha]_k$ is holomorphic at $\infty$ as we did for $f$. Moreover, we now know that it makes sense to talk about the Fourier expansion of $f[\alpha]_k$.

**Definition 4.26.** Let $k \in \mathbb{Z}$, and let $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. A function $f : \mathcal{H} \to \mathbb{C}$ is said to be a **modular form of weight $k$ with respect to** $\Gamma$ if

(1). $f$ is holomorphic,

(2). $f[\gamma]_k = f$ for all $\gamma \in \Gamma$,

(3). $f[\alpha]_k$ is holomorphic at $\infty$ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

If moreover we have that $a_0(f[\alpha]_k) = 0$ for all $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ in the Fourier expansion of $f[\alpha]_k$, we say that $f$ is a **cusp form of weight $k$ with respect to** $\Gamma$.

Let $f$ be a modular form of weight $k$ with respect to $\Gamma$, we now make the subtle observation that the holomorphy of $f[\alpha]_k$ at $\infty$ is in fact related to the holomorphy at the cusps of $\Gamma$. For any cusp $s$, we will write $s = \alpha(\infty)$ for some $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ (this notation can actually be justified so it is not that we *will* write but actually we *can* write). We then consider the weight-$k$ operator $[\alpha]_k$ which is given by

$$f[\alpha]_k(z) = j(\alpha, z)^{-k} f(\alpha(z)).$$

Observe that the factor of automorphy $j(\alpha, z)$ cannot be 0 nor $\infty$ anywhere. So as $z \to i\infty$, we see that $f[\alpha]_k(z) \to \infty$ on the left-hand side, and $f(\alpha(z)) \to f(s)$ on the right-hand side. So the condition (3) in the definition above is equivalent to saying that $f$ is holomorphic at all cusps.

### 4.5.2 Modularity theorem and its consequences

We are now finally in a position to state the modularity theorem. We give a rather simplified version (based on Silverman-Tate [26]) suitable with the exposure we have given so far. For the more complete version, we invite the readers to see Chapter 8 and 9 of Diamond-Shurman [5]. For the brave readers seeking the original research papers, see [3], [30] and [31].

**Theorem 4.11** (Modularity theorem, version 1)**.** *Let $E/\mathbb{Q}$ be an elliptic curve with $L$-function*

$L(E,s) = \sum a_n/n^s$ . *Define the function* $f : \mathcal{H} \to \mathbb{C}$ *by its Fourier expansion*

$$f_E(z) = \sum_{n=1}^{\infty} a_n q^n,$$

*where* $q = q(z) = e^{2\pi i z}$ *and the* $a_n$ *are the same coefficients in* $L(E,s)$*. Then there exists an integer* $N_E$*, called the conductor of* $E$*, so that* $f_E(z)$ *is a cusp form of weight* $2$ *with respect to* $\Gamma_0(N_E)$*.*

**Definition 4.27.** We say that an elliptic curve $E/\mathbb{Q}$ is **modular** if the $L$-function of $E$ has the property as described in Theorem 4.11.

**Theorem 4.12** (Modularity theorem, version 2)**.** *Every elliptic curve* $E/\mathbb{Q}$ *is modular.*

As mentioned before, the modularity theorem has the following powerful consequences.

**Theorem 4.13.** *Let* $E/\mathbb{Q}$ *be an elliptic curve and let* $L(E,s)$ *be its* $L$*-function. Then the function*

$$\Lambda(E,s) = N_E^{s/2}\,(2\pi)^{-s}\,\Gamma(s)\,L(E,s),$$

*where* $N_E$ *is the conductor of* $E$ *(as in Theorem 4.11) and* $\Gamma$ *is the gamma function, satisfies the functional equation*

$$\Lambda(E,s) = \pm\Lambda(E, 2-s)$$

*for all* $s \in \mathbb{C}$*. Furthermore,* $L(E,s)$ *has an analytic continuation to all of* $\mathbb{C}$*.*

For a reference of this theorem, see Theorem 8.8.3 and Section 5.10 of Diamond-Shurman [5]. In particular, see Theorem 5.10.2 in that section.

Let $E/\mathbb{Q}$ be an elliptic curve. We know by the Mordell-Weil theorem that $E(\mathbb{Q}) \cong E(\mathbb{Q})_{\text{tors}} \oplus \mathbb{Z}^r$ for some positive integer $r$. Mazur's torsion theorem characterizes the torsion points and the Nagell-Lutz theorem gives an effective method to compute them. However, we have yet to discuss the rank of $E(\mathbb{Q})$. This is because of the fact that the rank is still an unsolved problem. For a start, we do not know how big can the rank be i.e. we do not have an upper bound at all.

**Conjecture 1.** *There exist elliptic curves* $E/\mathbb{Q}$ *such that its Mordell-Weil group* $E(\mathbb{Q})$ *has arbitrarily large rank.*

Here are some evidences of why some mathematicians believe this conjecture is true.

**Example** (Nagao-Kouya [14])**.** The elliptic curve $E_{21}/\mathbb{Q}$ given by the minimal model

$$\begin{aligned} y^2 + xy + y = x^3 + x^2 &- 215843772422443922015169952702159835x \\ &- 19474361277787151947255961435459054151501792241320535 \end{aligned}$$

has rank $\geqslant 21$. That is, $E_{21}$ has at least 21 *independent* rational points of infinite order.

**Example** (Elkies [16])**.** The elliptic curve $E_{28}/\mathbb{Q}$ given by the Weierstrass equation

$$\begin{aligned} &y^2 + xy + y = x^3 - x^2 - 20067762415575526585033208209338542750930230312178956502x \\ &\quad + 3448161179503055646703298569039072037485594435931918036126600829629193944873224342 9 \end{aligned}$$

has trivial torsion and rank $\geqslant 28$. That is, $E_{28}$ has only 1 finite order rational point, and at least 28 *independent* rational points of infinite order.

Now, this is where the modularity theorem comes into the bigger picture. By Theorem 4.13, we now know that it is sensible to talk about the behavior of (the continued) $L(E,s)$ at $s = 1$. This is very important as there is a conjecture which relies on this fact that says that the order of vanishing of $L(E,s)$ at $s = 1$ is exactly the rank of $E(\mathbb{Q})$. This is the famous conjecture of

Birch and Swinnerton-Dyer which is one of the seven Millenium Problems that pays \$1,000,000 to anyone that solves it (only one has been proven so far — the Poincaré Conjecture).

**Conjecture 2** (Birch and Swinnerton-Dyer Conjecture)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Then the order of vanishing of $L(E,s)$ at $s=1$ is the rank of the Mordell-Weil group $E(\mathbb{Q})$. That is,*

$$L(E,s) = (s-1)^r g(s),$$

*for some function $g$ such that $g(1) \neq 0, \infty$ implies that $E(\mathbb{Q})$ has rank $r$.*

The most remarkable thing about the Birch and Swinnerton-Dyer conjecture is that it relates something analytic (the order of vanishing) to something completely algebraic (the free rank of the Mordell-Weil group), whereas the connection is far from obvious before this conjecture was made. Coates and Wiles proved in [4] that $L(E,1) = 0$ for a special class of elliptic curves $E/\mathbb{Q}$ (called elliptic curves with *complex multiplication*) whenever their Mordell-Weil group $E(\mathbb{Q})$ is infinite. This is one of the many evidences that the conjecture might indeed be true.

For the study of elliptic curves, one important corollary of the conjecture is that it implies that $E(\mathbb{Q})$ is infinite if and only if $L(E,1) = 0$. That is, $E(\mathbb{Q})$ is infinite whenever its rank is $> 0$. Moreover, if the Birch and Swinnerton-Dyer conjecture is proven to be true, we would finally have an algorithm to compute all the rational points on $E/\mathbb{Q}$ [5]. So solving this conjecture will really be a huge milestone for the problem of rational points on elliptic curves, and hence, on nonsingular cubic curves.

# 5 Beyond the Cubics

When introducing the notion of birational equivalence, we mentioned that it is a good candidate for classifying curves up to some sort of isomorphism in algebraic geometry. This isomorphism is called *birational isomorphism.* Based on our discussions throughout the paper, we see that the degree of a curve is not a good criteria for classification in this sense since it is not an invariant. For example, we have proved that the unit circle is birationally equivalent to a line, where the former has degree 2 and the latter has degree 1. However, we can classify curves by defining an invariant which is dependent on the degree. This invariant is known as the *genus* of a curve.

**Definition 5.1.** Let $C$ be a nonsingular curve in $\mathbb{P}^2(\mathbb{C})$ with degree $d$. Then the **genus** of $C$ is defined by

$$g = g(C) = \frac{(d-1)(d-2)}{2},$$

which is a non-negative integer. Of course, this definition has a natural extension by replacing $\mathbb{C}$ with a general field $K$.

The genus has many intepretations. A nonsingular curve $C$ defined over $\mathbb{Q}$ can be viewed as a curve defined over $\mathbb{C}$ since $\mathbb{Q} \subseteq \mathbb{C}$. The graph of the curve $C$ over $\mathbb{C}$ defines a compact one dimensional complex manifold (i.e. a Riemann surface). The same graph can be viewed as a two dimensional compact, orientable surface over $\mathbb{R}$. So when viewed in this way, the genus can be interpreted as the number of *holes* or *handles* of the surface defined by $C$. We shall not prove that the genus is really an invariant, but for the curious readers, we suggest reading Section II.5 of Silverman [24]. This section formulates the Riemann-Roch theorem which captures the existence and uniqueness of the genus, and implies the invariance of the genus.

**Example.** All the curves below are defined over $\mathbb{Q}$.

(i). The line $L : y = mx + c$ is a nonsingular curve with degree 1. Its genus is computed to be $g(L) = 0$.

(ii). The unit circle $C : x^2 + y^2 = 1$ is a nonsingular curve with degree 2. Its genus is computed to be $g(C) = 0$. Since $L$ and $C$ are birationally equivalent, this is one verification that indeed the genus is an invariant.

(iii). Any nonsingular (i.e. nondegenerate) conic $C$ has degree 2 and so has genus $g(C) = 0$.

(iv). An elliptic curve $E$ (which is nonsingular by definition) has degree 3 and so has genus

$$g(E) = \frac{(3-1)(3-2)}{2} = 1.$$

Viewed as a curve defined over $\mathbb{C}$, i.e., as a compact Riemann surface, an elliptic curve thus corresponds to a torus.

## 5.1 Faltings' theorem

We have proved that one can get infinitely many rational points from one rational point on a nonsingular conic. We have also showed examples of an elliptic curve having infinitely many rational points (when there is a point of infinite order). This tells us that curves of genus 0 and 1 can have infinitely many distinct rational points on them. How about for curves of genus 2 and higher? Mordell conjectured in 1922 that this is not possible. It was not until 61 years later that someone managed to prove this conjecture. This was the work of Gerd Faltings for which he won a Fields Medal for it.

**Theorem 5.1** (Faltings, 1983)**.** *Let $C$ be a nonsingular curve defined over $\mathbb{Q}$ of genus $g > 1$. Then $C$ has only a finite number of rational points.*

This gives a hint towards super hard problems like Fermat's last theorem.

**Example.** It is a well-known fact that the Fermat equation $X^3 + Y^3 = Z^3$ has no nontrivial integer solutions so let us look at the case where the exponent is $> 3$. The projective curve

$$\widehat{C} : F(X, Y, Z) = X^n + Y^n - Z^n = 0$$

defined by the Fermat equation, where $n > 3$, is the projective closure of the affine curve

$$C : x^n + y^n = 1.$$

Accordingly, the (primitive) integer solutions to the Fermat equation corresponds bijectively to the rational points on the Fermat curve. If we compute the tangent vector of $F$, we see that

$$\nabla F = \begin{pmatrix} nX^{n-1} \\ nY^{n-1} \\ nZ^{n-1} \end{pmatrix},$$

which vanishes only at $(X, Y, Z) = \mathbf{0}$. But by definition, $\mathbf{0} \notin \mathbb{P}^2$ where $\widehat{C}$ is defined on and so $\widehat{C}$ is a nonsingular curve. So it has a genus computed to be

$$g = \frac{(n-1)(n-2)}{2} > 1,$$

where the bound is true for all $n > 3$. Accordingly, Faltings' theorem can be applied and so we conclude that $C$ has only finitely many rational points. The correspondence between solutions thus implies that the Fermat equation $X^n + Y^n = Z^n$ has only finitely many integer solutions for $n > 3$.

## 5.2 Fermat's last theorem

Let us talk about a bit of history in proving Fermat's last theorem. Somewhere around 1637, Fermat claimed that he had a proof that the Fermat equation

$$X^n + Y^n = Z^n, \tag{5.1}$$

has no solution $(x, y, z) \in \mathbb{Z}^3$ with $xyz \neq 0$ for $n \geqslant 3$; although no evidence of the proof had actually been found. Observe that since $n \geqslant 3$, it is enough to consider only two cases for this problem: when $n = 4$ or when $n = p$ an odd prime. This is because $n$ is divisible by either 4 or $p$ (or both). To see this more clearly, suppose $(x, y, z) \in \mathbb{Z}^3$ is a solution to (5.1). If $n = 4\ell$ for some $\ell \in \mathbb{Z}$, then we have that $(x^\ell, y^\ell, z^\ell)$ is a solution to the Fermat equation of exponent 4. If $n = pk$ for some $k \in \mathbb{Z}$, then we that $(x^k, y^k, z^k)$ is a solution to the Fermat equation of exponent $p$. So proving that (5.1) has no nontrivial solutions for the case $n = 4$ or $n = p$ an odd prime is sufficient due to the contrapositive statement. The case when $n = 4$ is settled by Fermat himself and can be considered the simplest case of Fermat's last theorem (yes, not $n = 3$). This was done by proving the following stronger result using Fermat's method of infinite descent.

**Proposition 5.1.** *There are no integer solutions $(x, y, z)$ with $xyz \neq 0$ such that $x^4 + y^4 = z^2$.*

So the problem of Fermat's last theorem becomes a problem of showing that there is no nontrivial integer solution to the equation $X^p + Y^p = Z^p$ for all $p \geqslant 3$ an odd prime.

In the year 1770, the case with $p = 3$ was solved by Euler which also utilized the method of infinite descent. One of the major steps in this case was to find cubes of the form $p^2 + 3q^2$. Euler cleverly showed that given any two integers $a, b \in \mathbb{Z}$, defining

$$p = a^3 - 9ab^2, \quad q = 3(a^2b - b^3)$$

ensures that $p^2 + 3q^2 = (a^2 + 3b^2)^3$ which is a perfect cube. The problem comes when he tried to prove the converse, concluding that if $p^2 + 3q^2$ is a cube, then there must exist integers $a, b \in \mathbb{Z}$ such that $p^2 + 3q^2 = (a^2 + 3b^2)^3$. To do so, he introduced the ingenious idea of extending $\mathbb{Z}$ to include the algebraic numbers $x + y\sqrt{-3}$ where $x, y \in \mathbb{Z}$ (today, this extended integers should be familiar as $\mathbb{Z}[\sqrt{-3}]$ and has been fairly well-studied). Using this new idea, he can then for example write $p^2 + 3q^2 = (p + q\sqrt{-3})(p - q\sqrt{-3})$ and work from there. His initial attempt contains serious logical gaps such as implicitly assuming that the extended integers has unique factorization (which coincidentally it has) just like $\mathbb{Z}$. Nevertheless, Euler eventually published a complete and correct proof using techniques he developed in his work on the sums of two squares.

The next case $p = 5$ was initially "solved" by Dirichlet in September 1825, but it turns out that this was an incomplete proof and it only takes care of what is called the *first case* of $p = 5$. Dirichlet's approach to the first case is basically to try mimicking Euler's proof in the case $p = 3$, ending up with needing to prove a similar looking claim: if $p^2 - 5q^2$ is a fifth power, then there must exist integers $a, b \in \mathbb{Z}$ such that $p + q\sqrt{5} = (a + b\sqrt{5})^5$. However, it turns out that directly copying Euler's method was not enough and some additional conditions were required. The complete proof which covers both the first case and the *second case* of $p = 5$ was eventually given by Legendre (independently of Dirichlet) in September 1825. The case $p = 7$ was harder; and this is reflected by the fact that in 1832, Dirichlet managed to proved the case $p = 14$ (not a prime, but we want to be consistent in notation) but not $p = 7$. The case $p = 7$ ended up being proved by Lamé in 1839 using completely new techniques (which explains why Dirichlet failed).

In 1847, Lamé announced at the Paris Academy that he had found a proof to solve the general case of Fermat's last theorem. The idea is again based on infinite descent. He introduced the usage of the complex $p$-th roots of unity $\zeta = e^{2\pi i/p}$ which generates all $p$ distinct complex solutions $1, \zeta, \ldots, \zeta^{p-1}$ to the polynomial equation $x^p - 1 = 0$. Such a polynomial thus have a factorization into linear factors

$$x^p - 1 = \prod_{k=1}^{p} (x - \zeta^{k-1}).$$

If we put $x = -X/Y$, multiply through by $(-Y)^p$, and make use of the condition that $p$ is odd, we end up with the factorization

$$Z^p = X^p + Y^p = \prod_{k=1}^{p} (X + \zeta^{k-1}Y) = (X + Y)(X + \zeta Y) \cdots (X + \zeta^{p-1}Y),$$

which is the equation that forms the basis of Lamé's proof. The idea then is to show that each of the factors $(X + \zeta^{k-1}Y)$ are *relatively prime* to each other which implies that these factors must be a $p$-th power, leading to an infinite descent. The problem in this proof, however, is that it requires unique factorization i.e. it needs the factors appearing above to be *prime* factors — which is true if we are in $\mathbb{Z}$ but unknown (at that time) in this extended $\mathbb{Z}$ plus some complex numbers setting. Liouville, whom Lamé himself acknowledged to be the one suggesting the entire idea, suspects that unique factorization breaks in general (which turns out to be true), making the argument fail.

It was finally Kummer, in 1850, who managed to prove the first general case of Fermat's last theorem — the case for which $p \geqslant 3$ is a *regular prime*. Assuming (very) little algebraic number theory, we define briefly what this mysterious prime is (for a proper treatment, see [27] or any algebraic number theory textbook). Let $p$ be an odd prime, let $\zeta_p$ be a $p$-th root of unity and consider the $p$-th cyclotomic field

$$\mathbb{Q}(\zeta_p) = \left\{a_0 + a_1\zeta_p + \cdots + a_{p-1}\,\zeta_p^{p-1} : a_i \in \mathbb{Q}\right\}.$$

This is a number field as it is a finite field extension of $\mathbb{Q}$. In fact, $\mathbb{Q}(\zeta_p)$ is a Galois extension of $\mathbb{Q}$. We can then define a so-called *class group* $\mathcal{H}_p$, which is a quotient group, associated to the ring of integers $\mathcal{O}$ of $\mathbb{Q}(\zeta_p)$. Further define the *class-number* $h_p$ to be the order of this class group $\mathcal{H}_p$. We then say that $p$ is *regular* if $h_p$ is not divisible by $p$. It is known that there are infinitely many irregular (not regular) primes, the smallest being 37. However, it is not yet known whether there are infinitely many regular primes but this is a conjecture that mathematicians believe to be true. The primes $3, 5, 7$ are all regular primes so the result of Kummer agrees with what we knew from the work of Euler, Dirichlet, Legendre and Lamé.

The breakthrough that ultimately lead towards the correct proof of Fermat's last theorem is the discovery of a certain elliptic curve by Frey in 1984. This elliptic curve, now known as the *Frey curve*, relies on assuming the falsehood of Fermat's last theorem and is constructed as follows: suppose that $(a, b, c) \in \mathbb{Z}^3$ with $abc \neq 0$ is a solution to the Fermat equation of exponent $p \geqslant 5$, then the Frey curve is defined to be the rational elliptic curve

$$E_{a,b,c} : y^2 = x(x - a^p)(x + b^p).$$

Frey realized that this elliptic curve obeys some interesting properties: one of which is that it is *semistable* (i.e. if $E_{a,b,c}$ has bad reduction at a prime $p$, then the reduction is multiplicative) with conductor

$$N_{E_{a,b,c}} = \prod_{\ell \mid abc} \ell,$$

where the product is over all distinct primes $\ell \mid abc$. Most importantly, Frey suggested that the Frey curve $E_{a,b,c}$ cannot be modular. This conjecture was made precise as statements about modular forms and *Galois representations* by Serre and was then known as the *$\varepsilon$-conjecture*. Ribet eventually proved this conjecture in 1986 using his *level-lowering theorem*. The work by Serre and Ribet forms an important bridge between the Taniyama-Shimura-Weil conjecture (now Theorem 4.12) and Fermat's last theorem. How? Well, the existence of the non-modular Frey curve requires Fermat's last theorem to be false. So, if the Taniyama-Shimura-Weil conjecture is true, then no such curve should exist and so by the contrapositive, Fermat's last theorem is true. Therefore, proving Fermat's last theorem reduces to proving the Taniyama-Shimura-Weil conjecture. Wiles, with a little help from his former student Taylor, finally proved the Taniyama-Shimura-Weil conjecture for semistable rational elliptic curves in the year of 1995. Despite not proving the full conjecture, this was sufficient to imply the impossible existence of the non-modular semistable Frey curve, thus proving Fermat's last theorem. Building on Wiles' work, the full Taniyama-Shimura-Weil conjecture (as in the statement of Theorem 4.12) was eventually proved by Breuil, Conrad, Diamond and Taylor.

## Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Payman Kassaei for his helpful suggestions and invaluable guidance throughout this project. Working with him has been a really great pleasure. Thanks to my personal tutor Dr. Dmitri Panov for his wise advices and endless support throughout my undergraduate life at King's. I would also like to thank him and Prof. Fred Diamond for their helpful answers to my questions regarding cubic curves and the BSD conjecture. I am also grateful to Prof. Fred Diamond for his help in explaining a section in his book to me.

Thanks to Dr. James Newton for his patience answering the many, many questions I had when I first indulged in elementary number theory. I feel like there has not been a week in my second year of undergraduate without a visit to his office discussing some mathematics. I would also like to express my gratitude to my mentor, Dr. Kwok-Wing Tsoi, whose enthusiasm inspired me to study number theory. I would most probably be doing theoretical physics now if not for him.

Thanks to my friend and roommate Anas Razak for all our fun mathematical discussions, some of which turns out to be important for this project. Thanks also to my other flatmates and friends for keeping me sane throughout the Covid-19 pandemic. Last but not least, I would like to thank my parents for raising me into the person I am today, for the freedom of choice in whatever I want to pursue, and for always believing in me.

Finally, I am grateful for the financial support given by Majlis Amanah Rakyat (MARA) which provided funding for my three years of BSc studies at King's.

# References

[1] Wayne Aitken and Franz Lemmermeyer. Counterexamples to the Hasse Principle: An Elementary Introduction. *The American Mathematical Monthly*, 118(7):610–628, 2011. doi: 10.4169/amer.math.monthly.118.07.610.

[2] Michael Artin. *Algebra*. Pearson, 2004.

[3] Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor. On the modularity of elliptic curves over $\mathbb{Q}$: Wild 3-adic exercises. *Journal of the American Mathematical Society*, 14(4):843–939, 2001. doi: 10.1090/S0894-0347-01-00370-8.

[4] John Coates and Andrew Wiles. On the conjecture of Birch and Swinnerton-Dyer. *Inventiones mathematicae*, 39:223–251, 1977. doi: 10.1007/BF01402975.

[5] Fred Diamond and Jerry Shurman. *A First Course in Modular Forms*. Springer, 2005.

[6] Harold M. Edwards. *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*. Springer, 1977.

[7] Larry J. Gerstein. *Basic Quadratic Forms*. American Mathematical Society, 2008.

[8] Fernando Q. Gouvêa. *p-adic Numbers: An Introduction*. Springer, 2020.

[9] Kenneth Ireland and Michael Rosen. *A Classical Introduction to Number Theory*. Springer-Verlag New York, 1990.

[10] Neal Koblitz. *p-adic Numbers, p-adic Analysis and Zeta-Functions*. Springer, 1984.

[11] Álvaro Lozano-Robledo. *Elliptic Curves, Modular Forms, and Their L-Functions*. American Mathematical Society, 2011.

[12] Álvaro Lozano-Robledo. *Number Theory and Geometry: An Introduction to Arithmetic Geometry*. American Mathematical Society, 2019.

[13] Louis Joel Mordell. *Diophantine Equations*. Academic Press, London, 1969.

[14] Koh-ichi Nagao and Tomonori Kouya. An example of elliptic curve over $\mathbb{Q}$ with Rank $> 21$. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 70(4):104–105, 1994. doi: 10.3792/pjaa.70.104.

[15] Ivan Niven, Herbert S. Zuckerman, and Hugh Montgomery. *An Introduction to the Theory of Numbers*. Wiley, 1991.

[16] Noam D. Elkies. $\mathbb{Z}^{28}$ in $E(\mathbb{Q})$. *Number Theory Listserver*, May 2006.

[17] Miles Reid. *Undergraduate Algebraic Geometry*. Cambridge University Press, 1988.

[18] Paulo Ribenboim. *Fermat's Last Theorem for Amateurs*. Springer, 1999.

[19] Kenneth H. Rosen. *Elementary Number Theory and Its Applications*. Addison-Wesley, 2011.

[20] René Schoof. Counting points on elliptic curves over finite fields. *Journal de Théorie des Nombres de Bordeaux*, 7(1):219–254, 1995.

[21] Ernst S. Selmer. The Diophantine equation $ax^3 + by^3 + cz^3 = 0$. *Acta Mathematica*, 85:203 – 362, 1951. doi: 10.1007/BF02395746.

[22] Jean-Pierre Serre. *A Course in Arithmetic*. Springer, 1973.

[23] Igor R. Shafarevich. *Basic Algebraic Geometry 1: Varieties in Projective Space, Third Edition.* Springer, 2010.

[24] Joseph H. Silverman. *The Arithmetic of Elliptic Curves.* Springer, 2009.

[25] Joseph H. Silverman. *A Friendly Introduction to Number Theory.* Pearson, 2012.

[26] Joseph H. Silverman and John T. Tate. *Rational Points on Elliptic Curves, Second Edition.* Springer, 2015.

[27] Ian Stewart and David Tall. *Algebraic Number Theory and Fermat's Last Theorem, Third Edition.* A K Peters, 2002.

[28] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.2)*, 2020. `https://www.sagemath.org`.

[29] Lawrence C. Washington. *Elliptic Curves: Number Theory and Cryptography.* CRC Press, 2008.

[30] Andrew Wiles. Modular elliptic curves and Fermat's Last Theorem. *Annals of Mathematics*, 141(3):443–551, 1995. doi: 10.2307/2118559.

[31] Andrew Wiles and Richard Taylor. Ring-theoretic properties of certain Hecke algebras. *Annals of Mathematics*, 141(3):553–572, 1995. doi: 10.2307/2118560.