ulm university universität ulm

Diophantine Equations

Summer semester 2016

Universität Ulm

Stefan Wewers

Notes taken by **Tudor Micu**

Institut für Reine Mathematik

Version: **October 4, 2016**

Chapter 1

Introduction

There is no universally accepted definition of what a **Diophantine equation** is. In this course, we will assume that it is a single polynomial equation

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

where $F \in \mathbb{Z}[x_1, \ldots, x_n]$ is a polynomial in $n \geq 2$ variables x_i. Depending on the specific example at hand, we are interested in its *integral solutions* (i.e. the set of all $x = (x_i) \in \mathbb{Z}^n$ such that $F(x) = 0$) or its *rational solutions* (with $x_i \in \mathbb{Q}$). Mostly, the number n of variables will be $n = 2$ or $n = 3$, and then we will use the variable names x, y, z instead of $x_1, x_2, \ldots$.

The term *Diophantine equation* is derived from the ancient greek mathematician *Diophantus of Alexandria*, who lived during the third century AD and was the author of a series of books called *Arithmetica*. Many of these books are now lost, and the remaining ones where translated into Latin during the 17th century, and became a great source of inspiration for the mathematicians of that time. [1]

For the last 300 years, Diophantine equations have been a major and very active part of pure mathematics. To briefly explain where the fascination for them comes from we may mention the following point:

[1] From today's point of view the *Arithmetica* contains mainly *cooking-book mathematics* and does not rank among the great works of the ancient greek school of mathematics – like, for instance, Euclid's *Elements*. Even though it is has some historical importance, it has little bearing on what we mean today by the theory of Diophantine equations.

- Several parts of modern mathematics (algebra, number theory, algebraic geometry) have to a large extent been developed in order to solve Diophantine equations, and this influence continues to the present day.
- Diophantine equations present a particularly simple way to pose hard and interesting mathematical problems.

An excellent example illustrating both points above is *Fermat's Last Theorem*:

Theorem 1.1 (Fermat's Last Theorem, FLT)**.** *Let $n \geq 3$. Then the Diophantine equation*

$$x^n + y^n = z^n \tag{1.1}$$

has no solutions $x, y, z \in \mathbb{Z}$ with $xyz \neq 0$

This theorem has been formulated by Pierre de Fermat in 1638 and later became one of the most famous and notorious open problems in mathematics. Building on the work of many people, it was finally proved in 1995 by Andrew Wiles. You can read more about the fascinating and amusing story of this proof in Simon Singh's book *Fermat's Last Theorem* ([4]; see also the BBC documentary with the same name).

Remark 1.2. (i) A solution (x, y, z) of Fermat's equation (1.1) with $xyz = 0$ is called *trivial*. Obviously, the trivial solutions are of the form $(x, 0, x)$, $(0, y, y)$ or $(x, -x, 0)$ (for n odd).

(ii) A solution $(x, y, z) \in \mathbb{Z}^3$, with $z \neq 0$, of Equation (1.1) gives rise to a rational solution $\left(\frac{x}{z}, \frac{y}{z}\right)$ of the inhomogenous equation

$$x^n + y^n = 1. \tag{1.2}$$

This is a special instance of the following general phenomenon. *Integral* solutions of *homogenous* equations in n variables correspond to *rational* solutions of arbitrary equations in $n - 1$ variables.

(iii) The condition $n \geq 3$ in Theorem 1.1 is essential. For $n = 2$ the equation has infinitely many nontrivial solutions, as the following Theorem 1.4 shows.

Definition 1.3. A **pythagorean triple** (P.T.) is a triple $(x, y, z) \in \mathbb{Z}^3$ such that

- $x, y, z > 0$,
- $\gcd(x, y, z) = 1$,
- $x^2 + y^2 = z^2$.

Theorem 1.4. *Let (x, y, z) be a pythagorean triple. Then*

(a) z is odd and x and y have different parities.

(b) If we assume that x is odd then there exist $a, b \in \mathbb{N}$ with $a > b$, $\gcd(a, b) = 1$, $a \not\equiv b \ (2)$ and

$$x = a^2 - b^2,$$
$$y = 2ab,$$
$$z = a^2 + b^2.$$

Proof. Let (x, y, z) be a pythagorean triple. The two conditions $\gcd(x, y, z) = 1$ and $x^2 + y^2 = z^2$ combined imply

$$\gcd(x, y) = \gcd(y, z) = \gcd(z, x) = 1.$$

This means in particular that x and y can't both be even. But if they were both odd, then we'd have $z^2 \equiv x^2 + y^2 \equiv 2 \ (4)$, which is impossible. Therefore x and y have different parities and z is odd. This proves (a).

For the proof of (b) we rewrite the condition $x^2 + y^2 = z^2$ in the form

$$y^2 = z^2 - x^2 = (z - x)(z + x). \tag{1.3}$$

Using $2x = -(z - x) + (z + x)$ and $2z = (z - x) + (z + x)$ we see that

$$\gcd(z - x, z + x) = \gcd(2x, 2z) = 2 \cdot \gcd(x, z) = 2.$$

Therefore, there exist $u, v, w \in \mathbb{Z}$ so that $\gcd(u, v) = 1$ and

$$z + x = 2u,$$
$$z - x = 2v,$$
$$y = 2w.$$

Hence we may rewrite (1.3) as

$$w^2 = uv. \tag{1.4}$$

Since $\gcd(u, v) = 1$ we conclude from (1.4) that there exist $a, b \in \mathbb{Z}$ so that $a > b$, $\gcd(a, b) = 1$ and

$$u = a^2,$$
$$v = b^2,$$
$$w = ab.$$

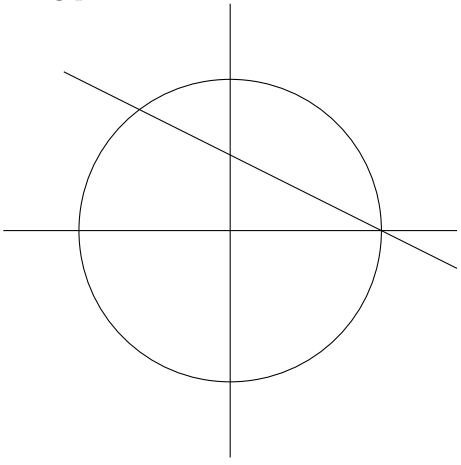
Plugging this into the equations defining u, v, w we obtain

$$x = \frac{(z + x) - (z - x)}{2} = u - v = a^2 - b^2,$$
$$y = 2ab,$$
$$z = a^2 + b^2.$$

We also see that $a \not\equiv b \pmod 2$, otherwise x and z would be even. This completes the proof of the theorem. □

The proof above use only elementary number theory; it heavily relies on the unique factorization theorem. Here is another proof with a more geometric flavour.

As we have seen in Remark 1.2 (ii), we may first look for *rational* solutions of the equation $x^2 + y^2 = 1$, i.e. for points on the unit circle with rational coordinates. For this we can use the fact that the unit circle has a *rational parametrization*. The underlying geometric construction is obvious from the following picture.



The point $(1,0)$ is clearly a rational solution, and every other rational solution can be connected to the point $(1,0)$ by a line $y = t(x-1)$ with $t \in \mathbb{Q}$. Conversely, given any such a line, its intersection with the unit circle consists of exactly two points, the point $(1,0)$ and another rational solution of $x^2+y^2=1$.

So, to find a parametrisation for the rational solutions it suffices to solve the system

$$x^2 + y^2 = 1,$$
$$y = t(x-1).$$

So we have $x^2 + t^2(x-1)^2 = 1$, which gives the equation $(1+t^2)x^2 - 2t^2x + (t^2-1) = 0$.

Its solutions are $(x,y) = (1,0)$ (which we already know is a rational solution) and

$$(x,y) = \left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right).$$

Thus, the parametrisation for the rational solutions of $x^2+y^2=1$ is:

$$x = \frac{1-t^2}{1+t^2},$$
$$y = \frac{2t}{1+t^2},$$

with $t \in \mathbb{Q}$. If we write $t = \frac{a}{b},\ a,b \in \mathbb{Z},\ ggT(a,b) = 1,$ we get:

$$x = \frac{a^2-b^2}{a^2+b^2},$$
$$y = \frac{2ab}{a^2+b^2}.$$

We obtain integral solutions of the homogenous equation $x^2+y^2=z^2$ by 'multiplying with the denominator'. If we do the carefully we obtain all pythagorean triples as follows:

- $(a^2-b^2, 2ab, a^2+b^2)$ if $a \not\equiv b\ (2)$,
- $\left(\frac{a^2-b^2}{2}, ab, \frac{a^2+b^2}{2}\right)$ if $a \equiv b \equiv 1\ (2)$.

The first case corresponds precisely to Theorem 1.4 (b).

Theorem 1.5. *The diophantine equation*

$$x^4 + y^4 = z^2$$

has no solutions $x, y, z \in \mathbb{Z}$ *with* $xyz \neq 0$

Corollary 1.6. *Fermat's Last Theorem is true for* $n = 4$

Proof. We will proceed by contradiction. Let (x, y, z) be a solution for $x^4 + y^4 = z^2$ with

(a) $x, y, z > 0$ mutually prime

(b) $x, z \equiv 1\ (2),\ y \equiv 0\ (2)$

(c) z minimal

The idea of the proof is to construct another solution (x_1, y_1, z_1) that satisfies $(a) - (c)$ for which $z_1 < z$. This will actually contradict the minimality of z and prove the theorem.

We have $(x^2)^2 + (y^2)^2 = z^2$ and hence (x^2, y^2, z) is a pythagorean triple. By Theorem 1.4 this means that we can find $a, b \in \mathbb{N}$ such that $a > b > 0$, $a \not\equiv b \pmod 2$, $\gcd(a, b) = 1$ and

$$\begin{aligned} x^2 &= a^2 - b^2, \\ y^2 &= 2ab, \\ z &= a^2 + b^2. \end{aligned}$$

Here we have assumed that x is odd and y even, which we may. But now $x^2 + b^2 = a^2$ and $a \equiv 1,\ b \equiv 0 \pmod 2$, i.e. (x, b, a) is a pythagorean triple, satisfying the conditions from Theorem 1.4. Using the theorem again we can find $c, d \in \mathbb{N}$ such that $c > d > 0$, $c \not\equiv d \pmod 2$, $\gcd(c, d) = 1$ and

$$\begin{aligned} x &= c^2 - d^2, \\ b &= 2cd, \\ a &= c^2 + d^2. \end{aligned}$$

Because $y^2 = 2ab$, $a \not\equiv b \pmod 2$ and $\gcd(a, b) = 1$ we can write

$$\begin{aligned} b &= 2w^2, \\ a &= z_1^2. \end{aligned}$$

So $w^2 = \dfrac{b}{2} = cd$ and because $\gcd(c, d) = 1$ we write

$$c = x_1^2,$$
$$d = y_1^2.$$

We remark that (x_1, y_1, z_1) is a solution for $x^4 + y^4 = z^2$:

$$z_1^2 = a = c^2 + d^2 = x_1^4 + y_1^4$$

and also that $z = a^2 + b^2 = z_1^4 + 4w^4 > z_1$.

Therefore the minimality of z is contradicted. □

This proof strategy is called "descente infinie" (infinite descent) and is based on the idea that one cannot build an infinite decreasing sequence of natural numbers.

Chapter 2

Rational solutions of quadratic equations

In this chapter we will consider *quadratic diophantine equations*, of the form

$$F(x, y) = ax^2 + bxy + cy^2 + dx + ey + f = 0,$$

with integral coefficients $a, \ldots, f \in \mathbb{Z}$ and $(a, b, c) \neq (0, 0, 0)$. We are interested in its *rational* solutions $x, y \in \mathbb{Q},\ F(x, y) = 0$.

As we will see below, it suffices to consider equations of the form

$$ax^2 + by^2 = 1, \tag{2.1}$$

with $a, b \in \mathbb{Z}$ both positive, $a, b > 0$.

Notation 2.1. We fix a quadratic polynomial $F \in \mathbb{Z}[x, y]$ as above. For any commutative ring R we note

$$X(R) := \left\{(x, y) \in R^2 \mid F(x, y) = 0\right\}.$$

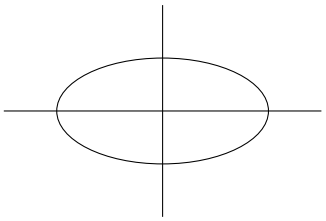
As we said before, we are mainly interested in $X(\mathbb{Q})$. To study this set we may use the fact that it is contained in $X(\mathbb{R})$, which is a *quadric*. See e.g. [6], §3.3.

Remark 2.2. A quadric is called *degenerate* if it has at most one point, or if it is the union of two lines. Otherwise, we call the quadric *nondegenerate*. We

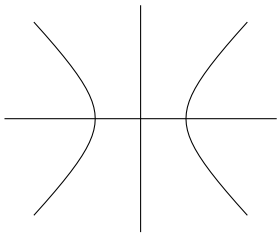
will assume from now on that the quadric defined by our quadratic Diophantine equation $F(x, y) = 0$ is nondegenerate.

There are three types of nondegenerate quadrics. You have learned in your linear algebra course that each of them can be brought, by a change of coordinate system representing a symmetry of the plane, into the following *normal form*.

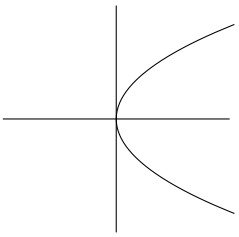
1. Ellipse (normal form): $ax^2 + by^2 = 1$, $a, b > 0$



2. Hyperbola (normal form): $ax^2 - by^2 = 1$, $a, b > 0$



3. Parabola (normal form): $ax = by^2$, $a, b > 0$



The proof of the existence of the normal form relies on the spectral theorem. It only works over the real numbers but in general not over the rationals (recall that one may have to solve a quadratic equation!). This is not acceptable for us because we are interested in the rational solutions of the equation. Fortunately, there is another way to obtain a normal form with works over the rationals (but which may not corresponds to a euclidean symmetry).

Proposition 2.3. Every quadratic diophantine equation

$$F(x, y) = 0$$

that represents a non-degenerate quadric can be brought to one of the two normal forms

$$ax^2 \pm by^2 = c,$$

or

$$ax = by^2,$$

with $a, b, c \in \mathbb{N}$, by a coordinate change over $\mathbb{Q}$.

The proof is very simple and left as an exercise. We only give an example.

Example 2.4. Consider the equation $F(x, y) = x^2 + xy + y^2 - 1 = 0$. We substitute $x = x_1 - \frac{1}{2}y$ and obtain:

$$\begin{aligned} F(x, y) &= x_1^2 - x_1 y + \frac{1}{4}y^2 + x_1 y - \frac{1}{2}y^2 + y^2 - 1 \\ &= x_1^2 + \frac{3}{4}y^2 - 1 = \frac{1}{4}(4x_1^2 + 3y^2 - 4). \end{aligned}$$

So there is a bijection between the rational solutions of the equation $x^2 + xy + y^2 = 1$ and $4x_1^2 + 3y^2 = 4$, given by

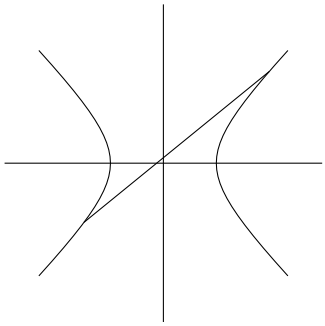
$$(x, y) \mapsto (x + \frac{1}{2}y, y).$$

Theorem 2.5. *Let $F(x, y) = 0$ be a quadratic diophantine equation that represents a non-degenerated quadric over $\mathbb{R}$. Then the following assertions are equivalent:*

(a) There is at least one rational solution.

(b) There are infinity many rational solutions.

Proof. The proof is a straightforward generalization of the parametrization of the unit circle constructed on page 4f: starting from one rational solution we obtain a parametrization of all rational solutions with respect to a parameter t.



□

Example 2.6 (the parable case). Let $F(x, y) = ax - by^2 = 0$. Because we always have the rational solution $(x, y) = (0, 0)$ we have an infinite number of rational solutions. To be more precise, we can parametrize: $x = abt^2, y = at,\ t \in \mathbb{Q}$.

It thus remains to consider the case $ax^2 \pm by^2 = c$ and see if the equation has any rational solutions. To this end, it makes sense to homogenize:

$$ax^2 \pm by^2 = cz^2$$

and look for the integral solutions $(x, y, z) \in \mathbb{Z}^3$ with $z \neq 0$ We have now arrived at the following formulation of our original question.

Problem. We are given $a, b, c \in \mathbb{Z} \setminus \{0\}$. We wish to find $(x, y, z) \in \mathbb{Z}^3$ satisfying

- $(x, y, z) \neq (0, 0, 0)$
- $gcd(x, y, z) = 1$
- $ax^2 + by^2 - cz^2 = 0$

Remark 2.7. Continue with the notation introduced above. Without loss of generality we may also assume:

(i) $a, b, c > 0$

(ii) $c = 1$,

(iii) a, b squarefree.

To see (i), let us write the equation first in the more symmetric form

$$ax^2 + by^2 + cz^2 = 0,$$

without making any prior assumption on the sign of the coefficients a, b, c. If a, b, c all have the same sign then obviously there is no solution $(x, y, z) \neq (0, 0, 0)$. It is therefore no restriction to assume that two out of three coefficients are positive and the third negative. After a suitable permutation of the variables x, y, z we may then assume that $a, b > 0$ and $c < 0$. But then we can rewrite our equation in the less symmetric form

$$ax^2 + by^2 = cz^2, \tag{2.2}$$

with $a, b, c > 0$. This shows (i).

If (x, y, z) is a solution of (2.2) then $(x_1, y_1, z_1) := (x, y, cz)$ is a solution to

$$(ac)x_1^2 + (bc)y_1^2 = z_1^2.$$

We may therefore assume that $c = 1$, (ii).

To prove (iii) we write $a = a_0 d^2$ with $a_0, d \in \mathbb{Z}$ and a_0 squarefree. If (x, y, z) is a solution of

$$ax^2 + by^2 = z^2,$$

then $(x_1, y_1, z_1) := (dx, y, z)$ is a solution to

$$a_0 x_1^2 + by_1^2 = z_1^2.$$

Hence we may assume that a (and for the same reason, b) is squarefree.

Example 2.8.

- $5x^2 + 7y^2 = z^2$ has no solutions
- $5x^2 + 11y^2 = z^2$ has solution $(1, 1, 4)$

Definition 2.9. Let $a, b \in \mathbb{Z},\ b > 0$.
We say that a is a **quadratic residue modulo b** and write $a\ QR\ b$ if there exists an $x \in \mathbb{Z}$ so that $a \equiv x^2\ (b)$.

Example 2.10. $7 \equiv 2\ (5)$ but the squares in $(\mathbb{Z}/5\mathbb{Z})^\times$ are $\bar{1}$ and $\bar{4}$, so 7 is not a quadratic residue modulo 5. On the other hand, $9 \equiv 4\ (5)$ so $9\ QR\ 5$.

Theorem 2.11 (Legendre)**.** *Let $a, b \in \mathbb{N}$ be squarefree. Then the diophantine equation*

$$ax^2 + by^2 = z^2 \tag{2.3}$$

has a solution $(x, y, z) \in \mathbb{Z}^3 \setminus \{(0,0,0)\}$ *if and only if the following conditions are satisfied:*

(i) a *QR* b*,*

(ii) b *QR* a*,*

(iii) $-\dfrac{ab}{d^2}$ *QR* d*, with* $d = gcd(a, b)$*.*

Proof. For now we will only prove the direct implication. Let $(x, y, z) \neq (0, 0, 0)$ be a non-trivial solution of $ax^2+by^2 = z^2$. We may assume that $\gcd(x, y, z) = 1$. From this and from the assumption that a, b are squarefree we deduce:

- $\gcd(x, y) = \gcd(y, z) = \gcd(z, x) = 1$,
- $\gcd(x, b) = 1$.

From (2.3) we have that $ax^2 \equiv z^2$ (b) and, because $\gcd(x, b) = 1$, we deduce

$$a \equiv \left(\frac{z}{x}\right)^2 \ (b)$$

This proves (i) and, by symmetry, (ii). The proof of (iii) is left as an exercise. □

The converse implication in Theorem 2.11 is more difficult; its proof can be found on p14f. In it we will use the following theorem:

Theorem 2.12 (Fermat's two squares theorem)**.** *Let* p *be a prime number. Then the following assertions are equivalent:*

(a) There exist $x, y \in \mathbb{Z}$ *such that* $x^2 + y^2 = p$*.*

(b) -1 *QR* p*.*

(c) $p = 2$ *or* $p \equiv 1$ (4)*.*

Proof. (a)⇒(c): If $p \neq 2$ then $x \not\equiv y$ (2) and hence $x^2 + y^2 \equiv 1$ (4).

$(c) \Rightarrow (b)$: We use the fact that $(\mathbb{Z}/p\mathbb{Z})^\times$ is a cyclic group of order $p - 1$. A solution of $x^2 \equiv -1$ (p) corresponds to an element $\bar{x} \in (\mathbb{Z}/p\mathbb{Z})^\times$ of order 4. This gives $4|p - 1$ and hence $p \equiv 1$ (4).

(b)⇒ (a): Let $a \in \mathbb{Z}$ with $a^2 \equiv -1$ (p). Then $p|a^2 + 1 = (a + i)(a - i)$. We use the fact that $\mathbb{Z}[i]$ is a euclidean ring, so it's a unique factorisation domain. If p were irreducible then it would be prime and from $p|(a + i)(a - i)$ we would

have $p|a+i$ or $p|a-i$. This would mean that $p|1$, which is impossible. Therefore p is not irreducible, so p can be factorised. Because $N(p) = p^2$ this factorisation can have at most two irreducible factors, due to the multiplicativity of the norm in $\mathbb{Z}[i]$. Moreover, these two factors are both of norm p. Let $p = \pi \cdot \pi'$ be the factorization. Then $\bar{\pi} \cdot p = p \cdot \pi'$, so $\pi' = \bar{\pi}$. Now, if $\pi = x + iy$ then we have

$$p = \pi \cdot \bar{\pi} = (x + iy)(x - iy) = x^2 + y^2.$$

□

Remark 2.13. The equivalence (b)⇔(c) is equivalent to the *first supplement to the law of quadratic reciprocity*, i.e. the formula

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

See e.g. [1], §5.1.

Corollary 2.14. *Let $b \in \mathbb{N}$ so that -1 QR b. Then there are $x, y \in \mathbb{Z}$ so that $b = x^2 + y^2$*

Remark 2.15. The converse is not true. For instance, $b = 9 = 3^2 + 0^2$, but -1 is not a quadratic residue modulo 9.

Proof. Let $b = \prod_{i=1}^{r} p_i^{e_i}$ the prime factor decomposition. Because $p_i|b$ we deduce -1 QR p_i for all i. From theorem 2.12 for all i we have $x_i, y_i \in \mathbb{Z}$ so that

$$p_i = x_i^2 + y_i^2 = |x_i + iy_i|^2.$$

Then

$$b = \prod_{i=1}^{r} (|x_i + iy_i|^2)^{e_i} = \left| \prod_{i=1}^{r} (x_i + iy_i)^{e_i} \right|^2 = x^2 + y^2,$$

because the norm of an element in $\mathbb{Z}[i]$ is always a sum of squares. □

We can now finish the proof of Legendre's theorem (Theorem 2.11).

Proof. We have already proven one implication. To prove the remaining one, let $a, b \in \mathbb{N}$ be square-free and assume that the following three conditions hold:

(i) a QR b,

(ii) b QR a,

(iii) $-\frac{ab}{d^2}$ QR d, $d = gcd(a, b)$.

We have to show that the equation

$$ax^2 + by^2 = z^2 \tag{2.4}$$

has a solution $(x, y, z) \in \mathbb{Z}^3 \backslash \{(0, 0, 0)\}$.

One idea used in the proof is the *norm trick*. Assume for simplicity that $b > 1$. Since we assume b to be square-free, this means that $\sqrt{b}$ is irrational. Then $\mathbb{Z}[\sqrt{b}]$ is a quadratic ring extension of $\mathbb{Z}$ (just like $\mathbb{Z}[i]$), and we can define the norm map as follows:

$$\begin{aligned} N : \mathbb{Z}[\sqrt{b}] &\to \mathbb{Z} \\ u + \sqrt{b}v &\mapsto u^2 - bv^2 \end{aligned}$$

As one can easily check, this map is *multiplicative*, i.e. $N(\alpha\beta) = N(\alpha)N(\beta)$. This shows that the product of two numbers of the form $Z^2 - bY^2$ is again of this form. More explictly, let x_1 and x_2 be norms:

$$\begin{aligned} x_1 = z_1^2 - by_1^2 = (z_1 - \sqrt{b}y_1)(z_1 + \sqrt{b}y_1) \\ x_2 = z_2^2 - by_2^2 = (z_2 - \sqrt{b}y_2)(z_2 + \sqrt{b}y_2) \end{aligned}$$

Then $x_1x_2 = z_3^2 - by_3^2$, where

$$\begin{aligned} y_3 &= z_1y_2 + z_2y_1 \\ z_3 &= z_1z_2 + by_1y_2 \end{aligned}$$

In fact, these identities can be checked by a direct calculation, and they work also for $b = 1$. We used the norm map only to *motivate* them.

The second idea of the proof is to use descent. More precisely we use induction with respect to $\max(a, b)$.

If $a = 1$ or $b = 1$, then $(x, y, z) = (1, 0, 1)$ or $(x, y, z) = (0, 1, 1)$, respectively, is a solution to (2.4). We may therefore assume $a, b > 1$ and, by symmetry, $a \geq b$.

We treat the case $a = b$ separately. From (iii) we get -1 QR a and from Corollary 2.14 $a = r^2 + s^2$. Then (r, s, a) is a solution for (2.4).

We may now assume that $a > b > 1$. By (ii) there is a $u \in \mathbb{Z}$ so that

$$u^2 \equiv b \pmod{a}.$$

We can choose u such that $|u| \leq a/2$. Then, there exists $T \in \mathbb{Z}$ so that $u^2 - b = aT$. Since b is assumed to be square-free, $A \neq 0$. Write $T = Am^2$ with $A \in \mathbb{Z}$ squarefree and $m \in \mathbb{N}$. We will prove that $0 < A < a$. Indeed, from

$$0 \leq u^2 = aAm^2 + b < a(Am^2 + 1)$$

we deduce that $A > 0$. Thus $u^2 - b = aAm^2 < u^2 < a^2/4$, and hence

$$0 < A \leq Am^2 \leq \frac{a}{4} < a.$$

Let us consider the equation

$$Ax^2 + by = z^2. \tag{2.5}$$

Since $A < a$, we can apply the induction hypothesis, which says that Legendre's theorem holds for (2.5).

Claim: Equation (2.5) satisfies (i)-(iii).

The proof of this claim is not difficult but a bit tedious. We omit it and instead refer to [1], §17, proof of Proposition 17.3.2.

Applying Legendre's theorem, we see that there exists a solution $(x_1, y_1, z_1) \in \mathbb{Z}^3 \setminus \{(0,0,0)\}$ for (2.5). Then

$$Ax_1^2 = z_1^2 - by_1^2,$$
$$aAm^2 = u^2 - b.$$

Multiplying both equations and using the norm trick, we conclude that

$$a(Amx_1)^2 = z_2^2 - by_2^2,$$

for some $y_2, z_2 \in \mathbb{Z}$. This means that (Amx_1, y_2, z_2) is a solution to (2.4). The theorem is proved. □

Remark 2.16. There is an equivalent formulation of the theorem in which all three variables x, y, z play symmetric roles. Let $a, b, c \in \mathbb{Z} \setminus \{0\}$ be squarefree, mutually prime, and not all of the same sign. Then the equation

$$ax^2 + by^2 + cz^2 = 0$$

has a solution $(x, y, z) \in \mathbb{Z}^3 \setminus \{(0,0,0)\}$ if and only if the following are true:

(i) $-ab\ QR\ c$,

(ii) $-bc\ QR\ a$,

(iii) $-ca\ QR\ b$.

Corollary 2.17. *Let $a, b \in \mathbb{N}$ squarefree. Then the equation*

$$ax^2 + by^2 = z^2$$

has a non-trivial solution if and only if for any prime p and $m \in \mathbb{N}$ the congruence equation

$$ax^2 + by^2 \equiv z^2 \pmod{p^m}$$

has a solution $(x, y, z) \in \mathbb{Z}^3$ with $(x, y, z) \not\equiv (0, 0, 0) \pmod{p}$.

Proof. The direct implication is clear. To prove the converse we may use Legendre's theorem. This means that it suffices to prove that Conditions (i)-(iii) in Theorem 2.11 hold.

We will only prove (ii); the proof of the other two conditions is similar. If $a = 1$ then $b\ QR\ a$ and (ii) holds. We may therefore assume that $a > 1$. Let $p|a$ be a prime divisor. We choose $m = 2$. By our assumption there exists $(x, y, z) \not\equiv (0, 0, 0)\ (p)$ with

$$ax^2 + by^2 \equiv z^2 \pmod{p^2}.$$

This gives $by^2 \equiv z^2 \pmod{p}$. If $p|y$ then $z \equiv 0 \pmod{p}$ and hence $p^2 \mid ax^2$. But a is assumed to be square-free, so $p|x$ as well. This contradicts the assumption $(x, y, z) \not\equiv (0, 0, 0) \pmod{p}$. Therefore we have $\gcd(p, y) = 1$ and thus

$$b \equiv \left(\frac{z}{y}\right)^2 \pmod{p}.$$

In other words, b is a quadratic residue modulo p.

Let $a = p_1 \dots p_r$ be the prime factor decomposition of a. The argument from above shows that $b \equiv x_i^2 \pmod{p_i}$, for all i. By the Chinese Remainder Theorem there exists x with $x \equiv x_i \pmod{p_i}$ for all i. Then $b \equiv x^2 \pmod{p_i}$ for all i. Using again the Chinese Remainder Theorem we conclude that $b \equiv x^2 \pmod{a}$. This means that b is a quadratic residue modulo a, and (ii) is proved. □

Chapter 3

The p-adic numbers

In this chapter we will introduce a new kind of numbers, called the *p-adic numbers*. Here p is a fixed prime number, and the set $\mathbb{Q}_p$ of all p-adic numbers is a field which contains the field $\mathbb{Q}$ of rational numbers. Thus every rational number can be considered as a p-adic number, just as every rational number can be considered as a real number. In fact, there is a surprising similarity in the way the real numbers are defined as limits of rational numbers and the parallel construction of p-adic numbers.

Here is the definition of the p-adic numbers in a nutshell. Given a prime number p, every nonzero rational number $x \in \mathbb{Q}^\times$ can be written in a unique way in the form

$$x = p^n \frac{a}{b},$$

with $a, b, n \in \mathbb{Z}$, $p \nmid a, b$ and $b > 0$. Then

$$|x|_p := p^{-n} \in \mathbb{Q}_{\geq 0}$$

is called the *p-adic absolute value* of x. We will see that the map

$$|\cdot|_p : \mathbb{Q} \to \mathbb{Q}_{\geq 0}$$

has quite similar properties to the usual (euclidean) absolute value. Rather strangely, $|x|_p \leq 1$ for all integers $x \in \mathbb{Z}$, and $|p|_p = 1/p < 1$. Therefore, it seems to make sense to consider power series in p with integral coefficients,

$$\sum_{k=0}^{\infty} a_k p^k, \tag{3.1}$$

with $a_k \in \mathbb{Z}$, and to ask whether such series converge. Well, the series does certainly not converge to a real number. Essentially, we will define p-adic numbers in such a way that (3.1) converges to a p-adic number.

This may appear to be a rather esoteric construction at first sight. We will try to convince you that, quite to the contrary, p-adic numbers are both useful and natural objects, and in particular so in the context of Diophantine equations.

3.1 The p-adic valuation

Let p be a prime number. Then every nonzero integer $a \in \mathbb{Z}\backslash\{0\}$ can be written, in a unique way, as

$$a = up^n, \tag{3.2}$$

where $u \in \mathbb{Z}$ is prime to p and $n \in \mathbb{N}_0$. This fact is a direct consequence of the *fundamental theorem of arithmetic* which states that every natural number can be factored in a unique way as a product of primes. We define the *p-adic valuation* of a as the exponent n of p in (3.2),

$$v_p(a) := n.$$

Then the prime factorization of a can be written as

$$a = \pm \prod_p p^{v_p(a)}. \tag{3.3}$$

For consistency, we set $v_p(0) = \infty$. Also, if $x = a/b$ is a rational number we set

$$v_p(x) := v_p(a) - v_p(b) \in \mathbb{Z} \cup \{\infty\}.$$

It is easy to see that this is well defined.

Lemma 3.1. The function $v_p : \mathbb{Q} \to \mathbb{Z} \cup \{\infty\}$ has the following properties (for any $a, b \in \mathbb{Q}$):

(i) $v_p(ab) = v_p(a) + v_p(b)$.

(ii) $v_p(a + b) \geq \min(v_p(a), v_p(b))$.

(iii) If $v_p(a) \neq v_p(b)$ then we actually have $v_p(a + b) = \min(v_p(a), v_p(b))$.

(iv) If $x \in \mathbb{Z}$ then $v_p(x) \geq 0$.

Proof. Easy and left as an exercise. □

Definition 3.2. Let p be a prime number and $x \in \mathbb{Q}$. The *p-adic absolute value* of x is defined as

$$|x|_p := p^{-v_p(x)}.$$

For $x = 0$ this has to be understood as $|0|_p := 0$.

Lemma 3.3. The function $|\cdot|_p : \mathbb{Q} \to \mathbb{Q}_{\geq 0}$ has the following properties:

(i) $|xy|_p = |x|_p \cdot |y|_p$.

(ii) $|x + y|_p \leq \max(|x|_p, |y|_p)$.

(iii) If $|x|_p \neq |y|_p$ then $|x + y|_p = \max(|x|_p, |y|_p)$.

(iv) For $x \in \mathbb{Z}$ we have $|x|_p \leq 1$.

Proof. This follows directly from Lemma 3.1. □

3.2 The definition of $\mathbb{Z}_p$ and $\mathbb{Q}_p$

Recall from your first analysis lecture that the field $\mathbb{R}$ of real numbers may be defined as the *completion* of the field $\mathbb{Q}$ with respect to the usual euclidean absolute value $|\cdot|$. It is possible to define the field $\mathbb{Q}_p$ of p-adic numbers in exactly the same way, i.e. by saying that the field $\mathbb{Q}_p$ is the completion of $\mathbb{Q}$ with respect to the p-adic absolute value. We will, however, follow a slightly different path which ultimately leads to the same result. To motivate our definition we prove the following lemma.

Lemma 3.4. Let p be an odd prime number and $a \in \mathbb{Z}$, $a \not\equiv 0 \pmod{p}$. Assume that the congruence equation

$$x^2 \equiv a \pmod{p}$$

has a solution $x \in \mathbb{Z}$ (i.e. that a is a quadratic residue modulo p). Then the congruence equation

$$x^2 \equiv a \pmod{p^m}$$

has a solution for all $m \in \mathbb{N}$.

Proof. We build inductively a sequence of integers $x_0, x_1, x_2, \ldots$ with

(a) $x_i^2 \equiv a \pmod{p^{i+1}},\ i \geq 0.$

(b) $x_i \equiv x_{i+1} \pmod{p^{i+1}},\ i \geq 0.$

Since a QR p there exists an x_0 so that $x_0^2 \equiv a \pmod{p}$. Note that $x_0 \not\equiv 0\ (p)$ because $a \not\equiv 0 \pmod{p}$.

Now assume that $x_0, \ldots, x_{i-1}$ have already been constructed. Our Ansatz is to write $x_i = x_{i-1} + p^i y$, with $y \in \mathbb{Z}$. Then (b) holds automatically, and we need to choose y such that (a) holds as well. But

$$\begin{aligned} x_i^2 = x_{i-1}^2 + p^i(2x_{i-1}y) + p^{2i}y^2 &\equiv a + p^i z + p^i(2x_{i-1}y) \pmod{p^{i+1}} \\ &\equiv a + p^i(z + 2x_{i-1}y) \pmod{p^{i+1}}. \end{aligned}$$

Hence $x_i^2 \equiv a \pmod{p^{i+1}}$ if and only if $z + 2x_{i-1}y \equiv 0 \pmod{p}$ if and only if

$$y \equiv -\frac{z}{2x_{i-1}} \pmod{p}.$$

We can choose y so that this works. □

Example 3.5.
$p = 7,\ a = 2,\ \left(\frac{2}{7}\right) = 1$
$x_0 = 3,\ x_0^2 \equiv 9 \equiv 2\ (7)$
$x_1 = 3 + 7a_1$ so $x_1^2 = 9 + 42a_1 + 7^2a_1^2 \equiv 9 + 7 \cdot (6a_1) \equiv 2\ (7^2)$
So $x_1 = 3 + 7 = 10,\ x_2 = x_1 + 7^2a_2$

We now give the formal definition of p-adic numbers. Let us fix a prime p. We will work with the rings of residue classes modulo p^k,

$$\mathbb{Z}/p^{k+1}\mathbb{Z},$$

for $k = 0, 1, 2, \ldots$. Note that we have a sequence of natural surjective ring homomorphisms

$$\cdots \to \mathbb{Z}/p^4\mathbb{Z} \to \mathbb{Z}/p^3\mathbb{Z} \to \mathbb{Z}/p^2\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}.$$

Definition 3.6. The *p-adic integers* are the elements of the set

$$\mathbb{Z}_p := \left\{ (x_k)_{k\in\mathbb{N}_0} \in \prod_{k=0}^{\infty} \mathbb{Z}/p^{k+1}\mathbb{Z} \mid x_{k+1} \equiv x_k \pmod{p^{k+1}} \right\}.$$

Proposition 3.7. Let $+, \cdot$ be the binary operations on $\mathbb{Z}_p$ defined by componentwise addition and multiplication. Then $(\mathbb{Z}_p, +, \cdot)$ is a commutative ring with zero element $(0)_{k\in\mathbb{N}_0}$ and unit element $(1)_{k\in\mathbb{N}_k}$. Moreover, the map

$$\epsilon_p : \mathbb{Z} \to \mathbb{Z}_p, \qquad \epsilon_p(a) := (a)_{k\in\mathbb{N}_0},$$

is an injective ring homomorphism.

Proof. This is a routine and rather boring exercise. □

Example 3.8. Let $p = 7$. Then

$$\epsilon_7(173) = (173, 173, 173, 173, \ldots) = (5, 26, 173, 173, \ldots),$$

and

$$\epsilon_7(-1) = (-1, -1, -1, -1, \ldots) = (6, 48, 342, 2400, \ldots).$$

If we apply the proof of Lemma 3.4 to the case $p = 7$ and $a = 2$ then we obtain a square root of 2 as an element of $\mathbb{Z}_7$ (i.e. the element $x \in \mathbb{Z}_7$ so that $x^2 = 2$):

$$\sqrt{2} := (3, 10, 108, 3166, \ldots).$$

Note that

$$-\sqrt{2} = (-3, -10, -108, -3166, \ldots) = (4, 39, 235, 235, \ldots)$$

is also a square root of 2, and that there is no natural way to distinguish between the two.

From now on we will consider $\mathbb{Z}$ as a subring of $\mathbb{Z}_p$ (via ϵ_p), and we will not distinguish between $a \in \mathbb{Z}$ and $\epsilon_p(a) \in \mathbb{Z}_p$, if the prime p to use is clear from the context.

Remark 3.9. Let $x = (x_k)_{k\in\mathbb{N}_0} \in \mathbb{Z}_p$. Without loss of generality we may assume that

$$0 \le x_k \le p^{k+1} - 1.$$

Then there is a unique sequence $a_0, a_1, a_2, \ldots$ with $0 \le a_i < p$ such that

$$x_k = a_0 + a_1 \cdot p + a_2 \cdot p^2 + \ldots + a_k \cdot p^k,$$

for all k. We write (formally)

$$x = \sum_{k=0}^{\infty} a_k \cdot p^k$$

and call this the *p-adic expansion* of $x \in \mathbb{Z}_p$.

For the moment it is just a notational device which is useful for calculations. Note that you have to mind the *carry-over* when doing computations with p-adic expansions.

Theorem 3.10. *(i) The ring $\mathbb{Z}_p$ is an integral domain (i.e. it has no zero divisors).*

(ii) The group of units of $\mathbb{Z}_p$ is the subset

$$\mathbb{Z}_p^\times = \mathbb{Z}_p \backslash p \cdot \mathbb{Z}_p = \{(x_k) \mid x_0 \neq 0\}.$$

(iii) Every nonzero element $x \in \mathbb{Z}_p$ has a unique representation of the form

$$x = u \cdot p^n,$$

with $u \in \mathbb{Z}_p^\times$ and $n \in \mathbb{N}_0$.

(iv) The only ideals in $\mathbb{Z}_p$ are 0 and $p^n\mathbb{Z}_p$, for $n \in \mathbb{N}_0$. We have

$$\cap_{n\in\mathbb{N}_0} p^n\mathbb{Z}_p = 0$$

and

$$\mathbb{Z}_p/p^n\mathbb{Z}_p \cong \mathbb{Z}/p^n\mathbb{Z}.$$

In particular, $\mathbb{Z}_p$ is a principal ideal domain with a unique maximal ideal $p\mathbb{Z}_p$.

Remark 3.11. We can say informally that by passing from $\mathbb{Z}$ to the ring $\mathbb{Z}_p$ we have 'eliminated' the primes $\ell \neq p$. This is because in principal ideal domains prime ideals correspond to prime elements, and $\mathbb{Z}$ and $\mathbb{Z}_p$ are principal. Thus, prime decomposition in $\mathbb{Z}_p$ becomes very simple as p remains the only prime.

Definition 3.12. The field of p-adic numbers is the fraction field of $\mathbb{Z}_p$,

$$\mathbb{Q}_p := \mathrm{Frac}(\mathbb{Z}_p) = \{\, \frac{x}{y} \mid x, y \in \mathbb{Z}_p,\, y \neq 0 \,\}.$$

If we write $x, y \in \mathbb{Z}_p\backslash\{0\}$ as

$$x = u \cdot p^n, \quad y = v \cdot p^m,$$

with $u, v \in \mathbb{Z}_p^\times$ and $n, m \in \mathbb{Z}$ (Theorem 3.10 (iii)), then

$$\frac{x}{y} = (uv^{-1}) \cdot p^{n-m}.$$

It follows that every element $x \in \mathbb{Q}_p^\times$ can be written in a unique way as

$$x = u \cdot p^n, \tag{3.4}$$

with $u \in \mathbb{Z}_p^\times$ and $n \in \mathbb{Z}$. From the p-adic expansion of the unit u we can then derive a presentation of x as a 'p-adic Laurent series' of the form

$$x = \sum_{k \gg -\infty} a_k \cdot p^k, \tag{3.5}$$

with $0 \leq a_k < p$. However, at this stage this is again just a convenient notation.

Example 3.13. For $p = 5$ we compute the p-adic expansion of $1/10 \in \mathbb{Q} \subset \mathbb{Q}_p$:

$$\begin{aligned}\frac{1}{10} &= 5^{-1}\left(1 + 2 \cdot \frac{1}{1-5}\right) \\ &= 5^{-1}\big(1 + 2 \cdot (1 + 5 + 5^2 + 5^3 + \ldots)\big) \\ &= 3 \cdot 5^{-1} + 2 + 2 \cdot 5 + 2 \cdot 5^2 + 2 \cdot 5^3 + \ldots.\end{aligned}$$

As an exercise,

- justify this computation, and
- find the general method for computing the p-adic expansion of a rational number.

Definition 3.14. Let $x \in \mathbb{Q}_p$. We know that for $x \neq 0$ there is a unique representation $x = u \cdot p^n,\ u \in \mathbb{Z}_p^\times,\ n \in \mathbb{Z}$. In analogy with the rational case, we define:

$$v_p(x) = \begin{cases} n & \text{if } x \neq 0 \\ \infty & \text{if } x = 0 \end{cases}$$

which we call the *p-adic valuation* and

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

which we call the *p-adic absolute value*.

From the definition of the p-adic absolute value it is very easy to prove the following facts.

Lemma 3.15. Let $x, y \in \mathbb{Q}_p$. Then

(i) $|xy|_p = |x|_p \cdot |y|_p$,

(ii) $|x + y|_p \leq \max\{|x|_p, |y|_p\}$,

(iii) If $|x|_p \neq |y|_p$ then $|x + y|_p = \max\{|x|_p, |y|_p\} \leq |x|_p + |y|_p$,

(iv) $|x|_p \leq 1 \Leftrightarrow x \in \mathbb{Z}_p$,

(v) $|x|_p < 1 \Leftrightarrow x \in p\mathbb{Z}_p$.

Statements (ii) and (iii) are called the *strong triangle inequality*, because they imply (and are strictly stronger than) the usual triangle inequality

$$|x + y|_p \leq |x|_p + |y|_p .$$

Theorem 3.16. *$(\mathbb{Q}_p, |\cdot|_p)$ is a complete valued field and is the completion of $\mathbb{Q}$ with respect to the absolute value $|\cdot|_p$. In particular:*

(i) Every Cauchy sequence in $\mathbb{Q}_p$, $(x_n)_{n\in\mathbb{N}}$ has a unique limit $x = \lim_{n\to\infty} x_n$

(ii) Every element $x \in \mathbb{Q}_p$ is the limit of a sequence of rational numbers.

Proof. (i) Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathbb{Q}_p$, i.e. for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ so that for all $n, m \geq N$ we have

$$|x_n - x_m|_p < \epsilon.$$

We will start by proving that we can assume that $x_n \in \mathbb{Z}_p$. For this we wish to prove that the sequence $|x_n|_p$ is bounded.

If we fix m then for all $n > N$ we have that $|x_n|_p \leq |x_m|_p$ and the upper bound is $|x_m|_p$.

If not, then we take n so that $|x_n|_p > |x_m|_p$.

Then the strong triangle inequality gives

$|x_n|_p = |x_n - x_m + x_m|_p \leq max(|x_n - x_m|_p, |x_m|_p)$

As we have assumed that $|x_n|_p > |x_m|_p$ the only option left is $|x_n|_p \leq |x_n - x_m|_p$.

Choose $\epsilon = p^l$. From the triangle inequality we have that

$$|x_n|_p \leq |x_n - x_m|_p \leq p^l$$

As multiplying by p^{-l} does not change the convergence, we can replace (x_n) with $(p^{-l}x_n)$ and then we have $|p^{-l} \cdot x_n|_p \leq p^l$, so $|x_n|_p \leq 1$, i.e. $x_n \in \mathbb{Z}_p$.

Write $x_n = (x_n^{(k)})_{k\in\mathbb{N}_0}$ where $x_n^{(k)} \in \mathbb{Z}/p^{k+1}\mathbb{Z}$. Let $k \in \mathbb{N}_0$. Then there exists an N_k so that for all $n, m \geq N_k$

$$|x_n - x_m|_p \leq p^{-k-1}.$$

This is equivalent to

$$x_n^{(k)} = x_m^{(k)} \text{ (in } \mathbb{Z}/p^{k+1}\mathbb{Z})$$

for all $n, m \geq N_k$. Without loss of generality we can assume that $N_0 \leq N_1 \leq N_2 \leq \ldots$. Then $|x_{N_k} - x_{N_{k+1}}|_p \leq p^{-k-1}$ and thus for all $k \in \mathbb{N}_0$

$$x_{N_k} \equiv x_{N_{k+1}} \pmod{p^{k+1}}.$$

Put $x := (x_{N_k}^{(k)}) \in \mathbb{Z}_p$. Then we have

$$|x - x_{N_k}|_p \leq p^{-k-1} \overset{k\to\infty}{\longrightarrow} 0$$

and this means that $x = \lim\limits_{n\to\infty} x_n$

(ii) Let $x \in \mathbb{Q}_p$ with p-adic expansion

$$x = \sum_{k=-N}^{\infty} a_k p^k, \ 0 \leq a_k < p$$

Then we can write $x = x_n + y_n$, where $x_n = \sum\limits_{k=-N}^{n} a_k p^k \in \mathbb{Q}$ and

$$y_n = \sum_{k=n+1}^{\infty} a_k p^k \in \mathbb{Q}_p.$$

Since

$$|y_n|_p \leq p^{-n-1}$$

we have $\lim_n y_n = 0$ and $\lim_n x_n = x$. □

Remark 3.17. (i) The proof of (ii) shows that the p-adic expansion can be indeed considered as a convergent series.

(ii) We can generally ask whether a series

$$\sum_{n=1}^{\infty} x_n \tag{3.6}$$

in $\mathbb{Q}_p$ converges (i.e. whether the finite partial sums converge to a p-adic number). Using the strong triangle inequality (Lemma 3.15 (iii)) it is easy to see that (3.6) is convergent if and only if $x_n \overset{n\to\infty}{\longrightarrow} 0$.

Example 3.18. Let p be a prime number and $a \in \mathbb{Z}$ with $a \equiv 1 \pmod{p}$. Let $n \in \mathbb{N},\ n \not\equiv 0 \pmod{p}$. Then there exists $b \in \mathbb{Z}_p$ with $b^n = a$ (an nth root of a in $\mathbb{Q}_p$).

Proof. We can write $a = 1 + pc$, with $c \in \mathbb{Z}$. We define

$$b := (1+pc)^{1/n} = \sum_{k=0}^{\infty} \binom{1/n}{k} c^k p^k.$$

If we can show that this series converges, then the standard proof from analysis that $b^n = a$ goes through. To prove the convergence it suffices by Remark 3.17 (ii) to show that $\binom{1/n}{k} \in \mathbb{Z}_p$. This is an interesting exercise. □

3.3 Hensel's Lemma

The most important theorem about p-adic numbers is called *Hensel's Lemma*[1]. There are many different formulations of this result, with varying generality. We will only need its most basic form.

Theorem 3.19 (Hensel's lemma)**.** *Let p be a prime number, $f \in \mathbb{Z}_p[t]$ and $x_0 \in \mathbb{Z}_p$ so that*

$$v_p(f(x_0)) \geq 2l + 1,$$

where $l := v_p(f'(x_0))$. Then there exists a uniquely determined root $x \in \mathbb{Z}_p$ of f, $f(x) = 0$, so that

$$x \equiv x_0 \pmod{p^{l+1}}.$$

Proof. It suffices to build a sequence $x_0, x_1, x_2 \ldots \in \mathbb{Z}_p$ with

- $x_k \equiv x_{k-1} \pmod{p^{k+l}}$, for all $k \geq 1$, and
- $f(x_k) \equiv 0 \pmod{p^{k+2l+1}}$, for all $k \geq 0$.

[1] named after Kurt Hensel, 1861-1941, who discovered (or invented) the p-adic numbers in 1897

Then $x := \lim x_k$ will give us the right solution.

Our assumption says that x_0 satisfies the second condition. We may therefore assume that $k \geq 1$ and that $x_0, \ldots, x_{k-1}$ have already been determined. To make sure that the first condition holds we have to choose $x_k = x_{k-1} + y \cdot p^{k+l}$ for a suitable element $y \in \mathbb{Z}_p$. Then, by Taylor expansion there exists a $z \in \mathbb{Z}_p$ so that:

$$\begin{aligned} f(x_k) &= f(x_{k-1} + yp^{k+l}) = f(x_{k-1}) + f'(x_{k-1}) \cdot yp^{k+l} + z \cdot p^{2k+2l} \\ &= p^{k+2l}(w + zp^k + uy), \end{aligned}$$

with

$$w := \frac{f(x_{k-1})}{p^{k+2l}}, \qquad u := \frac{f'(x_{k-1})}{p^l}.$$

By the induction hypothesis we have $v_p(w) \geq 0$, i.e. $w \in \mathbb{Z}_p$. We claim that $u \in \mathbb{Z}_p^\times$. To prove this claim we have to show that $v_p(f'(x_{k-1})) = l$. Using the strong triangle inequality, this follows easily from the fact that f' is a polynomial with coefficients in $\mathbb{Z}_p$ and that $x_{k-1} \equiv x_0 \pmod{p^l}$.

By choosing $y := -u^{-1}w$ we get $f(x_k) = p^{2k+2l}z \in p^{k+2l+1}\mathbb{Z}_p$. This completes the proof of the theorem. □

Remark 3.20. The proof is a form of Newtonian approximation

$$x_k := x_{k-1} - \frac{f(x_{k-1})}{f'(x_{k-1})}.$$

One can see that the convergence is 'quadratic', as in the classical case.

Example 3.21. Let $p = 5$, $f = x^3 - 2$ and $x_0 = 3$. Then

$$l = v_5(f'(3)) = v_5(3^3) = 0$$

and

$$f(3) = 3^3 - 2 = 25 \equiv 0 \pmod{5^2}.$$

We see that the conditions in Theorem 3.19 hold. Therefore, the series $x_0, x_1, \ldots$ defined inductively by

$$x_{k+1} = x_k - \frac{x_k^3 - 2}{3x_k^2}, \ k = 0, 1, \ldots,$$

converges to an element $x = \lim_k x_k \in \mathbb{Z}_5$ with $x^3 = 2$.

We calculate x with accuracy $\mathcal{O}(5^4)$. The first step of our Newton approximation scheme gives

$$\begin{aligned} x_1 =& 3 - \frac{27-2}{27} = 3 - \frac{1}{2+25} \cdot 5^2 \equiv 3 - \frac{1}{2}5^2 \pmod{5^4} \\ \equiv& 3 + \frac{2}{1-5}5^2 \pmod{5^4} \\ \equiv& 3 + 2(1 + 5 + 5^2 + \ldots) \cdot 5^2 \pmod{5^4} \\ \equiv& 3 + 2 \cdot 5^2 + 2 \cdot 5^3 \pmod{5^4}. \end{aligned}$$

Then

$$\begin{aligned} x_1^3 \equiv& (3 + 2 \cdot 5^2 + 2 \cdot 5^3)^3 \pmod{5^4} \\ \equiv& 27 + 3 \cdot 3^2 \cdot (2 \cdot 5^2 + 2 \cdot 5^3) \pmod{5^4} \\ \equiv& 2 + 5^2(1 + 3^3 \cdot 2 \cdot 6) \pmod{5^4} \\ \equiv& 2 + 5^2(1 + 54 \cdot 6) \pmod{5^4} \\ \equiv& 2 \pmod{5^4}. \end{aligned}$$

This means that $f(x_1) \equiv 0 \pmod{5^4}$. Hence we can stop here and obtain an approximation of x with accuracy $\mathcal{O}(5^4)$:

$$x = 3 + 2 \cdot 5^2 + 2 \cdot 5^3 + \ldots.$$

Chapter 4

The theorem of Hasse-Minkowski

The goal of this chapter is to prove the famous *Theorem of Hasse-Minkowski*.

Theorem (Hasse-Minkowski)**.** *A quadratic Diophantine equation has a rational solution if and only if it has solutions in* $\mathbb{R}$ *and* $\mathbb{Q}_p$ *for every prime* p*.*

We start with some preparation.

4.1 The equation $x^2 = a$

Let K be a field. Then the set $(K^\times)^2 = \left\{x^2 \mid x \in K^\times\right\}$, called the set of *squares*, is a subgroup of $K^\times$.

We wish to determine $(K^\times)^2$ for

$$K = \mathbb{Q},\ \mathbb{R},\ \mathbb{Q}_p,\ \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}.$$

For $K = \mathbb{R}, \mathbb{Q}$ this is easy, for $K = \mathbb{F}_p$ the result is well known from elementary number theory.

Remark 4.1. (i) The squares in $\mathbb{R}^\times$ are precisely the positive reals:

$$(\mathbb{R}^\times)^2 = \left\{a \in \mathbb{R} \mid a > 0\right\}.$$

(ii) Using the prime factorization, i.e. the representation of $a \in \mathbb{Q}^\times$ in the form

$$a = \pm \prod_p p^{v_p(a)},$$

we see immediately that a rational number a is a square if and only if it is positive and all p-valuations $v_p(a)$ are even:

$$(\mathbb{Q}^\times)^2 = \left\{ a \in \mathbb{Q} \mid a > 0,\ v_p(a) \equiv 0 \pmod 2 \text{ for every prime } p \right\}.$$

(iii) The multiplicative group of the finite field $\mathbb{F}_2$ is trivial, hence

$$(\mathbb{F}_2^\times)^2 = \mathbb{F}_2^\times = \{1\}.$$

Let $p \neq 2$ be an odd prime. Then $\mathbb{F}_p^\times$ is a cyclic group of order $p-1$ (which is even). Therefore, $(\mathbb{F}_p^\times)^2$ is the unique subgroup of order $(p-1)/2$. It follows that there are exactly $(p-1)/2$ squares and $(p-1)/2$ non-squares.

To decide whether an actual element of $\mathbb{F}_p^\times$ is a square or not it is useful to work with the *Legendre symbol*.

Definition 4.2. For $p \neq 2$ and $a \in \mathbb{Z}$ the *Legendre symbol* is:

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p|a \\ 1 & \text{if } p \not| a,\ a\ QR\ p \\ -1 & \text{if } p \not| a,\ a\ QNR\ p \end{cases}$$

Proposition 4.3. The Legendre symbol has the following properties:

(i) $\left(\frac{a}{p}\right)$ only depends on the residue class of a modulo p.

(ii) $\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod p$.

(iii) $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right) \cdot \left(\frac{b}{p}\right)$.

(iv) $\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} = \begin{cases} 1 & \text{if } p \equiv 1 \pmod 4, \\ -1 & \text{if } p \equiv 3 \pmod 4. \end{cases}$.

(v) $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} = \begin{cases} 1 & \text{if } p \equiv 1, 7 \pmod 8, \\ -1 & \text{if } p \equiv 3, 5 \pmod 8. \end{cases}$.

(vi) $\left(\frac{q}{p}\right) \cdot \left(\frac{p}{q}\right) = (-1)^{\frac{p-1}{2} \cdot \frac{q-1}{2}} = \begin{cases} 1 & \text{if } q \equiv 1 \text{ or } p \equiv 1 \pmod 4, \\ -1 & \text{if } p, q \equiv 3 \pmod 4. \end{cases}$

Lemma 4.4. Let $p \neq 2$, $a \in \mathbb{Q}_p^\times$. Write $a = u \cdot p^n$, $u \in \mathbb{Z}_p^\times$, $n = v_p(a) \in \mathbb{Z}$. Then $a \in (\mathbb{Q}_p^\times)^2$ if and only if $n \equiv 0 \pmod 2$, $\left(\frac{u}{p}\right) = 1$

Example 4.5. $p = 3$, $a = 18 = 2 \cdot 3^2$, $\left(\frac{2}{3}\right) = -1$, so $a \notin (\mathbb{Q}_p^\times)^2$

Proof. For the direct implication take $a = x^2$ with $x = v \cdot p^m$. Then $a = v^2 \cdot p^{2n}$, so $u = v^2$, $u = 2m \equiv 0 \pmod 2$. So u QR p and $\left(\frac{u}{p}\right) = 1$.

For the converse take $a = u \cdot p^{2m}$ with $\left(\frac{u}{p}\right) = 1$. p^{2m} is already a square in $\mathbb{Q}_p^\times$, so we can assume that $m = 0$. Then $a = u$. From $\left(\frac{a}{p}\right) = \left(\frac{u}{p}\right) = 1$ there exists $x_0 \in \mathbb{Z}$ so that $x_0^2 \equiv a \pmod p$.

Let $f := X^2 - a \in \mathbb{Z}_p[X]$. Then $f(x_0) \equiv 0 \pmod p$, $f'(x_0) = 2x_0 \not\equiv 0 \pmod p$, so $l := v_p(f'(x_0)) = 0$.

From Hensel's lemma there is an $x \in \mathbb{Z}_p$ with $f(x) = x^2 - a$, so $a \in (\mathbb{Q}_p^\times)^2$ □

Lemma 4.6. Let $p = 2$, $a \in \mathbb{Q}_2^\times$, $a = u \cdot 2^n$. Then

$$a \in (\mathbb{Q}_2^\times)^2 \text{ if and only if } n \equiv 0 \pmod 2 \text{ and } u \equiv 1 \pmod 8$$

Proof. Like in lemma 4.4 we can assume that $n = 0$ and $a \in \mathbb{Z}_2^\times$.

For the direct implication we have $a = x^2$.

Then $x \in \mathbb{Z}_2^\times$, so $x \equiv 1, 3, 5, 7 \pmod 8$, so $a \equiv 1 \pmod 8$.

For the converse let $a \equiv 1 \pmod 8$, $f := X^2 - a$, $x_0 := 1$. Then

$f(x_0) = 1 - a \equiv 0 \pmod 8$, so $v_2(f(x_0)) \geq 3$.

$f'(x_0) = 2$, so $l = v_2(f'(x_0)) = 1$

Applying Hensel's lemma gives us that there exists such an $x \in \mathbb{Z}_2$ so that $x^2 = a$ □

Corollary 4.7 (trivial case of Hasse-Minkowski)**.** *For $a \in \mathbb{Q}^\times$ we have that*

$$a \in (\mathbb{Q}^\times)^2 \textit{ if and only if } a \in (\mathbb{R}^\times)^2 \textit{ and } a \in (\mathbb{Q}_p^\times)^2 \textit{ for every } p \in \mathbb{P}$$

Notation 4.8. We can note $\mathbb{R} =: \mathbb{Q}_\infty$, so the corollary can be written as

$$a \in (\mathbb{Q}^\times)^2 \text{ if and only if } a \in (\mathbb{Q}_p^\times)^2 \text{ for every p} \in \mathbb{P} \cup \{\infty\}$$

4.2 The quadratic residue classes

For $K = \mathbb{R}$ we have $(\mathbb{R}^\times)^2 = \{a \mid a > 0\} \subseteq \mathbb{R}^\times$, so the map

$$\begin{aligned} \mathbb{R}^\times/(\mathbb{R}^\times)^2 &\to \{\pm 1\} \\ a \cdot (\mathbb{R}^\times)^2 &\mapsto \operatorname{sgn}(a) = \frac{a}{|a|} \end{aligned}$$

is a group isomorphism. Also, for $p \neq 2$ we have the group isomorphism

$$\begin{aligned} \overline{\phi}_p : \mathbb{F}_p^\times/(\mathbb{F}_p^\times)^2 &\to \{\pm 1\} \\ a \cdot (\mathbb{F}_p^\times)^2 &\mapsto \left(\frac{a}{p}\right) \end{aligned}$$

What does this look like for $K = \mathbb{Q}$?

Theorem 4.9. *The map*

$$\tilde{\phi} : \left\{ \begin{array}{ccc} \mathbb{Q}^\times & \to & \{\pm 1\} \times \bigoplus_{p\in\mathbb{P}} \mathbb{Z}/2\mathbb{Z} \\ a & \mapsto & (\operatorname{sgn}(a);\ (v_p(a) \pmod 2)_{p\in\mathbb{P}}) \end{array} \right.$$

induces a group isomorphism

$$\mathbb{Q}^\times/(\mathbb{Q}^\times)^2 \to \{\pm 1\} \times \bigoplus_{p\in\mathbb{P}} \mathbb{Z}/2\mathbb{Z}.$$

Remark 4.10. For $(\epsilon, (\overline{c}_p)_{p\in\mathbb{P}}) \in \{\pm 1\} \times \bigoplus_{p\in\mathbb{P}} \mathbb{Z}/2\mathbb{Z}$ we have that $\overline{c}_p = 0$ for almost all $p \in \mathbb{P}$, this is why we have a direct sum on the right hand side, instead of a Cartesian product.

Proof. The map $\tilde{\phi}$ is a group homomorphism because $sgn(ab) = sgn(a) \cdot sgn(b)$ and $v_p(ab) = v_p(a) + v_p(b)$. Its kernel is the subgroup of all $a \in \mathbb{Q}^\times$ such that

$$\operatorname{sgn}(a) = 1 \quad \text{and } v_p(a) \equiv 0 \pmod 2, \quad \forall\, p \in \mathbb{P}.$$

But this condition means that a is a square. Therefore,

$$\ker(\tilde{\phi}) = (\mathbb{Q}^\times)^2.$$

It remains to prove that $\tilde{\phi}$ is surjective. Take

$$(\epsilon, (\overline{c}_p)_{p\in\mathbb{P}}) \in \{\pm 1\} \times \bigoplus_{p\in\mathbb{P}} \mathbb{Z}/2\mathbb{Z}.$$

We set

$$c_p := \begin{cases} 1 & \text{if } \overline{c}_p = 1 \\ 0 & \text{if } \overline{c}_p = 0 \end{cases}$$

and $a := \epsilon \prod_{p\in\mathbb{P}} p^{c_p} \in \mathbb{Z}_p$. Then $\tilde{\phi}(a) = (\epsilon, \overline{c}_p)$. Hence $\tilde{\phi}$ is surjective.

From the first isomorphism theorem we deduce that $\tilde{\phi}$ induces the isomorphism:

$$\mathbb{Q}^\times/(\mathbb{Q}^\times)^2 \to \{\pm 1\} \times \bigoplus_{p\in\mathbb{P}} \mathbb{Z}/2\mathbb{Z}.$$

□

Remark 4.11. From the proof we see that the square free integers

$$a := \epsilon \prod_{p\in\mathbb{P}} p_i \in \mathbb{Z}_p \text{ with } p_i \neq p_j$$

form a system of representatives for $\mathbb{Q}^\times/(\mathbb{Q}^\times)^2$.

Theorem 4.12. *1. If $p \neq 2$ we have the group isomorphism:*

$$\begin{aligned} \phi_p : \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 &\to \{\pm 1\} \times \mathbb{Z}/2\mathbb{Z} \\ a(\mathbb{Q}_p^\times)^2 &\mapsto \left(\left(\frac{u}{p}\right), v_p(a) \mod 2\right) \end{aligned}$$

where $a = p^{v_p(a)} \cdot u$.

In particular, for $\epsilon \in \mathbb{Z}$ with $\left(\frac{\epsilon}{p}\right) = -1$ $\{1, \epsilon, p, \epsilon p\}$ is a system of representatives for $\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2$.

2. If $p = 2$ we have the group isomorphism:

$$\begin{aligned} \phi_2 : \mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2 &\to (\mathbb{Z}/8\mathbb{Z})^\times \times \mathbb{Z}/2\mathbb{Z} \\ a(\mathbb{Q}_2^\times)^2 &\mapsto (u \mod 8, v_2(a) \mod 2) \end{aligned}$$

where $a = 2^{v_2(a)} \cdot u$.

In particular, $\{\pm 1, \pm 5, \pm 2, \pm 10\}$ is a system of representatives for $\mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2$.

Proof. We use Lemmas 4.4 and 4.6, the proof is similar to Theorem 4.9.

□

4.3 The Hilbert symbol

We have seen that $x^2 = a$ is solvable in $\mathbb{Q}$ if and only if it is solvable in $\mathbb{Q}_p$ for every $p \in \mathbb{P} \cup \{\infty\}$.

For $p \neq 2$ and $a = p^n u$, solvability of $x^2 = a$ in $\mathbb{Q}$ is equivalent to $n \equiv 0 \pmod 2$ and $\left(\frac{u}{p}\right) = 1$. Hence the Legendre symbol measures the solvability of the equation $x^2 = a$. For the more general equation

$$ax^2 + by^2 = z^2 \tag{4.1}$$

the tool we are going to use is the *Hilbert symbol*.

Definition 4.13. Let $p \in \mathbb{P} \cup \{\infty\}$ and $a, b \in \mathbb{Q}_p^\times$. Then we note:

$$\left(\frac{a,b}{p}\right) = \begin{cases} 1 & \text{if (4.1) has solutions in } \mathbb{Q}_p, \\ -1 & \text{otherwise.} \end{cases}$$

Proposition 4.14. For $p = \infty$, $a, b \in \mathbb{R}^\times$ we have

$$\left(\frac{a,b}{p}\right) = 1 \iff a > 0 \text{ or } b > 0$$

Proof. If $\left(\frac{a,b}{p}\right) = 1$ then we have $x, y, z \in \mathbb{R} \setminus \{0\}$ that satisfy (4.1). Then a and b can't be both negative. Conversely, if $a > 0$, then $a = c^2$ then $(x, y, z) = (c^{-1}, 0, 1) \in \mathbb{R} \setminus \{0\}$ is a solution of (4.1). This completes the proof. □

Lemma 4.15. Let $p \in \mathbb{P} \cup \{\infty\}$, $a, b, c, d \in \mathbb{Q}_p^\times$

(i) $\left(\frac{a,b}{p}\right) = \left(\frac{b,a}{p}\right)$

(ii) $\left(\frac{a,1}{p}\right) = \left(\frac{a,-a}{p}\right) = 1$

(iii) $\left(\frac{a,1-a}{p}\right) = 1$ for $a \neq 1$

(iv) $\left(\frac{a,b}{p}\right) = \left(\frac{ac^2,bd^2}{p}\right)$

(v) $\left(\frac{ab,ac}{p}\right) = \left(\frac{ab,-bc}{p}\right)$

(vi) if $\left(\frac{a,c}{p}\right) = 1$ then $\left(\frac{a,bc}{p}\right) = \left(\frac{a,b}{p}\right)$

Proof. $\left(\frac{a,b}{p}\right) = 1$ if and only if $ax^2 + by^2 = z^2$ has solutions.

(i) We use the definition of the Hilbert symbol and switch the role of x and y.

(ii) For $b = 1$ $(0, 1, 1)$ is a solution of (4.1), so $\left(\frac{a,1}{p}\right) = 1$. For $b = -a$ (1,1,0) is a solution of (4.1), so $\left(\frac{a,-a}{p}\right) = 1$

(iii) We need to prove $ax^2 + (1 - a)y^2 = z^2$ is solvable. We remark that (1,1,1) is a solution.

(iv) (x, y, z) is a solution of $ax^2 + by^2 = z^2$ if and only if $(c^{-1}x, d^{-1}y, z)$ is a solution of $(ac^2)x^2 + (bd^2)y^2 = z^2$.

(v) Because a, b are invertible, multiplying $abx^2 + acy^2 = z^2$ by ab gives an equivalent equation:

$$abx^2 + acy^2 = z^2 \Leftrightarrow (abx)^2 + bc(ay^2) = abz^2 \Leftrightarrow (ab)z^2 - (bc)(ay^2) = (abx)^2.$$

Again, due to the invertibility of a and b the equation $(ab)z^2 - (bc)(ay^2) = (abx)^2$ is solvable if and only if $(ab)z^2 - (bc)\tilde{y}^2 = \tilde{x}^2$ is.

This gives $\left(\frac{ab,ac}{p}\right) = \left(\frac{ab,-bc}{p}\right)$.

(vi) Let us prove first that $G := \left\{ b \in \mathbb{Q}_p^\times \mid \left(\frac{a,b}{p}\right) = 1 \right\}$ is a subgroup of $\mathbb{Q}_p^\times$.

If $a \in (\mathbb{Q}_p^\times)^2$ then $G = \mathbb{Q}_p^\times$.

If not, then we can build an extension K of $\mathbb{Q}_p$ of degree 2 by adjoining a square root of a. It will be endowed with a norm

$$\begin{aligned} N_{K/\mathbb{Q}_p} : K &\to \mathbb{Q}_p \\ x + \sqrt{a}y &\mapsto x^2 - ay^2 \end{aligned}$$

By (ii) we have that $1 \in G$.

If $u \in G$ then the equation $ax^2 + uy^2 = z^2$ has a nontrivial solution. Then, if we divide by u^2, we get $a\left(\frac{x}{u}\right)^2 + \frac{1}{u}y^2 = \left(\frac{z}{u}\right)^2$ and so $u^{-1} \in G$.

Take now $u, v \in G$, so we have $x_1, y_1, z_1, x_2, y_2, z_2 \in \mathbb{Q}_p$ so that $ax_1^2 + uy_1^2 = z_1^2$ and $ax_2^2 + vy_2^2 = z_2^2$.

This gives $uv(y_1y_2)^2 = (z_1^2 - ax_1^2)(z_2^2 - ax_2^2) = N_{K/\mathbb{Q}_p}(z_1 + \sqrt{a}x_1)N_{K/\mathbb{Q}_p}(z_2 + \sqrt{a}x_2) = N_{K/\mathbb{Q}_p}((z_1z_2 + ax_1x_2) + \sqrt{a}(z_1x_2 + x_1z_2)) = Z^2 - aX^2$. So we have $aX^2 + uv(y_1y_2)^2 = Z^2$, so $\left(\frac{a,uv}{p}\right) = 1$ and $uv \in G$.

Thus $G \leq \mathbb{Q}_p^\times$. We have that $c \in G$ and we have to prove that $b \in G$ if and only if $bc \in G$. But this is obvious, due to the group structure of G.

□

Remark 4.16. $\left(\frac{a,b}{p}\right)$ depends only on the image of a, b in $\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2$

Theorem 4.17. *Let $p \in \mathbb{P}$, $p \neq 2$. Let $a = p^n u$ and $b = p^n v$, $n, m \in \mathbb{Z}$, $u, v \in \mathbb{Z}_p^\times$. Then*

$$\left(\frac{a,b}{p}\right) = (-1)^{nm\frac{p-1}{2}} \left(\frac{u}{p}\right)^m \left(\frac{v}{p}\right)^n$$

In particular:

(i) $a, b \in \mathbb{Z}_p^\times : \left(\frac{a,b}{p}\right) = 1$

(ii) for $n = 0$, $m = 1$: $\left(\frac{u,pv}{p}\right) = \left(\frac{u}{p}\right)$

(iii) for $n = m = 1$: $\left(\frac{up,vp}{p}\right) = \left(\frac{-uv}{p}\right)$

Proof. We first show $(i), (ii), (iii)$.

(i) We will use the following lemma:

Lemma 4.18. For $a, b \in \mathbb{F}_p^\times$ the equation $ax^2 + b = z^2$ is solvable in $\mathbb{F}_p$

Proof. Let $M_1 = \left\{ax^2 + b \mid x \in \mathbb{F}_p\right\}$ and $M_2 = \left\{z^2 \mid z \in \mathbb{F}_p\right\}$

Then $|M_1| = |M_2| = (|(\mathbb{F}_p^\times)^2| + 1 = \frac{p-1}{2} + 1 = \frac{p+1}{2}$.

But then $|M_1| + |M_2| > |\mathbb{F}_p| = p$, so $M_1 \cap M_2 \neq \emptyset$, so that there exist $x, z \in \mathbb{F}_p$ with $ax^2 + b = z^2$. □

Now let $a, b \in \mathbb{Z}_p^\times$ and $\overline{a}, \overline{b}$ their images in $\mathbb{F}_p$.

Let $(\overline{x}_0, \overline{z}_0)$ be a solution for $\overline{a}\ \overline{x}^2 + \overline{b} = \overline{z}^2$.

We take $f := aX^2 + b - z_0^2 \equiv 0 \pmod{p}$

For $x_0 \not\equiv 0 \pmod{p}$ we have $f'(x_0) = 2ax_0 \not\equiv 0 \pmod{p}$, because $p \neq 2$.

From Hensel's lemma we deduce the existence of an $x \in \mathbb{Z}_p$ with $f(x) = ax^2 + b - z_0^2 = 0$, so (4.1) has the solution $(x, 1, z_0)$ and thus $\left(\frac{a,b}{p}\right) = 1$.

For $x_0 \equiv 0 \pmod{p}$, because $b \in \mathbb{Z}_p^\times$ we have $z_0 \not\equiv 0 \pmod{p}$ and the same argument for $g := ax_0^2 + b - Z^2$ gives $z \in \mathbb{Z}_p$ with $ax_0^2 + b = z^2$, so $\left(\frac{a,b}{p}\right) = 1$.

(ii) We assume $\left(\frac{a,pv}{p}\right) = 1$, so we have a nontrivial solution (x, y, z) of $ux^2 + vpy^2 = z^2$

We may assume $x, y, z \in \mathbb{Z}_p$ and that at least one of the numbers x, y, z is in $\mathbb{Z}_p^\times$ (if it's not the case, we just eliminate p factors until it is).

We have $ux^2 \equiv z^2 \pmod{p}$. If $x \equiv 0 \pmod{p}$ then $z \equiv 0 \pmod{p}$, so $vpy^2 \equiv 0 \pmod{p^2}$. Because $v \in \mathbb{Z}_p^\times$ this gives the contradiction $y \equiv 0 \pmod{p}$. So $x \in \mathbb{Z}_p^\times$.

Therefore $u \equiv \left(\frac{z}{x}\right)^2 \pmod{p}$ and thus $\left(\frac{u}{p}\right) = 1$.

Conversely, if we have $\left(\frac{u}{p}\right) = 1$ then by Hensel's Lemma there is a $v \in \mathbb{Z}_p^\times$ so that $a = v^2$. This means that $(v^{-1}, 0, 1)$ is a solution of $ax^2 + vpy^2 = z^2$.

(iii) $\left(\frac{up, vp}{p}\right) \overset{4.15(v)}{=} \left(\frac{up, -uv}{p}\right) \overset{4.15(i)}{=} \left(\frac{-uv, up}{p}\right) \overset{4.17(ii)}{=} \left(\frac{-uv}{p}\right)$

We will now prove the main formula:

$$\left(\frac{a, b}{p}\right) = (-1)^{nm\frac{p-1}{2}} \left(\frac{u}{p}\right)^m \left(\frac{v}{p}\right)^n$$

As we have seen in remark 4.16 $\left(\frac{a,b}{p}\right)$ depends only on the classes $a(\mathbb{Q}_p^\times)^2$ and $b(\mathbb{Q}_p^\times)^2$ in $\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2$, so we can find $a_0, b_0 \in \{1, \epsilon, p, \epsilon p\}$ so that $\left(\frac{a,b}{p}\right) = \left(\frac{a_0,b_0}{p}\right)$. This means that we can actually restrict the discussion to the cases $n, m \in \{0, 1\}$ and $u, v \in \{1, \epsilon\}$

For (n, m) we have, ignoring symmetry, the three cases $(0, 0)$, $(1, 0)$, $(1, 1)$, corresponding to the formulas (i),(ii),(iii), which have already been proven. □

Theorem 4.19. *For $a = 2^n u$, $b = 2^m v$, $u, v \in \mathbb{Z}_2^\times$ we have*

$$\left(\frac{a, b}{2}\right) = (-1)^{\frac{u-1}{2}\cdot\frac{v-1}{2}} (-1)^{n\frac{v^2-1}{8}} (-1)^{m\frac{u^2-1}{8}}$$

In particular:

(i) $a, b \in \mathbb{Z}_p^\times : \left(\frac{a, b}{2}\right) = (-1)^{\frac{a-1}{2}\cdot\frac{b-1}{2}}$

(ii) $\left(\frac{u, 2v}{2}\right) = (-1)^{\frac{u-1}{2}\cdot\frac{v-1}{2}} \cdot (-1)^{\frac{u^2-1}{8}}$

(iii) $\left(\frac{2u, 2v}{2}\right) = (-1)^{\frac{u-1}{2}\cdot\frac{v-1}{2}} \cdot (-1)^{n\frac{v^2-1}{8}} (-1)^{m\frac{u^2-1}{8}}$

Proof. See [2], Theorem 14.13. □

Remark 4.20. We have the following values for $\left(\frac{a,b}{p}\right)$ for every choice of representatives:

For $p \equiv 1 \pmod{4}$:

$\left(\frac{a,b}{p}\right)$	1	ϵ	p	ϵp
1	$+1$	$+1$	$+1$	$+1$
ϵ	$+1$	$+1$	-1	-1
p	$+1$	-1	$+1$	-1
ϵp	$+1$	-1	-1	$+1$

For $p \equiv 3 \pmod 4$:

$\left(\frac{a,b}{p}\right)$	1	ϵ	p	ϵp
1	$+1$	$+1$	$+1$	$+1$
ϵ	$+1$	$+1$	-1	-1
p	$+1$	-1	-1	$+1$
ϵp	$+1$	-1	$+1$	-1

One can see from the tables that $\left(\frac{a,\cdot}{p}\right) : \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \to \{\pm 1\}$ is multiplicative, so $\left(\frac{a,bc}{p}\right) = \left(\frac{a,b}{p}\right) \cdot \left(\frac{a,c}{p}\right)$.

This also means that $\left(\frac{\cdot,\cdot}{p}\right) : \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \times \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2 \to \{\pm 1\}$ is bimultiplicative.

This also works for $p = 2$ and $p = \infty$.

Example 4.21.

Let us compute $\left(\frac{5,6}{p}\right)$ for $p \in \mathbb{P} \cup \{\infty\}$

- $p = \infty : \left(\frac{5,6}{\infty}\right) = 1$
- $p = 2 : \left(\frac{5,6}{2}\right) = \left(\frac{5,2}{2}\right) \cdot \left(\frac{5,3}{2}\right) = (-1)^{\frac{5^2-1}{8}}(1) = -1$
- $p = 3 : \left(\frac{5,6}{3}\right) = \left(\frac{5,2}{3}\right) \cdot \left(\frac{5,3}{3}\right) = \left(\frac{5}{3}\right)(1) = -1$
- $p = 5 : \left(\frac{5,6}{5}\right) = \left(\frac{5,2}{5}\right) \cdot \left(\frac{5,3}{5}\right) = \left(\frac{6}{5}\right) = \left(\frac{1}{5}\right) = 1$
- $p > 5 : \left(\frac{5,6}{p}\right) = 1$, because $5, 6 \in \mathbb{Z}_p^\times$

So the equation $5x^2 + 6y^2 = z^2$ is solvable in $\mathbb{R}$ and $\mathbb{Q}_p$ for $p > 3$, but not in $\mathbb{Q}_2$ and $\mathbb{Q}_3$. This means that it is not solvable in $\mathbb{Q}$ either.

Legendre's theorem says that $5x^2 + 6y^2 = z^2$ is solvable in $\mathbb{Q}$ if and only if 5 QR 6 and 6 QR 5. We see that 6 QR 5, but 5 QNR 6.

Theorem 4.22 (the product formula)**.** *Let $a, b \in \mathbb{Q}^\times$. Then*

$$\prod_{p \in \mathbb{P} \cup \{\infty\}} \left(\frac{a,b}{p}\right) = 1$$

Remark 4.23. For almost all $p \in \mathbb{P}$, $p \neq 2$, if $a, b \in \mathbb{Z}_p$ then $\left(\frac{a,b}{p}\right) = 1$ (finite product) The theorem says that the number of p with $\left(\frac{a,b}{p}\right) = -1$ is even.

Corollary 4.24. *If* $\left(\frac{a,b}{p}\right) = 1$ *for all* $p \in \mathbb{P} \cup \{\infty\}$, $p \neq q$, *then* $\left(\frac{a,b}{q}\right) = 1$

Proof. Due to the bimultiplicativity it suffices to prove the assertion in the following cases, for $p, q \in \mathbb{P} \setminus \{2\}$, $p \neq q$:

$(a, b) = (-1, -1), (-1, 2), (-1, q), (2, 2), (2, q), (q, q), (q, p)$

Because $\left(\frac{2,2}{p}\right) = \left(\frac{-1,2}{p}\right)$, $\left(\frac{q,q}{p}\right) = \left(\frac{-1,q}{p}\right)$ we are left with 5 cases:

(a,b)	$\left(\frac{a,b}{\infty}\right)$	$\left(\frac{a,b}{2}\right)$	$\left(\frac{a,b}{q}\right)$	$\left(\frac{a,b}{r}\right)$
$(-1,-1)$	-1	-1	$+1$	$+1$
$(-1,2)$	$+1$	$+1$	$+1$	$+1$
$(-1,q)$	$+1$	$(-1)^{\frac{q-1}{2}}$	$\left(\frac{-1}{q}\right)$	$+1$
$(2,q)$	$+1$	$(-1)^{\frac{q^2-1}{8}}$	$\left(\frac{2}{q}\right)$	$+1$
(q,r)	$+1$	$(-1)^{\frac{q-1}{2}\frac{r-1}{2}}$	$\left(\frac{r}{q}\right)$	$\left(\frac{q}{r}\right)$

□

4.4 Proof of the Theorem of Hasse and Minkowski

THIS SECTION STILL NEEDS TO BE PROOFREAD

Theorem 4.25 (Hasse-Minkowski, rank 3)**.** *For* $a, b \in \mathbb{Q}^\times$ *the equation*

$$ax^2 + by^2 = z^2 \tag{4.2}$$

has solutions in $\mathbb{Q}$ *if and only if it has solutions in* $\mathbb{Q}_p$, *for all* $p \in \mathbb{P} \cup \{\infty\}$

Proof. The direct implication is obvious. We will prove the converse, we have solutions in $\mathbb{Q}_p$ for all p and we wish to prove that there are rational solutions.

We may assume that $a, b \in \mathbb{Z}$ and are squarefree, as we know that $\left(\frac{a,b}{p}\right) = \left(\frac{ac^2,ad^2}{p}\right)$ (that is, the Hilbert symbol only depends on the residue classes of a and b in $\mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2$).

We will use induction with respect to $|a| + |b|$.

If $|a| + |b| = 2$ then, because of symmetry, we only have three cases:

- $x^2 + y^2 = z^2$, having $(1, 0, 1)$ as a solution
- $x^2 - y^2 = z^2$, having $(1, 1, 0)$ as a solution

- $-x^2 - y^2 = z^2$ not solvable in $\mathbb{R}$.

If $|a| + |b| > 2$, let $|a| \leq |b|$ so we have $|b| \geq 2$

We first prove that a is a quadratic residue modulo b.

We have $b = \pm \prod_i p_i$, so it suffices to prove that a QR p_i for all i.

For $p_i = 2$ or $p_i | a$ it is clear.

For $p_i \neq 2$ and $p_i \not| a$ we have, due to $p_i | b$: $1 = \left(\frac{a,b}{p_i}\right) = \left(\frac{a}{p_i}\right)$, so a QR p_i.

Therefore we have a QR b. Then we can write $a + bc = t^2$ with $|t| \leq \left|\frac{b}{2}\right|$ and $|c| < |b|$.

If $c = 0$ then $a = t^2$ and the equation (4.2) has the solution $(1, 0, t)$.

If $c \neq 0$ then for all $p \in \mathbb{P} \cup \{\infty\}$ we have $1 = \left(\frac{a,b}{p}\right)$ from the hypothesis. We also have $\left(\frac{a,bc}{p}\right) = \left(\frac{a,t^2-a}{p}\right) = 1$, because the equation $ax^2 + (t^2 - a)y^2 = z^2$ has the solution $(1, 1, t)$.

So by multiplicativity of the Hilbert symbol we have $\left(\frac{a,c}{p}\right) = 1$ for all $p \in \mathbb{P} \cup \{\infty\}$

Because $|a| + |c| < |a| + |b|$ we deduce from the induction hypothesis that there exist $(x_1, y_1, z_1) \in \mathbb{Z}^3 \setminus \{0\}$ so that

$$ax_1^2 + cy_1^2 = z_1^2$$

If $y_1 = 0$ then $(x_1, 0, z_1)$ is a solution for

$$ax^2 + by^2 = z^2$$

If $y_1 \neq 0$ then without loss of generality we can take $y_1 = 1$, so

$$\begin{cases} c = z_1^2 - ax_1^2 \\ bc = t^2 - a \cdot 1^2 \end{cases}$$

We use the norm trick and get $b = z_2^2 - ax_2^2$, so $(x_2, 1, z_2)$ is a solution of (4.2). □

Theorem 4.26 (The Hasse-Minkowski theorem)**.** *Let $a_1, \ldots, a_r \in \mathbb{Q}^\times$, $r \geq 2$.*

Then the quadratic ecuation

$$Q(X) := a_1 X_1^2 + \ldots a_r X_r^2$$

has a solution in $\mathbb{Q}$ if and only if it has a solution in $\mathbb{Q}_p$ for all $p \in \mathbb{P} \cup \{\infty\}$

Remark 4.27. The theorem is true in general for quadratic polynomials $Q(X) = Q(X_1, \ldots, X_n)$

Proof. We will use induction with respect to r and Dirichlet's theorem on arithmetic progressions: For $a, m \in \mathbb{N}$, $gcd(a, m) = 1$ there are infinitely many prime numbers p so that $p \equiv a \pmod{m}$.

For $r = 2$: $a_1X_1^2 + a_2X_2^2 = 0$ if and only if $\left(\frac{X_1}{X_2}\right)^2 = -\frac{a_2}{a_1}$

But $-\frac{a_2}{a_1}$ is in $(\mathbb{Q}^\times)^2$ if and only if $-\frac{a_2}{a_1} \in (\mathbb{Q}_p^\times)^2$ for all p, according to corollary 4.7.

For $r = 3$: $a_1X_1^2 + a_2X_2^2 + a_3X_3^2 = 0$ if and only if $-\frac{a_1}{a_3}X_1^2 + -\frac{a_2}{a_3}X_2^2 = X_3^2$

The statement follows by the preceding theorem.

For $r \geq 4$: We can assume without loss of generality that $a_1, \ldots, a_r \in \mathbb{Z}$ and are squarefree and we can write:

$Q(X) = a_1X_1^2 + \ldots a_rX_r^2 = f - g$

with

- $f(X_1, X_2) = a_1X_1^2 + a_2X_2^2$
- $g(X_3, \ldots, X_r) = -a_3X_3^2 - \ldots - a_rX_r^2$

So we may assume that $a_1, \ldots, a_r > 0$

MISSING TWO LINES

The idea is to determine $b \in \mathbb{Q}$ so that $f(X_1, X_2) = g(X_3, \ldots, X_r) = b$ has solutions in $\mathbb{Q}$.

Let S be the set of odd primes p so that $p|a_1 \ldots a_r$.

For $p \in S \cup \{2\}$ let $X^{(p)} = (X_1^{(p)}, \ldots, X_r^{(p)}) \in \mathbb{Z}_p \setminus \{0\}$ be a p-adic solution for $Q(X) = 0$

Put $b_p := f(X_1^{(p)}, X_2^{(p)}) = g(X_3^{(p)}, \ldots, X_r^{(p)})$

By the Chinese Remainder Theorem there exists a $b_0 \in \mathbb{Z}$ with

- $b_0 > 0$
- $b_0 \equiv b_2 \pmod{16}$
- $b_0 \equiv b_p \pmod{p^2}$ for all $p \in S$

Then for $m := 16 \prod_{p \in \mathbb{P}} p$ all the $b \in B := \{b_0 + km | k \in \mathbb{N}_0\}$ fulfill the condition. Because $gcd(b_0, m) = 1$, through the Dirichlet theorem of arithmetic progressions there exists a prime number $q \in B$.

$$\text{We choose } F(Y, X_1, X_2) = -qY^2 + f(X_1, X_2)$$
$$G(Z, X_3, \ldots, X_r) = -qZ^2 + g(X_3, \ldots, X_r)$$

We want to prove that $F() = 0$ and $G() = 0$ have solutions in $\mathbb{Q}_p$ for all $p \in \mathbb{P} \cup \{\infty\}$

For $p = \infty$ this is clear.

For $p \in S$: $f(X_1^{(p)}, X_2^{(p)}) = bp \equiv q \pmod{p}$, so $\frac{f(X_1^{(p)}, X_2^{(p)})}{q} \equiv 1 \pmod{p}$, so by the Hensel lemma there exists a $Y^{(p)}$, $(Y^{(p)})^2 \equiv \frac{f(X_1^{(p)}, X_2^{(p)})}{q} \pmod{p}$ and $-(Y^{(p)})^2 + f(X_1^{(p)}, X_2^{(p)}) = 0$.

So $(Y^{(p)}, X_1^{(p)}, X_2^{(p)})$ is a solution of $F = 0$.

In the same way $G = 0$ is solvable in $\mathbb{Q}_p$.

For $p = 2$ we reason as for $p \in S$.

If $p \notin S \cup \{2\}$ then the coefficients of F, G are in $\mathbb{Z}_p^\times$, so $F, G = 0$ solvable in $\mathbb{Q}_p$.

From the induction hypothesis $F = G = 0$ is solvable in $\mathbb{Q}$.

MISSING LINES

So there exist $X_1, \ldots, X_r \in \mathbb{Q}$ so that $f(X_1, X_2) \equiv q \equiv g(X_3, \ldots, X_r)$.

Thus $Q(X_1, \ldots, X_n) = 0$, $X \neq 0$

□

Chapter 5

Elliptic curves

5.1 Motivation: rational points on plane curves

In the previous two chapters we were concerned with homogenous quadratic Diophantine equations. We saw that the crucial case to consider were equations in three variables, e.g. the equation

$$ax^2 + by^2 = z^2, \tag{5.1}$$

where $a, b \in \mathbb{Q}^\times$. If we dehomogenize a homogenous equation, we end up with a nonhomogenous equation in one variable less. For instance, a solution of (5.1) with $z \neq 0$ may be normalized to $z = 1$. Hence if we restrict attention to solutions with $z \neq 0$ we may replace (5.1) by the nonhomogenous equation

$$ax^2 + by^2 = 1. \tag{5.2}$$

Depending on the signs of a, b the set of real solutions of (5.2) is either empty, an ellipse or a parabola.

Our two main results about the set of rational solutions to (5.2) were:

(a) If there is one rational solution, then there are infinitely many solutions, parametrized by rational functions in one parameter t (Theorem 2.5).

(b) The existence of a rational solutions is equaivalent to the existence of solutions in the real numbers and in all p-adic number fields $\mathbb{Q}_p$ (Theorem 4.25).

What happens if we look at equations of higher degree, e.g. cubic equations? The general answer is that both (a) and (b) fail to be true. For instance, we will see later that the cubic equation

$$y^2 = x^3 + 1 \tag{5.3}$$

has exactly 5 rational solutions, namely $(x, y) = (-1, 0), (0, \pm 1), (2, \pm 3)$. This means that (a) is not true for cubic equations. Also, a famous example due to Selmer is the cubic equation

$$3x^3 + 4y^3 + 5z^3 = 0. \tag{5.4}$$

It can be shown that (5.4) has a solution in $\mathbb{R}$ and $\mathbb{Q}_p$, for all primes p, but no solution in $\mathbb{Q}$. So (b) is false as well for cubic equations.

5.2 Plane curves

Let us fix a field k. For any $n \in \mathbb{N}$ we denote by

$$\mathbb{A}^n_k := k^n$$

the *affine space* of dimension n over k. In these lectures we will be exclusively interested in the cases $n = 1, 2$. We call $\mathbb{A}^1_k = k$ the *affine line* and $\mathbb{A}^2_k = k^2$ the *affine plane* over k. An element $P = (a, b) \in \mathbb{A}^2_k$ is called a (k-rational) *point*.

Let $k[x, y]$ denote the ring of polynomials in two variables x, y over k. A polynomial $f \in k[x, y]$ gives rise to a function

$$f : \mathbb{A}^2_k \to k, \qquad P = (a, b) \mapsto f(P) := f(a, b).$$

We will typically not make any formal distinction between a polynomial and its associated function.[1]

Definition 5.1. A subset $X \subset \mathbb{A}^2_k$ is called an *affine plane curve* if there exists a nonconstant polynomial $f \in k[x, y]$ such that

$$X = \{P \in \mathbb{A}^2_k \mid f(P) = 0\}.$$

We will often write this as

$$X : \ f(x, y) = 0.$$

[1] This is potentially dangerous: if k is a finite field, then the polynomial f is not uniquely determined by its associated function!

Example 5.2. A *line* is given by a linear equation:

$$L: \ ax + by + c = 0.$$

Here $a, b, c \in k$ and $(a, b) \neq (0, 0)$. A *quadric* is given by a quadratic equation

$$Q: \ a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6 = 0,$$

with $a_1, \ldots, a_6 \in k$ and $(a_1, a_2, a_3) \neq (0, 0, 0)$.

Remark 5.3. Definition 5.1 is a bit problematic for several reasons. For instance, if $k = \mathbb{R}$ then the empty set $X(\mathbb{R}) = \emptyset$ is an algebraic curve (a quadric) because it is the set of solutions to the equation

$$x^2 + y^2 + 1 = 0.$$

This is certainly not what one understands by an algebraic curve! There are several ways to give better definitions. See e.g. the lecture notes [5] or any book on algebraic geometry. For the present course we may ignore this problems because we will only treat a limit number of special cases where they do not cause trouble.

Definition 5.4. The *projective space* of dimension n over k is the set

$$\mathbb{P}^n_k := \big(k^{n+1} - \{(0, \ldots, 0)\}\big)/\sim,$$

where $\sim$ is the equivalence relation defined by

$$(a_1, \ldots, a_{n+1}) \sim (ta_1, \ldots, ta_{n+1}), \qquad \text{for } t \in k^\times.$$

The equivalence class of $(a_1, \ldots, a_{n+1}) \in k^{n+1} - \{0\}$ is denoted by

$$P = [a_1 : \ldots : a_{n+1}] \in \mathbb{P}^n_k$$

and called a *point* with *projective coordinates* $a_1, \ldots, a_{n+1}$.

We will only be concerned with the cases $n = 1, 2$. We call $\mathbb{P}^1_k$ the *projective line* and $\mathbb{P}^2_k$ the *projective plane*.

By definition, a point $[a : b] \in \mathbb{P}^1_k$ corresponds to a line in the affine plane through the origin:

$$[a : b] \mathrel{\widehat{=}} \{\, (ta, tb) \in \mathbb{A}^2_k \mid t \in k\}.$$

This line has nonzero slope if and only if $b \neq 0$. If this is the case, then $[a : b] = [a/b : 1]$. If the line has zero slope then $b = 0$ and $[a : b] = [1 : 0]$. We see that the map

$$\mathbb{A}^1_k \hookrightarrow \mathbb{P}^1_k, \qquad a \mapsto [a : 1] \tag{5.5}$$

is injective, and the complement of its image consists of a single point. We write suggestively $\infty := [0 : 1]$. If we identify the affine line $\mathbb{A}^1_k = k$ with its image under the map (5.5) then we can write

$$\mathbb{P}^1_k = \mathbb{A}^1_k \cup \{\infty\}.$$

This suggests that the projective line is a *compactification* of the affine line. Indeed, for $k = \mathbb{R}$ or $\mathbb{C}$ one can check that $\mathbb{P}^1_k$ carries a natural topology such that $\mathbb{P}^1_k$ is compact and $\mathbb{A}^1_k$ is a dense open subset.

Similarly, the projective plane can be seen as a compactification of the affine plane. We have a natural embedding

$$\mathbb{A}^2_2 \hookrightarrow \mathbb{P}^2_k, \qquad (a, b) \mapsto [a : b : 1].$$

The complement consists of all points of the form $[a : b : 0]$, where $(a, b) \neq (0, 0)$. Note that this set can be identified with the projective line. It is called the *line at infinity.*

Definition 5.5. A subset $X \subset \mathbb{P}^2_k$ is called a *projective plane curve* if there exists a homogenous polynomial

$$F = \sum_{i+j+k=d} a_{i,j,k}\, x^i y^j z^k$$

of degree $d > 0$ such that

$$X = \{[a : b : c] \mid F(a, b, c) = 0\}.$$

We often write this as:

$$X : \ F(x, y, z) = 0.$$

Note that the condition $F(a, b, c) = 0$ depends only on the class $[a : b : c]$ because F is homogenous.

Example 5.6. A subset $L \subset \mathbb{P}^2_k$ is called a *(projective) line* if it is defined by a linear equation,

$$L : \ ax + by + cz = 0.$$

Here $(a, b, c) \neq (0, 0, 0)$. If, moreover, $(a, b) \neq 0$, then the intersection of L with the affine plane $\mathbb{A}^2_k \subset \mathbb{P}^2_k$ is a line in the usual sense, given by the equation

$$ax + by + c = 0.$$

If $a = b = 0$, then we may assume that $c = 1$, and L is the line at infinity, given by the equation $z = 0$.

Proposition 5.7. Let $L_1, L_2 \subset \mathbb{P}^2_k$ be two projective lines. Then either $L_1 = L_2$, or L_1 and L_2 intersect in a unique point.

Proof. For $i = 1, 2$ the line L_i is given by an equation

$$L_i : \ a_i x + b_i y + c_i z = 0.$$

The intersection of L_1 and L_2 consists of the points $[x : y : z] \in \mathbb{P}^2_k$ where $(x, y, z) \in k^3 - \{(0, 0, 0)\}$ is a nontrivial solution of the system of linear equations

$$a_1 x + b_1 y + c_1 z = 0,$$
$$a_2 x + b_2 y + c_2 z = 0.$$

By linear algebra we know that the set of solutions is a linear subspace of k^3 of dimension 1 or 2. Moreover, if the dimension is two then the two equations are linearly dependent and each of them define the full set of solutions. This would mean that $L_1 = L_2$. So if $L_1 \neq L_2$ then the set of solutions is of the form

$$\{ (tx_0, ty_0, tz_0) \mid t \in k\},$$

where $(x_0, y_0, z_0) \neq (0, 0, 0)$. It follows that the intersection

$$L_1 \cap L_2 = \{[x_0 : y_0 : z_0]\}$$

consists of a single point. □

Remark 5.8. The proposition says that in the projective plane parallel lines do not exist. To reconcile this with our geometric intuition (which is trained on working with the affine plane) we consider the following example. Let

$$L_1 : \ 2x - y = 1, \qquad L_2 : \ 2x - y = 2,$$

be two lines in the affine plane (with coordinates x, y). Obviously, L_1 and L_2 are parallel, i.e. they do not intersect. However, L_1, L_2 can be extended uniquely to lines in the projective plane,

$$\tilde{L}_1 : \ 2x - y - z = 0, \qquad \tilde{L}_2 : \ 2x - y - 2z = 0.$$

The intersection $\tilde{L}_1 \cap \tilde{L}_2$ is computed by the system of linear equations

$$2x - y - z = 0,$$
$$2x - y - 2z = 0.$$

Its solution set is one dimensional and is spanned by the point $(1, 2, 0)$. It follows that the lines $\tilde{L}_1$ and $\tilde{L}_2$ intersect in the point $[1 : 2 : 0]$ on the 'line at infinity'.

So far, we have identified the affine plane with the subset of the projective plane defined by the condition $z = 1$. Surely, the choice of the variable z is somewhat arbitrary. To be more impartial we define subsets

$$U_x, U_y, U_z \subset \mathbb{P}^2_k$$

by the conditions $x = 1$, $y = 1$ and $z = 1$. Then we have natural identifications

$$U_x \cong \mathbb{A}^2_k, \quad [1 : y : z] \mapsto (y, z),$$
$$U_y \cong \mathbb{A}^2_k, \quad [x : 1 : z] \mapsto (x, z),$$
$$U_z \cong \mathbb{A}^2_k, \quad [x : y : 1] \mapsto (x, y).$$

Accordingly, we call U_x the *y-z-plane*, U_y the *x-z-plane* and U_z the *x-y-plane*. Note that

$$\mathbb{P}^2_k = U_x \cup U_y \cup U_z.$$

Let $X \subset \mathbb{P}^2_k$ be a projective curve defined by the equation

$$X : \ F(x, y, z) = 0$$

where $F \in k[x, y, z]$ is homogenous of degree d. Then

$$X_x := X \cap U_x = \{[1 : y : z] \mid F(1, x, y) = 0\}$$

may be considered, via the identification $U_x \cong \mathbb{A}^2_k$, as an affine plane curve, defined by the equation

$$X_x : \ F(1, y, z) = 0.$$

Similarly, $X_y := X \cap U_y$ is defined by $F(x, 1, z) = 0$ and $X_z := X \cap U_z$ is defined by $F(x, y, 1) = 0$.

5.3 The group law on an elliptic curve

Definition 5.9. Let K be a field of characteristic $\neq 2, 3$ and $a, b \in K$. We consider the projective curve $E = E_{a,b}$ with equation

$$E: \ y^2 z = x^3 + axz^2 + bz^3.$$

We have $E \subseteq \mathbb{P}^2_K$ and we say that it is in *Weierstrass normal form*.

Remark 5.10. (i) The affine curve $E \cap \mathbb{A}^2_K$ (where $\mathbb{A}^2_K$ is the (x, y)-plane) is given by the dehomogenized equation

$$y^2 = f(x) := x^3 + ax + b.$$

It uniquely determines the projective curve $E \subseteq \mathbb{P}^2_K$. We will mostly write down the affine equation above to define a curve E in Weierstrass normal form. But actually, we will always mean by E the *projective* curve given by the homogenized equation.

(ii) The unique point on E "at infinity" is $\mathcal{O} := [0 : 1 : 0]$. This is a smooth point of E. To see this, we consider the the affine (x, z)-plane, where $y = 1$. The affine equation for E is then

$$E: \ z = x^3 + axz^2 + bz^3,$$

and the point $\mathcal{O}$ is the origin, $\mathcal{O} = (0, 0)$. It is now an easy exercise to show that $\mathcal{O}$ is a smooth point of E, and that the tangent to E at $\mathcal{O}$ is the line $z = 0$ (formerly known as the line at infinity). Note, however, that the proof uses the assumption $\mathrm{char}(K) \neq 2$!

(iii) The curve E is smooth if and only if the polynomial $f = x^3 + ax + b$ has no double root. By basic algebra this holds if and only if the discriminant of f is not zero, that is

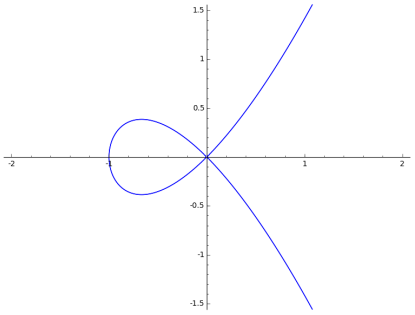
$$\Delta(f) = -4a^3 - 27b^2 \neq 0.$$

You should prove this as an exercise. You will have to use the assumption that $\mathrm{char}(K) \neq 2$.

Example 5.11. (i) Consider the curve E/K with Weierstrass equation

$$E: \ y^2 = x^3 + x^2.$$

For $K = \mathbb{R}$, its affine part looks like this:

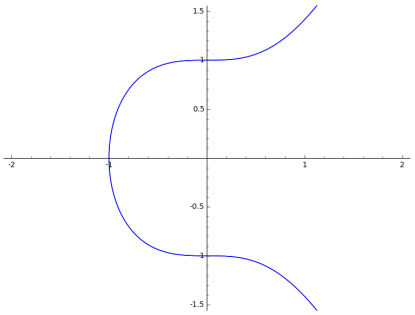


One checks that $f(0) = f'(0) = 0$, so $(0, 0)$ is a singular point, confirming the picture. But note that the algebraic calculation shows that E is singular over any field K (whereas the picture is only meaningful for $K = \mathbb{R}$).

(ii) Consider the curve E/K in Weierstrass equation

$$E: \ y^2 = x^3 + 1.$$

The discriminant of $f = x^3 + 1$ is $\Delta = -27 \neq 0$ (because we assume that $\mathrm{char}(K) \neq 3$!), and hence E/K is smooth. For $K = \mathbb{R}$ we get the following (affine) picture of E:



Definition 5.12. The curve $E : y^2 = f(x) = x^3 + ax + b$ is called an *elliptic curve over* K if E is smooth (i.e. if $\Delta = -4a^3 - 27b^2 \neq 0$).

Let us fix an elliptic curve E/K. For the moment, let us also assume that K is algebraically closed. We want to define a binary operation $\oplus : E \times E \to E$ (called *addition*) such that $(E, \oplus)$ an abelian group with neutral element $\mathcal{O}$.

The first key point is the following lemma.

Lemma 5.13. Let $E \subset \mathbb{P}^2_K$ be an elliptic curve over an algebraically closed field K, and let $L \subset \mathbb{P}^2_K$ be a projective line. Then there are exactly three points of intersection of E and L, *counted with multiplicity*. More precisely, there are three cases as follows.

(a) The intersection $E \cap L$ consists of three pairwise distinct points P, Q, R. We write this as

$$E \cap L = P + Q + R.$$

(b) The intersection $E \cap L$ consists of two distinct points P, Q, and for exactly one of them, say P, L is the tangent to E at P. We write this as

$$E \cap L = 2 \cdot P + Q.$$

(c) The intersection $E \cap L$ consists of a unique point P. If this is the case, then L is the tangent to E at P, and P is an *inflection point* of E. We write this as

$$E \cap L = 3 \cdot P.$$

Proof. This is a special case of *Bezout's Theorem*, see [3], Appendix A.3-4. See also Construction 5.18 below. □

Construction 5.14. Let $P, Q \in E$. If $P \neq Q$ then we let $L \subseteq \mathbb{P}^2_k$ be the unique projective line through P and Q. If $P = Q$ then we let $L := T_{E,P}$ be the tangent to E at $P = Q$. Then by Lemma 5.13 there exists a unique point $R \in E$ such that

$$E \cap L = P + Q + R.$$

We write R' for the point symmetric to R with respect to the x-axis. Explicitely,

$$R' := \begin{cases} (x, -y), & \text{if } R = (x, y) \in \mathbb{A}^2_K, \\ \mathcal{O}, & \text{if } R = \mathcal{O}. \end{cases}$$

Note that R' is the unique point such that

$$E \cap L' = R + R' + \mathcal{O}.$$

We define

$$P \oplus Q := R'.$$

See Example 5.15 and Figure 5.1 for an example.

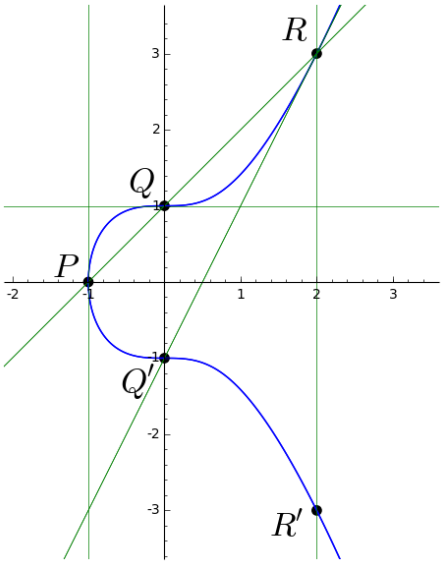


Figure 5.1: Addition on the elliptic curve $E:\ y^2 = x^3 + 1$.

Example 5.15. Let us consider the elliptic curve $E:\ y^2 = x^3 + 1$ over $\mathbb{Q}$. We immediately find two rational points

$$P := (-1, 0), \quad Q := (0, 1)$$

on E. Let us compute $P \oplus Q$. The unique line L_1 through P and Q is given by the affine equation $L_1:\ x - y + 1 = 0$. To compute the third intersection point R we have to solve the system of equations

$$y^2 = x^3 + 1,$$

$$y = x - 1.$$

Eliminating y we obtain the following cubic equation in x:

$$x^3 - x^2 + 2x = x(x + 1)(x - 2) = 0.$$

It has three solutions $x = -1, 0, 2$, corresponding to the x-coordinate of the three points of intersection of E with L_1. Hence

$$E \cap L = P + Q + R, \qquad \text{where } R := (2, 3).$$

It follows that

$$P \oplus Q = R' = (2, -3).$$

As an exercise, you should now compute $P \oplus P$, $Q \oplus Q$ and $R \oplus Q'$.

Remark 5.16. The binary operation $\oplus : E \times E \to E$ defined above is commutative and has $\mathcal{O}$ as a neutral element. Moreover, for all $P \in E$ the point P' (the mirror image of P after reflection at the x-axis) is an inverse to P with respect to $\oplus$. These claims follow easily from the construction of $\oplus$ and the fact that $\mathcal{O}$ is an inflection point of E.

Theorem 5.17. *The binary operation $\oplus$ is associative, and hence $(E, \oplus)$ is an abelian group with neutral element $\mathcal{O}$.*

Proof. See e.g. [5], §1.7. □

Construction 5.18. In what follows we will work out an explicit formula for the group law on the elliptic curve

$$E : \ y^2 = f(x) = x^3 + ax + b.$$

Let $P, Q \in E$ be arbitrary points. If one of them is the neutral element $\mathcal{O}$, say $Q = \mathcal{O}$, then $P \oplus Q = P \oplus \mathcal{O} = P$. We may therefore assume that $P, Q \neq \mathcal{O}$, and write $P = (x_1, y_1)$ and $Q = (x_2, y_2)$.

Let L denote the line through P and Q (or the tangent in P, in the case $P = Q$). Also, let R be the third point of intersection, such that

$$E \cap L = P + Q + R.$$

It suffices to compute R, because $P \oplus Q = R'$, by definition. If $Q = P'$, i.e. if $x_1 = x_2$, $y_1 = -y_2$, then $R = \mathcal{O}$. Otherwise, $R \neq \mathcal{O}$, and we can write $R = (x_3, y_3)$. Moreover, the line L has finite slope and may be written as

$$L : \ y = \lambda x + \nu,$$

with $\lambda, \nu \in K$. Explicitly, we have

$$\lambda = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1}, & \text{if } P \neq Q, \\ \dfrac{-f'(x_1)}{2y_1}, & \text{if } P = Q, \end{cases}$$

and

$$\nu = y_1 - \lambda x_1.$$

We calculate $E \cap L$ by substituting $y = \lambda x + \nu$ into the equation for E. We obtain a cubic equation in x with three roots $x = x_1, x_2, x_3$ (compare Example 5.15):

$$\begin{aligned} & (\lambda x + \nu)^2 = x^3 + ax + b \\ \Leftrightarrow \quad & x^3 - \lambda^2 x^2 + (a - 2\lambda\nu)x + b - \nu^2 = 0 \\ \Leftrightarrow \quad & (x - x_1)(x - x_2)(x - x_3) = 0. \end{aligned}$$

If we expand the product in the third line and compare the coefficients of x^2 with the same coefficient in the second line we obtain the identity

$$x_1 + x_2 + x_3 = \lambda^2.$$

This gives an explicit formula for x_3 in terms of $P = (x_1, y_1)$ and $Q = (x_2, y_2)$. An explict formula for y_3 is then obtained from the identity $y_3 = \lambda x_3 + \nu$. We have proved:

Lemma 5.19. Let $P = (x_1, y_2)$, $Q = (x_2, y_2)$ be affine points on the elliptic curve $E: \ y^2 = x^3 + ax + b$. Then

$$P \oplus Q = \begin{cases} (x_3, y_3), & \text{if } x_1 \neq x_2 \text{ or } P = Q, \\ \mathcal{O}, & \text{if } x_1 = x_2 \text{ and } y_2 = -y_1, \end{cases}$$

where

$$x_3 := \lambda^2 - x_1 - x_2, \qquad y_3 := \lambda x_3 + \nu.$$

Here

$$\lambda := \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1}, & \text{if } x_2 \neq x_1, \\ \dfrac{-3x^2 - a}{2y_1}, & \text{if } x_1 = x_2, \end{cases}$$

and

$$\nu := y_1 - \lambda x_1.$$

Example 5.20. We come back to the elliptic curve $E: y^2 = x^3 + 1$ from Example 5.15. Let $P = (-1, 0)$ and $Q = (0, 1)$. You should check again, using Lemma 5.19, that

$$P \oplus Q = R' = (2, -3).$$

You will then also find that

$$2 \cdot P := P \oplus P = \mathcal{O},$$

so P is a *point of order* 2,

$$3 \cdot Q := Q \oplus Q \oplus Q = \mathcal{O},$$

i.e. Q is a point of order 3 (because it is an inflection point), and

$$2 \cdot R = Q.$$

It follows that P is a point of order 6, i.e. it generates a cyclic subgroup of order 6 of E:

$$\langle P \rangle_{\mathbb{Z}} = \{\mathcal{O},\, R,\, 2 \cdot R = Q,\, 3 \cdot R = P,\, 4 \cdot R = R',\, 5 \cdot R = Q'\}.$$

Note that all these points are $\mathbb{Q}$-rational. We will prove later that these are in fact *all* $\mathbb{Q}$-rational points on E.

Example 5.21. $E/\mathbb{Q} : y^2 = x^3 - 2$
$P = (3, 5)$, $-P = (3, -5)$
$2P = \ldots$

5.4 Points of finite order

5.5 The Theorem of Mordell-Weil

Theorem 5.22. *Let $E/\mathbb{Q}$ be an elliptic curve over $\mathbb{Q}$, $\Gamma = E(\mathbb{Q})$ the group of rational points. Then Γ is finitely generated, that is, there exist $P_1, \ldots, P_N \in \Gamma$ so that $\Gamma =< P_1, \ldots, P_n >_{\mathbb{Z}}$*

Remark 5.23. (i) This is the Mordell theorem and has a more general form, proven by Weil, wherein we replace $\mathbb{Q}$ by an extension $K/\mathbb{Q}$ and E by an abelian variety A over K.

If $dim(A) = 1$ we get back the case for elliptic curves.

(ii) Γ is a finitely generated abelian group and can be decomposed into its free part and its torsion part: $\Gamma = \Gamma_{free} \oplus \Gamma_{tor}$

$\Gamma_{free} \cong < P_1, \ldots, P_n >_{\mathbb{Z}} \cong \mathbb{Z}^r$

$r := rang(\Gamma)$ =Mordell-Weil Rank of $E/\mathbb{Q}$

Proof. in Silverman-Tate (with an additional hypothesis $E[2](\mathbb{Q}) \neq \{0\}$) □

Definition 5.24. (i) For $x = \frac{m}{n} \in \mathbb{Q}$, $gcd(m,n) = 1$ we note $H(x) := max(|m|, |n|)$ the *height* of x and $h(x) := \log H(x)$ the *logarithmic height*. $h(x)$ can be interpreted as the number of bits needed to represent x

(ii) For $P = (x,y) \in \Gamma = E(\mathbb{Q}) \setminus \{0\}$ define $h(P) := h(x)$ the *height* of P and by convention $h(\mathcal{O}) = 0$

Lemma 5.25. For $c > 0$ we have the number of points whose height is upper bounded by c is finite, that is

$$| \left\{ P \in \Gamma \mid h(P) < c \right\} | < \infty$$

Proof. The number of rational numbers $\frac{a}{b}$ of (logarithmic) height smaller than p is finite (because there is a finite number of possibilities for a and b).

Because $h(P) = h(x)$ and once we choose x we have two possibilities for y, this means that

$$| \left\{ P \in \Gamma \mid h(P) < c \right\} | < \infty$$

□

Lemma 5.26. Let $P_0 \in \Gamma$ be a fixed point. Then there exists a constant $k_0 = k_0(E, P_0)$ so that for all $P \in \Gamma$ we have

$$h(P + P_0) \leq 2h(P) + k_0$$

Lemma 5.27. There exists a constant $k = k(E)$ so that for all $P \in \Gamma$

$$h(2P) \geq 4h(P) - k$$

(i.e. $h : \Gamma \to \mathbb{R}_{\geq 0}$ is "almost quadratic")

Lemma 5.28 (Weak form of the Mordell-Weil theorem)**.**

$$|\Gamma / 2\Gamma| < \infty$$

This means that there exist $Q_1, \ldots Q_m \in \Gamma$ so that for every $P \in \Gamma$ there is a $Q \in \Gamma$ and $i \in \{1, 2, \ldots, m\}$ so that $P = 2Q + Q_i$.

We will prove that these four lemmas imply the theorem.

Proposition 5.29. Let Γ be an abelian group, $h : \Gamma \to \mathbb{R}_{\geq 0}$ so that

(a) $|\{P \in \Gamma \mid h(P) < c\}| < \infty$ for all $c \in \mathbb{R}_{>0}$

(b) for all $P_0 \in \Gamma$ there exists a constant k_0 so that $h(P + P_0) \leq 2h(P) + k_0$

(c) there exists a constant k so that for all $P \in \Gamma$: $h(2P) \geq 4h(P) - k$

(d) $|\Gamma/2\Gamma| < \infty$

Then Γ is finitely generated.

Proof. We will use a descent argument.

Let $Q_1, \ldots, Q_m$ be a system of representatives for $\Gamma/2\Gamma$ (according to (d)).

For $P \in \Gamma$ there exists a sequence $P_1, P_2, \ldots$, $i_1, i_2, \ldots \in \{1, \ldots, m\}$

$P = Q_{i_1} + 2P_1$

$P_1 = Q_{i_2} + 2P_2$

$\ldots$

$P_{m-1} = Q_{i_m} + 2P_m$

So $P = Q_{i_1} + 2Q_{i_2} + 4Q_{i_3} + \ldots + 2^{m-1}Q_{i_m} + 2^m P_m$ for all $m \in \mathbb{N}$.

It suffices to prove that there exists a constant $c > 0$ so that $h(P_m) < c$ for a $m \in \mathbb{N}$

In this case Γ will be generated by $\{Q_1, \ldots, Q_m\} \cup \{P \in \Gamma \mid h(P) \leq c\}$

Proof. We apply (b):

$h(P - Q_i) \leq 2h(P) + k_i$ for all $P \in \Gamma$, $i \in \{1, \ldots, m\}$

We note $k' := \max_i k_i$.

From (c) and $2P_j = P_{j-1} - Q_{i_j}$ we have $4h(P_j) \leq h(2P_j) + k = h(P_{j-1} - Q_{i_j}) + k \leq 2h(P_{j-1}) + k' + k$

We note $k'' := k' + k$

So $h(P_j) \leq \frac{1}{2}h(P_{j-1}) + \frac{k''}{4} = \frac{3}{4}h(P_{j-1}) - \frac{1}{4}(h(P_{j-1}) - k'')$

So for $h(P_{j-1}) \geq k'' =: c$ we have $h(P_j) \leq \frac{3}{4}h(P_{j-1})$

We can't have $h(P_j) > c$ for all $j \in \mathbb{N}$ because $h(P_j) \leq \left(\frac{3}{4}\right)^j h(P) \to 0$

So there exists an m so that $h(P_m) \leq c$. □

□

We will see the proof of lemma 5.28 in the following section.

For now, we will give a sketch of the proof of lemmas 5.26 and 5.27, that can be found in section III.2. of [3]

Lemma 5.30. There exists a constant $k = k(E)$ so that for all $P \in \Gamma$

$$h(2P) \geq 4h(P) - k$$

Proof. We may assume that $P \notin E[2]$, so that $2P \neq 0$.

If $P = (x, y)$ and $2P = (\xi, \eta)$, where

$\xi = \frac{f'(x)^2 - (8x+4a)f(x)}{4f(x)} = \frac{x^4+\ldots}{4x^3+\ldots} = \frac{\phi(X)}{\psi(X)}$

with $\phi, \psi \in \mathbb{Z}[X]$, $\deg(\phi) = 4$, $\deg(\psi) = 3$, $gcd(\phi, \psi) = 1$ (in $\mathbb{C}[X]$)

Lemma 5.27 follows from the following:

Lemma 5.31. Let $\phi, \psi \in \mathbb{Z}[X]$, $gcd(\phi, \psi) = 1$, $d = \max(deg(\phi), deg(\psi))$.

(a) there exists an $R = R(\phi, \psi) \in \mathbb{Z} \setminus \{0\}$ so that for all $x = \frac{m}{n} \in \mathbb{Q}$

$$gcd\left(n^d \phi\left(\frac{m}{n}\right), n^d \psi\left(\frac{m}{n}\right)\right) | R$$

(b) there exist $k_1 = k_1(\phi, \psi), k_2 = k_2(\phi, \psi)$ so that for all $x = \frac{m}{n} \in \mathbb{Q}$ with $\psi(x) \neq 0$,

$$dh(x) - k_1 \leq h\left(\frac{\phi(x)}{\psi(x)}\right) \leq dh(x) + k_2$$

Proof. We note

$\Phi(X, Y) = Y^d \phi(\frac{X}{Y})$

$\Psi(X, Y) = Y^d \psi(\frac{X}{Y})$

For $x = \frac{m}{n}$ we have $n^d \phi(x) = \Phi(m, n)$, $n^d \psi(x) = \Psi(m, n)$

Because $gcd(\phi, \psi) = 1$ there exist $F, G \in \mathbb{Z}[X]$:

$$F\phi + G\psi = A \in \mathbb{Z} \setminus \{0\}$$

Let $D := \max(deg(F), deg(G))$ and

$\Phi(m, n) = a_0 m^d + a_1 m^{d-1} n + \ldots$

Then $n^D F(\frac{m}{n})\Phi(m, n) + n^D G(\frac{m}{n})\Psi(m, n) = A n^{D+d}$ for all $x = \frac{m}{n} \in \mathbb{Q}$

$\gamma := gcd(\Phi(m, n), \Psi(m, n)) | A n^{D+d}$

Then γ divides $A n^{D+d-1} \Phi(m, n) = A a_0 m^d n^{D+d-1} + \ldots$

So $\gamma | A a_0 m^d n^{D+d-1}$ and finally $\gamma | R$

□

□

Remark 5.32. The proof of the Mordell-Weil theorem is not effective, i.e. it does not give a constant $c > 0$ so that

$$\Gamma =< P_1, \ldots, P_N >_{\mathbb{Z}},\ h(P_i) \leq c$$

There is no known algorithm that computes the Mordell-Weil group Γ for any elliptic curve E.

There exists one, but it works under the hypothesis of the Birch and Swinnerton-Dyer conjecture.

We will now prove the weak version of the Mordell-Weil theorem:

Lemma 5.33 (Weak Mordell-Weil theorem)**.** $|\Gamma/2\Gamma| = (\Gamma : 2\Gamma) < \infty$

In order to simplify the presentation, we will make the assumption that there exists a point $T \in E[2](\mathbb{Q})$, $T \neq 0$. This is equivalent to the existence of a $x_0 \in \mathbb{Q}$ so that $E : y^2 = f(x) = (x - x_0)(x^2 + \ldots)$.

After substituting x by $x + x_0$ we have $x_0 = 0$ and $T := (0, 0)$, which gives $E : y^2 = x^3 + ax^2 + bx$.

In this case $\Delta = b^2(a^2 - 4b)$ and this is nonzero if and only if $b \neq 0$ and $a^2 \neq 4b$.

We consider the map

$[2] : E \to E,\ P \mapsto 2P$

$$P = (x, y) \mapsto \begin{cases} \left(\frac{(x^2-b)^2}{4y^2}, \frac{(x^2-b)(x^4+2ax^3+6bx^2+2abx+b^2)}{8y^3}\right) \text{ if } y \neq 0 \\ \mathcal{O} \text{ if } P \in E[2] \end{cases}$$

This map has the following properties:

(i) $[2]$ is a group homomorphism;

(ii) $[2]$ is surjective ($E = E(\mathbb{C})$);

(iii) $Ker[2] = E[2] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$;

(iv) $Q \in E$, $[2]^{-1} = \{P \in E \mid 2P = \mathcal{O}\}$ has exactly 4 elements.

Proposition 5.34. We define $\overline{E} : y^2 = x^3 + \overline{a}x^2 + \overline{b}x$

where $\overline{a} = -2a$, $\overline{b} = a^2 - 4b$, $\overline{T} = (0, 0)$

(i) The map $\phi : E \to \overline{E}$ $\phi(P) = \begin{cases} \left(\frac{y^2}{x^2}, \frac{y(x^2-b)}{x^2}\right) \text{ if } P = (x,y) \neq \mathcal{O}, T \\ \overline{\mathcal{O}} \text{ if } P = \mathcal{O}, T \end{cases}$ is a group homomorphism with $Ker(\phi) = \{\mathcal{O}, T\}$

(ii) The map $\psi : \overline{E} \to E$ $\psi(\overline{P}) = \begin{cases} \left(\frac{\overline{y}^2}{\overline{x}^2}, \frac{\overline{y}(\overline{x}^2-\overline{b})}{\overline{x}^2}\right) \text{ if } \overline{P} \neq \overline{\mathcal{O}}, \overline{T} \\ \overline{\mathcal{O}} \text{ if } P = \mathcal{O}, T \end{cases}$

(iii) We have $\psi \circ \phi = [2]_E$ and $\phi \circ \psi = [2]_{\overline{E}}$.

Proof. Calculations □

Remark 5.35. The existence of $\phi : E \to \overline{E}$ over $\mathbb{C}$ can be seen in the following way:

$\Delta =< \omega_1, \omega_2 >_{\mathbb{Z}}$

$E \cong \mathbb{C}/\Delta$

$\overline{E} := \mathbb{C}/\overline{\Delta}$, $\overline{\Delta} =< \frac{1}{2}\omega_1, \omega_2 >$

$\phi : E \to \overline{E},\ \phi(z + \Delta) := z + \overline{\Delta}$

$\psi : \overline{E} \to E,\ \phi(z + \overline{\Delta}) := 2z + \Delta$

$Ker(\phi) = \overline{\Delta}/\Delta$

Lemma 5.36. Let A, B be abelian groups and consider group homomorphisms $\phi : A \to B$ and $\psi : B \to A$ so that $\phi \circ \psi = [2]_B$, $\psi \circ \phi = [2]_A$ and the indexes $(B : \phi(A))$ and $(A : \psi(B))$ are finite.

Then $(A : 2A) \leq (A : \psi(A))(B : \phi(A)) < \infty$

Proof. quite easy □

Corollary 5.37. *Note* $\Gamma := E(\mathbb{Q})$ *and* $\overline{\Gamma} := \overline{E}(\mathbb{Q})$

If $(\Gamma : \psi(\overline{\Gamma}))$, $(\overline{\Gamma} : \phi(\Gamma)) < \infty$ *then* $(\Gamma : 2\Gamma) < \infty$

We will prove that $(\Gamma : \psi(\overline{\Gamma})) < \infty$.

$(\overline{\Gamma} : \phi(\Gamma)) < \infty$ *will work the same way.*

Proposition 5.38. 1. The function

$$\begin{aligned} \alpha : \Gamma &\to \mathbb{Q}^\times/(\mathbb{Q}^\times)^2 \\ P = (x,y) &\mapsto \tilde{x} := x(\mathbb{Q}^\times)^2 \text{ for } P \neq \mathcal{O}, \mathcal{T} \\ \mathcal{O} &\mapsto \tilde{1} \\ T = (0,0) &\mapsto \tilde{b} \end{aligned}$$

is a group homomorphism with $Ker(\alpha) = \psi(\overline{\Gamma})$

2. Let $p_1, \ldots, p_r$ be the prime divisors of b.

   Then $\alpha(\Gamma)$ is the subgroup of $\mathbb{Q}^\times/(\mathbb{Q}^\times)^2$ generated by $\pm 1, p_1, \ldots, p_r$

We will prove this proposition in a bit but, first, we will use it to prove that:

Corollary 5.39. $(\Gamma : \psi(\overline{\Gamma})) < \infty$

Proof. $\Gamma/\psi(\overline{\Gamma}) \overset{\overline{\alpha}}{\hookrightarrow} G :=< \pm\tilde{1}, \tilde{p}_1, \ldots, \tilde{p}_r > \subseteq \mathbb{Q}^\times/(\mathbb{Q}^\times)^2$

□

Proof. We can now prove the proposition 5.38.

(i) $\alpha(-P) = \alpha(P) = \alpha(P)^{-1}, \forall P \in \Gamma$.

Let $P_1, P_2, P_3 \in \Gamma$, $P_i = (x_i, y_i)$ with $P_1 + P_2 + P_3 = \mathcal{O}$.

It suffices to show that $\alpha(P_1)\alpha(P_2)\alpha(P_3) = \tilde{1}$.

If $P_1 + P_2 + P_3 = \mathcal{O}$ then P_1, P_2, P_3 are on a line $L : y = \lambda x + \nu$.

This means that x_1, x_2, x_3 are the roots of the equation

$L \cap E : x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x - \nu^2 = 0$

Then $x_1x_2x_3 = \nu^2$, so $\alpha(P_1)\alpha(P_2)\alpha(P_3) = \tilde{x}_1\tilde{x}_2\tilde{x}_3 = \tilde{1}$.

Thus α is a homomorphism.

(ii) $P = (x, y) \in Ker(\alpha)$ if and only if $x \in (\mathbb{Q}^\times)^2$ and we will prove that this is equivalent to $P \in \psi(\overline{\Gamma})$

Because $\psi(\overline{P}) = \begin{cases} \left(\frac{\overline{y}^2}{\overline{x}^2}, \frac{\overline{y}(\overline{x}^2 - \overline{b})}{\overline{x}^2}\right) \text{ if } \overline{P} \neq \overline{\mathcal{O}}, \overline{T} \\ \overline{\mathcal{O}} \text{ if } P = \mathcal{O}, T \end{cases}$

If $x \in (\mathbb{Q}^\times)^2$ then $x = \omega^2 \neq 0$. This gives:

$\overline{x} := \frac{1}{2}(\omega^2 - a + \frac{y}{\omega})$

$\overline{y} := \pm\overline{x}\omega$.

(iii) Let $P = (x, y) \in \Gamma$.

Then $x = \frac{m}{e^2}$, $y = \frac{n}{e^3}$, $m, n, e \in \mathbb{Z}$, $e > 0$, $(m, e) = (n, e) = 1$

Then $n^2 = m(m^2 + ame^2 + be^4)$

Let $p|m$ be a prime divisor. Then $v_p(m) \equiv 0 \pmod 2$ or $p|b$

Then $\alpha(P) \in < \pm\tilde{1}, \tilde{p}_1, \ldots, \tilde{p}_r > \subseteq \mathbb{Q}^\times/(\mathbb{Q}^\times)^2$.

□

Question: Does the proof of the Mordell-Weil theorem provide an algorithm for determining Γ?

Answer: Unfortunately not.

The problem is:

$$\overline{\alpha} : \Gamma/\psi(\overline{\Gamma}) \hookrightarrow \left\{\tilde{b}_1 \in \mathbb{Q}^\times/(\mathbb{Q}^\times)^2 | b_1 | b\right\}$$
$$P = (x, y) \mapsto \tilde{x}, P \neq \mathcal{O}, T$$
$$T = (0, 0) \mapsto \tilde{b}$$
$$\mathcal{O} \mapsto \tilde{1}$$

Let $b_1|b$. Can we decide if there exists a point $P = (x, y) \in \Gamma$ with $\tilde{x} = \tilde{b}_1$ in $\mathbb{Q}^\times/(\mathbb{Q}^\times)^2$ so that $x = b_1 x_1^2, x_1 \in \mathbb{Q}^\times$?

Write $x = \frac{m}{e^2}$, $y = \frac{n}{e^3}$

$n^2 = m(m^2 + ame^2 + be^4)$

We have $b_1 = \pm gcd(m, b)$, $mb_1 > 0$

Then $m = b_1 m_1$, $b = b_1 b_2$, $gcd(m', b_2) = 1$, $m_1 > 0$

Then $n^2 = b_1^2 m_1(b_1 m_1^2 + am_1 e^2 + b_2 e^4)$

$m_1 = M^2$, $n = b_1 MN$, so $N^2 = b_1 M^4 + aM^2 e^2 + b_2 e^4$ $(*)$

The proof shows that

$\mathrm{Im}(\alpha) = \left\{\tilde{x} \in \mathbb{Q}^\times/(\mathbb{Q}^\times)^2 | x = b_1 \frac{M^2}{e^2}, (N, M, e) \text{ is a solution for } (*), M \neq e\right\}$

Bibliography

[1] K. Ireland and M. Rosen. *A Classical Introduction to Modern Number Theory*. Springer, 2. edition, 1990.

[2] S. Müller-Stach and J. Piontkowski. *Elementare und algebraische Zahlentheorie*. Viehweg, 2006.

[3] J.H. Silverman and J. Tate. *Rational points on elliptic curves*. Springer, 1992.

[4] S. Singh. *Fermat's Last Theorem*. Fourth Estate Ltd, 1997.

[5] S. Wewers. Einführung in die Algebraische Geometrie. Vorlesungsskript, SS 15.

[6] S. Wewers. Lineare Algebra II. Vorlesungsskript, SS 13.