A Cyclotomic Proof of Catalan's Conjecture

Jeanine Daems

29th September 2003

Contents

Chapter 1

Introduction

Recently, Catalan's conjecture, one of the famous classical problems in number theory, has been proven. This means that within ten years after Wiles' proof of Fermat's last theorem, another classical diophantine equation has been proven to have no "non-trivial" solutions. This time the proof is due to Preda Mihăilescu, so we might say that now Catalan's conjecture has become Mihăilescu's theorem. Catalan's conjecture is not very difficult to understand: it says that the difference between two perfect powers (where we ignore 0 and 1) is always more than 1, unless these powers are equal to $8 = 2^3$ and $9 = 3^2$.

Suppose we have two perfect powers that are only 1 apart, then there are also two perfect powers with prime exponents that are only 1 apart. (If, for instance, x^8 and y^{15} differ by 1, then $(x^4)^2$ and $(y^3)^5$ are powers with prime exponents that differ by 1.) It follows that it suffices to prove the following theorem.

Theorem 1.1 (Mihăilescu). *Let p and q be prime numbers. Then the equation*

$$x^p - y^q = 1 \tag{1.1}$$

has no solutions in positive integers x and y, other than $3^2 - 2^3 = 1$.

The aim of this thesis is to give a proof of theorem 1.1. Before we start, we will present some of the history of Catalan's conjecture, in particular the history of the results we will use. Mihăilescu's proof uses the fact that we may reduce the theorem to the case with odd prime exponents. The cases in which one of p and q is 2 had been treated before. We will deal with these cases in chapter 3.

In the chapters 4 and 5 we give the setting in which the proofs in the subsequent chapters will take place, and we prove or formulate some preliminary results. For instance, we will show that for any solution of equation (1.1) with p and q odd primes, we have that q^2 divides x. Because the proof in chapter 6 only works for primes p and q that are at least 5 and the proof in chapter 8 uses that p and q are at least 7, we still need to take care of the cases in which one of p and q is smaller than 7. In section 5.3 we show that for these exponents there exist no solutions to the Catalan equation.

After that, we will make a separation in cases. Without loss of generality, we may

assume that $p > q$. The first case then is the case in which q does not divide $p-1$, the second case is the case in which q divides $p-1$.

The case in which q does divide $p-1$, had already been solved using results of Baker, Tijdeman, Mignotte and Roy. The final part of this proof consisted of electronic computations that exclude a certain number of possible exponents. Mihăilescu has now found a new proof of this case, using algebraic number theory. Computations on computers are no longer needed. We will give R. Schoof's version of Mihăilescu's proof of this case in chapter 6.

We will also give the proof of the second case, in which q does not divide $p-1$. This part of the proof had been found by Mihăilescu before he found the new proof of the first case. We will give H.W. Lenstra, Jr.'s simpler version of it. Further, in chapter 7 we will prove one of the theorems we use, using Runge's method. Finally, in chapter 8 we give the main argument that brings all these ingredients together. It starts by assuming that there exists a solution of equation (1.1) for odd prime exponents p and q that are at least 7 and eventually derives a contradiction from this.

Leiden, 29th September 2003

Chapter 2

History

In this chapter we present some of the history of Catalan's conjecture. The conjecture has been open for more than 150 years and a fair number of people have made efforts to solve it. This chapter is based on the survey of the history of Catalan's conjecture in Ribenboim's book [14], but we give some more details on certain developments and we will concentrate on the people who contributed to the proof as it stands now.

The story of Catalan's conjecture starts in the year 1844, when Crelle's Journal [4] published an extract from a letter from the Belgian mathematician Eugène Charles Catalan (1814–1894) to the editor. The extract was the following.

Note

extraite d'une lettre adressée à l'éditeur par Mr. *E. Catalan*,
Répétiteur à l'école polytechnique de Paris.

Je vous prie, Monsieur, de vouloir bien énoncer, dans votre recueil, le théorème suivant, que je crois vrai, bien que je n'aie pas encore réussi à le démontrer complètement: d'autres seront peut-être plus heureux:

> Deux nombres entiers consécutifs, autres que 8 et 9, ne peuvent être des puissances exactes; autrement dit: l'équation $x^m - y^n = 1$, dans laquelle les inconnues sont entières et positives, n'admèt qu'une seule solution.

According to Dickson [7] p.731, this was not the first time that people thought about this subject, but this was the first time the conjecture was stated in this general form. Philippe de Vitry (1291–1361), who is better known as a composer and music theorist than as a mathematician, posed the question as follows: all powers of 2 and 3 differ by more than unity except the pairs 1 and 2, 2 and 3, 3 and 4, 8 and 9. Levi ben Gerson (1288–1344), who was also known as Gersonides, solved the problem by proving that $3^m \pm 1$ always has an odd prime factor if $m > 2$, so $3^m \pm 1$ can not be a power of 2. Euler solved the equation $x^2 - y^3 = 1$; already in 1738 he showed [8] that the only positive solution is $x = 3, y = 2$. In chapter 3 we give a modern version of Euler's proof, as well as his own version. We will show that these proofs are essentially the same.

Only six years after Catalan's publication, the first result on the question he posed appeared in print. The French mathematician Victor Amédée Lebesgue (1791–1875) showed [10] that the equation $x^p - y^2 = 1$, where p is a prime number, has no solutions in positive integers x and y. He used Gaussian integers to do this. We will give a proof of his theorem in chapter 3. This proof is essentially the same proof as Lebesgue's, but we give it in a more modern way. Note that this Lebesgue is *not* the same person as the much better-known mathematician Henri Léon Lebesgue (1875–1941), after whom the Lebesgue measure on the real numbers is named.

At the end of his article, Lebesgue says that the other cases of the equation $x^m = y^n + 1$ seem to present more difficulties and that he does not know what Mr. Catalan has found on the subject so far. But Catalan had not found much. The only results he ever found [5] were not published until 1885. By this time he had become a Professor at the University of Liege in Belgium. In this article he tells us about the time when he was trying to prove his conjecture, and how hard it turned out to be:

> Après avoir perdu près d'une année à la recherche d'une démonstration qui fuyait toujours, j'abandonnai cette recherche fatigante.

He only made some empirical observations, which he stated without proof, hoping that other people might find them useful. The observations he mentions all are special cases of the general conjecture, for example the equations $(x+1)^x - x^x = 1$, $x^y - y^x = 1$ and $x^p - q^y = 1$, where p and q are prime.

After Lebesgue's result, for some time all progress consisted in dealing with the small exponents. Nagell showed in 1921 that the difference between a third power and an other perfect power never is equal to 1. In 1932, Selberg proved that $x^4 - y^n = 1$ has no solution in positive integers when $n > 1$. We do not need this result in this thesis, however, because in 1965 Ko Chao [9] showed that the equation $x^2 - y^q = 1$ has no solutions in positive integers when $q \geq 5$, which is of course stronger than Selberg's result. In 1976 Chein [6] gave a simpler proof of Chao's theorem, using that if $x^2 - y^q = 1$ with q prime and $x \geq 1$, $y \geq 1$, then 2 divides y and q divides x. This also is a result of Nagell. The proof of H.W. Lenstra Jr. that we will give in chapter 3 is even simpler, since it does not use this result of Nagell.

The next result that did not just deal with small exponents was achieved by Le Veque in 1952 [11]. He looked at the number of solutions of the Catalan equation and showed that the equation $x^a - y^b = 1$ has at most one solution for given integers x and y, unless $x = 3, y = 2$, in which case there are exactly two.

In 1953 [2] and 1960 [3] Cassels published some findings on the equation $x^p - y^q = 1$, where p and q are odd primes. He proved that if this equality holds for positive integers x and y, then p divides y and q divides x. For the case $p = 2$ this had already been shown by Nagell. In this thesis we will derive an even stronger result using Cassels' findings, namely that q^2 divides x. This has been shown by Mihăilescu. From Cassels' theorem it almost immediately follows that three consecutive integers cannot be perfect powers, as A. Mąkowski showed in a very short article in 1962 [12].

Hyyrő also worked on the Catalan conjecture. He sharpened Cassels' results in 1964, when he gave several congruence relations that hold for integers x and y greater than 1 and primes p and q such that $x^p - y^q = 1$. What is useful for our purpose is that he obtained a large lower bound of the absolute value of x. However, we follow

Bilu's approach [1] and we derive a weaker lower bound: $|x| \geq q^{p-1}$. In both the case in which q does divide $p-1$ and the case in which q does not divide $p-1$ we use this lower bound.

Baker's theory on effective bounds for solutions of certain types of diophantine equations applies to the Catalan equation. In 1976, Tijdeman [17] used Baker's theory and he showed that there is an effectively computable upper bound on the sizes of p, q, x and y, where p and q are primes and x and y are integers such that $x^p - y^q = 1$. Of course, this implies that Catalan's equation only has a finite number of solutions. More developments of this analytic approach followed, but we will not go further into these.

Inkeri defined the concept of a *Wieferich pair* in the context of the Catalan equation as follows: a Wieferich pair is a pair (p, q) of primes such that $p^{q-1} \equiv 1 \pmod{q^2}$ and $q^{p-1} \equiv 1 \pmod{p^2}$. In 1990, he showed that if the Catalan equation (1.1) holds, then either (p, q) is a Wieferich pair, or q divides h_p, the class number of the cyclotomic field $\mathbb{Q}(\zeta_p)$, or p divides h_q, the class number of $\mathbb{Q}(\zeta_q)$. There were more developments in this direction also. Bugeaud and Hanrot [19], for instance, proved a class number criterion concerning Catalan's equation, which implies that the Catalan equation $x^p - y^q = 1$ has no solution in non-zero integers x and y if p and q are primes such that one of them is smaller than 43. Our proof in section 5.3 looks like their proof. Finally, Mihăilescu [13] succeeded in eliminating the class number criteria by showing that if equation (1.1) holds, then (p, q) is a Wieferich pair. In our thesis we will see this in corollary 5.8 and we use it in the proof of the case in which q divides $p-1$. This result rules out many pairs of exponents p and q.

Recently, Mihăilescu proved that the Catalan equation (1.1) has no solutions if p and q are odd and q does not divide $p-1$. By this result the Catalan conjecture became a theorem. And this year Mihăilescu succeeded in finding a more elegant proof of Catalan's conjecture in the case where q does divide $p-1$. So now Catalan's conjecture is a theorem with an algebraic proof in which no computer calculations are needed.

Chapter 3

Even exponents

Mihăilescu's proof of Catalan's conjecture deals with the cases in which both exponents are odd primes, as the cases in which one of the exponents is equal to 2 had been dealt with earlier. There are two cases to consider, namely the case that $q = 2$ and the case that $p = 2$. In this chapter we will give proofs that in both cases no "non-trivial" solutions to Catalan's equation exist.

3.1 The case $q = 2$: Victor Lebesgue

In chapter 2 we saw that the case $q = 2$ has been dealt with by the French mathematician V.A. Lebesgue in 1850. He proved that there are no solutions in positive integers to the equation $x^p - y^2 = 1$, where p is prime.

Theorem 3.1 (Lebesgue). *Let p be a prime number. Then the equation*

$$x^p - y^2 = 1$$

has no solutions in non-zero integers x and y.

Proof. Let p be a prime. Suppose there exists a solution of $x^p - y^2 = 1$ such that x and y are non-zero integers. If $p = 2$, then the relation $x^p - y^2 = 1$ implies that 1 is the difference of the two squares x^p and y^2, so the only solution we find here has $y = 0$, which we excluded from the beginning. So we may assume p to be odd.

Suppose that x is even, then we obtain $4|x^p$. Then we find that $y^2 \equiv 3 \pmod 4$, which of course leads to a contradiction. So x is odd and it immediately follows that y has to be even.

In the ring $\mathbb{Z}[i]$ we have $x^p = y^2 + 1 = (y - i)(y + i)$. It is known that $\mathbb{Z}[i]$ is a unique factorisation domain. Now there is no prime $\pi \in \mathbb{Z}[i]$ such that $\pi | y - i$ and $\pi | y + i$. For suppose there is such a prime, then it has to divide $y + i - (y - i) = 2i$, so π divides 2, since i is a unit. It follows that 2 divides x, which we have proven to be impossible. We may conclude that all primes of $\mathbb{Z}[i]$ that divide x, divide exactly one of $y + i$ or $y - i$. It follows that, up to units, $y + i$ and $y - i$ are both p-th powers in $\mathbb{Z}[i]$. Since p is odd, all units in $\mathbb{Z}[i]$ are p-th powers, so there are $a, b \in \mathbb{Z}$ such that $y - i = (a + bi)^p$ and $y + i = \overline{y - i} = (a - bi)^p$.

Using Newton's Binomial Theorem, we can write

$$y - i = (a + bi)^p = \sum_{j=0}^{p} \binom{p}{j} a^j (bi)^{p-j}.$$

Taking the imaginary part at both sides, we get

$$-1 = \sum_{j=0}^{\frac{p-1}{2}} \binom{p}{2j} a^{2j} b^{p-2j} i^{p-2j-1} = b \sum_{j=0}^{\frac{p-1}{2}} \binom{p}{2j} a^{2j} b^{p-2j-1} i^{p-2j-1}. \tag{3.1}$$

So b divides -1, which yields $b = \pm 1$.

Now we know that $x^p = (a+bi)^p(a-bi)^p = (a+i)^p(a-i)^p = (a^2+1)^p$, so $x = a^2+1$ and a is even. It is obvious that a can not be equal to 0, because if this were the case we would find the trivial solution $x = 1, y = 0$, which we already excluded.

Going back to (3.1), we have

$$-1 = b \sum_{j=0}^{\frac{p-1}{2}} \binom{p}{2j} a^{2j} (-1)^{\frac{p-1}{2}-j} = b(-1)^{\frac{p-1}{2}} \sum_{j=0}^{\frac{p-1}{2}} \binom{p}{2j} (-a^2)^j.$$

From this we obtain

$$\sum_{j=0}^{\frac{p-1}{2}} \binom{p}{2j} (-a^2)^j = \pm 1.$$

Viewing this equality modulo 4, all terms with $j > 0$ vanish as a is even, so if we take the sum modulo 4 we get 1. It follows that at the right-hand side we also have 1. Therefore, $\sum_{j=1}^{\frac{p-1}{2}} \binom{p}{2j}(-a^2)^j = 0$. We find

$$a^2 \binom{p}{2} = \sum_{j=2}^{\frac{p-1}{2}} \binom{p}{2j} (-a^2)^j. \tag{3.2}$$

Note that this equation also holds if $p = 3$, then it says that $a^2\binom{p}{2} = 0$, which implies that $a = 0$, so then $y = 0$, which we excluded.

From equation (3.2) we will derive a 2-adic contradiction. Define $v_2(\alpha) = \mathrm{ord}_2(\alpha)$ for $\alpha \in \mathbb{Q}^*$. If two numbers are equal, they have the same number of factors 2. We will show that

$$v_2\left(a^2 \binom{p}{2}\right) < v_2\left(\sum_{j=2}^{\frac{p-1}{2}} \binom{p}{2j} (-a^2)^j\right). \tag{3.3}$$

If we have proved this, we are done.

We will compare the number of factors 2 in each term of the sum with the number of factors 2 on the left-hand side. Let k be an integer greater than 1. We start by writing $\frac{\binom{p}{2k}}{\binom{p}{2}}$ in a different manner:

$$\frac{\binom{p}{2k}}{\binom{p}{2}} = \frac{2!(p-2)!}{(2k)!(p-2k)!} = \binom{p-2}{2k-2} \frac{2}{2k(2k-1)} = \binom{p-2}{2k-2} \frac{1}{k(2k-1)}.$$

As $\binom{p-2}{2k-2}$ is integral and $2k-1$ is odd,

$$v_2\left(\frac{\binom{p}{2k}}{\binom{p}{2}}\right) \geq v_2(\frac{1}{k}) = -v_2(k). \tag{3.4}$$

Since a is even,

$$v_2(a) \geq 1. \tag{3.5}$$

Further, $v_2(k) < 2k-2$ since $k < 2^{2k-2}$ for all integers $k > 1$. It follows that

$$2k - 2 - v_2(k) > 0. \tag{3.6}$$

If we put (3.4), (3.5) and (3.6) together, we come to the following conclusion:

$$v_2\left(\frac{a^{2k}\binom{p}{2k}}{a^2\binom{p}{2}}\right) \geq (2k-2)v_2(a) - v_2(k) \geq 2k - 2 - v_2(k) > 0.$$

It follows that $v_2(a^{2k}\binom{p}{2k}) > v_2(a^2\binom{p}{2})$ for all integers $k > 1$, which implies (3.3). This is what we wanted to prove. □

3.2 The case $p = 2$ and $q \geq 5$: Ko Chao

In 1965 Ko Chao [9] proved that if $q \geq 5$ is prime, then the equation $x^2 = y^q + 1$ has no solutions in positive integers x and y. In 1976 a simpler proof of Chao's result was given by E.Z. Chein [6]. Here we will give the proof by H.W. Lenstra, Jr., which is somewhat different from Chein's proof. In his proof, Chein uses Nagell's result that if $x^2 = y^q + 1$ holds, with q prime and $x, y \geq 1$, then 2 divides y and q divides x. Lenstra does not need this.

Theorem 3.2 (Ko Chao). *Let $q \geq 5$ be a prime number. Then there are no positive integers x and y such that $x^2 = y^q + 1$.*

It is sufficient to look at the solutions with $x, y > 0$, because if there is a solution with x negative, then $-x > 0$ also gives a solution, and if y would be negative, then $x^2 - y^q \geq 2$, which is impossible.

Now we are going to prove theorem 3.2. First we start by proving two lemmas.

Lemma 3.3. *If $a, b \in \mathbb{Z}$, $\gcd(a, b) = 1$ and p is a prime number, then*

$$\gcd(\frac{a^p - b^p}{a - b}, a - b) = 1 \text{ or } p.$$

Proof. Note that

$$\frac{a^p - b^p}{a - b} = \sum_{i=0}^{p-1} a^i b^{p-1-i} \equiv pb^{p-1} \pmod{a-b} \equiv pa^{p-1} \pmod{a-b}.$$

It follows that $\gcd(\frac{a^p-b^p}{a-b}, a-b)$ divides both pb^{p-1} and pa^{p-1}. Since a and b are coprime, we find that $\gcd(\frac{a^p-b^p}{a-b}, a-b) | p$. So $\gcd(\frac{a^p-b^p}{a-b}, a-b) = 1$ or p, because p is prime. □

Lemma 3.4. *If $a, b, c \in \mathbb{Z}$ and p and q are prime numbers not both equal to 2, $\gcd(a, b) = 1$, $p \nmid c$ and $a^p - b^p = c^q$, then $a - b$ is a q-th power.*

Proof. We already saw that $\frac{a^p-b^p}{a-b} = \sum_{i=0}^{p-1} a^i b^{p-1-i} \in \mathbb{Z}$, so $a - b$ divides $a^p - b^p$. Therefore, we can write

$$c^q = \frac{a^p - b^p}{a - b}(a - b).$$

First, we will show that the factors $\frac{a^p-b^p}{a-b}$ and $a - b$ are coprime. Suppose there exists a prime r such that $r|a - b$ and $r|\frac{a^p-b^p}{a-b} = \frac{c^q}{a-b}$. According to lemma 3.3 $\gcd(\frac{a^p-b^p}{a-b}, a - b) = 1$ or p. If it is 1, then there does not exist an r as above. If it is p, then $p|\frac{a^p-b^p}{a-b} = \frac{c^q}{a-b}$, so $p|c$, which is not true by assumption.

If q is odd, then -1 is a q-th power, so all units in $\mathbb{Z}$ (i.e. 1 and -1) are q-th powers. Then we are done, since it follows that both $\frac{a^p-b^p}{a-b}$ and $a - b$ are q-th powers.

We are left with the case $q = 2$, so p is not equal to 2 by assumption. Suppose that $a - b$ is not a q-th power, i.e. a square. Then we have $a - b = -d^2$ for an integer $d \neq 0$. (We may assume $d \neq 0$, since $a - b = 0$ is a square.) It follows that $\frac{a^p-b^p}{a-b} = -(\frac{c}{d})^2$, so $\frac{a^p-b^p}{a-b} \leq 0$. If $\frac{a^p-b^p}{a-b} = 0$, then the only solution is $a = b = 1$ and $a - b = 0$ is a square. If $\frac{a^p-b^p}{a-b} < 0$, then either $a - b$ or $a^p - b^p$ is negative, but not both. But since p is odd, $a - b$ and $a^p - b^p$ are both positive or both negative. So this case does not occur.

It follows that $a - b$ is a q-th power. □

Now we are ready to prove theorem 3.2.

Proof of theorem 3.2. Assume that x and y are positive integers and $q \geq 5$ is a prime such that $x^2 = y^q + 1$. Then $y^q = x^2 - 1 = (x + 1)(x - 1)$. Suppose that x is even, then y is odd, so $2 \nmid y$. Now we use lemma 3.4 and we find that $x - 1$ is a q-th power. But then $x + 1$ is also a q-th power, since $(x - 1)(x + 1) = c^q$. Let $x - 1 = s^q$ and $x + 1 = t^q$. So now we have $t^q - s^q = 2$ for some $s, t \in \mathbb{Z}$ and $q \geq 5$. So the only solution we find is $t = 1, s = -1$. But this implies $x = 0$, which leads to a contradiction. So x is odd and y is even.

Therefore, $\gcd(x + 1, x - 1) = \gcd(x + 1, 2) = 2$. Let $\varepsilon \in \{-1, 1\}$ be such that $x \equiv \varepsilon \pmod 4$. Then $2||x + \varepsilon = 2w^q$ and $2^{q-1}|x - \varepsilon = 2^{q-1}z^q$ for $w, z \in \mathbb{Z}$ such that $\gcd(w, 2z) = 1$, since the only prime that divides both $x + 1$ and $x - 1$ is 2. Now $y = 2wz$. (Note that until now, we have not yet used that $q \geq 5$; q is odd suffices.)

We know that $x + \varepsilon = 2w^q$ and $x - \varepsilon = 2^{q-1}z^q$ for some integers w and z. It follows that $(\frac{w}{z})^q = 2^{q-2}\frac{x+\varepsilon}{x-\varepsilon} > 1$ since $q \geq 5$. (So here we use that $q > 3$. Further, we have used that $x \neq 1$, but we are allowed to use that because $x = 1$ only yields a solution with $y = 0$. If we would not want to use here that $q \neq 3$, then we would have to exclude the case in which x is equal to 3 separately.) It follows that $w > z$, and $w^2 - 2\varepsilon z$ is not a square, because $|2\varepsilon z| = 2|z| < 2w$ and it is even, so $|2\varepsilon z| < 2w - 1$.

So $w^{2q} - (2\varepsilon z)^q = (\frac{x-\varepsilon}{2} + \varepsilon)^2 - 4\varepsilon\frac{x-\varepsilon}{2} = (\frac{x-\varepsilon}{2} - \varepsilon)^2 = (\frac{x-3\varepsilon}{2})^2$.

Assume that $q \nmid \frac{x-3\varepsilon}{2}$ and apply lemma 2. Then it follows that $w^2 - 2\varepsilon z$ is a square, which leads to a contradiction with what we have seen before. So $q|\frac{x-3\varepsilon}{2}$. Therefore, $q|x - 3\varepsilon$, so $x \equiv 3\varepsilon \pmod q$ and $x \not\equiv 0 \pmod q$, since $3\varepsilon = \pm 3$ and $q \geq 5$. Here we use $q \geq 5$ in an essential way. We find that q does not divide x.

We may conclude now that $x^2 = y^q - (-1)^q$ with $q \nmid x$ and $\gcd(y, -1) = 1$. Lemma 3.4 now says that $y+1$ is a square, $y+1 = s^2$, say. Now we have the two following relations:

$$s^2 - y \cdot 1^2 = 1 \tag{3.7}$$

$$x^2 - y(y^{\frac{q-1}{2}})^2 = 1. \tag{3.8}$$

Note that these equations give two different solutions to the Pell equation

$$u^2 - yv^2 = 1. \tag{3.9}$$

The solution $(s, 1)$ is a fundamental solution, so there exists $m \in \mathbb{Z}$ such that

$$x + y^{\frac{q-1}{2}}\sqrt{y} = (s + \sqrt{y})^m \text{ in } \mathbb{Z}[\sqrt{y}]. \tag{3.10}$$

It follows that $x \equiv s^m + ms^{m-1}\sqrt{y} \,(\mathrm{mod}\, y\mathbb{Z}[\sqrt{y}])$. So $ms^{m-1}\sqrt{y} \in \mathbb{Z} + y\mathbb{Z}[\sqrt{y}]$, so $ms^{m-1} \equiv 0 \,(\mathrm{mod}\, y)$. Since y is even and therefore s is odd, it follows that m is even, say $m = 2n$.

Taking (3.10) modulo s, we have $x + y^{\frac{q-1}{2}}\sqrt{q} \equiv \sqrt{y}^m = y^n \,(\mathrm{mod}\, s\mathbb{Z}[\sqrt{y}])$, so $y^{\frac{q-1}{2}}\sqrt{y} \in \mathbb{Z} + s\mathbb{Z}[\sqrt{y}]$, so $y^{\frac{q-1}{2}} \equiv 0 \,(\mathrm{mod}\, s\mathbb{Z}[\sqrt{y}])$. Since $s^2 = y+1$, $y \equiv -1 \,(\mathrm{mod}\, s)$. Therefore, $1 \equiv 0 \,(\mathrm{mod}\, s\mathbb{Z}[\sqrt{y}])$, so $s = 1$ ($s > 0$ because $x - \varepsilon > 0$). This leads to the solutions $y = 0, x = \pm 1$, but we assumed that $y \neq 0$. The conclusion is now that there are no solutions to the equation $x^2 = y^q + 1$ with $x, y \in \mathbb{Z}_{>0}$ and $q \geq 5$ a prime, which is what we wanted to prove. □

Now we still are left with the case in which $q = 3$. But Ko Chao had no need to look at this case, because it had already been dealt with by Euler in 1738. We will give a proof in the next section.

3.3 The case $p = 2$ and $q = 3$: Euler

Theorem 3.5 (Euler). *If x and y are positive rationals such that $x^2 = y^3 + 1$, then $x = 3$ and $y = 2$.*

We will give a modern proof of this theorem, using the theory of elliptic curves. After that, we will show that the proof Euler gave 265 years ago is essentially the same as this modern proof.

In modern terminology, Euler finds the points with positive rational coordinates on the elliptic curve $D : y^2 = x^3 + 1$. Let us view these points as affine points of the projective curve $y^2z = x^3 + z^3$. Together with the point $\mathcal{O} = (0 : 1 : 0)$ at infinity the affine points form an additive group with unit element $\mathcal{O}$. Even though it is an open problem to exhibit an algorithm that is guaranteed to find generators for the group of rational points on an elliptic curve, it can be done in most special cases. In the present case, Euler's theorem is implied by the following result.

Theorem 3.6. *The group of rational points on the elliptic curve $y^2 = x^3 + 1$ is a cyclic group of order 6 with elements*

$$\{(-1, 0), (0, 1), (0, -1), (2, 3), (2, -3), \mathcal{O}\},$$

where $\mathcal{O}$ denotes the point at infinity.

Proof. We assume some familiarity with the theory of elliptic curves over $\mathbb{Q}$, in particular the treatment of Silverman and Tate in chapter III of [15].

By the Mordell-Weil theorem, the group of rational points on an elliptic curve E is a finitely generated abelian group, i.e. it is of the form

$$E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus T, \tag{3.11}$$

with $T = E(\mathbb{Q})_{\text{tors}}$, the torsion subgroup of $E(\mathbb{Q})$, which is a finite abelian group. The number r is called the *rank* of the elliptic curve. The aim of the proof is to show that the rank of our elliptic curve in theorem 3.6 is 0, using a 2-descent, as finding T is easy.

The most important ingredient of the proof of the Mordell-Weil theorem is that the index $[E(\mathbb{Q}) : mE(\mathbb{Q})]$ is finite, for $m \in \mathbb{Z}$. In our case we choose $m = 2$. Let us assume this and let Q_1, Q_2, ..., Q_n be representatives for the cosets of $2E(\mathbb{Q})$. Then for any P in $E(\mathbb{Q})$, we can write

$$P - Q_i = 2P', \tag{3.12}$$

for some $i = 1, \ldots, n$ and for a point $P' \in E(\mathbb{Q})$. Now we can do the same with P', and so on. The basic idea is that the 'size' of the points $P, P', P'', \ldots$ we get in this way becomes smaller in every step.

There is a common notion of the size of a point in the case of elliptic curves, namely the height of a point. First, we define the height of a rational number. Let $x = \frac{v}{w}$ be a rational number written in lowest terms. Then the height $H(x)$ of x is defined as follows:

$$H(x) = H(\frac{v}{w}) = \max\{|v|, |w|\}.$$

We define the height of a point to be the height of the x-coordinate of the point.

Following the procedure indicated above, we always arrive at a point $P^{(j)}$, for some integer j, such that the height of $P^{(j)}$ is smaller than a certain given integer κ. Since there is only a finite number of points with height smaller than a given integer, it follows that all points in $E(\mathbb{Q})$ are generated by the finite set

$$\{Q_1, \ldots, Q_n\} \cup \{R \in E(\mathbb{Q}) : H(R) \leq \kappa\},$$

for some integer κ.

The standard algorithm that we use for our problem uses more details from the proof of the Mordell-Weil theorem, as it can be found in [15]. The homomorphism α we use below, for instance, plays an important part in the proof of the finiteness of the index $[E(\mathbb{Q}) : 2E(\mathbb{Q})]$.

Computing $E(\mathbb{Q})/2E(\mathbb{Q})$ is relatively easy if $E(\mathbb{Q})$ has a rational 2-torsion point. Assume, after a coordinate change $(x, y) \mapsto (x + e, y)$, that E is an elliptic curve given by the equation

$$E : y^2 = x^3 + ax^2 + bx$$

and construct the curve E' as follows: let it be given by the equation

$$E' : y^2 = x^3 + \bar{a}x^2 + \bar{b}x,$$

where $\bar{a} = -2a$ and $\bar{b} = a^2 - 4b$.

Then there is an isogeny $\varphi : E \to E'$ defined by

$$\varphi(P) = \begin{cases} (\frac{y^2}{x^2}, \frac{y(x^2-b)}{x^2}) & \text{if } P = (x,y) \neq \mathcal{O}, (0,0); \\ \overline{\mathcal{O}} & \text{if } P = \mathcal{O} \text{ or } P = (0,0). \end{cases}$$

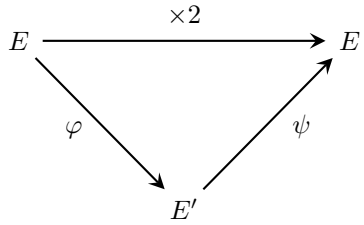
The kernel of φ is $\{\mathcal{O}, (0,0)\}$.

Similarly, construct the curve E'' from E' and define the map $\overline{\varphi} : E' \to E''$ similar to φ. The curve E'' is isomorphic to E via the map $(x,y) \to (\frac{x}{4}, \frac{y}{8})$. There is thus a dual isogeny $\psi : E' \to E$ defined by

$$\psi(\overline{P}) = \begin{cases} (\frac{\bar{y}^2}{4\bar{x}^2}, \frac{\bar{y}(\bar{x}^2-\bar{b})}{8\bar{x}^2}) & \text{if } \overline{P} = (\bar{x}, \bar{y}) \neq \overline{\mathcal{O}}, (\bar{0}, \bar{0}); \\ \mathcal{O} & \text{if } \overline{P} = \overline{\mathcal{O}} \text{ or } \overline{P} = (\bar{0}, \bar{0}). \end{cases}$$

The composition $\psi \circ \varphi : E \to E$ is multiplication by 2: $\psi \circ \varphi(P) = 2P$ for all points P in $E(\mathbb{Q})$.

The following diagram displays the situation we have.



Consider the elliptic curve D given by the Weierstrass equation

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

First, we change coordinates in such a way that we move the rational 2-torsion point $(-1, 0)$ to the origin $(0, 0)$. In these new coordinates the equation becomes

$$E : y^2 = x(x^2 - 3x + 3).$$

Let $\mathcal{O}$ denote the point on E at infinity. It is obvious that the group of rational points $E(\mathbb{Q})$ of this new elliptic curve is isomorphic to the group of rational points of D.

In our case the curve E' is defined by the equation

$$E' : y^2 = x(x^2 + 6x - 3).$$

In addition to all this we need the following map. Define the map $\alpha : E(\mathbb{Q}) \to \mathbb{Q}^*/\mathbb{Q}^{*^2}$ by

$$\begin{aligned} \alpha(\mathcal{O}) &= 1 \pmod{\mathbb{Q}^{*^2}} \\ \alpha(0,0) &= b \pmod{\mathbb{Q}^{*^2}} \\ \alpha(x,y) &= x \pmod{\mathbb{Q}^{*^2}} \quad \text{if } x \neq 0. \end{aligned}$$

The map α is a group homomorphism. It can be shown easily that the kernel of α equals the image of $\psi(E'(\mathbb{Q}))$. Therefore, α induces an injective homomorphism

$$E(\mathbb{Q})/\psi(E'(\mathbb{Q})) \hookrightarrow \mathbb{Q}^*/\mathbb{Q}^{*^2}.$$

We define the map $\bar{\alpha} : E'(\mathbb{Q}) \to \mathbb{Q}^*/\mathbb{Q}^{*^2}$ in the same way.

It follows that the image of α is isomorphic to $E(\mathbb{Q})/\psi(E'(\mathbb{Q}))$. Therefore, the index $[E(\mathbb{Q}) : \psi(E'(\mathbb{Q}))]$ is equal to $\#\alpha(E(\mathbb{Q}))$. Similarly, we find that the index $[E'(\mathbb{Q}) : \phi(E(\mathbb{Q}))]$ equals $\#\bar{\alpha}(E'(\mathbb{Q}))$.

Let us look at the quotient group $E(\mathbb{Q})/2E(\mathbb{Q})$. According to (3.11) this group is of the form

$$E(\mathbb{Q})/2E(\mathbb{Q}) \cong (\mathbb{Z}/2\mathbb{Z})^r \oplus T/2T \cong (\mathbb{Z}/2\mathbb{Z})^r \oplus T[2],$$

where $T[2]$ is the 2-torsion part of T. Therefore,

$$[E(\mathbb{Q}) : 2E(\mathbb{Q})] = 2^r \cdot \#T[2]. \tag{3.13}$$

For a 2-torsion point (x, y) we have $y = 0$, so we have $x(x^2 - 3x + 3) = 0$ and the only rational solution is $x = 0$. Since $\mathcal{O}$ also is a 2-torsion point, we obtain $\#T[2] = 2$.

From group theory it follows that

$$\begin{aligned} [E(\mathbb{Q}) : 2E(\mathbb{Q})] &= [E(\mathbb{Q}) : \psi(E'(\mathbb{Q}))] \cdot [\psi(E'(\mathbb{Q})) : \psi \circ \varphi(E(\mathbb{Q}))] \\ &= \frac{[E(\mathbb{Q}) : \psi(E'(\mathbb{Q}))] \cdot [E'(\mathbb{Q}) : \varphi(E(\mathbb{Q}))]}{[\ker(\psi) : \ker(\psi) \cap \varphi(E(\mathbb{Q}))]}. \end{aligned} \tag{3.14}$$

We know that $\ker(\psi) = \{\overline{\mathcal{O}}, (\bar{0}, \bar{0})\}$. We need to find out whether or not $(\bar{0}, \bar{0})$ is an element of $\varphi(E(\mathbb{Q}))$. The point $(\bar{0}, \bar{0})$ is an element of $\varphi(E(\mathbb{Q}))$ if and only if there is a rational point (x, y) on E with $x \neq 0$ and $y = 0$. But we saw that there is no such point. Therefore,

$$[\ker(\psi) : \ker(\psi) \cap \varphi(E(\mathbb{Q}))] = 2.$$

Putting (3.13) and (3.14) together, we find the following equality:

$$2^r = \frac{[E(\mathbb{Q}) : 2E(\mathbb{Q})]}{4} = \frac{\#\alpha(E(\mathbb{Q})) \cdot \#\bar{\alpha}(E'(\mathbb{Q}))}{4}. \tag{3.15}$$

So computing the number of elements in the images of α and $\bar{\alpha}$ suffices.

Let us see what these images look like. In order to determine the image of α, we have to find out which rational numbers, modulo squares, can occur as the x-coordinate of points in $E(\mathbb{Q})$. We start by writing

$$x = \frac{m}{e^2} \quad \text{and} \quad y = \frac{n}{e^3}$$

in lowest terms and with $e > 0$. If $m = 0$, then $(x, y) = (0, 0)$ and $\alpha(0, 0) = 3$. We look at the points with m and n not equal to 0. These points satisfy

$$n^2 = m(m^2 - 3me^2 + 3e^4). \tag{3.16}$$

Let $b_1 = \pm \gcd(m, b)$, where we choose the sign such that $mb_1 > 0$. Then we have $m = b_1 m_1$ and $b = b_1 b_2$, with $\gcd(m_1, b_2) = 1$ and $m_1 > 0$. If we substitute this in (3.16), we find that b_1 divides n, so $n = b_1 n_1$, say. So we have

$$n_1^2 = m_1(b_1 m_1^2 - 3m_1 e^2 + b_2 e^4).$$

Since $\gcd(b_2, m_1) = 1$ and $\gcd(e, m_1) = 1$, both factors at the right-hand side are squares. So we can factor $n_1 = MN$ and we find that $M^2 = m_1$ and $N^2 = b_1 m_1^2 - 3m_1 e^2 + b_2 e^4$. It follows that

$$N^2 = b_1 M^4 - 3M^2 e^2 + b_2 e^4. \tag{3.17}$$

Therefore, the point (x, y) we started with can be written as $(\frac{b_1 M^2}{e^2}, \frac{b_1 MN}{e^3})$, so modulo squares, the x-coordinate is a divisor of b, so it divides 3.

We start by showing that the number of elements in $\alpha(E(\mathbb{Q}))$ is equal to 2. In our case $b = 3$, so we have to take care of the divisors ± 1 and ± 3. From now on, by saying that the number s is an element of the image of α, we mean that the class in $\mathbb{Q}^*/\mathbb{Q}^{*^2}$ to which s belongs, is an element of the image of α. We already know that $1 \in \alpha(E(\mathbb{Q}))$, since $\alpha(\mathcal{O}) = 1$. Since $\alpha(0, 0) = b = 3$, we also know that 3 is contained in the image of α. The image of α is a subgroup of $\mathbb{Q}^*/\mathbb{Q}^{*^2}$, so if -1 is contained in the image of α, then -3 is also contained in it, and vice versa. Therefore, we only have to deal with one of them. Let us take $b_1 = -1$. The equation we now get is:

$$N^2 = -M^4 - 3M^2 e^2 - 3e^4. \tag{3.18}$$

Taking this equation modulo 3, we immediately see that there is no solution, since we are allowed to assume that $\gcd(M, N) = 1$. Therefore, -1 is not contained in the image of α. It follows that $\#\alpha(E(\mathbb{Q})) = 2$, which is what we wanted to prove.

Similarly, it can be shown that the number of elements in the image of $\bar{\alpha}$ is also equal to 2. From equation (3.15) it follows that the rank of E is 0.

Because the rank of E is 0, all rational points on E have finite order. For the computation of the rational points we can use the Nagell-Lutz theorem, which states that if $P = (x, y)$ is a rational point of finite order on an elliptic curve E, and Δ is the discriminant of the cubic polynomial that defines E, then x and y are integers, and either $y = 0$, or else y^2 divides Δ.

In our case $\Delta = -27$, so all rational points on E have $y = 0$ or $y^2 | -27$, where y is an integer. Trying all possible values for y yields the following points on D: $(-1, 0)$, $(0, 1)$, $(0, -1)$, $(2, 3)$ and $(2, -3)$. Therefore, the group $D(\mathbb{Q})$ has order 6 and consists of

$$\{(-1, 0), (0, 1), (0, -1), (2, 3), (2, -3), \mathcal{O}\}.$$

This is what we wanted to prove. □

Of course, Euler did not use all these theorems about elliptic curves. He gave an elementary proof, using Fermat's method of descent. Since both proofs use some kind of descent, we might expect them to be essentially the same. We give Euler's original proof [8] in Latin and we explain what happens in modern notation using the terminology of elliptic curves. We will see that indeed both proofs are much alike.

Euler's formulation of the theorem is:

Theorema
Nullus cubus, ne quidem numeris fractis exceptis, unitate auctus quadratum efficere potest praeter unicum casum, quo cubus est 8*.*

In other words: if we add 1 to a rational cube, then it never becomes a square, unless the cube is 8. Euler obviously assumes that the numbers he is talking about do not equal 0.

Euler's proof of theorem 3.5. ***Demonstratio***
Propositio ergo huc redit, ut $\frac{a^3}{b^3}+1$ *nunquam esse possit quadratum praeter casum, quo* $\frac{a}{b}=2$*. Quocirca demonstrandum erit hanc formulam* a^3b+b^4 *nunquam fieri posse quadratum, nisi sit* $a=2b$*.*

Consider the equation $y^2 = x^3 + 1$. Suppose this equation has a positive rational solution $(\frac{a}{b}, y)$, where a and b are coprime integers. This is equivalent to saying that $\frac{a^3}{b^3}+1$ is a square, which implies that a^3b+b^4 is a square. Note that this assumption already rules out three of the rational points we found in our first proof, namely the points $(0, 1)$, $(0, -1)$ and $(-1, 0)$. We need to show that $a = 2b$ provides the only other solution.

Haec autem expressio resolvitur in istos tres factores $b(a+b)(aa-ab+bb)$*, qui primo quadratum constituere possunt, si esse posset* $b(a+b)=aa-ab+bb$*, unde prodit* $a=2b$*, qui erit casus, quem excepimus. Pono autem, ut ulterius pergam,* $a+b=c$ *seu* $a=c-b$*, qua facta substitutione habebitur*

$$bc(cc-3bc+3bb),$$

quam demonstrandum est quadratum esse non posse, nisi sit $c=3b$*; sunt autem* b *et* c *numeri inter se primi. Hic autem duo occurrunt casus considerandi, prout* c *vel multiplum est ternarii vel secus; illo enim casu factores* c *et* $cc-3bc+3bb$ *communem divisorem habebunt* 3*, hoc vero omnes tres inter se erunt primi.*

The expression $a^3b + b^4$ equals $b(a + b)(a^2 - ab + b^2)$. Of course, this is a square if $b(a+b) = a^2-ab+b^2$, which gives us the solution $a = 2b$. Euler now applies the same change of coordinates we did: he introduces the new variable c, which is defined by $c = a + b$. This amounts to a transformation such that $(0, 0)$ becomes a rational 2-torsion point. Of course, $\gcd(b, c)$ is also 1. Now the equation $\frac{c}{b}((\frac{c}{b})^2 - 3\frac{c}{b} + 3) = y^2$ holds, so we have a rational point $P = (\frac{c}{b}, y)$ on the elliptic curve $y^2 = x(x^2-3x+3)$, which we called E in the previous proof. This is the same as saying that

$$bc(c^2 - 3bc + 3b^2) \tag{3.19}$$

is a square of a rational number, like Euler does. Note that the solution $a = 2b$ corresponds to $c = 3b$.

From now on Euler assumes that $\frac{a}{b}$ is not equal to 2, which is the same as assuming that c is not equal to $3b$. Since b and c are coprime, the only case in which two factors in expression (3.19) have a factor greater than 1 in common, is when c and $c^2-3bc+3b^2$ are not coprime. This implies that c is divisible by 3. Therefore, Euler distinguishes two cases, the first case being the case in which c is not a multiple of 3, and the second the case in which c is a multiple of 3.

Case 1: 3 *does not divide* c

Sit primo c *non divisibile per* 3*; necesse erit, ut singuli illi tres factores sint quadrata, scilicet* b *et* c *et* $cc-3bc+3bb$ *seorsim. Fiat ergo* $cc-3bc+3bb=(\frac{m}{n}b-c)^2$*; erit*

$$\frac{b}{c}=\frac{3nn-2mn}{3nn-mm} \quad vel \quad \frac{b}{c}=\frac{2mn-3nn}{mm-3nn},$$

cuius fractionis termini erunt primi inter se, nisi m sit multiplum ternarii.

In this case 3 does not divide c, so all three of b, c and $c^2 - 3bc + 3b^2$ are coprime and they must all be squares. So $c^2 - 3bc + 3b^2 = (\frac{m}{n}b - c)^2$, say, where we can take m and n to be coprime, positive integers. Since $\frac{b}{c} \neq 0$, this yields

$$\frac{b}{c} = \frac{3n^2 - 2mn}{3n^2 - m^2}.$$

The numbers $3n^2 - 2mn$ and $3n^2 - m^2$ are coprime, unless 3 divides m. So now we have a separation in cases again.

Case 1.1: 3 *does not divide* m

Sit ergo m *per* 3 *non divisibile; erit vel* $c = 3nn - mm$ *vel* $c = mm - 3nn$ *et vel* $b = 3nn - 2mn$ *vel* $b = 2mn - 3nn$. *At cum* $3nn - mm$ *quadratum esse nequeat, ponatur* $c = mm - 3nn$, *quod quadratum fiat radicis* $m - \frac{p}{q}n$, *hincque oritur* $\frac{m}{n} = \frac{3qq+pp}{2pq}$ *atque*

$$\frac{b}{nn} = \frac{2m}{n} - 3 = \frac{3qq - 3pq + pp}{pq}.$$

Quadratum ergo esset haec formula $pq(3qq - 3pq + pp)$, *quae omnino similis est propositae* $bc(3bb - 3bc + cc)$ *et ex multo minoribus numeris constat.*

Suppose 3 does not divide m. Then we either have $b = 3n^2 - 2mn$ and $c = 3n^2 - m^2$, or $b = 2mn - 3n^2$ and $c = m^2 - 3n^2$. Taking $3n^2 - m^2$ modulo 4, we find that it cannot be a square. Therefore, $c = n^2 - 3n^2$ and $b = 2mn - 3n^2$. Now $m^2 - 3n^2$ is a square, $(m - \frac{p}{q}n)^2$ say, where we take p and q to be coprime, positive integers. Then $\frac{m}{n} = \frac{3q^2+p^2}{2pq}$ and it follows that $\frac{b}{n^2} = \frac{3q^2-3pq+p^2}{pq}$. We already saw that b is a square, so $\frac{3q^2-3pq+p^2}{pq}$ is a square and $pq(3q^2 - 3pq + p^2)$ is a square too. Euler proceeds by saying that he has found these integers p and q such that $pq(p^2 - 3pq + 3q^2)$ is a square, just as $c^2 - 3bc + 3b^2$ is a square. However, p and q are smaller, which is not further specified by him, but we will get to that soon.

Now let us translate this argument to the language of elliptic curves. We can see immediately that these new integers also give a new rational point $(\frac{p}{q}, y') = P'$ on our elliptic curve E. We want to find out what this reduction actually means. Since we used a 2-descent in the previous proof, we might expect that it has something to do with multiplication by 2. This turns out to be true. Computing the x-coordinate of the point $2P'$ leads to the following expression:

$$x(2P') = \frac{((\frac{p}{q})^2 - 3)^2}{4y'^2} = \frac{p^4 - 6p^2q^2 + 9q^4}{4(p^3q - 3p^2q^3 + 3pq^3)}. \tag{3.20}$$

And if we now compute $\frac{c}{b}$ in terms of p and q we find:

$$\frac{c}{b} = \frac{m^2 - 3n^2}{2mn - 3n^2} = \frac{(3q^2 + p^2)^2 - 3(2pq)^2}{2(3q^2 + p^2)(2pq) - 3(2pq)^2} = \frac{9q^4 - 6p^2q^2 + p^4}{4(p^3q - 3p^2q^2 + 3pq^3)},$$

which is obviously equal to the expression in (3.20). Therefore, we have found a new point P', such that $2P'$ is equal to the point P from which we started. (Note the similarity of the equation $P = 2P'$ to equation (3.12).) This does not yet show that there are no other positive rational points, since it could be that after a while we find a point that we already had.

So now we are left with the problem of the integers becoming 'smaller'. In our modern proof we also used some sort of size of points on an elliptic curve, namely the height of a point. Let us find out whether this notion of size is sufficient for Euler's purpose.

We take a look at the heights of the points in question. The height of $P = (\frac{c}{b}, y)$ is equal to $\max\{|b|, |c|\}$ and the height of $P' = (\frac{p}{q}, y')$ is equal to $\max\{|p|, |q|\}$. We will show that the height of P' is smaller than the height of P. We know that $b = 2mn - 3n^2$, so n divides b, which implies that $|n| \leq |b|$. Further, $n = 2pq$ or $n = pq$, so $|p|$ and $|q|$ are both smaller than or equal to $|n|$. Now we have proven that $\max\{|p|, |q|\}$ is smaller than or equal to $|b|$. Note that $\max\{|p|, |q|\}$ is only equal to b when $b = n = \max\{|p|, |q|\}$, which means that $b = 2mn - 3n^2 = n$ so $m = 2$ and $n = 1$. This is only the case when $b = 1$ and $c = 1$, which means that $a = 0$, which we excluded from the start. It follows that for all positive rational points P the height of P' is smaller than the height of P.

We may now conclude the following. If there exists a point $P = (\frac{c}{b}, y)$ on E as above, then there exists a sequence P', P'', P''', ... of points on E such that the height of each point is smaller than the height of its predecessor. But this is impossible, because the point P we started with has finite height $\max\{|b|, |c|\}$ and all the heights are integers larger than 0. Therefore, such a point does not exist.

Case 1.2: 3 divides m

At sit m multiplum ternarii, puta m = 3k; erit $\frac{b}{c} = \frac{nn-2kn}{nn-3kk}$*, unde erit vel* $c = nn - 3kk$ *vel* $c = 3kk - nn$*; quia autem* $3kk - nn$ *quadratum esse nequit, ponatur* $c = nn - 3kk$ *eiusque radix* $n - \frac{p}{q}k$*, unde fiet* $\frac{n}{k} = \frac{3qq+pp}{2pq}$ *seu* $\frac{k}{n} = \frac{2pq}{3qq+pp}$ *atque*

$$\frac{b}{nn} = 1 - \frac{2k}{n} = \frac{pp + 3qq - 4pq}{3qq + pp}.$$

Quadratum ergo esse deberet $(pp + 3qq)(p - q)(p - 3q)$*. Ponatur* $p - q = t$ *et* $p - 3q = u$*; erit* $q = \frac{t-u}{2}$ *et* $p = \frac{3t-u}{2}$ *illaque formula abit in hanc* $tu(3tt - 3tu + uu)$*, quae iterum similis est priori* $bc(3bb - 3bc + cc)$*.*

In this case, 3 divides m, so $m = 3k$, say. Then we find that

$$\frac{b}{c} = \frac{n^2 - 2kn}{n^2 - 3k^2}.$$

Again, it follows that $c = n^2 - 3k^2$ and $b = n^2 - 2kn$. We know that c is a square, so put $c = (n - \frac{p}{q})^2$, where we can take p and q to be coprime, positive integers. It follows that $\frac{n}{k} = \frac{3q^2+p^2}{2pq}$ and therefore $\frac{b}{n^2} = \frac{p^2+3q^2-4pq}{3q^2+p^2}$. Since b is a square, $\frac{p^2+3q^2-4pq}{3q^2+p^2}$ also is a square, and therefore $(p^2 + 3q^2)(p - q)(p - 3q)$ is a square. Now the substitutions $t = p - q$ and $u = p - 3q$ yield the following familiar relation:

$$tu(3t^2 - 3tu + u^2)$$

is a square.

Euler proceeds in the same way as he did in the previous case. He says that now he has found new integers t and u, such that the expression $tu(3t^2 - 3tu + u^2)$ is a square, in which the integers t and u are in some sense smaller than b and c.

We translate to the terminology of elliptic curves again. We have found a second rational point $P' = (\frac{u}{t}, y')$ on E and we wonder whether $2P'$ is equal to P again,

so we do the same computations. First, we compute the x-coordinate of the point $2P'$:

$$x(2P') = \frac{((\frac{u}{t})^2 - 3)^2}{4y'^2} = \frac{u^4 - 6u^2t^2 + 9t^4}{4(u^3t - 3u^2t^3 + 3ut^3)}. \tag{3.21}$$

And second, we compute $\frac{c}{b}$ in terms of t and u:

$$\begin{aligned} \frac{c}{b} &= \frac{n^2 - 3k^2}{n^2 - 2kn} \\ &= \frac{(3(\frac{t-u}{2})^2 + (\frac{3t-u}{2})^2)^2 - 12(\frac{3t-u}{2} \cdot \frac{t-u}{2})^2}{(3(\frac{t-u}{2})^2 + (\frac{3t-u}{2})^2)^2 - 2(3(\frac{t-u}{2})^2 + (\frac{3t-u}{2}))^2 \cdot 2\frac{t-u}{2} \cdot \frac{3t-u}{2}} \\ &= \frac{9t^4 - 6t^2u^2 + u^4}{4tu(3t^2 - 3tu + u^2)}, \end{aligned} \tag{3.22}$$

which is obviously equal to (3.21).

Now we still have to show that the height of P' is smaller than the height of P. The substitution $p - q = t$ and $p - 3q = u$ amounts to saying that $q = \frac{t-u}{2}$ and $p = \frac{3t-u}{2}$. Now we have two possibilities again, because t and u are either both positive or negative.

First, suppose that t and u are both positive. Note that n divides b, so $n \leq b$. Then from $q = \frac{t-u}{2} > 0$ it follows that $t > u$. For p we derive $p = \frac{3t-u}{2} > \frac{3t-t}{2} = t$, so we have proven that $u < t < p < n \leq b$. So in this case the height of $P' = (\frac{u}{t}, y'')$ is smaller than the height of the original point P.

We are left with the case in which t and u are both negative. Then $q - p = |u|$ and $3q - p = |t|$. Therefore, $q = \frac{|t|-|u|}{2}$, so $|t| > |u|$. We saw that $\frac{n}{k} = \frac{3q^2+p^2}{2pq}$. Note that the only possible common divisor of $3q^2 + p^2$ and $2pq$ is 2. (Here we use that m and n are coprime integers.) So the following inequality holds:

$$\begin{aligned} n \geq \frac{3q^2 + p^2}{2} &= \frac{3(|t|^2 - 2|t||u| + |u|^2) + |t|_6^2|t||u| + 9|u|^2}{8} \\ &= \frac{|t|^2 - 3|t||u| + 3|u|^2}{2} > \frac{|t|^2}{2} \geq |t|. \end{aligned} \tag{3.23}$$

This last inequality holds because t and u are integers such that $|t| > |u| > 0$, so $|t| \geq 2$. So now we have $|u| < |t| < n \leq b$ and in this case we also find that the height of the new point $P' = (\frac{u}{t}, y')$ is smaller than the height of the original point $P = (\frac{c}{b}, y)$. Similar to case 1.1 we get a contradiction from this. It follows that the point $P = (\frac{c}{b}, y)$ does not exist.

Case 2: 3 *does divide* c

Restat ergo posterior casus, quo est c *multiplum ternarii, puta* $c = 3d$*, atque quadratum esse debet* $bd(bb-3bd+3dd)$*; quae cum iterum similis sit priori, manifestum est utroque casu evenire non posse, ut formula proposita sit quadratum. Quamobrem praeter cubum* 8 *alius ne in fractis quidem datur, qui cum unitate faciat quadratum. Q.E.D.*

If 3 divides c, then we can write $c = 3d$, where d is a positive integer. Since we assumed that $\frac{c}{b}$ is not equal to 3, we use that d is not equal to b. We already know

that $bc(c^2 - 3bc + 3b^2)$ is a square, so it follows that $bd(b^2 - 3bd + 3d^2)$ is also a square. Now we find the new point $P' = (\frac{b}{d}, y')$ on E. Since b and c are coprime, 3 does not divide b. Now we can repeat the argument in case 1 for this point and derive a contradiction. It follows that the point we started with in this case also does not exist.

In terms of elliptic curves, what Euler says here is that if the rational point $P = (\frac{c}{b}, y)$ lies on E, then the point $P' = P + (0,0) = (3\frac{b}{c}, y') = (\frac{3b}{3d}, y') = (\frac{b}{d}, y')$ also lies on E. But case 1 applies to the point P', so it does not exist, and the point we started with does not exist, too. Again, compare the equation $P + (0,0) = P - (0,0) = 2P'$ to equation (3.12).

Note that in this part of the proof we really use the assumption that $\frac{c}{b} \neq 3$. If we would apply the previous argument to the case $\frac{c}{b} = 3$, we would find the point $(\frac{b}{d}, y') = (1,0)$, which is the point to which the argument of case 1 did not apply. This point corresponds to the solutions $(2,3)$ and $(2,-3)$ on the original elliptic curve D. □

If we compare Euler's proof to the proof of the Mordell-Weil theorem, we see that they are very much alike, but there is a small distinction. In Euler's proof, we start with a rational point on E with certain restrictions upon it (for instance, that it is not $(1,0)$), and using the method of descent it follows that such a point does not exist. In the proof of the Mordell-Weil theorem we start with any rational point on E, and using the method of descent we find that such a point always is generated by a finite number of points.

The way the method of descent works turns out to be the same. If we compare the procedure that writes a rational point P on E as $P - Q_i = 2P'$ to what happens in Euler's proof, we find that Euler does exactly the same: in case 1.1 and 1.2 he writes the point P that he starts with in the form $P = 2P'$, where P' is a rational point on E with height smaller than the height of P. In case 2, however, he first adds the point $(0,0)$ to it and than he applies case 1 again. In terms of equation (3.12) this amounts to saying that $(0,0)$ is a representative of a coset of $2E(\mathbb{Q})$ and that $P - (0,0) = 2P'$ for some rational point P' on E with height smaller than the height of $P - (0,0)$.

Our conclusion is that Euler's way of proving theorem 3.5 is essentially the same as our modern proof that uses the theory and terminology of elliptic curves. In our modern way of looking at things, the ingenious substitutions Euler invented have obtained a geometrical meaning.

Chapter 4

Cyclotomic fields

Before we start with the remaining part of the proof, we have to give some preliminary remarks. First we describe the setting in which we will work and after that we give some important definitions and facts, concerning for example cyclotomic units. For more details, see for instance [18].

4.1 Setting

Let p be an odd prime number. Let Φ_p be the p-th cyclotomic polynomial in $\mathbb{Q}[X]$, i.e. $\Phi_p = \frac{X^p-1}{X-1}$. Consider the field extension $\mathbb{Q}[X]/(\Phi_p) \cong \mathbb{Q}(\zeta)$ of $\mathbb{Q}$, where ζ denotes a primitive p-th root of unity. This is a field extension of degree $p-1$, since Φ_p is of degree $p-1$ and it is irreducible in $\mathbb{Q}[X]$. We denote $\mathbb{Q}(\zeta)$ by K.

This field extension is Galois with Galois group

$$G = \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \cong (\mathbb{Z}/p\mathbb{Z})^*,$$

since the map

$$\begin{array}{ccc} (\mathbb{Z}/p\mathbb{Z})^* & \xrightarrow{\sim} & \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \\ a \,(\mathrm{mod}\, p) & \longmapsto & (\sigma_a : \zeta \mapsto \zeta^a) \end{array} \tag{4.1}$$

is an isomorphism.

The automorphism σ_{p-1} acts in all embeddings as complex conjugation. Therefore, we call σ_{p-1} complex conjugation.

The fixed field of complex conjugation is $\mathbb{Q}(\zeta+\zeta^{-1})$, which is called the *maximal real subfield* of $\mathbb{Q}(\zeta)$. We denote $\mathbb{Q}(\zeta+\zeta^{-1})$ by K^+. The field extension $\mathbb{Q}(\zeta+\zeta^{-1})$ of $\mathbb{Q}$ has degree $\frac{p-1}{2}$ and it is Galois with Galois group

$$G^+ = \mathrm{Gal}(\mathbb{Q}(\zeta+\zeta^{-1})/\mathbb{Q}) \cong (\mathbb{Z}/p\mathbb{Z})^*/(\pm 1).$$

Some parts of the proof in chapter 8 consist of working with ideals in the rings of algebraic integers of $K = \mathbb{Q}(\zeta)$ and $K^+ = \mathbb{Q}(\zeta+\zeta^{-1})$. The ring of integers $\mathcal{O}_K$ of K is the ring $\mathbb{Z}[\zeta]$. The ring of integers $\mathcal{O}_{K^+}$ of K^+ is the ring $\mathbb{Z}[\zeta+\zeta^{-1}]$.

We formulate some lemma's that will be very useful in the following chapters.

Lemma 4.1. *The prime p is totally ramified in $\mathbb{Q}(\zeta)$ and $(p) = (1-\zeta)^{p-1}$, where $\mathfrak{P} = (1-\zeta)$ is a prime ideal in $\mathcal{O}_K = \mathbb{Z}[\zeta]$.*

Proof. We use the Kummer-Dedekind theorem. In $\mathbb{F}_p[X]$ we have $\frac{X^p-1}{X-1} = \frac{(X-1)^p}{X-1} = (X-1)^{p-1}$. Since $\frac{X^p-1}{X-1} = X^{p-1} + X^{p-2} + \ldots X + 1 \equiv p \,(\mathrm{mod}\, X-1)$, the remainder of $\frac{X^p-1}{X-1}$ upon division by $X-1$ is not divisible by p^2, so $(p, 1-\zeta)$ is invertible and we have the equality $(p, 1-\zeta)^{p-1} = (p)$. Therefore, the only prime ideal that lies above p is the ideal $(p, 1-\zeta) = (1-\zeta)$. This last equality holds since $p = (1-\zeta)\prod_{a=2}^{p-1}(1-\zeta^a)$. □

Lemma 4.2. *All primes p' in $\mathbb{Z}$ distinct from p do not ramify in $\mathbb{Q}(\zeta)$.*

Proof. We use the Kummer-Dedekind theorem again. If we take $\frac{X^p-1}{X-1}$ modulo p', it is obvious that this polynomial is separable and therefore p' does not ramify in $\mathbb{Q}(\zeta)$. □

Lemma 4.3. *Suppose r and s are integers with $\gcd(p, rs) = 1$. Then $\frac{\zeta^r-1}{\zeta^s-1}$ is a unit in $\mathbb{Z}[\zeta]$ and therefore the ideals $(1-\zeta^r)$ and $(1-\zeta^s)$ are equal.*

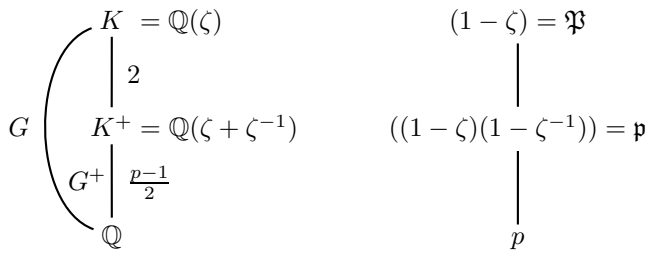
Proof. Writing $r \equiv st \,(\mathrm{mod}\, p)$ for some integer t, we have

$$\frac{\zeta^r - 1}{\zeta^s - 1} = \frac{\zeta^{st} - 1}{\zeta^s - 1} = 1 + \zeta^s + \ldots + \zeta^{s(t-1)} \in \mathbb{Z}[\zeta].$$

Similarly, we find that $\frac{\zeta^s-1}{\zeta^r-1} \in \mathbb{Z}[\zeta]$. It follows that the number $\frac{\zeta^r-1}{\zeta^s-1}$ is a unit in $\mathbb{Z}[\zeta]$, so the ideals $(1-\zeta^r)$ and $(1-\zeta^s)$ are equal. □

Since p is totally ramified in $\mathbb{Q}(\zeta)$, it follows that p is also totally ramified in $\mathbb{Q}(\zeta + \zeta^{-1})$. Therefore, we have $(p) = ((1-\zeta)(1-\zeta^{-1}))^{\frac{p-1}{2}}$. From now on, we denote the $\mathcal{O}_{K^+}$-ideal $((1-\zeta)(1-\zeta^{-1}))$ by $\mathfrak{p}$. Further, let λ denote the element $(1-\zeta)(1-\zeta^{-1})$ that generates $\mathfrak{p}$. We find that the ideals $((1-\zeta^a)(1-\zeta^{-a}))$ are the same for all $a = 1, \ldots, \frac{p-1}{2}$.

We summarize the previous remarks in a diagram.



4.2 Galois modules

Let L be a field extension of $\mathbb{Q}$ that is Galois. Then the Galois group $G_L = \mathrm{Gal}(L/\mathbb{Q})$ acts on L as an automorphism. Therefore, it acts on anything that is canonically defined in terms of L, such as the unit group L^*, the unit group $\mathcal{O}_L^*$ of the ring of integers of L, the group $\mathcal{I}_L$ of invertible $\mathcal{O}_L$-ideals, the group $\mathcal{P}_L$ of principal

fractional $\mathcal{O}_L$-ideals and the class group of L. All subgroups of these groups that are closed under the induced Galois action have a Galois action as well.

Abelian groups with a Galois action of Galois group G are $\mathbb{Z}[G]$-modules. Therefore, all groups mentioned above are $\mathbb{Z}[G]$-modules. Let q be a prime number. A $\mathbb{Z}[G]$-module that is annihilated by q is also an $\mathbb{F}_q[G]$-module.

There are two useful concepts concerning elements of group rings, namely the *weight* and the *size*. Let $\mathbb{Q} \subset L$ be a field extension that is Galois and let H be the Galois group $\mathrm{Gal}(L/\mathbb{Q})$. Then H acts on the multiplicative group L^*.

Definition 4.1. *Consider the group ring $\mathbb{Z}[H]$. Define the* weight *of $\theta = \sum_{\sigma \in H} n_\sigma \sigma$ in $\mathbb{Z}[H]$ by*

$$w(\theta) = \sum_{\sigma \in H} n_\sigma.$$

The weight function is additive and multiplicative, so it defines a homomorphism $\mathbb{Z}[H] \to \mathbb{Z}$. Therefore, its kernel is a $\mathbb{Z}[H]$-ideal.

Definition 4.2. *The kernel of the weight homomorphism is called the* augmentation ideal *of $\mathbb{Z}[H]$.*

The other important property of elements of $\mathbb{Z}[H]$ is the size.

Definition 4.3. *Consider the group ring $\mathbb{Z}[H]$. Define the* size *of an element $\theta = \sum_{\sigma \in H} n_\sigma \sigma \in \mathbb{Z}[H]$ by*

$$\|\theta\| = \sum_{\sigma \in H} |n_\sigma|.$$

It is easy to see that for all elements $\theta_1 = \sum_{\sigma \in G} n_\sigma \sigma$ and $\theta_2 = \sum_{\sigma \in G} m_\sigma \sigma$ in $\mathbb{Z}[H]$ we have that $\|\theta_1 \theta_2\| = \sum_{\sigma \in G} \sum_{\varphi\psi = \sigma} |n_\varphi m_\psi| \leq (\sum_{\sigma \in G} |n_\sigma|) \cdot (\sum_{\sigma \in G} |m_\sigma|) = \|\theta_1\| \cdot \|\theta_2\|$. From the triangle inequality in $\mathbb{Z}$ it follows that for all elements θ_1 and θ_2 in H we have $\|\theta_1 + \theta_2\| \leq \|\theta_1\| + \|\theta_2\|$.

In our case, where we take group rings over the Galois groups G and G^+, the group rings $\mathbb{F}_q[G]$ and $\mathbb{F}_q[G^+]$ have a nice structure if q does not divide $p-1$, as we show in the following lemma.

Lemma 4.4. *If q does not divide $p-1$, the the group ring $\mathbb{F}_q[G]$ equals a finite product of finite fields, i.e.*

$$\mathbb{F}_q[G] \cong \prod_i^{<\infty} F_i,$$

for finite fields F_i. The same statement holds for the group ring $\mathbb{F}_q[G^+]$.

Proof. The group $G = \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$ is isomorphic to $(\mathbb{Z}/p\mathbb{Z})^*$. Therefore, G is a cyclic group and it has a generator σ, say. Note that $\sigma^{p-1} = 1$. It follows that all elements of $\mathbb{F}_q[G]$ are of the form $\sum_{i=1}^{p-1} n_\sigma \sigma^i$. Now the identification $\sigma \mapsto X$ gives rise to an isomorphism $\mathbb{F}_q[G] \xrightarrow{\sim} \mathbb{F}_q[X]/(X^{p-1} - 1)$.

The polynomial $X^{p-1} - 1$ is separable in $\overline{\mathbb{F}_q}$, since its derivative equals $(p-1)X^{p-2}$, which does not equal 0, because q does not divide $p-1$ by assumption. The Chinese Remainder Theorem tells us that $\mathbb{F}_q[X]/(X^{p-1} - 1) \cong \prod_g \mathbb{F}_q[X]/(g(X))$, where the $g \in \mathbb{F}_q[X]$ are the irreducible polynomials that divide X^{p-1}. Since $X^{p-1} - 1$ is

separable, all g's occur with multiplicity 1. Therefore, all the $\mathbb{F}_q[X]/(g(X))$ are finite field extensions of $\mathbb{F}_q$, so they are finite fields themselves.

For the group ring $\mathbb{F}_q[G^+]$, the proof is completely similar:

$$\mathbb{F}_q[G^+] \xrightarrow{\sim} \mathbb{F}_q[X]/(X^{\frac{p-1}{2}} - 1).$$

□

4.3 Cyclotomic units

One of the concepts we use in the new part of the proof is the concept of *cyclotomic units*. In this section we give a definition and we formulate Thaine's theorem.

Definition 4.4. *Let E be $\mathbb{Z}[\zeta]^*$, the group of units of $\mathcal{O}_K$. Let V be the multiplicative group generated by $\{\pm\zeta, 1-\zeta^a : 1 < a \leq p-1\}$. We define the* cyclotomic units *C of $K = \mathbb{Q}(\zeta)$ by*

$$C = V \cap E.$$

We also define the cyclotomic units C^+ of $K^+ = \mathbb{Q}(\zeta + \zeta^{-1})$. Let E^+ be $(\mathbb{Z}[\zeta + \zeta^{-1}])^*$, the group of units of $\mathcal{O}_{K^+}$. Then we define C^+ by

$$C^+ = E^+ \cap C.$$

In order to give a better idea of what these cyclotomic units look like, we state the following lemma. For a proof, see [18] again.

Lemma 4.5. *The cyclotomic units C^+ of K^+ are generated by -1 and the units*

$$\xi_a = \zeta^{\frac{1-a}{2}} \frac{1-\zeta^a}{1-\zeta},$$

where $1 < a \leq \frac{p-1}{2}$. The cyclotomic units C of K are generated by ζ and the cyclotomic units of K^+.

There exists a connection between the cylotomic units of K^+ and the class number h^+ of K^+, as we can see in the following theorem. For a proof, see chapter 8 in [18].

Theorem 4.6. *The cyclotomic units C^+ of $K^+ = \mathbb{Q}(\zeta + \zeta^{-1})$ are of finite index in the full unit group $E^+ = \mathcal{O}_{K^+}$, and*

$$h^+ = [E^+ : C^+].$$

This theorem states that the number of elements of the class group of K^+ equals the number of elements in E^+/C^+, but the equality does not come from some canonical isomorphism of the groups Cl_{K^+} and E^+/C^+. Even though Cl_{K^+} and E^+/C^+ need not be isomorphic as $\mathbb{Z}[G^+]$-modules, they do share certain properties as Galois-modules. The following theorem, that was proven by Thaine [16], states that their Sylow-q-subgroups have an important propery in common. Let $(E^+/C^+)_q$ denote the q-Sylow subgroup of E^+/C^+ and let $(Cl_{K^+})_q$ denote the q-Sylow subgroup of the class group of K^+. Now Thaine's theorem states the following:

Theorem 4.7 (Thaine). *If ε is an element of $\mathbb{Z}[G^+]$ that annihilates $(E^+/C^+)_q$ and q does not divide $p-1$, then the element ε also annihilates $(Cl_{K^+})_q$.*

4.4 Stickelberger's theorem

In the proof of a theorem of Mihăilescu we want to use Stickelberger's theorem. To be able to do this, we first have to give some definitions. For more details on this subject, see [18], chapter 6, for example.

Definition 4.5. *Define the* Stickelberger element θ_S *as follows:*

$$\theta_S = \sum_{a=1}^{p-1} \frac{a}{p} \sigma_a^{-1} \in \mathbb{Q}[G].$$

Definition 4.6. *Define the* Stickelberger ideal *by:*

$$I_S = \mathbb{Z}[G] \cap \theta_S \mathbb{Z}[G].$$

According to [18], section 6.2, the following lemma holds.

Lemma 4.8. *Let I' be the ideal of $\mathbb{Z}[G]$ generated by elements of the form $c - \sigma_c$, with $\gcd(c, p) = 1$. Let $\beta \in \mathbb{Z}[G]$. Then*

$$\beta\theta_S \in \mathbb{Z}[G] \Leftrightarrow \beta \in I'.$$

Proof. For a real number x, let $\lfloor x \rfloor$ denote the entier of x, i.e. $\lfloor x \rfloor$ is the largest integer that is smaller than or equal to x. Let $\{x\}$ denote the fractional part of x, i.e. $\{x\} = x - \lfloor x \rfloor$.

Let us compute $(c - \sigma_c)\theta_S$:

$$\begin{aligned}
(c - \sigma_c)\theta_S &= (c - \sigma_c) \sum_{a=1}^{p-1} \frac{a}{p} \sigma_a^{-1} \\
&= \sum_{a=1}^{p-1} \frac{ca}{p} \sigma_a^{-1} - \sum_{a=1}^{p-1} \frac{a}{p} \sigma_{c^{-1}a}^{-1} \\
&= \sum_{a=1}^{p-1} \frac{ca}{p} \sigma_a^{-1} - \sum_{a=1}^{p-1} \left\{ \frac{ca}{p} \right\} \sigma_a^{-1} \\
&= \sum_{a=1}^{p-1} \left(\frac{ca}{p} - \left\{ \frac{ca}{p} \right\} \right) \sigma_a^{-1}. \qquad (4.2)
\end{aligned}$$

It follows that $(c - \sigma_c)\theta_S \in \mathbb{Z}[G]$.

Suppose that $(\sum_{a=1}^{p-1} x_a \sigma_a)\theta_S \in \mathbb{Z}[G]$, with the x_a elements of $\mathbb{Z}$. Then we have:

$$\begin{aligned}
(\sum_{a=1}^{p-1} x_a \sigma_a)(\sum_{c=1}^{p-1} \frac{c}{p} \sigma_c^{-1}) &= \sum_{a=1}^{p-1} \sum_{c=1}^{p-1} x_a \frac{c}{p} \sigma_a \sigma_c^{-1} \\
&= \sum_{b=1}^{p-1} \sum_{a=1}^{p-1} x_a \left\{ \frac{ab}{p} \right\} \sigma_b^{-1}. \qquad (4.3)
\end{aligned}$$

The coefficient of σ_1 is equal to

$$\sum_{a=1}^{p-1} x_a \left\{ \frac{a}{p} \right\} = \sum_{a=1}^{p-1} x_a \frac{a}{p} = \frac{1}{p} \sum_{a=1}^{p-1} x_a a,$$

so p divides $\sum_{a=1}^{p-1} x_a a$. Note that $p = p\sigma_1 = (p+1) - \sigma_{p+1}$ in $\mathbb{Z}[G]$. It follows that p is an element of I'. Therefore, $\sum_{a=1}^{p-1} x_a a$ is an element of I'. We obtain that

$$\sum_{a=1}^{p-1} x_a \sigma_a = \sum_{a=1}^{p-1} x_a(\sigma_a - a) + \sum_{a=1}^{p-1} x_a a \in I'.$$

This is what we wanted to prove. □

We can determine a set of generators of the Stickelberger ideal I_S.

Lemma 4.9. *The Stickelberger ideal I_S is generated by elements θ_c, where*

$$\theta_c = \sum_{a=1}^{p-1} \left\lfloor \frac{ac}{p} \right\rfloor \sigma_a^{-1},$$

for all integers c with $\gcd(c,p) = 1$.

Proof. From lemma 4.8 it follows that $I_S = (\theta_S)I'$. So we have to show that $(\theta_S)I'$ is generated by elements θ_c as above. Since I' is defined as the ideal generated by elements $c - \sigma_c$, it suffices to show that for all c that are coprime with p we have $(c - \sigma_c)\theta_S = \theta_c$.

Now let us determine $(c - \sigma_c)\theta_S$:

$$\begin{aligned}(c-\sigma_c)\theta_S &= \sum_{a=1}^{p-1} \left(\frac{ca}{p} - \left\{ \frac{ca}{p} \right\} \right) \sigma_a^{-1} \\ &= \sum_{a=1}^{p-1} \left\lfloor \frac{ca}{p} \right\rfloor \sigma_a^{-1} \\ &= \theta_c. \end{aligned} \tag{4.4}$$

So indeed $(c - \sigma_c)\theta_S = \theta_c$, and we obtain that I is generated over $\mathbb{Z}$ by all the θ_c with $\gcd(p,c) = 1$. □

We get an important property of the Stickelberger ideal from Stickelberger's theorem. For a proof, see [18] again.

Theorem 4.10 (Stickelberger). *Let J be a fractional ideal of $\mathbb{Q}(\zeta)$ and suppose that θ is an element of the Stickelberger ideal I_S. Then J^θ is a principal ideal. In other words: the Stickelberger ideal annihilates the class groep of $\mathbb{Q}(\zeta)$.*

Like there is a connection between the group E^+/C^+ and the class number h^+ of K^+, there also is a connection between the Stickelberger ideal and the class numbers of K and K^+. The ideal I_S^- is an ideal of the ring $\mathbb{Z}[G]$, but it is also a subgroup of $\mathbb{Z}[G]^- = \mathbb{Z}[G](1-\iota)$. Iwasawa's theorem states that the index of I_S^- in $\mathbb{Z}[G]^-$ equals the number h^-, which is defined as the quotient $\frac{h}{h^+}$ of the class number h of K and the class number h^+ of K^+. It follows that the index of I_S^- in $\mathbb{Z}[G]^-$ is finite. We will use this in chapter 6. For a proof, see [18, chapter 6]. Theorem 4.10 in [18] states that h^+ indeed divides h.

Theorem 4.11 (Iwasawa). *The index $[\mathbb{Z}[G]^- : I_S^-]$ equals h^-.*

Let $\iota = \sigma_{p-1}$ denote complex conjugation. By I_S^- we denote the $\mathbb{Z}[G]$-ideal that is obtained by multiplying the Stickelberger ideal I_S by the element $1 - \iota$, i.e. $I_S^- = I_S(1-\iota)$. Using the generators of the Stickelberger ideal we found in lemma 4.9, we construct a $\mathbb{Z}$-basis for the ideal I_S^-.

Lemma 4.12. *For all integers c that are coprime to p, let θ_c denote the element $(c-\sigma_c)\theta_S$ as in lemma 4.9. For the integers $k = 1, \ldots, \frac{p-1}{2}$, define elements $\tilde{\theta}_k$ of I_S^- as follows:*

$$\tilde{\theta}_k = (\theta_{k+1} - \theta_k)(1-\iota).$$

Then $\tilde{\theta}_1, \ldots, \tilde{\theta}_{\frac{p-1}{2}}$ form a $\mathbb{Z}$-basis of $I_S^- = I_S(1-\iota)$ and the elements $\tilde{\theta}_k$ all satisfy

$$\|\tilde{\theta}_k\| \leq p-1.$$

Proof. In lemma 4.9 we saw that the Stickelberger ideal I_S is generated over $\mathbb{Z}$ by all the θ_c. We have the following equality:

$$\begin{aligned}
\theta_c + p\theta_S &= \sum_{a=1}^{p-1} \left\lfloor \frac{ac}{p} \right\rfloor \sigma_a^{-1} + p\sum_{a=1}^{p-1} \frac{a}{p}\sigma_a^{-1} \\
&= \sum_{a=1}^{p-1} \left(\left\lfloor \frac{ac}{p} \right\rfloor + a \right) \sigma_a^{-1} \\
&= \sum_{a=1}^{p-1} \left\lfloor \frac{ac}{p} + a \right\rfloor \sigma_a^{-1} \\
&= \sum_{a=1}^{p-1} \left\lfloor \frac{a(c+p)}{p} \right\rfloor \sigma_a^{-1} \\
&= \theta_{c+p}. \qquad (4.5)
\end{aligned}$$

It follows that the ideal I_S is generated over $\mathbb{Z}$ by the finite set $\theta_1 = 0$, $\theta_2, \ldots, \theta_{p-1}$, $p\theta_S$. We have an other useful equality:

$$\begin{aligned}
\theta_c(1-\iota) + \theta_{p-c}(1-\iota) &= (c-\sigma_c)\theta_S(1-\iota) + (p-c-\sigma_{p-c})\theta_S(1-\iota) \\
&= \theta_S(p(1-\iota) - \sigma_c + \sigma_{c(p-1)} - \sigma_{-c} + \sigma_{-c(p-1)}) \\
&= \theta_S(p(1-\iota) - \sigma_c + \sigma_{-c} - \sigma_{-c} + \sigma_c) \\
&= p\theta_S(1-\iota). \qquad (4.6)
\end{aligned}$$

From this equality it follows that the ideal I_S^- is generated by the elements $\theta_1(1-\iota)$, $\theta_2(1-\iota), \ldots, \theta_{\frac{p+1}{2}}(1-\iota)$. Since $\theta_1 = 0$, the ideal I_S^- is also generated by the elements $\tilde{\theta}_1$, $\tilde{\theta}_2, \ldots, \tilde{\theta}_{\frac{p-1}{2}}$, where $\tilde{\theta}_k = (\theta_{k+1} - \theta_k)(1-\iota)$.

These elements even form a basis of I_S^-, because the $\mathbb{Z}$-rank of I_S^- is $\frac{p-1}{2}$. We can see this as follows. The $\mathbb{Z}$-rank of $\mathbb{Z}[G]$ is equal to $p-1$. This is immediately clear from the definition of $\mathbb{Z}[G]$. In $\mathbb{Z}[G]$, we have $\sigma(1-\iota) = \sigma - \sigma\iota$ and $\sigma\iota(1-\iota) = \sigma\iota + \sigma$. It follows that the $\mathbb{Z}$-rank of $\mathbb{Z}[G](1-\iota)$ is equal to $\frac{p-1}{2}$. The index of $I_S^- = I_S(1-\iota)$ in $\mathbb{Z}[G](1-\iota)$ is finite. This is a consequence of the Iwasawa class number formula, which we saw in theorem 4.11. It says that the index $[\mathbb{Z}[G](1-\iota) : I_S^-]$ equals h^-.

So we know that the index of I_S^- in $\mathbb{Z}[G](1-\iota)$ is finite and therefore, the $\mathbb{Z}$-rank of I_S^- has to be equal to $\frac{p-1}{2}$ as well.

Now we are left with the size of the $\tilde{\theta}_k$. First, we compute the weight of θ_S:

$$w(\theta_S) = \sum_{a=1}^{p-1} \frac{a}{p} = \frac{1+2+3+\ldots+p-1}{p} = \frac{1}{p}\frac{p-1}{2}p = \frac{p-1}{2}.$$

Since the weight function is a ring homomorphism, for all integers c that are coprime to p we have:

$$w(\theta_c) = w(c - \sigma_c) \cdot w(\theta_S) = (c-1)\frac{p-1}{2}.$$

As we saw in lemma 4.9, we also have $\theta_c = \sum_{a=1}^{p-1} \left\lfloor \frac{ac}{p} \right\rfloor \sigma_a^{-1}$. This implies that all coefficients of $\theta_{c+1} - \theta_c$ are positive or equal to 0. Therefore

$$\|\theta_{c+1} - \theta_c\| = w(\theta_{c+1} - \theta_c) = \frac{p-1}{2}.$$

It follows that

$$\|\tilde{\theta}_k\| = \|(\theta_{k+1} - \theta_k)(1-\iota)\| \leq \|\theta_{k+1} - \theta_k\| \cdot \|1 - \iota\| = p - 1.$$

This is what we wanted to show. □

Chapter 5

Results by Cassels, Mihăilescu, Bugeaud and Hanrot

In this chapter we derive some results concerning possible solutions of the Catalan equation (1.1) using Stickelberger's theorem. In the first section, we state Cassels' theorem and some of its consequences. In the second section, we state a result of Mihăilescu that will be a very important ingredient of the proofs in chapter 6 and chapter 8.

5.1 Cassels' theorem and some consequences

We start by a theorem that Cassels proved in 1962. For a proof, see [3] or [14].

Theorem 5.1 (Cassels). *Let p and q be odd primes and let x and y be positive integers such that $x^p - y^q = \pm 1$. Then p divides y and q divides x.*

If x and y both are negative, then we can write $(-x)^p - (-y)^q = -x^p + y^q = -1$. From Cassels' theorem it follows that p divides $-y$ and q divides $-x$ and therefore p divides y and q divides x. So it follows that for all non-zero integers x and y such that $x^p - y^q = 1$, p divides y and q divides x.

Cassels' theorem yields the following useful lemma.

Lemma 5.2. *Let p and q be odd primes and let x and y be non-zero integers such that $x^p - y^q = 1$. Then there exist non-zero integers a and b, and positive integers u and v such that*

$$\begin{cases} x - 1 = p^{q-1}a^q \\ \frac{x^p-1}{x-1} = pu^q, \end{cases} \quad \text{where } p \nmid u,\ \gcd(a,u) = 1 \text{ and } y = pau, \tag{5.1}$$

and

$$\begin{cases} y + 1 = q^{p-1}b^p \\ \frac{y^q+1}{y+1} = qv^p, \end{cases} \quad \text{where } q \nmid v,\ \gcd(b,v) = 1 \text{ and } x = qbv. \tag{5.2}$$

Proof. Note that $y^q = x^p - 1 = (x-1)\frac{x^p-1}{x-1}$. From lemma 3.3 we know that $\gcd(x-1, \frac{x^p-1}{x-1}) = 1$ or p. In our case p divides y and therefore p^q divides $y^q = (x-1)\frac{x^p-1}{x-1}$.

If p divides $\frac{x^p-1}{x-1}$, then $x^p - 1 \equiv 0 \pmod p$ and also $x^p - 1 \equiv x - 1 \pmod p$, so p divides $x-1$. Conversely, we have $\frac{x^p-1}{x-1} = x^{p-1} + x^{p-2} + \ldots + x + 1 \equiv p \pmod{x-1}$, so if p divides $x-1$, then p divides $\frac{x^p-1}{x-1}$ as well. It follows that $\gcd(x-1, \frac{x^p-1}{x-1}) = p$.

Let i be an integer greater than or equal to 1 such that p^i divides $x-1$ and p^{i+1} does not. Say, $x - 1 = p^i d$, where $\gcd(p, d) = 1$. Then the following congruence holds:

$$\begin{aligned} x^p - 1 &= (p^i d + 1)^p - 1 \\ &\equiv p^{i+1} d + \frac{1}{2}(p-1)p^{2i+1}d^2 \pmod{p^{3i+1}}. \end{aligned} \tag{5.3}$$

Therefore,

$$\begin{aligned} \frac{x^p - 1}{x - 1} &\equiv \frac{p^{i+1}d + \frac{1}{2}(p-1)p^{2i+1}d^2}{p^i d} \pmod{p^{2i+1}} \\ &\equiv p + \frac{1}{2}(p-1)p^{i+1}d \pmod{p^{2i+1}} \\ &\equiv p \pmod{p^2} \end{aligned} \tag{5.4}$$

and p^2 does not divide $\frac{x^p-1}{x-1}$.

It follows that there exists a non-zero integer a and a positive integer u such that

$$x - 1 = p^{q-1}a^q \tag{5.5}$$

$$\frac{x^p - 1}{x - 1} = pu^q \tag{5.6}$$

and

$$y = pau, \tag{5.7}$$

where p does not divide u and $\gcd(a, u) = 1$. The proof of the second part of the lemma is completely similar to that of the first part. □

Let p and q be odd primes. From now on, we assume that x and y are non-zero integers such that $x^p - y^q = 1$.

Cassels' theorem also can be used to make estimates for the sizes of x and y. From lemma 5.2, for instance, it follows immediately that

$$|x| \geq p^{q-1} - 1 \tag{5.8}$$

and

$$|y| \geq q^{p-1} - 1. \tag{5.9}$$

However, we need stronger estimates than these. From the following theorem we derive a stronger estimate for $|x|$.

Theorem 5.3. *Let p, q and v be defined as in lemma 5.2. If p does not divide $q-1$, then q^{p-2} divides $v-1$.*

Proof. In lemma 5.2 we saw that there is a positive integer v such that $\frac{y^q+1}{y+1} = qv^p$. Therefore,

$$\begin{aligned} q(v^p-1) &= \frac{y^q+1}{y+1} - q = y^{q-1} - y^{q-2} + y^{q-3} - \ldots - y + 1 - q \\ &= y^{q-1} - 1 + -y^{q-2} - 1 + y^{q-3} - 1 - \ldots - y - 1 \\ &= ((-y)^{q-1} - 1) + ((-y)^{q-2} - 1) + \ldots + (-y - 1). \end{aligned} \tag{5.10}$$

Since $y+1$ divides all of the $(-y)^i - 1$, we find that $y+1$ divides $q(v^p-1)$. From lemma 5.2, we also know that there exists a non-zero integer b such that $y + 1 = q^{p-1}b^q$. It follows that $q^{p-1}b^q$ divides $q(v^p-1)$. Therefore, q^{p-2} divides $v^p - 1$, so $v^p \equiv 1 \pmod{q^{p-2}}$. The order of the group $(\mathbb{Z}/q^{p-2}\mathbb{Z})^*$ equals $\varphi(q^{p-2}) = q^{p-3}(q-1)$. According to our assumption, p does not divide this order. It follows that $v \equiv 1 \pmod{q^{p-2}}$, which is what we wanted to show. □

From this theorem, we derive the following estimate.

Corollary 5.4. *The inequality*

$$|x| \geq q^{p-1}$$

holds.

Proof. If p divides $q-1$, then p is smaller than q. Since $x^p = y^q + 1$, we have $|x|^p \geq |y|^q - 1$. Because $p < q$, we find $|x| > |y|$. We use equation (5.9), which states that $|y| \geq q^{p-1} - 1$. Therefore, $|x| > |y| \geq q^{p-1} - 1$, so $|x| \geq q^{p-1}$.

If p does not divide $q-1$, then we use theorem 5.3 and we find that q^{p-2} divides $v - 1$. Since v is greater than 1, this implies that $v \geq q^{p-2} + 1$. Because we have that $x = qvb$, it follows that $|x| = |qvb| \geq qv \geq q^{p-1} + q > q^{p-1}$. □

The lemma we prove next also is a very useful corollary of lemma 5.2.

Lemma 5.5. *The number* $\lambda := \frac{x-\zeta}{1-\zeta}$ *is an element of* $\mathcal{O}_K = \mathbb{Z}[\zeta]$. *The principal ideal* (λ) *is a* q*-th power of an* $\mathcal{O}_K$*-ideal.*

Proof. From (5.1) we know that p divides $x - 1$. Therefore, $x \equiv 1 \pmod{\mathfrak{P}^{p-1}}$. It follows that $x - \zeta \equiv 1 - \zeta \equiv 0 \pmod{\mathfrak{P}}$, so $\mathfrak{P}$ divides $(x - \zeta)$. But $\mathfrak{P}^2$ does not, since $x - \zeta = x - 1 + 1 - \zeta \equiv 1 - \zeta \pmod{\mathfrak{P}^2} \not\equiv 0 \pmod{\mathfrak{P}^2}$. Hence $\lambda = \frac{x-\zeta}{1-\zeta}$ is an algebraic integer such that (λ) is not divisible by $\mathfrak{P}$. Note that $\lambda^\sigma \in \mathcal{O}_K = \mathbb{Z}[\zeta]$ for all $\sigma \in G$, because λ is an element of $\mathcal{O}_K$. Also, for all $\sigma \in G$, $\mathfrak{P}$ does not divide (λ^σ): the same argument applies since $(1-\zeta) = (1-\zeta^\sigma)$ for all $\sigma \in G$.

We will look at the greatest common divisor of the two ideals (λ^σ) and (λ^τ) for σ and τ in G. For all elements λ and τ in G we have the identity

$$(1-\zeta^\sigma)\lambda^\sigma - (1-\zeta^\tau)\lambda^\tau = (1-\zeta^\sigma)\frac{x-\zeta^\sigma}{1-\zeta^\sigma} - (1-\zeta^\tau)\frac{x-\zeta^\tau}{1-\zeta^\tau} = \zeta^\tau - \zeta^\sigma.$$

Therefore, for different elements σ and τ in G, the greatest common divisor of (λ^σ) and (λ^τ) is a divisor of the ideal $(\zeta^\tau - \zeta^\sigma) = \mathfrak{p}$. Since $\mathfrak{P}$ does not divide (λ^σ) or (λ^τ), it follows that $\gcd(\lambda^\sigma, \lambda^\tau) = 1$. So the numbers λ^σ, where $\sigma \in G$, are pairwise coprime.

Let us consider the following product:

$$\prod_{\sigma\in G} \lambda^\sigma = \prod_{\sigma\in G} \frac{x-\zeta^\sigma}{1-\zeta^\sigma} = \frac{x^p-1}{x-1} \prod_{\sigma\in G} \frac{1}{1-\zeta^\sigma} = \frac{1}{p}\frac{x^p-1}{x-1}.$$

From (5.6) we know that $\frac{x^p-1}{x-1} = pu^q$ for a positive integer u, so

$$\prod_{\sigma \in G} \lambda^\sigma = u^q$$

for a positive integer u. Since all principal ideals (λ^σ) are pairwise coprime, each of them is a q-th power of an $\mathcal{O}_K$-ideal. □

5.2 Results by Mihăilescu

In this section we prove several results of Mihăilescu. The most important one is that q^2 divides x. The following lemma will be useful for this purpose. An element a of a ring R is called *nilpotent* if there exists an integer n such that $a^n = 0$.

Lemma 5.6. *The ring $\mathcal{O}_K/(q)$ does not contain nilpotent elements, and if $\alpha, \beta \in \mathcal{O}_K = \mathbb{Z}[\zeta]$ satisfy the congruence $\alpha^q \equiv \beta^q \pmod{q}$, then $\alpha^q \equiv \beta^q \pmod{q^2}$.*

Proof. First, we show that the ring $\mathcal{O}_K/(q)$ does not have nilpotent elements. As we showed in lemma 4.2, q is unramified in $K = \mathbb{Q}(\zeta)$. So (q) is of the form $(q) = \mathfrak{q}_1 \cdot \mathfrak{q}_2 \cdot \ldots \cdot \mathfrak{q}_s$, where the $\mathfrak{q}_i$ are distinct prime ideals of $\mathcal{O}_K = \mathbb{Z}[\zeta]$. The Chinese Remainder Theorem now tells us that

$$\mathcal{O}_K/(q) \cong \mathcal{O}_K/\mathfrak{q}_1 \times \mathcal{O}_K/\mathfrak{q}_2 \times \ldots \mathcal{O}_K/\mathfrak{q}_s.$$

Since all $\mathfrak{q}_i$ are prime ideals, it follows that $\mathcal{O}_K/(q)$ has no nilpotent elements other than 0.

Let $\alpha, \beta \in \mathbb{Z}[\zeta]$ such that $\alpha^q \equiv \beta^q \pmod{q}$. Then we have:

$$(\alpha - \beta)^q \equiv \alpha^q - \beta^q \equiv 0 \pmod{q}.$$

Since $\mathcal{O}_K/(q)$ does not have nilpotent elements other than 0, it follows that $\alpha - \beta \equiv 0 \pmod{q}$. Therefore, $\alpha = \beta + kq$, with $k \in \mathcal{O}_K$. It follows that $\alpha^q = (\beta + kq)^q \equiv \beta^q + q\beta^{q-1}kq + \frac{1}{2}q(q-1)\beta^{q-2}(kq)^2 + \ldots + (kq)^q \equiv \beta^q \pmod{q^2}$. □

Let $\iota = \sigma_{p-1}$ denote complex conjugation. By I_S^- we denote the $\mathbb{Z}[G]$-ideal that is obtained by multiplying the Stickelberger ideal I_S by the element $1 - \iota$, i.e. $I_S^- = I_S(1 - \iota)$. Let α be an element of K^*, then $\alpha^{1-\iota} = \frac{\alpha}{\bar{\alpha}}$, so for all embeddings σ of K in $\mathbb{C}$ we have that $|\sigma(\alpha^{1-\iota})| = 1$.

Now we have the following theorem, which is due to Mihăilescu.

Theorem 5.7 (Mihăilescu). *For any $\theta \in I_S^-$, the element $(x-\zeta)^\theta$ is a q-th power in $K = \mathbb{Q}(\zeta)$. We also have that q^2 divides x and p^2 divides y.*

Proof. Let θ be an element of I_S^- and write $\theta = (1-\iota)\theta'$, with $\theta' \in I_S$. Put $\lambda = \frac{x-\zeta}{1-\zeta}$, as we did in lemma 5.5. This lemma says that (λ) is a q-th power of an $\mathcal{O}_K$-ideal, $(\lambda) = \mathfrak{a}^q$, say. From Stickelberger's theorem it follows that $\mathfrak{a}^{\theta'}$ is a principal ideal, $\mathfrak{a}^{\theta'} = (\alpha)$, say, with $\alpha \in K = \mathbb{Q}(\zeta)$. So $(\lambda^{\theta'}) = (\mathfrak{a}^q)^{\theta'} = (\mathfrak{a}^{\theta'})^q = (\alpha)^q$. Therefore, $\lambda^{\theta'} = \eta\alpha^q$, where η is a unit in $\mathcal{O}_K = \mathbb{Z}[\zeta]$. We obtain a useful equality for the element $(x-\zeta)^\theta$:

$$\begin{aligned}
(x-\zeta)^\theta &= (\lambda(1-\zeta))^\theta = \lambda^{(1-\iota)\theta'}(1-\zeta)^{(1-\iota)\theta'} \\
&= \left(\tfrac{\lambda}{\bar{\lambda}}\right)^{\theta'} \left(\tfrac{1-\zeta}{1-\bar{\zeta}}\right)^{\theta'} = \left(\tfrac{\lambda^{\theta'}}{\bar{\lambda}^{\theta'}}\right) \left(\tfrac{1-\zeta}{1-\bar{\zeta}}\right)^{\theta'} \\
&= \tfrac{\eta}{\bar{\eta}} \left(\tfrac{\alpha}{\bar{\alpha}}\right)^q \left(\tfrac{1-\zeta}{1-\bar{\zeta}}\right)^{\theta'}.
\end{aligned}$$

We will show that $\frac{\eta}{\bar{\eta}}$ is a root of unity. We use lemma 1.6 of [18], which says that if ν is an algebraic integer all of whose conjugates have absolute value 1, then ν is a root of unity. We already know that $\frac{\eta}{\bar{\eta}} \in \mathbb{Z}[\zeta]$ and it is clear that all conjugates of η have absolute value 1. Therefore, $\frac{\eta}{\bar{\eta}}$ is a root of unity. Of course, $\frac{1-\zeta}{1-\zeta} = -\zeta$ is a root of unity as well.

It follows that $(x-\zeta)^\theta$ is equal to a q-th power times a root of unity. All roots of unity in $\mathbb{Q}(\zeta)$ are of the form $\pm\zeta^a$, with a an integer. Since p and q are coprime and q is odd, all these roots of unity are q-th powers in $\mathbb{Q}(\zeta)$. Therefore, the element $(x-\zeta)^\theta$ is a q-th power in $K = \mathbb{Q}(\zeta)$. This settles the first part of the theorem.

Since $(1-\zeta^{-1}x)^\theta$ is equal to $(-\zeta^{-1}(x-\zeta))^\theta$, it is equal to $(x-\zeta)^\theta$ times a root of unity. Since $(x-\zeta)^\theta$ is a q-th power in K, we obtain that $(1-\zeta^{-1}x)^\theta$ is a q-th power in K as well. Since $(x-\zeta)^\theta$ is an element of $\mathbb{Z}[\zeta]$, and $\frac{\eta}{\bar{\eta}}$ and $\left(\frac{1-\zeta}{1-\zeta}\right)^\theta$ are elements of $\mathbb{Z}[\zeta]^*$, we have $\left(\frac{\alpha}{\bar{\alpha}}\right)^q \in \mathcal{O}_K = \mathbb{Z}[\zeta]$. Therefore, $\frac{\alpha}{\bar{\alpha}} \in \mathcal{O}_K = \mathbb{Z}[\zeta]$. So we have that $(x-\zeta)^\theta$ is a q-th power in $\mathbb{Z}[\zeta]$ and so is $(1-\zeta^{-1}x)^\theta$, let us say that $(1-\zeta^{-1}x)^\theta = a^q$.

Using Cassels' result that q divides x (theorem 5.1), we find that $(1-\zeta^{-1}x)^\theta \equiv 1 \pmod q$. Now we use lemma 5.6, applied to $\alpha = a$ and $\beta = 1$, and we obtain

$$(1-\zeta^{-1}x)^\theta \equiv 1 \pmod{q^2}.$$

If $\theta = \sum_{\sigma\in G} n_\sigma \sigma$, then we have

$$\begin{aligned}(1-\zeta^{-1}x)^\theta &= \prod_{\sigma\in G}(1-\sigma(\zeta^{-1})x)^{n_\sigma} \\ &\equiv \prod_{\sigma\in G}(1-n_\sigma\sigma(\zeta^{-1})x) \pmod{q^2} \\ &\equiv 1 - x\sum_{\sigma\in G} n_\sigma\sigma(\zeta^{-1}) \pmod{q^2}.\end{aligned}$$

Therefore, $1-x\sum_{\sigma\in G} n_\sigma\sigma(\zeta^{-1}) \equiv 1 \pmod{q^2}$. Now either q^2 divides x, or q divides $\sum_{\sigma\in G} n_\sigma\sigma(\zeta^{-1})$. In the latter case, q divides n_σ for all $\sigma \in G$. But if, for instance,

$$\theta = (1-\iota)\theta_2 = -\sigma_1^{-1} - \ldots - \sigma_{\frac{p-1}{2}}^{-1} + \sigma_{\frac{p+1}{2}}^{-1} + \ldots + \sigma_{p-1}^{-1},$$

this is not true. We conclude that q^2 divides x, which is what we wanted to prove. If (x,y,p,q) is a solution of the Catalan equation with p and q odd primes and x and y non-zero integers, then so is $(-y,-x,q,p)$. Because we have shown that q^2 divides x, we also know that p^2 divides y. □

From this theorem we derive the following corollary.

Corollary 5.8. *Let p and q be odd primes. If $x^p - y^q = 1$ has a solution in non-zero integers x and y, then the congruences*

$$p^{q-1} \equiv 1 \pmod{q^2}$$

and

$$q^{p-1} \equiv 1 \pmod{p^2}$$

hold.

This result is called the double Wieferich *relation.*

Proof. In theorem 5.7 we saw that q^2 divides x. Together with equation (5.1) this yields

$$p^{q-1}a^q \equiv -1 \pmod{q^2}, \tag{5.11}$$

with a a non-zero integer. Since $p^{q-1} \equiv 1 \pmod{q}$, it follows that $a^q \equiv -1 \pmod{q}$. Therefore, $a \equiv -1 \pmod{q}$ and $a^q \equiv -1 \pmod{q^2}$. Together with equation (5.11) this implies that

$$p^{q-1} \equiv 1 \pmod{q^2}.$$

By symmetry, we also find that $q^{p-1} \equiv 1 \pmod{p^2}$. □

5.3 Small p and q

Our proof in chapter 6.1 only works if p and q are at least 5, and our proof in chapter 8 needs p and q to be at least 7. So we have to deal with the cases in which p or q are small separately.

In 2000, Bugeaud and Hanrot [19] published an article in which they proved that the Catalan equation has no solution in non-negative integers x and y if p or q is smaller than 43.

The proof we give for the small cases, depends on the following theorem. The proof is similar to that of Bugeaud and Hanrot and our version is based on the treatment of René Schoof.

First, we prove the following lemma.

Lemma 5.9. *The following equality holds:*

$$\sum_{i=1}^{p-1} \frac{\zeta^i}{(1-\zeta^i)^2} = \frac{1-p^2}{12}.$$

Proof. Define the function $f(X)$ as the p-th cyclotomic polynomial, i.e.

$$f(X) = \prod_{\zeta \neq 1} (X - \zeta) = X^{p-1} + X^{p-2} + \ldots + X + 1.$$

We compute the logarithmic derivative of f and we obtain:

$$\frac{f'(X)}{f(X)} = \sum_{\zeta \neq 1} \frac{1}{X-\zeta} = \sum_{\zeta \neq 1} \frac{\zeta^{-1}}{X\zeta^{-1} - 1} = \sum_{\zeta \neq 1} \frac{\zeta}{\zeta X - 1}.$$

Therefore,

$$\begin{aligned} \sum_{\zeta \neq 1} \frac{\zeta}{\zeta - X} &= \frac{1}{X} \sum_{\zeta \neq 1} \frac{\zeta}{\zeta \frac{1}{X} - 1} = \frac{1}{X} \frac{f'(\frac{1}{X})}{f(\frac{1}{X})} = \frac{X^{p-2} f'(\frac{1}{X})}{X^{p-1} f(\frac{1}{X})} \\ &= \frac{X^{p-2} + 2X^{p-3} + \ldots + (p-2)X + p - 1}{X^{p-1} + X^{p-2} + \ldots + X + 1}. \end{aligned} \tag{5.12}$$

So define $g(X) = X^{p-2} + 2X^{p-3} + \ldots + (p-2)X + p - 1$, and we find

$$\sum_{\zeta \neq 1} \frac{\zeta}{(1-\zeta)^2} = -\left(\frac{g}{f}\right)'(1) = \frac{g(1)f'(1) - f(1)g'(1)}{(f(1))^2}.$$

We can compute these numbers: $f(1) = p$, $f'(1) = 1 + 2 + \ldots + p - 1 = \frac{1}{2}p(p-1)$, $g(1) = p - 1 + p - 2 + \ldots + 2 + 1 = \frac{1}{2}p(p-1)$ and $g'(1) = \sum_{i=1}^{p-1}(p-i)(i-1) = \sum_{i=1}^{p-1}((p+1)i - p + i^2) = (p+1)\frac{1}{2}p(p-1) - p(p-1) + \frac{1}{6}p(p-1)(2p-1)$. So we find:

$$\begin{aligned}\sum_{\zeta \neq 1} \frac{\zeta}{(1-\zeta)^2} &= \frac{\frac{1}{4}p^2(p-1)^2 - p\sum_{i-1}^{p-1}(p-i)(i-1)}{p^2} \\ &= \frac{\frac{1}{4}p^2(p-1)^2 - \frac{1}{6}p^2(p-1)(2p-1)}{p^2} \\ &= \frac{1-p^2}{12}. \end{aligned} \tag{5.13}$$

□

Theorem 5.10. *Let p and q be odd primes. Let h_p^- be the relative class number of $\mathbb{Q}(\zeta)$. If q does not divide h_p^-, then the Catalan equation $x^p - y^q = 1$ has no solution in non-zero integers x and y.*

Proof. Define π as the generator $1 - \zeta$ of the ideal $\mathfrak{P}$ above p. From equation (5.1) we know that $x \equiv 1 \pmod{p^{q-1}} \equiv 1 \pmod{\pi^4}$.

In lemma 5.5 we saw that

$$\left(\frac{x-\zeta}{1-\zeta}\right) = \mathfrak{a}^q,$$

with $\mathfrak{a}$ an ideal in $\mathcal{O}_K = \mathbb{Z}[\zeta]$. We see that $\mathfrak{a}^q$ is a principal ideal, so q divides the class number h_p of $\mathbb{Q}(\zeta)$. By assumption, q does not divide h_p^-, so q divides h_p^+, which is the class number of $K^+ = \mathbb{Q}(\zeta + \zeta^{-1})$.

From theorem 14 in [18], we know that the natural map $\psi : Cl_{K^+} \longrightarrow Cl_K$ is injective. Therefore, $\#(Cl_K/\psi(Cl_{K^+})) = \frac{h_p}{h_p^+} = h_p^-$. Consider the class $[\mathfrak{a}]^q\psi(Cl_{K^+}) = [1]$, since $\mathfrak{a}^q$ is a principal ideal. But on the other hand, $[\mathfrak{a}]^q\psi(Cl_{K^+}) = ([\mathfrak{a}]\psi(Cl_{K^+}))^q$. Since $\gcd(h_p^-, q) = 1$, it follows that $[\mathfrak{a}]\psi(Cl_{K^+}) \in \psi(Cl_{K^+})$. Therefore, there exists an ideal class $[\mathfrak{b}] \in Cl_{K^+}$ such that $\psi([\mathfrak{b}]) = [\mathfrak{a}]$. So there exists an $\gamma \in K^* = \mathbb{Q}(\zeta)^*$ such that $\mathfrak{a} = (\gamma)\mathfrak{b}'$, where we view $\mathfrak{b}' = \mathfrak{b}\mathcal{O}_K$ as an $\mathcal{O}_K$-ideal. From this equality we see that $\mathfrak{b}'^q$ is a principal ideal as well. We find that there exist an element $\beta \in K^* = \mathbb{Q}(\zeta)^*$ and a unit η_0 such that $\frac{x-\zeta}{1-\zeta} = \eta_0\beta\gamma^q$.

Put $\mu = \frac{x-1}{\zeta-1} = \frac{x-1}{\pi}$. Note that μ is an algebraic integer, as we saw in lemma 5.5. We have

$$\mu + 1 = \frac{x-\zeta}{1-\zeta}$$

and

$$\bar{\mu} + 1 = \frac{x-\bar{\zeta}}{1-\bar{\zeta}}.$$

It follows that $\frac{\mu+1}{\bar{\mu}+1} = \frac{x-\zeta}{1-\zeta}\frac{1-\bar{\zeta}}{x-\zeta} = -\zeta^{-1}\frac{x-\zeta}{x-\zeta}$, so there exists $\alpha \in \mathbb{Q}(\zeta)$ such that

$$\frac{\mu+1}{\bar{\mu}+1} = \alpha^q.$$

Consider the element

$$\eta = (\sqrt[q]{1+\mu} + \zeta^{\frac{-1}{q}}\sqrt[q]{1+\bar{\mu}})^q = (1+\bar{\mu})(\alpha + \zeta^{\frac{-1}{q}})^q.$$

This last equality is an equality of elements of $\mathbb{Q}(\zeta)$. Of course, $(1+\bar{\mu})\alpha^i$ is integral for all integers $i \leq q$. Since $\zeta^{\frac{-1}{q}} = \zeta^r$ for some integer r, we find that η is integral, i.e. is an element of $\mathbb{Z}[\zeta]$. It is obvious that $\sqrt[q]{1+\mu} + \zeta^{\frac{-1}{q}}\sqrt[q]{1+\bar{\mu}}$ divides $(1+\mu+\zeta^{-1}(1+\bar{\mu}))^q$, since for all $a, b \in \mathbb{Q}(\zeta)$ we have $a+b = (a^{\frac{1}{q}} + b^{\frac{1}{q}})\sum_{i=0}^{q-1} x^{\frac{q-1-i}{q}} y^{\frac{i}{q}}$. We have

$$1+\mu+\zeta^{-1}(1+\bar{\mu}) = \frac{x-\zeta}{1-\zeta} + \zeta^{-1}\frac{x-\bar{\zeta}}{1-\bar{\zeta}} = \frac{x-\zeta}{1-\zeta} + \frac{\bar{\zeta}-x}{1-\zeta} = \zeta^{-1}\frac{1-\zeta^2}{1-\zeta},$$

which is a unit as we saw in lemma 4.3. Therefore, η is a unit in $\mathbb{Z}[\zeta]$ itself, so

$$N_{\mathbb{Q}(\zeta)/\mathbb{Q}}(\eta) = 1.$$

Note that $\mu = \frac{x-1}{1-\zeta} = \frac{x-1}{\pi}$ has at least 3 factors π, so $\mu+1$ is very close to 1 under the π-adic valuation. Now Hensel's lemma applied to the valuation ring $\mathbb{Z}_p[\zeta]$ and polynomial $X^q - (\mu+1)$ implies that $u = \sqrt[q]{1+\mu} + \zeta^{\frac{-1}{q}}\sqrt[q]{1+\bar{\mu}} \in \mathbb{Z}_p[\zeta]$. In $\mathbb{Z}_p[\zeta]$, the norm $N(u)$ of u equals 1.

Using the Taylor series expansion of $\sqrt[q]{1+\mu}$, we find the following equalities. Note that μ depends on ζ.

$$\begin{aligned}
N(u) &= \prod_{\zeta\neq 1}(1 + \frac{\mu}{q} + \zeta^{-\frac{1}{q}} + \frac{\zeta^{-\frac{1}{q}}\bar{\mu}}{q} + (\mathrm{mod}\,\mu^2)) \\
&= \prod_{\zeta\neq 1}(1+\zeta^{-\frac{1}{q}}) \prod_{\zeta\neq 1}\left(1 + \frac{\frac{\mu}{q} + \frac{\zeta^{-\frac{1}{q}}\bar{\mu}}{q}}{1+\zeta^{-\frac{1}{q}}} + (\mathrm{mod}\,\mu^2)\right) \\
&= N(-1-\zeta)\prod_{\zeta\neq 1}\left(1 + \frac{\frac{\mu}{q} + \frac{\zeta^{-\frac{1}{q}}\bar{\mu}}{q}}{1+\zeta^{-\frac{1}{q}}} + (\mathrm{mod}\,\mu^2)\right) \\
&= 1 + \sum_{\zeta\neq 1}\frac{\mu + \zeta^{-\frac{1}{q}}\bar{\mu}}{q(1+\zeta^{-\frac{1}{q}})} + (\mathrm{mod}\,\mu^2). \qquad (5.14)
\end{aligned}$$

This last equality holds because $N(-1-\zeta) = (-1)^{p-1} + (-1)^{p-2} + \ldots + (-1) + 1 = 1$.

Writing μ as $\frac{x-1}{1-\zeta}$ we find:

$$1 = N(u) = 1 + \sum_{\zeta\neq 1}\frac{x-1}{q}\frac{\frac{1}{1-\zeta} + \frac{\zeta^{-\frac{1}{q}}}{1-\zeta}}{1+\zeta^{-\frac{1}{q}}} + (\mathrm{mod}\left((x-1)\left(\frac{x-1}{\pi^2}\right)\right)). \qquad (5.15)$$

Since q and q are coprime integers, there exists $r \in \{0, 1, \ldots, p-1\}$ such that $r \equiv -\frac{1}{q}\,(\mathrm{mod}\,p)$. Using this, we rewrite one of the terms in the expression above and we obtain:

$$\begin{aligned}
\frac{\frac{1}{1-\zeta} + \frac{\zeta^{-\frac{1}{q}}}{1-\zeta}}{1+\zeta^{-\frac{1}{q}}} &= \frac{1-\zeta^{1+r}}{(1-\zeta)(1+\zeta^r)} = \frac{1-(1+\pi)^{r+1}}{-\pi(1+(1+\pi)^r)} \\
&= \frac{1-1-(r+1)\pi + (\mathrm{mod}\,\pi^2)}{-\pi(2+r\pi + (\mathrm{mod}\,\pi^2))} = \frac{r+1}{2} + (\mathrm{mod}\,\pi). \qquad (5.16)
\end{aligned}$$

Note that $\frac{r+1}{2}$ is independent of ζ.

It follows that

$$\sum_{\zeta \neq 1} \frac{1-\zeta^r}{(1-\zeta)(1+\zeta^r)} \equiv (p-1)\frac{r-1}{2} \,(\mathrm{mod}\, \pi). \tag{5.17}$$

Since we work in $\mathbb{Z}_p[\zeta]$, we obtain also

$$\sum_{\zeta \neq 1} \frac{1-\zeta^r}{(1-\zeta)(1+\zeta^r)} \equiv (p-1)\frac{r-1}{2} \,(\mathrm{mod}\, p). \tag{5.18}$$

From (5.15) and (5.18) we get:

$$1 \equiv 1 + \frac{(x-1)(p-1)(r+1)}{2q} \,(\mathrm{mod}(x-1)p).$$

So p divides $r+1$, so $q \equiv 1 \,(\mathrm{mod}\, p)$. Therefore,

$$\mu + \zeta^{-\frac{1}{q}}\bar{\mu} = \frac{x-1}{1-\zeta} + \zeta^{-1}\frac{x-1}{1-\bar{\zeta}} = 0,$$

so the linear term in the Taylor series of $\sqrt[q]{1+\mu} + \zeta^{-\frac{1}{q}}\sqrt[q]{1+\bar{\mu}}$ equals 0. So,

$$\begin{aligned} N(u) &= \prod_{\zeta \neq 1}(1 + \binom{\frac{1}{q}}{2}\mu^2 + \zeta^{-\frac{1}{q}} + \zeta^{-\frac{1}{q}}\binom{\frac{1}{q}}{2}\bar{\mu}^2) + \,(\mathrm{mod}\, \mu^3) \\ &= 1 + \sum_{\zeta \neq 1}\binom{\frac{1}{q}}{2}(x-1)^2\frac{\frac{1}{(1-\zeta)^2} + \frac{\zeta^{-1}}{(1-\bar{\zeta})^2}}{1+\zeta^{-1}} + \,(\mathrm{mod}\, \mu^3). \end{aligned} \tag{5.19}$$

We compute the term $\frac{\frac{1}{(1-\zeta)^2} + \frac{\zeta^{-1}}{(1-\bar{\zeta})^2}}{1+\zeta^{-1}}$, using the identity $-\bar{\zeta}(1-\zeta) = 1-\bar{\zeta}$:

$$\begin{aligned} \frac{\frac{1}{(1-\zeta)^2} + \frac{\zeta^{-1}}{(1-\bar{\zeta})^2}}{1+\zeta^{-1}} &= \frac{1+\zeta^{-1}(-\zeta)^2}{(1-\zeta)^2(1+\zeta^{-1})} = \frac{1+\zeta}{(1-\zeta)^2(1+\zeta^{-1})} \\ &= \frac{(1+\zeta)\zeta}{(1-\zeta)^2(\zeta+1)} = \frac{\zeta}{(1-\zeta)^2}. \end{aligned} \tag{5.20}$$

So we conclude:

$$N(u) = 1 + \sum_{\zeta \neq 1}\binom{\frac{1}{q}}{2}(x-1)^2\frac{\zeta}{(1-\zeta)^2} + \,(\mathrm{mod}\, \mu^3). \tag{5.21}$$

Lemma 5.9 implies that

$$N(u) = 1 + \binom{\frac{1}{q}}{2}(x-1)^2\frac{1-p^2}{12} + \,(\mathrm{mod}\left((x-1)^2\frac{x-1}{\pi^3}\right)).$$

It follows that $\frac{x-1}{\pi^3}$ divides $\binom{\frac{1}{q}}{2}\frac{1-p^2}{12}$ in $\mathbb{Z}_p[\zeta]$. Therefore, $x-1$ divides $\frac{(q-1)(1-p^2)\pi^3}{3}$ in $\mathbb{Z}_p[\zeta]$, so p^{q-1} divides $q-1$, which is a contradiction. Therefore, the equation $x^p - y^q = 1$ does not have any solutions in non-zero integers x and y. □

By symmetry, it follows that if p does not divide h_q^-, then the Catalan equation $x^p - y^q = 1$ has no solution in non-zero integers x and y as well.

Corollary 5.11. *If p or q is smaller than 7, then the Catalan equation $x^p - y^q = 1$ has no solution in non-zero integers x and y.*

Proof. Let us say that $p < 7$. Then $h_p^- = 1$, so q does not divide h_p^-, since q is prime. Now we apply theorem 5.10 and we are done. □

Chapter 6

The first case: q divides $p-1$

In this chapter we deal with the case in which q divides $p-1$. Assuming that a solution of the Catalan equation exists in this case, we finally arrive at a contradiction, so actually we prove in this chapter that if x and y are integers and p and q are odd primes such that $x^p - y^q = 1$, then q does not divide $p-1$. We follow the treatment of René Schoof.

We noted already that if (x, y, p, q) is a solution of the Catalan equation, then so is $(-y, -x, q, p)$. Thus if we show that there are no solutions to the Catalan equation in which p divides $q-1$, then we have also shown that there are no solutions to the Catalan equation in which q divides $p-1$.

First, we deal with the case in which p and q are both at least 5. The following theorem implies the result of this chapter.

Theorem 6.1 (Mihăilescu). *Let p and q be primes that are at least equal to 5. Put $s = \left\lfloor \frac{3q}{2(p-1)^2} \right\rfloor$. If $\binom{s+\frac{p-1}{2}}{s} > \frac{s+1}{3}(p-1)^2 + 1$, then the equation $x^p - y^q = 1$ has no solutions in non-zero integers x and y.*

What this theorem tells us is that if we have two primes p and q that are at least 5, then the Catalan equation has no solution if q is 'much larger' than p. We will show that if p divides $q-1$, then the condition of the theorem is fulfilled.

For the rest of this chapter, let p and q be primes with p, $q \geq 5$ and assume that x and y are non-zero integers such that $x^p - y^q = 1$.

Before we start proving theorem 6.1, we prove the following lemma.

Lemma 6.2. *If $\binom{s+\frac{p-1}{2}}{s} > \frac{s+1}{3}(p-1)^2 + 1$, then there exist more than q distinct elements $\theta \in I_S^-$ such that*

$$\|\theta\| \leq \frac{3}{2}\frac{q}{p-1}.$$

Proof. We use lemma 4.12. It tells us that the elements $\tilde{\theta}_1, \ldots, \tilde{\theta}_{\frac{p-1}{2}}$ form a $\mathbb{Z}$-basis

of I_S^- and that they all satisfy $\|\tilde{\theta}_i\| \leq p-1$. Now consider the elements of the form

$$\theta = \sum_{i=1}^{\frac{p-1}{2}} \lambda_i \tilde{\theta}_i,$$

with $\lambda_i \in \mathbb{Z}_{\geq 0}$. If the λ_i satisfy

$$\sum_{i=1}^{\frac{p-1}{2}} \lambda_i \leq s = \left\lfloor \frac{3}{2} \frac{q}{(p-1)^2} \right\rfloor,$$

then θ has the propery that $\|\theta\| \leq \frac{3}{2}\frac{q}{p-1}$.

It is easy to see that there are $\binom{s+\frac{p-1}{2}}{s}$ elements θ like this. Choosing $\frac{p-1}{2}$ dots in a row of $s + \frac{p-1}{2}$ is the same as choosing non-negative integers $\lambda_1, \ldots, \lambda_{\frac{p-1}{2}+1}$ that add up to s. Of course, this is the same as choosing non-negative integers $\lambda_1, \ldots, \lambda_{\frac{p-1}{2}}$ such that $\sum_{i=1}^{\frac{p-1}{2}} \lambda_i \leq s$.

Further, there are as many elements $\theta = \sum_{i=1}^{p-1} \lambda_i \tilde{\theta}_i$ that satisfy $\|\theta\| \leq \frac{2}{3}\frac{q}{p-1}$ with all $\lambda_i \in \mathbb{Z}_{\leq 0}$. Therefore, there are at least $2\binom{s+\frac{p-1}{2}}{s} - 1$ elements θ in I_S^- that satisfy $\|\theta\| \leq \frac{3}{2}\frac{q}{p-1}$.

By the assumption that $\binom{s+\frac{p-1}{2}}{s} > \frac{s+1}{3}(p-1)^2 + 1$, this number is strictly larger than $\frac{2}{3}(s+1)(p-1)^2$, which is at least q. It follows that the number of elements θ in I_S^- that satisfy $\|\theta\| \leq \frac{3}{2}\frac{q}{p-1}$ is greater than q, which is what we wanted to show. $\square$

Proof of theorem 6.1. In this proof, we will use the definitions and results from section 4.4 and chapter 5. We work in the group ring $\mathbb{Z}[G]$ and the field $\mathbb{Q}(\zeta)$.

We assumed that there exists a solution in non-zero integers x and y to the Catalan equation $x^p - y^q = 1$. We will show that the absolute value of x is very small, which is contradictory to corollary 5.4. We do this by constructing a special element θ in the ideal I_S^-, and manipulating with the element $\alpha \in K^*$ that satisfies $(x-\zeta)^\theta = \alpha^q$. Therefore, such a solution does not exist.

By theorem 5.7 we know that for all $\theta \in I_S^-$ we have $(x-\zeta)^\theta = \alpha^q$ for some $\alpha \in K^*$. For a given θ, this element α is unique, since $K = \mathbb{Q}(\zeta)$ does not contain any primitive q-th roots of unity.

The map $\varphi : I_S^- \to K^*$ given by $\theta \mapsto \alpha$, where α is the element of K^* such that $(x-\zeta)^\theta = \alpha^q$, is an injective homomorphism. It is easy to see that φ indeed is a homomorphism. and we want to show that this homomorphism φ is injective. Suppose we have two elements θ and θ' of I_S^- such that $\varphi(\theta) = \varphi(\theta')$. That means that $(x-\zeta)^\theta = (x-\zeta)^{\theta'}$, which is equivalent to $(x-\zeta)^{\theta-\theta'} = 1$. Therefore, it suffices to show that $x-\zeta$ and all its conjugates $x-\zeta^\sigma$ are multiplicatively independent, i.e. if $\prod_{\sigma\in G}(x-\zeta^\sigma)^{n_\sigma} = 1$, then all the n_σ are equal to 0.

So assume we have $\prod_{\sigma\in G}(x-\zeta^\sigma)^{n_\sigma} = 1$. As we did in lemma 5.5, put $\lambda = \frac{x-\zeta}{1-\zeta}$. We know that (λ^σ) and (λ^τ) are coprime for distinct σ and τ in G. We also saw before that the ideal $(1-\zeta)^2$ does not divide $(x-\zeta)$, nor its conjugates. It follows that the ideals $(x-\zeta^\sigma)$ and $(x-\zeta^\tau)$ only have one factor $(1-\zeta)$ in common. We show that

all ideals $(x-\zeta^\sigma)$ have a prime factor distinct from $(1-\zeta)$. It is known that $(1-\zeta)$ divides $(x-\zeta^\sigma)$ for all $\sigma \in G$, so we have to show that $(x-\zeta^\sigma)$ does not divide $(1-\zeta)$. It is easy to see that this is the case, because the norm $N(x-\zeta^\sigma) > p$ and $N(1-\zeta) = p$. Let $\mathfrak{p}_\sigma$ be a prime divisor of $x-\zeta^\sigma$, that is distinct from $(1-\zeta)$. Let $e \geq 1$ be the multiplicity of $\mathfrak{p}_\sigma$ in $(x-\zeta^\sigma)$. It follows that the ideal $\mathfrak{p}_\sigma^{en_\sigma}$ divides 1, which of course implies that $n_\sigma = 0$. So all the n_σ are equal to 0, which is what we wanted to prove. It follows that the homomorphism φ is injective.

The α we find for $\theta \in I_S^-$ lies on the unit circle under all embeddings $\tau : \mathbb{Q}(\zeta) \hookrightarrow \mathbb{C}$. We can show this in the following way. If γ is an element of K^*, then we have for all embeddings τ:

$$|\tau(\gamma^{1-\iota})|^2 = |\tau(\gamma)^{1-\iota}| = 1.$$

Since $1-\iota$ divides $\theta = \theta'(1-\iota)$ for all $\theta \in I_S^-$, we have that $|\tau((x-\zeta)^\theta)| = 1$ for all embeddings τ, so $|\tau(\alpha^q)| = |\tau(\alpha)|^q = 1$ and therefore, $|\tau(\alpha)| = 1$.

All elements θ of the ideal I_S^- are of the form $\theta = \theta'(1-\iota)$, where $\theta' \in I_S$. Therefore, $w(\theta) = w(\theta'(1-\iota)) = w(\theta')w(1-\iota) = 0$. It follows that the ideal I_S^- is contained in the augmentation ideal I_{aug} of $\mathbb{Z}[G]$. In particular, for any $\theta \in I_S^-$ we have $x^\theta = x^{w(\theta)} = 1$, and therefore,

$$(1-\frac{\zeta}{x})^\theta = (\frac{x-\zeta}{x})^\theta = (x-\zeta)^\theta. \tag{6.1}$$

Choose an embedding $\sigma : \mathbb{Q}(\zeta) \hookrightarrow \mathbb{C}$.

Consider the principal branch of the complex logarithm, i.e. $\log(re^{ia}) = \log r + ia$ for all elements re^{ia}, with $r \in \mathbb{R}$ and $-\pi < a \leq \pi$. For an element $z \in \mathbb{C}$, define $\text{Arg}(z)$ as the principal value of the argument of z, i.e. $|z|e^{\text{Arg}(z)} = z$ and $-\pi < \text{Arg}(z) \leq \pi$. With this convention, we have for all $z_1, z_2 \in \mathbb{C}$:

$$|\log(z_1 z_2)| \leq |\log(z_1)| + |\log(z_2)|.$$

Equation (6.1) implies that $\alpha^q = (1-\frac{\zeta}{x})^\theta$. Therefore, $\sigma(\alpha^q)$ is rather close to 1 and we can use the Taylor expansion of the principal branch of the logarithm. We estimate $|\log(\sigma(\alpha)^q)|$:

$$\begin{aligned} |\log(\sigma(\alpha)^q)| &= |\log(\sigma(\alpha^q))| = |\log(\sigma((1-\frac{\zeta}{x})^\theta))| \\ &= |\log(\prod_{\tau\in G}(1-\frac{\sigma(\zeta^\tau)}{x})^{n_\tau})| \leq \sum_{\tau\in G}|n_\tau| \cdot |\log(1-\frac{\sigma(\zeta^\tau)}{x})|. \end{aligned} \tag{6.2}$$

To be able to estimate this expression further, we use the Taylor expansion of the logarithm to estimate the $|\log(1-\frac{\sigma(\zeta^\tau)}{x})|$. For all $\tau \in G$ we have:

$$|\log(1-\frac{\sigma(\zeta^\tau)}{x})| = \left|-\frac{\sigma(\zeta^\tau)}{x} - \frac{1}{2}\left(\frac{\sigma(\zeta^\tau)}{x}\right)^2 - \frac{1}{3}\left(\frac{\sigma(\zeta^\tau)}{x}\right)^3 - \ldots\right|$$

$$
\begin{aligned}
&\leq \left|\frac{\sigma(\zeta^\tau)}{x}\right| + \frac{1}{2}\left|\frac{\sigma(\zeta^\tau)^2}{x^2}\right| + \frac{1}{3}\left|\frac{\sigma(\zeta^\tau)^3}{x^3}\right| + \dots \\
&= \frac{1}{|x|} + \frac{1}{2}\frac{1}{|x|^2} + \frac{1}{3}\frac{1}{|x|^3} + \dots \\
&\leq \frac{1}{|x|}\left(1 + \frac{1}{|x|} + \frac{1}{|x|^2} + \frac{1}{|x|^3} + \dots\right) \\
&= \frac{1}{|x|}\frac{1}{1-\frac{1}{|x|}}. \qquad (6.3)
\end{aligned}
$$

Putting the equations (6.2) and (6.3) together, we obtain:

$$
\begin{aligned}
|\log(\sigma(\alpha)^q)| &\leq \sum_{\tau\in G} |n_\tau| \cdot \left(\frac{1}{|x|}\frac{1}{1-\frac{1}{|x|}}\right) \\
&\leq \sum_{\tau\in G} |n_\tau| \left(\frac{1}{|x|}\frac{1}{1-\frac{1}{3}}\right) \\
&= \sum_{\tau\in G} |n_\tau| \frac{3}{2|x|} \\
&= \|\theta\| \frac{3}{2|x|}. \qquad (6.4)
\end{aligned}
$$

Here we used that $|x| \geq 3$. This is obviously true, see for instance (5.1).

Since $|\sigma(\alpha)^q| = 1$, we know that

$$|\log(\sigma(\alpha)^q)| = |\log|\sigma(\alpha)^q| + i\operatorname{Arg}(\sigma(\alpha)^q)| = |\operatorname{Arg}(\sigma(\alpha)^q)|.$$

Together with (6.4) this implies that

$$|\operatorname{Arg}(\sigma(\alpha)^q)| = |\log(\sigma(\alpha)^q)| \leq \frac{3}{2}\frac{\|\theta\|}{|x|}. \qquad (6.5)$$

Now we know that $-\frac{3}{2}\frac{\|\theta\|}{|x|} \leq \operatorname{Arg}(\sigma(\alpha)^q) \leq \frac{3}{2}\frac{\|\theta\|}{|x|}$. Also, we have that $\operatorname{Arg}(\sigma(\alpha)^q) \equiv q\operatorname{Arg}(\sigma(\alpha)) \pmod{2\pi}$. We conclude that there exists an integer k associated to θ satisfying $-\frac{q}{2} < k < \frac{q}{2}$ and

$$|\operatorname{Arg}(\sigma(\alpha)) - \frac{2k\pi}{q}| \leq \frac{3}{2}\frac{\|\theta\|}{q|x|}. \qquad (6.6)$$

Note that this inequality tells us that for all $\theta \in I_S^-$, the corresponding $\sigma(\alpha)$ is close to a q-th root of unity, since we already saw that $|\sigma(\alpha)| = 1$.

If $\|\theta\| \leq \frac{q}{p-1}$, then there is exactly one integer k such that inequality (6.6) is satisfied. We can see this as follows. Suppose there are two integers k and l, such that both $|\operatorname{Arg}(\sigma(\alpha)) - \frac{2k\pi}{q}| \leq \frac{3}{2}\frac{\|\theta\|}{q|x|}$ and $|\operatorname{Arg}(\sigma(\alpha)) - \frac{2l\pi}{q}| \leq \frac{3}{2}\frac{\|\theta\|}{q|x|}$. Then we obtain:

$$|\frac{2k\pi}{q} - \frac{2l\pi}{q}| \leq |\operatorname{Arg}(\sigma(\alpha)) - \frac{2k\pi}{q}| + |\operatorname{Arg}(\sigma(\alpha)) - \frac{2l\pi}{q}| \leq \frac{3\|\theta\|}{q|x|}. \qquad (6.7)$$

Therefore,

$$|k-l| \leq \frac{3}{2\pi}\frac{\|\theta\|}{|x|} \leq \frac{3}{2\pi}\frac{q}{(p-1)|x|} \leq \frac{3}{2\pi}\frac{q}{(p-1)q^{p-1}} < 1,$$

where we use corollary 5.4. It follows that $k = l$.

In lemma 6.2 we saw that there are more than q elements θ in I_S^- with size $\|\theta\| \leq \frac{q}{p-1}$. The box principle now implies that there exist two distinct elements in I_S^- of size at most $\frac{3}{2}\frac{q}{p-1}$ to which the same integer k is associated. This means that these elements have corresponding α's that are close to the same q-th root of unity. Let us call them $\bar{\theta}_1$ and $\bar{\theta}_2$. For $i = 1, 2$, let α_i denote the element such that $(x-\zeta)^{\bar{\theta}_i} = \alpha_i^q$. Define $\bar{\theta} = \bar{\theta}_1 - \bar{\theta}_2$ and let α denote the number such that $(x-\zeta)^{\bar{\theta}} = \alpha^q$. It follows that $\alpha = \frac{\alpha_1}{\alpha_2}$. Since $\bar{\theta}_1$ and $\bar{\theta}_2$ are close to the same q-th root of unity, we expect $\sigma(\alpha)$ to be close tp 1. We make this explicit in the following estimations. Note that $|\log(\sigma(\alpha))|$ equals $|\operatorname{Arg}(\sigma(\alpha))|$, because $|\sigma(\alpha)| = 1$. We estimate $|\log(\sigma(\alpha))|$:

$$\begin{aligned}
|\log(\sigma(\alpha))| &= |\operatorname{Arg}(\sigma(\alpha))| \\
&\leq |\operatorname{Arg}(\sigma(\alpha_1)) - \operatorname{Arg}(\sigma(\alpha_2))| \\
&= |\operatorname{Arg}(\sigma(\alpha_1)) - \frac{2k\pi}{q} + \frac{2k\pi}{q} - \operatorname{Arg}(\sigma(\alpha_2))| \\
&\leq |\operatorname{Arg}(\sigma(\alpha_1)) - \frac{2k\pi}{q}| + |\operatorname{Arg}(\sigma(\alpha_2)) - \frac{2k\pi}{q}| \\
&\leq \frac{3}{2}\frac{\|\theta_1\|}{q|x|} + \frac{3}{2}\frac{\|\theta_2\|}{q|x|} \\
&\leq 2 \cdot \frac{3}{2}\frac{\frac{3q}{2(p-1)}}{q|x|} = \frac{3^2}{2}\frac{1}{|x|(p-1)}.
\end{aligned} \tag{6.8}$$

Using this inequality, we estimate $|\sigma(\alpha) - 1|$. We also use the Taylor expansion of the exponential function.

$$\begin{aligned}
|\sigma(\alpha) - 1| &= |e^{\log(\sigma(\alpha))} - 1| \\
&= |-1 + 1 + \log(\sigma(\alpha)) + \frac{1}{2!}\log(\sigma(\alpha))^2 + \frac{1}{3!}\log(\sigma(\alpha))^3 + \dots| \\
&\leq |\log(\sigma(\alpha))| + \frac{1}{2!}|\log(\sigma(\alpha))|^2 + \frac{1}{3!}|\log(\sigma(\alpha))|^3 + \dots \\
&\leq \frac{3^2}{2}\frac{1}{|x|(p-1)} + \frac{1}{2!}\left(\frac{3^2}{2}\frac{1}{|x|(p-1)}\right)^2 + \frac{1}{3!}\left(\frac{3^2}{2}\frac{1}{|x|(p-1)}\right)^3 + \dots \\
&\leq \frac{3^2}{2}\frac{1}{|x|(p-1)}\left(1 + \frac{1}{|x|} + \frac{1}{|x|^2} + \dots\right) \\
&= \frac{3^2}{2}\frac{1}{|x|(p-1)|}\frac{1}{1-\frac{1}{|x|}} \\
&\leq \frac{3^2}{2}\frac{1}{|x|(p-1)}\frac{1}{1-\frac{1}{3}} = \frac{3^3}{2^2}\frac{1}{|x|(p-1)}.
\end{aligned} \tag{6.9}$$

We find that $\sigma(\alpha - 1)$ is very small, which indeed implies that $\sigma(\alpha)$ is close to 1. Of course, since $|\sigma(\alpha) - 1| = \sigma(\alpha - 1)\overline{\sigma(\alpha - 1)}$, the same is true for its complex conjugate. For all other embeddings $\tau : \mathbb{Q}(\zeta) \hookrightarrow \mathbb{C}$ we also have that $|\tau(\alpha)| = 1$, so $|\tau(\alpha - 1)| = |\tau(\alpha) - 1| \leq |\tau(\alpha)| + 1 = 2$. We obtain the following estimate for the norm of $\alpha - 1$:

$$|N(\alpha - 1)| = |\prod_{\sigma \in G} \sigma(\alpha - 1)|$$

$$\leq \left(\frac{3^3}{2^2}\frac{1}{|x|(p-1)}\right)^2 2^{p-3}. \tag{6.10}$$

Note that the number $\alpha - 1$ does not equal 0, because $\alpha = 1$ would imply $\bar{\theta} = 0$, which is not true, because we defined $\theta = \bar{\theta}_1 - \bar{\theta}_2$, with $\bar{\theta}_1$ and $\bar{\theta}_2$ distinct elements of I_S^-.

Consider the principal ideal (α). Let J be the denominator of (α) and let J' be the numerator. Since $|N(\alpha)| = \prod_{\sigma\in G} |\sigma(\alpha)| = 1$, we have $N(J) = N(J')$. Since $(\alpha)^q = (x-\zeta)^{\bar{\theta}}$, the ideal J^q is the denominator of $(x-\zeta)^{\bar{\theta}}$ and the ideal J'^q is de numerator of the same ideal. In fact, $J^q = \prod_{n_\sigma \leq 0}(x-\zeta^\sigma)^{n_\sigma}$ and $J'^q = \prod_{n_\sigma>0}(x-\zeta^\sigma)^{|n_\sigma|}$. We obtain that the ideal $(JJ')^q$ equals $\prod_{\tau\in G}(x-\zeta^\tau)^{|n_\tau|}$. This yields:

$$N((JJ')^q) = N(J)^{2q} \leq N(\prod_{\tau\in G}(x-\zeta^\tau)^{|n_\tau|}). \tag{6.11}$$

We estimate this latter norm. Computing the norm of $x - \zeta^\tau$ we find

$$\begin{aligned} N(x-\zeta^\tau) &= \prod_{\sigma\in G}\sigma(x-\zeta^\tau) \\ &= \prod_{\sigma\in G}(x-\sigma(\zeta)) \\ &= x^{p-1} + x^{p-2} + \ldots + x + 1 \\ &\leq |x|^{p-1} + \binom{p}{1}|x|^{p-2} + \binom{p}{2}|x|^{p-3} + \ldots + \binom{p}{p-1}|x| + 1 \\ &= (|x|+1)^{p-1}. \end{aligned} \tag{6.12}$$

Since $\|\bar{\theta}\| = \|\bar{\theta}_1 - \bar{\theta}_2\| \leq \|\bar{\theta}_1\| + \|\bar{\theta}_2\| \leq 3\frac{q}{p-1}$, we have

$$N(\prod_{\tau\in G}(x-\zeta^\tau)^{|n_\tau|}) \leq (|x|+1)^{\|\bar{\theta}\|(p-1)} \leq (|x|+1)^{3q}.$$

It follows that $N(J)$ is at most $(|x|+1)^{\frac{3}{2}}$.

We defined J as being the denominator of α. That is equivalent to saying that $J\cdot(\alpha)$ is an $\mathcal{O}_K$-ideal. If $J\cdot(\alpha)$ is an $\mathcal{O}_K$-ideal, then $J\cdot(\alpha-1)$ is an $\mathcal{O}_K$-ideal as well. Therefore, the denominator of (α) equals the denominator of $(\alpha-1)$.

All this leads to the following inequality:

$$(|x|+1)^{-\frac{3}{2}} \leq N(J)^{-1} \leq |N(\alpha-1)| \leq \left(\frac{3^3}{2^2}\frac{1}{|x|(p-1)}\right)^2 2^{p-3}.$$

This yields:

$$\sqrt{|x|+1} \leq \frac{(|x|+1)^2}{|x|^2}\left(\frac{3^3}{2^2}\frac{1}{p-1}\right)^2 2^{p-3} \leq \left(\frac{3^3}{2^2}\frac{1}{p-1}\right)^2 2^{p-2} \leq \frac{46}{(p-1)^2}2^{p-2}. \tag{6.13}$$

But, using corollary 5.4, we also have:

$$\sqrt{|x|+1} \geq \sqrt{q^{p-1}} = \sqrt{q}^{p-1} \geq \sqrt{5}^{p-1}.$$

However, $\sqrt{5}^{p-1} > \frac{46}{(p-1)^2}2^{p-2}$ for all primes $p \geq 5$. This is impossible, so we have proven the theorem. □

Before proving that in the case of this chapter the Catalan equation has no solutions, we prove the following lemma.

Lemma 6.3. *Let $k \geq 2$ and $s \geq 4$ be integers. If the inequality*

$$\binom{s+k}{s} > \frac{4}{3}(s+1)k^2 + 1$$

holds, then it also holds for all pairs of integers s' and k' for which $s' \geq s$ and $k' \geq k$.

Proof. This lemma can be proven by induction. It suffices to take care of the steps $s \mapsto s+1$ and $k \mapsto k+1$.

First, we explain the step $s \mapsto s+1$. We want to show that $\binom{s+1+k}{s+1} > \frac{4}{3}(s+2)k^2+1$, given that $\binom{s+k}{s} > \frac{4}{3}(s+1)k^2+1$. Let us estimate $\binom{s+1+k}{s+1}$:

$$\begin{aligned}
\binom{s+1+k}{s+1} &= \frac{s+1+k}{s+1}\binom{s+k}{s} \\
&> \frac{s+k+1}{s+1}(\frac{4}{3}(s+1)k^2+1) \\
&> \frac{s+k+1}{s+1}\frac{4}{3}(s+1)k^2+1 \\
&= (s+1+k)\frac{4}{3}k^2+1. \qquad (6.14)
\end{aligned}$$

Therefore, it suffices to show that $s+k+1 > s+2$, which is obviously true, since $k \geq 2$.

Next, we explain the step $k \mapsto k+1$. We want to show that $\binom{s+k+1}{s} > \frac{4}{3}(s+1)(k+1)^2+1$, given that $\binom{s+k}{s} > \frac{4}{3}(s+1)k^2$. We obtain the following inequalities:

$$\begin{aligned}
\binom{s+k+1}{s} &= \frac{s+k+1}{k+1}\binom{s+k}{s} \\
&> \frac{s+k+1}{k+1}(\frac{4}{3}(s+1)k^2+1) \\
&> \frac{s+k+1}{k+1}\frac{4}{3}(s+1)k^2+1. \qquad (6.15)
\end{aligned}$$

So it suffices to show that $\frac{s+k+1}{k+1}k^2 > (k+1)^2$, which is equivalent to $(s+k+1)k^2 > (k+1)^3$. Therefore, since $s \geq 4$, it suffices to show that $(k+5)k^2 > (k+1)^3$. Because we have $k \geq 2$, this is indeed true. □

The corollary we now derive from theorem 6.1 is the main result of this chapter.

Corollary 6.4. *For any pair of primes p, q with p and q both greater than or equal to 5 such that p divides $q-1$, the equation $x^p - y^q = 1$ has no solution in non-zero integers x and y.*

Proof. By assumption, we have $q \equiv 1 \pmod{p}$, therefore there exists an integer k such that $q = 1 + kp$. Also, we saw in corollary 5.8 that $q^{p-1} \equiv 1 \pmod{p^2}$. If we

compute q^{p-1}, we find:

$$\begin{aligned} q^{p-1} &= (1+kp)^{p-1} \\ &= 1+(p-1)kp+\binom{p-1}{2}(kp)^2+\ldots+(p-1)(kp)^{p-2}+(kp)^{p-1} \\ &\equiv 1+(p-1)kp \,(\mathrm{mod}\, p^2). \end{aligned} \tag{6.16}$$

It follows that $1+(p-1)kp \equiv 1 \,(\mathrm{mod}\, p^2)$. Since p does not divide $p-1$, we have that p divides k. So $q = 1+kp \equiv 1 \,(\mathrm{mod}\, p^2)$.

We conclude that q is of the form $q = 1+lp^2$, for some positive integer l. Of course, q cannot be equal to $1+p^2$ or $1+3p^2$, because these numbers are even. It cannot be equal to $1+2p^2$ either, because this number is divisible by 3. It follows that $q \geq 1+4p^2 > 4(p-1)^2$.

The number $s = \left\lfloor \frac{3}{2}\frac{q}{(p-1)^2} \right\rfloor$ in the statement of theorem 6.1 thus satisfies $s \geq 6$. The inequality of lemma 6.3 is satisfied for the pairs $(s,k) = (6,4)$, $(7,3)$ and $(9,2)$. Therefore, according to the lemma, it is satisfied for all pairs (s,k) with $s \geq 4$ and $k \geq 4$, except for the pair $(s,k) = (6,3)$ and the pairs (s,k) with $k = 2$ and $s \leq 8$.

Now put $k = \frac{p-1}{2}$ and apply lemma 6.3. It follows that the inequality in the statement of theorem 6.1 is satisfied for all primes $p \geq 5$, except this with $\frac{p-1}{2} = 3$ and $s = 6$ and those with $\frac{p-1}{2} = 2$ and $s \leq 8$. However, these primes do not correspond to pairs of primes $p, q \geq 5$ such that $q \equiv 1 \,(\mathrm{mod}\, p^2)$. We can see this in the following way. Suppose that $\frac{p-1}{2} = 3$ and $s = 6$. Then $p = 3$ and $\frac{3}{2}\frac{q}{(p-1)^2} < 7$, which implies that $q < 168$. But we already saw that $q \geq 1+4p^2 = 197$, so this is impossible. Now consider the other case, in which $\frac{p-1}{2} = 2$ and $s \leq 8$. Then we have $p = 5$ and from $s \leq 8$ it follows that $q < 96$. But again, $q > 1+4p^2 = 101$, so this case also is impossible. Therefore, the condition of theorem 6.1 is satisfied for all primes p and q that are greater than or equal to 5 such that $q \equiv 1 \,(\mathrm{mod}\, p)$. $\square$

Chapter 7

A Runge-type theorem

From now on, we assume that p and q are primes such that $p > q \geq 7$ and p does not divide $q-1$, and that x and y are non-zero integers such that $x^p - y^q = 1$. Define $G = \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q}) \cong (\mathbb{Z}/p\mathbb{Z})^*$, as we did in chapter 4.

In this chapter we will prove a theorem of Mihăilescu, using Runge's method. We follow the treatments of Bilu in [1] and Schoof. The theorem we prove is the following:

Theorem 7.1. *Suppose that $\theta \in \mathbb{Z}[G]$. If $1+\iota$ divides θ, q divides $w(\theta)$ and θ has the property that $(x-\zeta)^\theta = \alpha^q$ for some $\alpha \in K^* = \mathbb{Q}(\zeta)^*$, then $\theta \in q\mathbb{Z}[G]$.*

Let θ be $\sum_{\sigma\in G} n_\sigma \sigma$. We want to show that the coefficients n_σ are divisible by q. Of course, this is the same as showing that for all $\sigma \in G$ the integer $n_\sigma + qk_\sigma$ is divisible by q for an integer k_σ. So without loss of generality we may assume that n_σ lies between 0 and $q-1$ for all $\sigma \in G$. Note that the coefficients of $q\sum_{\sigma\in G}\sigma - \theta$ then lie between 1 and q.

The weight of $\theta + q\sum_{\sigma\in G}\sigma - \theta$ equals $q(p-1)$. Therefore, one of θ and $q\sum_{\sigma\in G}\sigma - \theta$ has weight mq, with m an integer such that $0 \leq m \leq \frac{p-1}{2}$. So by using $q\sum_{\sigma\in G}\sigma - \theta$ instead of θ if necessary, we may assume that θ has weight at most $q\frac{p-1}{2}$.

By assumption, there exists $\alpha \in \mathbb{Q}(\zeta)^*$ such that $(x-\zeta)^\theta = \alpha^q$.

Since the weight of θ equals mq, we have

$$\left(1-\frac{\zeta}{x}\right)^\theta = \left(\frac{1}{x}\right)^\theta (x-\zeta)^\theta = \left(\frac{1}{x}\right)^{w(\theta)} (x-\zeta)^\theta = \left(\frac{\alpha}{x^m}\right)^q. \tag{7.1}$$

We will analyse this element α taking a q-th root in $\mathbb{Q}(\zeta)^*$ using a power series. Runge's method consists of showing that a certain partial sum of this power series is so close to the root we have, that they are actually equal. We show that the difference of this partial sum and the root is an algebraic integer that has norm smaller than 1, so it has to be 0.

The power series we use is the following:

$$F(T) = (1-\zeta T)^{\frac{\theta}{q}} = \prod_{\tau \in G} (1-\zeta^\tau T)^{\frac{n_\tau}{q}},$$

where we define $(1-\zeta^\tau T)^{\frac{n_\tau}{q}}$ as $\sum_{k\geq 0} \binom{\frac{n_\tau}{q}}{k} (-\zeta^\tau T)^k \in \mathbb{Q}(\zeta+\zeta^{-1})[[T]]$. Viewed as a series in $\mathbb{R}$, it has radius of convergence 1. Since $1+\iota$ divides θ, this series has coefficients in $\mathbb{Q}(\zeta+\zeta^{-1})$. We write $F(T) = \sum_{k\geq 0} \alpha_k T^k$, with $\alpha_k \in \mathbb{Q}(\zeta+\zeta^{-1})$. As we see in equation (7.1), we use this series for $T = \frac{1}{x}$, which is small because $|x|$ is large, as we saw in corollary 5.4, for instance.

Let $F_l(T)$ denote the l-th partial sum of $F(T)$, i.e. $F_l(T) = \sum_{k=0}^{l} \alpha_k T^k$. Let $F^\tau(T)$ denote the power series obtained by applying τ to the coefficients of $F(t)$.

Lemma 7.2. *This power series $F(T)$ has the following properties.*

1. *The power series $F(T)$ has the form*

$$F(T) = \sum_{k\geq 0} \frac{a_k}{k! q^k} T^k$$

with $a_k \in \mathbb{Z}[\zeta]$ and $a_k \equiv (-\sum_{\tau\in G} n_\tau \zeta^\tau)^k \pmod q$.

2. *Let σ be an embedding $\sigma : \mathbb{Q}(\zeta+\zeta^{-1}) \hookrightarrow \mathbb{R}$. Under this embedding, $F(T)$ can be viewed as a power series in $\mathbb{R}[[T]]$. Then for any $t \in \mathbb{R}$ with $|t| < 1$ we have that*

$$|F(t) - F_l(t)| \leq \binom{m+l}{m+1} \frac{|t|^{l+1}}{(1-|t|)^{m+l}}.$$

If in addition $t \in \mathbb{Q}$ satisfies $|t| < 1$ and $F(t) = (1-\zeta t)^{\frac{\theta}{q}} \in K = \mathbb{Q}(\zeta)$, then for all $\tau \in G$ we have that $\tau(F(t)) = F^\tau(t)$.

Proof.

1. We have the following more general fact. Let R be a domain of characteristic 0 and let $I \subset R$ be an ideal. Suppose we have two power series $\sum_{k\geq 0} \frac{a_k}{k!} T^k$ and $\sum_{k\geq 0} \frac{b_k}{k!} T^k$ such that $a_k \equiv a^k \pmod I$ and $b_k \equiv b^k \pmod I$ for all $k \geq 0$. Then the product of these two power series is equal to $\sum_{k\geq 0} \frac{c_k}{k!} T^k$, with $c_k = \sum_{j=0}^{k} \binom{k}{j} a_j b_{k-j} \equiv \sum_{j=0}^{k} \binom{k}{j} a^j b^{k-j} \pmod q \equiv (a+k)^k \pmod q$. We apply this result with $R = \mathbb{Z}[\zeta]$ and the power series

$$\begin{aligned}(1-\zeta^\tau qT)^{n_\tau/q} &= \sum_{k\geq 0} \left(\frac{1}{q}\right)^k \frac{n_\tau(n_\tau-q)(n_\tau-2q)\dots(n_\tau-(k-1)q)}{k!} (\zeta^\tau qT)^k \\ &= \sum_{k\geq 0} \frac{n_\tau(n_\tau-q)(n_\tau-2q)\dots(n_\tau-(k-1)q)}{k!} (-\zeta^\tau T)^k \quad (7.2)\end{aligned}$$

for all $\tau \in G$.

The coefficients of this power series are indeed of the form $\frac{a_k}{k!}$ with $a_k = n_\tau(n_\tau-q)(n_\tau-2q)\dots(n_\tau-(k-1)q)(-\zeta^\tau)^k \in \mathbb{Z}[\zeta]$. We see immediately that $a_k \equiv (-n_\tau\zeta^\tau)^k \pmod q$. Therefore, $F(qT) = \sum_{k\geq 0} \frac{c_k}{k!} T^k$ with $c_k \equiv (-\sum_{\tau\in G} n_\tau \zeta^\tau)^k \pmod q$. This is what we wanted to show.

2. Compare the series $\sum_{k\geq 0}\binom{\frac{n_\tau}{q}}{k}(-\zeta^\tau T)^k$ with the series $\sum_{k\geq 0}\binom{\frac{-n_\tau}{q}}{k}(-T)^k$. For our embedding $\sigma:\overline{\mathbb{Q}(\zeta)}\hookrightarrow\mathbb{R}$ the absolute values of the coefficients of the first series are smaller than the coefficients of the second series, because the following inequality holds for all positive rational numbers a and all integers k.

$$\begin{aligned}\left|\binom{a}{k}\right| &= \left|\frac{a(a-1)(a-2)\dots(a-(k-1))}{k!}\right| \\ &\leq (-1)^k\frac{-a(-a-1)(-a-2)\dots(-a-(k-1))}{k!} \\ &= (-1)^k\binom{-a}{k}. \end{aligned} \tag{7.3}$$

It follows that

$$\left|\sigma\left(\binom{\frac{n_\tau}{q}}{k}(-\zeta^\tau)^k\right)\right| = \left|\binom{\frac{n_\tau}{q}}{k}(-\sigma(\zeta^\tau))^k\right| = \left|\binom{\frac{n_\tau}{q}}{k}\right| \leq (-1)^k\binom{\frac{-n_\tau}{q}}{k}.$$

Note that we use here that the $n_\tau\geq 0$ for all $\tau\in G$.

So the absolute values of the coefficients of $F(T)$ are smaller than the coefficients of $\prod_{\tau\in G}(1-T)^{-n_\tau/q}=(1-T)^{-\frac{1}{q}\sum_{\tau\in G}n_\tau}=(1-T)^{-m}$. So for $t\in\mathbb{R}$ with $|t|<1$ we have:

$$|F(t)-F_l(t)|\leq|(1-t)^{-m}-s_l(t)|,$$

where $s_l(t)$ denotes the sum of the terms of degree at most l of the Taylor series expansion of $(1-t)^{-m}$.

Now from the standard estimate of the remainder term of a Taylor series we obtain:

$$(1-t)^{-m}-s_l(t)=\left.\frac{d^{l+1}(1-z)^{-m}}{dz^{l+1}}\right|_{z=\xi}\frac{t^{l+1}}{(l+1)!},$$

for some $\xi\in\mathbb{R}$ with $|\xi|<|t|$.

So,

$$\begin{aligned}(1-t)^{-m}-s_l(t) &= \frac{t^{l+1}}{(l+1)!}\frac{(m+l)!}{(m-1)!}(1-|\xi|)^{-m-l} \\ &= t^{l+1}\binom{m+l}{l+1}\frac{1}{(1-|\xi|)^{m+l}} \\ &\leq \frac{|t|^{l+1}}{(1-|t|)^{l+m}}\binom{m+l}{l+1}. \end{aligned} \tag{7.4}$$

This is what we wanted to prove.

Since $1+\iota$ divides θ, we know that $(1-\zeta t)^{\frac{\theta}{q}}=F(t)\in\mathbb{Q}(\zeta+\zeta^{-1})$ for $t\in\mathbb{Q}$. Under the embedding σ, view $F(T)$ as a series in $\mathbb{R}[[T]]$ again. For all embeddings $\tau:\mathbb{Q}(\zeta+\zeta^{-1})\hookrightarrow\mathbb{R}$, we find $\tau(F(t))=\tau((1-\zeta t)^{\frac{\theta}{q}})=(1-\tau(\zeta)t)^{\frac{\theta}{q}}\in\mathbb{R}$.

On the other hand, $F^\tau(t)=\prod_{\nu\in G}(1-\zeta^{\tau\nu}t)^{\frac{n_\nu}{q}}=(1-\zeta t)^{\tau\theta}$, which lies in $\mathbb{R}$ under the embedding σ, because if $1+\iota$ divides θ, then $1+\iota$ divides $\tau\theta$, too.

Now $\tau(F(t))$ and $F^\tau(t)$ are both a q-th root of the image of $(1-\zeta^\tau t)^\theta$ in $\mathbb{R}$ under the embedding σ. Since $\mathbb{R}$ does not contain primitive q-th roots of unity, this root is unique and therefore, $\tau(F(t))=F^\tau(t)$.

□

Consider equation 7.1 again, which says that

$$\left(1-\frac{\zeta}{x}\right)^{\theta} = \left(\frac{\alpha}{x^m}\right)^q.$$

Since $1+\iota$ divides θ, we know that $\left(1-\frac{\zeta}{x}\right)^{\theta} \in \mathbb{Q}(\zeta+\zeta^{-1})$ and is totally positive, i.e. under all embeddings $\sigma : \mathbb{Q}(\zeta+\zeta^{-1}) \hookrightarrow \mathbb{R}$ we have $\sigma\left(\left(1-\frac{\zeta}{x}\right)^{\theta}\right) > 0$.

Note that since $(x-\zeta)^{\theta}$ is an algebraic integer, α is an algebraic integer, too. Of course, since $(x-\zeta)^{\theta} \in \mathbb{Q}(\zeta+\zeta^{-1})$, α is contained in $\mathbb{Z}[\zeta+\zeta^{-1}]$. Because $\left(\frac{\alpha}{x^m}\right)^q = \left(1-\frac{\zeta}{x}\right)^{\theta}$ is totally positive, so is $\frac{\alpha}{x^m}$, because q is odd.

Now consider $x^m F(\frac{1}{x})$. We know that for $t \in \mathbb{R}$ with $|t|<1$, $F(t)$ converges under all embeddings of $\mathbb{Q}(\zeta)$ in $\mathbb{R}$. Because $|x| \geq q^{p-1}$, obviously $\frac{1}{x} < 1$. So $\sigma(F(\frac{1}{x}))$ converges in $\mathbb{R}$ for all embeddings σ. So $\sigma(x^m F(\frac{1}{x}))$ converges to a q-th root of $\sigma((1-\frac{\zeta}{x})^{\theta})$ in $\mathbb{R}$. But $\mathbb{R}$ does not have any q-th roots of unity other than 1, so under all embeddings σ we have $\sigma(x^m F(\frac{1}{x})) = \sigma(\alpha)$.

Remember that $F_l(T)$ denotes the l-th partial sum of the series $F(T)$. Define $\gamma = q^{m+\mathrm{ord}_q(m!)} x^m F_m(\frac{1}{x})$. Lemma 7.2.1 implies $\gamma \in \mathbb{Z}[\zeta+\zeta^{-1}]$.

The most important step in our Runge-type argument consists of showing that for every embedding $\sigma : \mathbb{Q}(\zeta+\zeta^{-1}) \hookrightarrow \mathbb{R}$ we have

$$|\sigma(q^{m+\mathrm{ord}_q(m!)}\alpha - \gamma)| < 1.$$

From the second statement in lemma 7.2.2 we know

$$\begin{aligned} |\sigma(q^{m+\mathrm{ord}_q(m!)}\alpha-\gamma)| &= |q^{m+\mathrm{ord}_q(m!)}x^m\sigma(\alpha)-\sigma(\gamma)| \\ &= q^{m+\mathrm{ord}_q(m!)}|x^m\sigma(F(\frac{1}{x}))-x^m\sigma(F_m(\frac{1}{x}))| \\ &= q^{m+\mathrm{ord}_q(m!)}|x^m F^{\sigma}(\frac{1}{x})-F_m^{\sigma}(\frac{1}{x})|. \end{aligned} \tag{7.5}$$

From the first statement in lemma 7.2.2 it follows that

$$\begin{aligned} |x^m F^{\sigma}(\frac{1}{x})-F_m^{\sigma}(\frac{1}{x})| &\leq \binom{2m}{m+1}\frac{|x|^m\frac{1}{|x|^{m+1}}}{(1-\frac{1}{|x|})^{2m}} \\ &= \binom{2m}{m+1}\frac{1}{|x|}\frac{1}{(1-|x|)^{2m}}. \end{aligned} \tag{7.6}$$

Putting inequalities (7.5) and (7.6) together, we find

$$\begin{aligned} |\sigma(q^{m+\mathrm{ord}_q(m!)}\alpha-\gamma)| &\leq q^{m+\mathrm{ord}_q(m!)}\binom{2m}{m+1}\frac{1}{|x|}\frac{1}{(1-|x|)^{2m}} \\ &= q^{m+\lfloor\frac{m}{q}\rfloor+\lfloor\frac{m}{q^2}\rfloor+\cdots}\binom{2m}{m+1}\frac{1}{|x|}\frac{1}{(1-|x|)^{2m}} \\ &\leq q^{m+\frac{m}{q-1}}\binom{2m}{m+1}\frac{1}{|x|}\frac{1}{(1-|x|)^{2m}} \\ &\leq q^{m+\frac{m}{q-1}}q^{m\frac{\log 4}{\log q}}\frac{1}{|x|}\frac{1}{(1-|x|)^{2m}}. \end{aligned} \tag{7.7}$$

We get this last inequality from the fact that $\binom{2m}{m+1} < (1+1)^{2m} = 4^m = q^{m\frac{\log 4}{\log q}}$. Now we estimate this last expression $q^{m+\frac{m}{q-1}} q^{m\frac{\log 4}{\log q}} \frac{1}{|x|} \frac{1}{(1-|x|)^{2m}}$ using $0 \leq m \leq \frac{p-1}{2}$, which implies $q > 2m$, and $|x| \geq q^{p-1}$, which we proved in corollary 5.4. From these bounds we obtain:

$$\begin{aligned} \frac{1}{(1-|x|)^{2m}} &\leq \frac{1}{(1-|x|)^{p-1}} \leq \frac{1}{(1-q^{p-1})^{p-1}} < \frac{1}{(1-q^{p-1})^{p-1}} \frac{(q^{p-1})p}{q^{p-1}-1} \\ &= \left(\frac{q^{p-1}}{q^{p-1}-1}\right)^p = \left(1-\frac{1}{q^{p-1}}\right)^{-p}. \end{aligned} \tag{7.8}$$

And, using again that $m \leq \frac{p-1}{2}$ and $|x| \geq q^{p-1}$, it follows from (7.7) and (7.8) that

$$|\sigma(q^{m+\mathrm{ord}_q(m!)}\alpha - \gamma)| \leq q^{\frac{p-1}{2}(-1+\frac{1}{q-1}+\frac{\log 4}{\log q})} \left(1-\frac{1}{q^{p-1}}\right)^{-p}. \tag{7.9}$$

Our aim is to show that this expression we found here is smaller than 1. To do this, we take the q-logarithm and we show that it is negative. Taking the q-logarithm, we obtain:

$$\frac{p-1}{2}\left(-1+\frac{1}{q-1}+\frac{\log 4}{\log q}\right) - p\frac{\log(1-\frac{1}{q^{p-1}})}{\log q}.$$

Now we use that q is at least 7. We find:

$$\frac{p-1}{2}\left(-1+\frac{1}{q-1}+\frac{\log 4}{\log q}\right) \leq \frac{p-1}{2}\left(-1+\frac{1}{6}+\frac{\log 4}{\log q}\right)$$

and

$$-\log(1-\frac{1}{q^{p-1}}) = \log(\frac{q^{p-1}-1}{q^{p-1}})^{-1} = \log(\frac{q^{p-1}}{q^{p-1}-1}) \leq \log(\frac{7^2}{7^2-1}) = \log(\frac{49}{48}) < \frac{1}{48}.$$

Therefore,

$$\begin{aligned} &\frac{p-1}{2}\left(-1+\frac{1}{q-1}+\frac{\log 4}{\log q}\right) - p\frac{\log(1-\frac{1}{q^{p-1}})}{\log q} \\ &\quad \leq \frac{p-1}{2}\left(-1+\frac{1}{6}\frac{\log 4}{\log q}\right) + \frac{p}{48\log 7} < 0 \end{aligned}$$

for all $p \geq 7$. This is what we wanted to show.

We conclude that for all embeddings $\sigma : \mathbb{Q}(\zeta+\zeta^{-1}) \hookrightarrow \mathbb{R}$, we have $|\sigma(q^{m+\mathrm{ord}_q(m!)}\alpha - \gamma)| < 1$. Remember that α and γ are algebraic integers, so $q^{m+\mathrm{ord}_q(m!)}\alpha - \gamma$ is an algebraic integer, so its norm is an integer. However, we showed that for all embeddings in $\mathbb{R}$ its absolute value is smaller than 1. It follows that $q^{m+\mathrm{ord}_q(m!)}\alpha - \gamma$ equals 0. We obtain:

$$q^{m+\mathrm{ord}_q(m!)}\alpha = \gamma = q^{m+\mathrm{ord}_q(m!)}x^m F_m(\frac{1}{x}) = \sum_{k=0}^{m} q^{m+\mathrm{ord}_q(m!)}\frac{a_k}{k!q^k}x^{m-k}.$$

We know that $q^{m+\mathrm{ord}_q(m!)}\alpha$ is an algebraic integer, and all terms of the series on the right-hand side are algebraic integers as well. If $m = 0$, then $\theta = 0$ and we are done. If $m \geq 1$, then $q^{m+\mathrm{ord}_q(m!)}\alpha$ is obviously divisible by q. For $k < m$, all terms $q^{m+\mathrm{ord}_q(m!)}\frac{a_k}{k!q^k}$ are divisible by q. It follows that the m-th term is also divisible by q. Note that the m-th term $q^{m+\mathrm{ord}_q(m!)}\frac{a_m}{m!q^m}$ has as many factors q as a_m has. Therefore, $a_m \equiv 0 \pmod q$.

On the other hand, from lemma 7.2.1 we know that

$$a_m \equiv \left(-\sum_{\tau \in G} n_\tau \zeta^\tau \right)^m \pmod{q}.$$

In lemma 5.6 however, we proved that the ring $\mathbb{Z}[\zeta]/(q)$ does not have any nilpotent elements. Therefore, it follows that q divides $\sum_{\tau \in G} n_\tau \zeta^\tau$. Since the ζ^τ are linearly independent over $\mathbb{Q}$, we find that q divides n_τ for all $\tau \in G$. We conclude that $\theta \in q\mathbb{Z}[G]$, which is what we wanted to show.

Actually, the result we use in chapter 8 does not concern the group ring $\mathbb{Z}[G]$, but it concerns $\mathbb{Z}[G^+]$. But the analogue of theorem 7.1 for the group ring $\mathbb{Z}[G^+]$ is a corollary of theorem 7.1, as we show now.

Corollary 7.3. *If $\theta \in \mathbb{Z}_{\geq 0}[G^+]$ such that $w(\theta) \equiv 0 \pmod{q}$ and if $((x-\zeta)(x-\zeta^{-1}))^\theta$ is a q-th power in $\mathbb{Q}(\zeta+\zeta^{-1})^*$, then $\theta \in q\mathbb{Z}[G^+]$.*

Proof. Suppose $\theta \in \mathbb{Z}_{\geq 0}[G^+]$ satisfies the conditions of the statement. Reducing the coefficients of θ modulo q, we find an element $\theta' = \sum_{a=1}^{\frac{p-1}{2}} n_{\sigma_a} \sigma_a$ of $\mathbb{F}_q[G^+]$. Here σ_a denotes the element of G^+ that maps $\zeta + \zeta^{-1}$ to $\zeta^a + \zeta^{-a}$.

Define $\tilde{\theta} \in \mathbb{F}_q[G]$ as follows: $\tilde{\theta} = \sum_{a=1}^{\frac{p-1}{2}} n_{\sigma_a}(\sigma_a + \sigma_{-a})$. Here σ_a denotes the element of G that maps ζ to ζ^a. Note that $\tilde{\theta}$ is divisible by $1+\iota$. Since $w(\theta) = \sum_{\sigma \in G^+} n_\sigma$ is divisible by q, the weight of $\tilde{\theta}$ is 0 in $\mathbb{F}_q$. Therefore, $\theta' \in (1+\iota)I_{\mathrm{aug}}$. We find that

$$\begin{aligned}
(x-\zeta)^{\tilde{\theta}} &= (x-\zeta)^{\sum_{a=1}^{\frac{p-1}{2}} n_{\sigma_a}(\sigma_a+\sigma_{-a})} \\
&= (x-\zeta)^{\sum_{a=1}^{\frac{p-1}{2}} n_{\sigma_a}\sigma_a} (x-\zeta)^{\sum_{a=1}^{\frac{p-1}{2}} n_{\sigma_a}\sigma_{-a}} \\
&= ((x-\zeta)(x-\zeta^{-1}))^{\sum_{a=1}^{\frac{p-1}{2}} n_{\sigma_a}\sigma_a} \\
&= ((x-\zeta)(x-\zeta^{-1}))^{\theta'}. \qquad (7.10)
\end{aligned}$$

It follows that $(x-\zeta)^{\tilde{\theta}}$ differs from $(x-\zeta)^\theta$ by a q-th power in $\mathbb{Q}(\zeta+\zeta^{-1})^*$. By assumption, $((x-\zeta)(x-\zeta^{-1}))^\theta$ is a q-th power in $\mathbb{Q}(\zeta+\zeta^{-1})^*$, so $(x-\zeta)^{\tilde{\theta}}$ is a q-th power in $\mathbb{Q}(\zeta+\zeta^{-1})^*$. Theorem 7.1 now tells us that $\theta' = 0$ in $\mathbb{F}_q[G]$, so all coefficients n_σ are divisible by q. Therefore, $\theta \in q\mathbb{Z}[G^+]$, which is what we wanted to show. □

Chapter 8

The second case: q does not divide $p-1$

In this chapter we put all ingredients from the previous chapters together. We assume there exists a solution in non-zero integers x and y of the equation $x^p - y^q = 1$, where p and q are distinct odd primes that are at least 7. Note that in section 5.3 we dealt with the cases in which p or q is smaller than 7. Without loss of generality, we assume $p > q$. Eventually, we derive a contradiction from all this.

We make a further assumption in this chapter, namely that q does not divide $p-1$. We are allowed to do this since the case in which q does divide $p-1$ has been dealt with in chapter 6.

So the theorem we prove in this chapter is the following.

Theorem 8.1. *Let p and q be distinct odd primes that are at least equal to 7, such that q does not divide $p-1$. Then the Catalan equation $x^p - y^q = 1$ has no solution in non-zero integers x and y.*

In this chapter, we will work mainly in the field $\mathbb{Q}(\zeta + \zeta^{-1})$.

8.1 An exact sequence

From the equation $x^p - y^q = 1$, it follows that

$$y^q = x^p - 1 = \prod_{i=0}^{p-1}(x - \zeta^i). \tag{8.1}$$

First, we define a very useful subgroup of $\mathbb{Q}(\zeta + \zeta^{-1})^*$.

Definition 8.1. *We define H as follows:*

$$H = \{\alpha \in \mathbb{Q}(\zeta + \zeta^{-1})^* : \forall \mathfrak{l} \neq \mathfrak{p} : \operatorname{ord}_{\mathfrak{l}} \alpha \equiv 0 \pmod{q}\},$$

where the $\mathfrak{l}$'s range over the prime ideals in $\mathcal{O}_{K^+}$.

It can be shown easily that H is a subgroup of $\mathbb{Q}(\zeta+\zeta^{-1})^*$. Since for all elements $\sigma \in G^+$ we have $\sigma(\mathfrak{p}) = \mathfrak{p}$, the subgroup H is closed under the Galois action and therefore, it is a $\mathbb{Z}[G^+]$-module. As we noted in section 4.2, a $\mathbb{Z}[G^+]$-module that is annihilated by q is an $\mathbb{F}_q[G^+]$-module as well. We find that the group $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ is an $\mathbb{F}_q[G^+]$-module. First, note that $\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \subset H$; since the elements of $\mathbb{Q}(\zeta+\zeta^{-1})^*$ are q-th powers itself, the ideals they generate will be q-th powers of fractional ideals.

We will show that H contains the element $(x-\zeta)(x-\zeta^{-1})$. From lemma 5.5, we know that in $\mathbb{Q}(\zeta)$, the element $\frac{x-\zeta}{1-\zeta}$ generates a q-th power of an $\mathcal{O}_K$-ideal. Therefore, in $\mathbb{Q}(\zeta+\zeta^{-1})$, the element $(x-\zeta)(x-\zeta^{-1})$ generates an ideal of the form $\mathfrak{p}^k\mathfrak{a}^q$, with $\mathfrak{a}$ an $\mathcal{O}_{K^+}$-ideal an k an integer. Therefore, $\mathrm{ord}_{\mathfrak{l}}((x-\zeta)(x-\zeta^{-1})) \equiv 0 \pmod q$ for all prime ideals $\mathfrak{l}$ distinct from $\mathfrak{p}$. So $(x-\zeta)(x-\zeta^{-1})$ is an element of H.

We define an element ξ of $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ as follows.

Definition 8.2. *The element ξ of $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ is the image of $(x-\zeta)(x-\zeta^{-1})$ under the projection map $H \longrightarrow H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$.*

We have to give some more definitions before we can state the main theorem of this section.

Definition 8.3. *Define the subgroup $\mathcal{E}$ of $\mathbb{Q}(\zeta+\zeta^{-1})^*$ as follows.*

$$\mathcal{E} = \{\alpha \in \mathbb{Q}(\zeta+\zeta^{-1})^* : \forall \mathfrak{l} \neq \mathfrak{p} : \mathrm{ord}_{\mathfrak{l}}\,\alpha = 0\},$$

where the $\mathfrak{l}$ ranges over the prime ideals of $\mathcal{O}_{K^+}$.

Just as for H, it is easy to see that $\mathcal{E}$ indeed is a subgroup of $\mathbb{Q}(\zeta+\zeta^{-1})^*$. Actually, it is isomorphic (as groups) to $\mathcal{E} = E^+ \times \langle\lambda\rangle$, where $\lambda = (1-\zeta)(1-\zeta^{-1})$. This is the case because if α is an element of $\mathcal{E}$, then for all prime ideals $\mathfrak{l} \neq \mathfrak{p}$ of $\mathcal{O}_{K^+}$ we have $\mathrm{ord}_{\mathfrak{l}}\,\alpha = 0$. Therefore, $(\alpha) = \mathfrak{p}^k$ for some integer k and it follows that $\alpha = u\lambda^k$, with u a unit of $\mathcal{O}_{K^+}$. So we find a group homomorphism $\mathcal{E} \longrightarrow E^+ \times \langle\lambda\rangle$ given by $u\lambda^k \mapsto (u, \lambda^k)$, that is bijective.

Definition 8.4. *$Cl_{K^+}[q]$ denotes the q-torsion part of the class group Cl_{K^+} of $K^+ = \mathbb{Q}(\zeta+\zeta^{-1})$, i.e.*

$$Cl_{K^+}[q] = \{[\mathfrak{a}] \in Cl_{K^+} : [\mathfrak{a}]^q = [1]\}.$$

From now on, we denote $\mathbb{F}_q[G^+]$ by R.

In this section we prove the following theorem.

Theorem 8.2. *Let the sequence*

$$1 \longrightarrow \mathcal{E}/\mathcal{E}^q \xrightarrow{\psi'} H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \xrightarrow{\psi} Cl_{K^+}[q] \longrightarrow 1 \tag{8.2}$$

of $\mathbb{F}_q[G^+]$-modules be defined as follows.

The map ψ' is induced by the embedding $\mathcal{E} \longrightarrow H$.

For α in $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$, we have $(\alpha) = \mathfrak{p}^k\mathfrak{a}^q$ for some fractional $\mathcal{O}_{K^+}$-ideal $\mathfrak{a}$ and an integer k. Now define the map ψ as follows: for each coset of $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$, choose a representative, α, say. We define $\psi(\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q}) = [\mathfrak{a}]Cl_{K^+}[q]$. Then this sequence is exact.

Proof. It is clear that all groups in the sequence are indeed $\mathbb{F}_q[G^+]$-modules.

Note that for all $\alpha \in H$ with $(\alpha) = \mathfrak{p}^k\mathfrak{a}^q$, we have that $\mathfrak{a}^q$ is a principal ideal. Therefore, the image of ψ is indeed contained in $Cl_{K^+}[q]$.

Since ψ' is induced by the embedding $\mathcal{E} \longrightarrow H$, it is a well-defined R-linear homomorphism which is injective. From the definition of ψ' it is obviously an R-linear homomorphism as well.

It is easy to see that ψ is a well-defined R-linear homomorphism too. For $[\mathfrak{a}]Cl_{K^+}[q] \in Cl_{K^+}[q]$, we know that $\mathfrak{a}^q$ is a principal ideal, with generator α, say. The image of $\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ under ψ now equals $[\mathfrak{a}]Cl_{K^+}[q]$. Therefore, ψ is surjective.

We need to show that the image of ψ' equals the kernel of ψ. Let us find out what the kernel of ψ looks like. The kernel of ψ consists of cosets $\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ such that $\psi(\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q}) = [\mathfrak{a}]Cl_{K^+}[q] = [1]Cl_{K^+}[q]$. This means that the ideal (α) equals $\mathfrak{p}^k(\beta)^q$, for some $\beta \in \mathbb{Q}(\zeta+\zeta^{-1})^*$. It follows that $\alpha = \lambda^k\beta^q$, so $\alpha H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ is the image under ψ' of the element $\lambda^k\mathcal{E}/\mathcal{E}^q$. It is obvious that every element in the image of ψ' maps to $[1]Cl_{K^+}[q]$ under ψ.

It follows that the sequence we defined above is indeed an exact sequence of $\mathbb{F}_q[G^+]$-modules. □

8.2 The module $\mathcal{E}/\mathcal{E}^q$ is isomorphic to $\mathbb{F}_q[G^+]$ as an $\mathbb{F}_q[G^+]$-module

In this section we show that $\mathcal{E}/\mathcal{E}^q$ is R-isomorphic to $R = \mathbb{F}_q[G^+]$ itself.

Theorem 8.3. *The $\mathbb{F}_q[G^+]$-module $\mathcal{E}/\mathcal{E}^q$ is free of rank 1 over $\mathbb{F}_q[G^+]$.*

First, we prove the following lemma which will help us in the proof of this theorem.

Lemma 8.4. *Suppose $L \subset V = \mathbb{Z}[G^+] \otimes_{\mathbb{Z}} \mathbb{R}$ is both a lattice of full rank and a $\mathbb{Z}[G^+]$-submodule. Then $L/qL \cong_{\mathbb{Z}[G^+]} \mathbb{F}_q[G^+]$ for all primes q that do not divide $\#G^+$.*

Proof. Note that $\mathbb{Z}[G^+] \otimes_{\mathbb{Z}} \mathbb{R}$ is isomorphic to $\prod_{\sigma\in G^+} \mathbb{R}$. As L contains an $\mathbb{R}$-basis for V and $\mathbb{Q}$ is dense in $\mathbb{R}$, we know that $\mathbb{Q} \cdot L$ is dense in V.

Define $e = (1, 0, \ldots, 0) \in V$. Note that the $\sigma(e)$ for $\sigma \in G^+$ are linearly independent over $\mathbb{R}$, since $\sigma(e) = (0, \ldots, 0, 1, 0, \ldots, 0)$ with the 1 on place σ^{-1}. Of course, it could be the case that $e \notin \mathbb{Q} \cdot L$. But we can find $e' \in V$ such that e' is as close to e as we want. In particular, we can choose e' such that the $\sigma(e')$ are linearly independent over $\mathbb{R}$ as well. Write $e' = (\varepsilon_\sigma)_{\sigma\in G^+}$.

Now choose $N \in \mathbb{Z}$ such that $Ne' \in L$. Such an N obviously exists, because $e' \in \mathbb{Q} \cdot L$. Define $L' = \mathbb{Z}[G^+] \cdot Ne'$. Because L is a $\mathbb{Z}[G^+]$-module, $L' \subset L$. Then $L' \cong_{\mathbb{Z}[G^+]} \mathbb{Z}[G^+]$, by the map $\theta Ne' \mapsto \theta$. In other words: L' is free of rank 1 as a $\mathbb{Z}[G^+]$-module.

We can see easily that $L/L' \cong_{\mathbb{Z}[G]} qL/qL'$. The map φ defined by

$$\varphi : L/L' \longrightarrow qL/qL'$$

$$l + L' \quad \longmapsto \quad ql + qL' \tag{8.3}$$

is an $\mathbb{Z}[G^+]$-linear isomorphism, since multiplication by q is $\mathbb{Z}[G^+]$-linear and it induces this isomorphism.

Our claim is now that $L'/qL' \cong_{\mathbb{Z}[G^+]} \mathbb{F}_q[G^+]$. We already know that $L' \cong_{\mathbb{Z}[G^+]} \mathbb{Z}[G^+]$. Therefore, $L'/qL' \cong_{\mathbb{Z}[G^+]} \mathbb{Z}[G^+]/q\mathbb{Z}[G^+] \cong_{\mathbb{Z}[G^+]} \mathbb{F}_q[G^+]$.

There exists a finite simple filtration of L/qL'. In lemma 4.4 we showed that $\mathbb{F}_q[G^+]$ is a finite product of finite fields, so L/qL' is isomorphic to a direct sum of vector spaces. Therefore, if we take a submodule that is not equal to 0 or to the module itself, then the dimension of this submodule is at least 1 smaller than the dimension of the module itself. It follows that each sequence $L/qL' \supset L_1 \supset L_2 \supset L_3 \supset \ldots$ of distinct submodules is finite.

We can construct a simple filtration in two ways, namely via L'/qL' or via qL/qL', let us say:

$$L/qL' \supset M_1 \supset \ldots \supset M_{i-1} \supset L'/qL' \supset \ldots \supset \{0\}$$

and

$$L/qL' \supset N_1 \supset \ldots \supset N_{j-1} \supset qL/qL' \supset \ldots \supset \{0\}.$$

Of course, L/qL' is Jordan-Hölder equivalent to itself, so there exists a permutation π of the indices such that $M_i/M_{i+1} \cong N_{\pi(i)}/N_{\pi(i)+1}$. Therefore, $(L'/qL')/\{0\} \cong L'/qL'$ and $(L'/qL')/(qL/qL') \cong L/qL$ are Jordan-Hölder equivalent as modules over $\mathbb{Z}[G^+]$ and over $\mathbb{F}_q[G^+]$.

We already saw in lemma 4.4 that if q does not divide the order $p-1$, then $\mathbb{F}_q[G^+]$ is isomorphic to a finite product of finite fields, so an $\mathbb{F}_q[G^+]$-module M is a finite product of vector spaces. Therefore, if M' and M are $\mathbb{F}_q[G^+]$-modules such that $M' \subset M$, then $M \cong_{\mathbb{F}_q[G^+]} M' \oplus (M/M')$. So every module is isomorphic to the direct sum of its Jordan-Hölder factors.

It follows that L'/qL' and L/qL are isomorphic to the same direct sum of the same quotients M_i/M_{i+1} and, therefore, they are $\mathbb{F}_q[G^+]$-isomorphic. □

Proof of theorem 8.3. We already noted that $\mathcal{E} \cong E^+ \times \langle\lambda\rangle$. The Dirichlet unit theorem tells us that $E^+ = \mathcal{O}_{K^+}^* \cong \mu_{K^+} \times \eta_1 \times \eta_2 \times \ldots \times \eta_{n-1}$, with μ_{K^+} the subgroup of $K^+ = \mathbb{Q}(\zeta + \zeta^{-1})$ consisting of roots of unity, the η_i elements of E^+ called *fundamental units* and $n = \frac{p-1}{2}$, the number of elements in G^+. Therefore,

$$E^+ \cong \{\pm 1\} \times \mathbb{Z}^{n-1}$$

and

$$\mathcal{E} \cong \{\pm 1\} \times \mathbb{Z}^n,$$

where the isomorphisms are isomorphisms of groups. Here we use that λ is not a unit.

It follows that

$$\mathcal{E}/\mathcal{E}^q \cong (\mathbb{Z}/q\mathbb{Z})^n.$$

The prime q does not divide $2n = p - 1$ by assumption.

Consider the following map:

$$\begin{aligned} L : E^+ = \mathcal{O}_{K^+}^* &\longrightarrow \prod_{\sigma \in G^+} \mathbb{R} \\ u &\longmapsto (\log|\sigma(u)|)_{\sigma \in G^+}. \end{aligned} \tag{8.4}$$

It is easy to see that L is a homomorphism. According to the Dirichlet unit theorem, the kernel $\ker L$ equals $\mu_{K^+} = \{\pm 1\}$ and the image $L(E^+)$ is a lattice of rank $n-1$ in $\prod_{\sigma \in G^+} \mathbb{R}$. In fact, $L(E^+)$ is contained in the hyperplane $\mathcal{H} = \{(x_\sigma)_{\sigma \in G^+} \in \prod_{\sigma \in G^+} \mathbb{R} : \sum_{\sigma \in G^+} x_\sigma = 0\}$, because $\sum_{\sigma \in G^+} \log|\sigma(u)| = \log|\prod_{\sigma \in G^+} \sigma(u)| = \log|N(u)| = 0$, since $|N(u)| = 1$ for all $u \in E^+ = \mathcal{O}_{K^+}^*$.

We have the map

$$\begin{aligned} L : \mathcal{E} &\longrightarrow \prod_{\sigma \in G^+} \mathbb{R} \\ x &\longmapsto (\log|\sigma(x)|)_{\sigma \in G^+}. \end{aligned} \tag{8.5}$$

This map L is a homomorphism as well. Of course, $L(E^+)$ is contained in $L(\mathcal{E})$. We have for all elements x of $\mathcal{E}$ that $x = u\lambda^k$ with $u \in E^+$ and $k \in \mathbb{Z}$. Note that $N(\lambda) = N((1-\zeta)(1-\zeta^{-1})) = \prod_{a=1}^{p}(1-\zeta^a) = p$, so $L(\mathcal{E}) \not\subset \mathcal{H}$. Therefore, $L(\mathcal{E})$ is a lattice of rank n in $\prod_{\sigma \in G^+} \mathbb{R}$.

It is clear that L is a $\mathbb{Z}[G^+]$-linear map. Therefore, $L(\mathcal{E})$ is a $\mathbb{Z}[G^+]$-submodule of $\prod_{\sigma \in G^+} \mathbb{R}$.

From lemma 8.4 it follows that $L(\mathcal{E})/qL(\mathcal{E}) \cong_{\mathbb{Z}[G^+]} \mathbb{F}_q[G^+]$.

The map $\mathcal{L}$ defined by

$$\begin{aligned} \mathcal{L} : \mathcal{E}/\mathcal{E}^q &\longrightarrow L(\mathcal{E})/qL(\mathcal{E}) \\ e + \mathcal{E}^q &\longmapsto L(e) + qL(\mathcal{E}). \end{aligned} \tag{8.6}$$

gives an isomorphism of $\mathbb{Z}[G^+]$-modules.

We conclude that $\mathcal{E}/\mathcal{E}^q \cong_{\mathbb{Z}[G^+]} L(\mathcal{E})/qL(\mathcal{E}) \cong_{\mathbb{Z}[G^+]} \mathbb{F}_q[G^+]$. Since $\mathcal{E}/\mathcal{E}^q$ is a $\mathbb{F}_q[G^+]$-module, it follows that $\mathcal{E}/\mathcal{E}^q \cong_{\mathbb{F}_q[G^+]} \mathbb{F}_q[G^+]$. □

8.3 All cyclotomic units belong to $\xi^{I_{\text{aug}}}$

In the previous section we saw that $\mathcal{E}/\mathcal{E}^q$ is R-isomorphic to $R = \mathbb{F}_q[G^+]$. In lemma 4.4 we saw what $\mathbb{F}_q[G^+]$ looks like: it is isomorphic to a finite product of finite fields.

Summarizing, we have the following situation:

$$R = \mathbb{F}_q[G^+] \cong \mathbb{F}_q[X]/(X^{\frac{p-1}{2}} - 1) \cong \prod_i^{<\infty} (\text{finite fields}).$$

Lemma 8.5. *The subset $C^+\mathcal{E}^q/\mathcal{E}^q \subset \mathcal{E}/\mathcal{E}^q$ is R-isomorphic to an R-ideal $\mathfrak{a}$.*

Proof. Since C^+ is closed under the Galois action, $C^+\mathcal{E}^q/\mathcal{E}^q$ is a sub-R-module of $\mathcal{E}/\mathcal{E}^q \cong R$. Therefore, it is R-isomorphic to an ideal of R. □

We are interested in the structure of the R-ideal $\mathfrak{a}$. The following lemma tells us more about it.

Lemma 8.6. *The R-ideal $\mathfrak{a}$ is a principal ideal (e), with $e^2 = e$.*

Proof. We saw that R is isomorphic to a finite product of finite fields, let us say there are n fields $F_1, \ldots, F_n$ in the product. Therefore, all ideals in R are isomorphic to an ideal of this finite product of finite fields. Of course, all ideals $I \subset \prod_{i=1}^n F_i$ are of the form $I = I_1 \times \ldots \times I_n$, with $I_j \subset F_j$ an ideal for all $j = 1, \ldots, n$. Since the only ideals of a field are the trivial ideals 0 and F_i itself, all ideals in R have a generator $e = (e_1, \ldots, e_n)$, where e_i equals 0 or 1. Of course, for such a generator we have that $e^2 = e$. □

This ideal $\mathfrak{a}$ has an other interesting property. Let I_{aug} be the augmentation ideal of the weight homomorphism on $R = \mathbb{F}_q[G^+]$.

Lemma 8.7. *The ideal $\mathfrak{a}$ is contained in the augmentation ideal I_{aug}.*

Proof. Remember that the augmentation ideal I_{aug} is defined as the kernel of the weight homomorphism. So we want to show that for all elements $a = \sum_{\sigma \in G^+} a_\sigma \sigma \in \mathfrak{a}$, the sum $\sum_{\sigma \in G^+} a_\sigma$ equals 0. From lemma 4.5 it is easy to see that for all cyclotomic units $c \in C^+$, we have that the norm $N(c) = c^{\sum_{\sigma \in G^+} \sigma}$ equals 1.

Let $a = \sum_{\sigma \in G^+} a_\sigma \sigma$ be an element of $\mathfrak{a}$ and let φ be the R-isomorphism between $C^+\mathcal{E}^q/\mathcal{E}^q$ and $\mathfrak{a}$. Define $c\mathcal{E}^q$ to be the inverse image of a under φ, i.e. $c\mathcal{E}^q = \varphi^{-1}(a)$. Then we have that $\varphi((c\mathcal{E}^q)^{\sum_{\sigma \in G^+} \sigma}) = \varphi(\mathcal{E}^q) = 0$ and, on the other hand, $\varphi((c\mathcal{E}^q)^{\sum_{\sigma \in G^+} \sigma}) = (\sum_{\sigma \in G^+} \sigma)\varphi(c\mathcal{E}^q) = (\sum_{\sigma \in G^+} \sigma) \cdot a$. It follows that $(\sum_\sigma \sigma)(\sum_{\sigma \in G^+} a_\sigma \sigma) = \sum_{\psi \in G^+} (\sum_{\tau\sigma = \psi} a_\sigma)\psi = 0$. Therefore, $\frac{p-1}{2} \sum_{\sigma \in G^+} a_\sigma = 0$, so $\sum_{\sigma \in G^+} a_\sigma = 0$, which is what we wanted to show. □

Lemma 8.8. *The ideal $\mathfrak{a}$ annihilates the q-torsion part $Cl_{K^+}[q]$ of the class group of $\mathbb{Q}(\zeta + \zeta^{-1})$.*

Proof. Since the isomorphism $\varphi : \mathcal{E}/\mathcal{E}^q \longrightarrow R$ also gives an isomorphism between $C^+\mathcal{E}^q/\mathcal{E}^q$ and $\mathfrak{a}$, the quotients $\mathcal{E}/C^+\mathcal{E}^q$ and $R/\mathfrak{a}$ are R-isomorphic as well. It follows that the R-ideal $\mathfrak{a}$ annihilates $\mathcal{E}/C^+\mathcal{E}^q$. Because E^+/C^+E^{+q} is contained in $\mathcal{E}/C^+\mathcal{E}^q$, the ideal $\mathfrak{a}$ also annihilates E^+/C^+E^{+q}. Of course, the generator e of $\mathfrak{a}$, that exists according to lemma 8.6, also annihilates E^+/C^+E^{+q}.

Choose a non-negative integer m such that the Sylow-q-subgroup $(E^+/C^+)_q$ of E^+/C^+ equals $E^+/(C^+E^{+q^m})$. Of course we have to show that such an m exists. We know that the group E^+/C^+ is finite and abelian, so it is isomorphic to the direct product of its Sylow-subgroups. In other words, we have

$$E^+/C^+ \cong S_{p_1} \times S_{p_2} \times \ldots \times S_{p_t},$$

where the p_i are the distinct prime divisors of $\#(E^+/C^+)$. Now consider $(E^+/C^+)^{q^m}$, where $q^m = \#(E^+/C^+)_q$. Then we find:

$$\cong S_{p_1}^{q^m} \times S_{p_2}^{q^m} \times \ldots \times S_{p_t}^{q^m}.$$

If p_i does not equal q, then $S_{p_i}^{q^m} \cong S_{p_i}$. If $p_i = q$, then $S_{p_i}^{q^m} = \{1\}$. It follows that $(E^+/C^+)_q \cong (E^+/C^+)/(E^+/C^+)^{q^m}$. Also, the group $(E^+/C^+)^{q^m}$ is isomorphic to

the group $E^{+q^m}/(C^+ \cap E^{+q^m})$ by the isomorphism $\varphi : u^{q^m}C^+ \mapsto u^{q^m}(C^+ \cap E^{+q^m})$. It follows that

$$\begin{aligned}(E^+/C^+)_q &\cong (E^+/C^+)/(E^+/C^+)^{q^m} \\ &\cong (E^+/C^+)/(E^{+q^m}/(C^+ \cap E^{+q^m})) \\ &\cong E^+/C^+E^{+q^m}. \end{aligned} \tag{8.7}$$

We have the following inclusions:

$$E^+ \supset C^+E^{+q} \supset C^+E^{+q^2} \supset C^+E^{+q^3} \supset \ldots \supset C^+E^{+q^m}.$$

Choose an $\varepsilon \in \mathbb{Z}[G^+]$ such that ε maps to e under the canonical map $\mathbb{Z}[G^+] \longrightarrow \mathbb{F}_q[G^+]$. We know that e annihilates E^+/C^+E^{+q}. It follows that ε also annihilates E^+/C^+E^{+q}.

The claim is now that ε also annihilates all quotients $C^+E^{+q}/C^+E^{+q^2}$, $C^+E^{+q^2}/C^+E^{+q^3}$, ..., $C^+E^{+q^{m-1}}/C^+E^{+q^m}$. We can see this as follows. For $k = 1, \ldots, m-1$, define the map

$$\begin{aligned}\varphi : C^+E^{+q^{k-1}}/C^+E^{+q^k} &\longrightarrow C^+E^{+q^k}/C^+E^{+q^{k+1}} \\ uC^+E^{+q^k} &\longmapsto u^qC^+E^{+q^{k+1}}. \end{aligned} \tag{8.8}$$

First, we show that φ is well-defined. Let u, u' be elements of E^+ such that $uC^+E^{+q^k} = u'C^+E^{+q^k}$, so $u = u'cv^{q^k}$ for some $c \in C^+$ and $v \in E^+$. Then

$$\begin{aligned}\varphi(uC^+E^{+q^k}) &= u^qC^+E^{+q^{k+1}} = (u'cv^{q^k})^qC^+E^{+q^{k+1}} \\ &= u'^qC^+E^{+q^{k+1}} = \varphi(u'C^+E^{+q^k}). \end{aligned} \tag{8.9}$$

It is obvious that φ is an R-linear group homomorphism. Also, we have that φ is surjective: the coset $u^{q^k}C^+E^{+q^{k+1}}$ has the coset $u^{q^{k-1}}C^+E^{+q^k}$ as an original.

It follows that $\varphi(uC^+E^{+q})^\varepsilon = \varphi((uC^+E^{+q})^\varepsilon) = \varphi(C^+E^{+q}) = C^+E^{+q^2}$, so since ε annihilates E^+/C^+E^{+q} and the map φ is surjective, it also annihilates $C^+E^{+q}/C^+E^{+q^2}$. Similarly, it follows that ε annihilates all quotients $C^+E^{+q}/C^+E^{+q^2}$, ..., $C^+E^{+q^{m-1}}/C^+E^{+q^m}$.

Now our claim is that ε^m annihilates $E^+/C^+E^{+q^m} = (E^+/C^+)_q$. Let $uC^+E^{+q^m}$ be an element of $E^+/C^+E^{+q^m}$. We know that ε annihilates E^+/C^+E^{+q}, so $(uC^+E^{+q})^\varepsilon = u^\varepsilon C^+E^{+q}$, so $u^\varepsilon \in C^+E^{+q}$. We know that ε annihilates $C^+E^{+q}/C^+E^{+q^2}$, so $(u^\varepsilon C^+E^{+q^2})^\varepsilon = u^{\varepsilon^2}C^+E^{+q^2} = C^+E^{+q^2}$, so $u^{\varepsilon^2} \in C^+E^{+q^2}$. If we go on like this, we find that ε^m annihilates $E^+/C^+E^{+q^m}$.

From Thaine's theorem 4.7, we obtain that ε^m annihilates $(Cl_{K^+})_q$ as well. Since all elements in $Cl_{K^+}[q]$ have order 1 or q, the prime q is an exponent of $Cl_{K^+}[q]$. From group theory, it follows that $\#Cl_{K^+}[q]$ divides q^n for some integer n. Therefore, $Cl_{K^+}[q]$ is a q-group and it is contained in the Sylow-q-subgroup $(Cl_{K^+})_q$. So ε^m also annihilates $Cl_{K^+}[q]$.

Under the map $\mathbb{Z}[G^+] \longrightarrow \mathbb{F}_q[G^+]$, the element $\varepsilon^m \mapsto e^m = e$. Suppose $\varepsilon^m = \sum_{\sigma \in G^+} \varepsilon_\sigma \sigma$, with $\varepsilon_\sigma \in \mathbb{Z}$. We have $e = \sum_{\sigma \in G^+} e_\sigma \sigma$ with $e_\sigma \in \mathbb{F}_q[G^+]$ such that

e_σ is a representative of the coset $\bar{\varepsilon}_\sigma$ for all $\sigma \in G^+$. Therefore, for an ideal class $[\mathfrak{b}] \in Cl_{K^+}[q]$ we have

$$\begin{aligned}[\mathfrak{b}]^{\varepsilon^m} &= [\mathfrak{b}]^{(\sum_{\sigma\in G^+}\varepsilon_\sigma\sigma)^m} = [\mathfrak{b}^{(\sum_{\sigma\in G^+}(e_\sigma+k_\sigma q)\sigma)^m} \\ &= [\mathfrak{b}]^{(\sum_{\sigma\in G^+}e_\sigma\sigma+q\sum_{\sigma\in G^+}k_\sigma\sigma)^m} = [\mathfrak{b}]^{(\sum_{\sigma\in G^+}e_\sigma\sigma)^m} \\ &= [\mathfrak{b}]^{e^m} = [\mathfrak{b}]^e, \end{aligned} \tag{8.10}$$

with $k_\sigma \in \mathbb{Z}$ for all $\sigma \in G^+$. On the other hand, $[\mathfrak{b}]^{\varepsilon^m} = [1]$. Therefore, e annihilates $Cl_{K^+}[q]$, which is what we wanted to show. □

We show that $\mathfrak{a}$ maps $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ into $C^+\mathcal{E}^q/\mathcal{E}^q$.

Lemma 8.9. *We have that*

$$(H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q})^{\mathfrak{a}} \subset C^+\mathcal{E}^q/\mathcal{E}^q.$$

Proof. Consider the exact sequence 8.2 of R-modules again. Let ψ' be the injective map $\mathcal{E}/\mathcal{E}^q \longrightarrow H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$. Then we have for all $\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \in H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$:

$$\psi'((\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q})^{\varepsilon^m}) = (\psi'(\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q}))^{\varepsilon^m} = [1],$$

because ε^m annihilates $Cl_{K^+}[q]$. Therefore, for all $\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \in H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ we have $(\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q})^{\varepsilon^m} \in \ker(\psi) = \mathrm{im}(\psi') = \mathcal{E}/\mathcal{E}^{\mathrm{II}}$. So ε^m maps $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ into $\mathcal{E}/\mathcal{E}^q$.

Now we show that ε^{m+1} maps $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ into $C^+\mathcal{E}^q/\mathcal{E}^{\mathrm{II}}$. Note that for all $\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \in H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ we have $(\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q})^{\varepsilon^k} = (\alpha\mathbb{Q}(\zeta+\zeta^{-1})^{*q})^{e^k}$ for all positive integers k. Therefore, $\alpha^{\varepsilon^m}\mathbb{Q}(\zeta+\zeta^{-1})^{*q} = \alpha^{e^m}\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \in \mathcal{E}/\mathcal{E}^q$, so $\alpha^{e^m} \in \mathcal{E}$. It follows that $\alpha^{\varepsilon^{m+1}} = \alpha^{e^{m+1}} = \alpha^{e^m}e \in C^+\mathcal{E}^q$, because $\alpha^{e^m} \in \mathcal{E}$ and e annihilates $\mathcal{E}/C^+\mathcal{E}^q$. We obtain that ε^{m+1} maps $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ inside $C^+\mathcal{E}^q/\mathcal{E}^q$. □

Before we can prove the main result of this section, we prove the following lemma, in which we use the Runge-type theorem we saw in chapter 7.

Remember that ξ is the element of $H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$ that is the image of $(x-\zeta)(x-\zeta^{-1})$ under the projection map $H \longrightarrow H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$.

Lemma 8.10. *We have the following R-isomorphism:*

$$\xi^{\mathfrak{a}} \cong \mathfrak{a}.$$

Proof. It is obvious that $\xi^{\mathfrak{a}} \subset H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$, because $\xi \in H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$.

Consider the map

$$\begin{aligned}\varphi : I_{\mathrm{aug}} &\longrightarrow H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \\ \theta &\longmapsto \xi^\theta. \end{aligned} \tag{8.11}$$

Of course, we also have the restriction map

$$\begin{aligned}\varphi' : \mathfrak{a} &\longrightarrow \xi^{\mathfrak{a}} \subset H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \\ \theta &\longmapsto \xi^\theta. \end{aligned} \tag{8.12}$$

It is obvious that φ and φ' are homomorphisms. We show that φ is R-linear. Let θ and θ' be elements of I_{aug}. Then $\varphi(\theta\cdot\theta') = \xi^{\theta\cdot\theta'} = (\xi^\theta)^{\theta'} = \varphi(\theta)^{\theta'}$. Of course, φ' is R-linear as well.

An equivalent formulation of corollary 7.3 is the following: the map given by

$$\begin{array}{rcl} I_{\text{aug}} & \longrightarrow & H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q} \\ \theta & \longmapsto & \xi^{\theta} \end{array} \tag{8.13}$$

is injective. Therefore, φ' is injective as well.

By definition of $\xi^{\mathfrak{a}}$, the map φ' is surjective. So we have shown that $\mathfrak{a}$ is R-isomorphic to $\xi^{\mathfrak{a}}$. □

We conclude that

$$\mathfrak{a} \cong \xi^{\mathfrak{a}} \subset (H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q})^{\mathfrak{a}} \subset C^{+}\mathcal{E}^{q}/\mathcal{E}^{q} \cong \mathfrak{a}. \tag{8.14}$$

It follows that all the '$\subset$' actually are '$=$'. Therefore,

$$C^{+}/(C^{+}\cap\mathcal{E}^{q}) \cong C^{+}\mathcal{E}^{q}/\mathcal{E}^{q} = (H/\mathbb{Q}(\zeta+\zeta^{-1})^{*q})^{\mathfrak{a}} = \xi^{\mathfrak{a}} \subset \xi^{I_{\text{aug}}}. \tag{8.15}$$

8.4 The contradiction

The aim of this section is to derive a contradiction from the facts we found in the previous section, together with the fact that q^2 divides x, which we proved in theorem 5.7.

By abuse of notation, let ξ denote the element $(x-\zeta)(x-\zeta^{-1}) \in H$. We start by proving the following lemma.

Lemma 8.11. *If $\gamma \in C^{+}$, then for all $\sigma \in G$ there exist non-negative integers n_σ and there exists $\delta \in \mathcal{O}_{K^+}$ such that $\delta^q\gamma = \prod_{\sigma\in G^+}(\sigma(\xi))^{n_\sigma}$. The element γ is a q-th power modulo $q^2\mathcal{O}_{K^+}$.*

Proof. At the end of the previous section, we concluded that $C^{+}/(C^{+}\cap\mathcal{E}^{q})$ is contained in $\xi^{I_{\text{aug}}} = (\xi\mathbb{Q}(\zeta+\zeta^{-1})^{*q})^{I_{\text{aug}}} = \xi^{I_{\text{aug}}}\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$. It follows that if $\gamma \in C^{+}$, then $\gamma \in \xi^{\theta}\mathbb{Q}(\zeta+\zeta^{-1})^{*q}$, for some $\theta \in I_{\text{aug}}$, so there exists $\theta' \in \mathbb{Z}[G^{+}]$ and $\delta \in \mathbb{Q}(\zeta+\zeta^{-1})^{*}$ such that $\gamma\delta^{q} = \xi^{\theta'}$. We can choose $\theta' = \sum_{\sigma\in G^+} n_\sigma\sigma$ such that the coefficients n_σ of θ' are non-negative integers, since the difference between ξ^{θ} and $\xi^{\theta'}$ is a q-th power, no matter what lift of θ we choose.

We need to show that the δ we find in this way is an algebraic integer, i.e. that it is an element of $\mathcal{O}_{K^+}$. Of course, $\gamma \in \mathcal{O}_{K^+}^{*}$ and also $\gamma^{-1} \in \mathcal{O}_{K^+}^{*}$. Let σ be an element of G^{+}. Then we find that $\sigma(\xi) = \sigma((x-\zeta)(x-\zeta^{-1})) = (x-\zeta^{a})(x-\zeta^{-a})$ for some $a \in \mathbb{Z}$. Since $(x-\zeta^{a})(x-\zeta^{-a}) = x^2-(\zeta^{a}+\zeta^{-a})x+1 \in \mathbb{Z}[\zeta+\zeta^{-1}] = \mathcal{O}_{K^+}$, we obtain $\xi^{\theta'} \in \mathcal{O}_{K^+}$. It follows that $\delta^{q} = \gamma^{-1}\xi^{\theta'} \in \mathbb{Z}[\zeta+\zeta^{-1}] = \mathcal{O}_{K^+}$. Since δ is integral over $\mathcal{O}_{K^+}$, it is an element of $\mathcal{O}_{K^+}$, which is what we wanted to show.

We have $\sigma(\xi) = (x-\zeta^{a})(x-\zeta^{-a}) = x^2-(\zeta^{a}+\zeta^{-a})x+1$ for some integer a. We saw in theorem 5.7 that q^2 divides x. Therefore, $\sigma(\xi) \equiv 1 \pmod{q^2}$. It follows that $\delta^{q}\gamma = \prod_{\sigma\in G^+}\sigma(\xi)^{n_\sigma} \equiv 1 \pmod{q^2}$. So $\delta^{q-1}\gamma$ is the inverse of δ modulo $q^2\mathcal{O}_{K^+}$ and $\gamma \equiv (\delta^{q-1}\gamma)^{q} \pmod{q^2}$. This means that γ is a q-th power modulo $q^2\mathcal{O}_{K^+}$. □

Lemma 8.12. *The element $1+\zeta$ of $\mathbb{Q}(\zeta)$ is a q-th power modulo $q^2\mathbb{Z}[\zeta]$.*

Proof. First, we show that $(1+\zeta)^p \in C^+$. Note that $1+\zeta = \frac{1-\zeta^2}{1-\zeta}$, so $(1+\zeta)^p = \left(\frac{1-\zeta^2}{1-\zeta}\right)^p = \xi_2^p \in C^+$, with ξ_2 defined as we did in theorem 4.5. Applying lemma 8.11 we obtain that $(1+\zeta)^p$ is a q-th power modulo $q^2\mathcal{O}_{K^+}$. We need to show that $1+\zeta$ is a q-th power modulo $q^2\mathcal{O}_K$. Since all elements of $\mathcal{O}_{K^+} = \mathbb{Z}[\zeta+\zeta^{-1}]$ are also elements of $\mathcal{O}_K = \mathbb{Z}[\zeta]$, we have the following identity:

$$(1+\zeta)^p = \alpha^q + q^2\mathbb{Z}[\zeta],$$

for some $\alpha \in \mathbb{Z}[\zeta]$. Of course, $(1+\zeta)^p \in C^+ \in \mathbb{Z}[\zeta+\zeta^{-1}]^*$, so there exists $a \in \mathbb{Z}[\zeta+\zeta^{-1}]^*$ such that $a(1+\zeta)^p = 1$. Therefore, we have that $a(1+\zeta)^{p-1}$ is the inverse of $1+\zeta$, so $1+\zeta \in \mathbb{Z}[\zeta+\zeta^{-1}]^*$.

Since p and q are coprime, there exists $s \in \mathbb{Z}$ such that $ps \equiv 1 \pmod q$, let us say that $ps = 1+kq$. We find that:

$$(1+\zeta)(1+\zeta)^{kq} = (1+\zeta)^{1+kq} = (1+\zeta)^{ps} = ((1+\zeta)^p)^s \equiv \alpha^{qs} \pmod{q^2\mathbb{Z}[\zeta]}.$$

Therefore, $1+\zeta \equiv (\alpha^s)^q((1+\zeta)^{-k})^q \pmod{q^2\mathbb{Z}[\zeta]}$, so $1+\zeta$ is a q-th power modulo $q^2\mathbb{Z}[\zeta]$. □

We conclude that $1+\zeta \equiv \alpha^q \pmod{q^2\mathbb{Z}[\zeta]}$ for some $\alpha \in \mathbb{Z}[\zeta]$.

Of course, the Galois group $G = \mathrm{Gal}(\mathbb{Q}(\zeta)/\mathbb{Q})$ acts on $\mathbb{Q}(\zeta)$ and also on $\mathbb{Z}[\zeta]$. For $a \in \mathbb{Z}$, let σ_a denote the element of G that sends ζ to ζ^a. Of course, if two integers a and b are congruent modulo p, then $\sigma_a = \sigma_b$. The following lemma tells us more about σ_q.

Lemma 8.13. *For all $\alpha \in \mathbb{Z}[\zeta]$ we have that $\sigma_q(\alpha) \equiv \alpha^q \pmod{q\mathbb{Z}[\zeta]}$.*

Proof. Since q is unramified in $\mathbb{Q}(\zeta)$, we know that there exists a Frobenius automorphism $\tau_q \in G$, i.e. $\tau_q(\alpha) \equiv \alpha^q \pmod q$ for all $\alpha \in \mathbb{Z}[\zeta]$. Since this congruence holds for $\alpha = \zeta$, we find that τ_q equals σ_q. Therefore, for all $\alpha \in \mathbb{Z}[\zeta]$ we have $\sigma_q(\alpha) \equiv \alpha^q \pmod q$. □

Now we apply σ_q to $1+\zeta$. On the one hand, this yields $\sigma_q(1+\zeta) = 1+\zeta^q$, and on the other hand $\sigma_q(1+\zeta) \equiv \sigma_q(\alpha^q) \equiv (\sigma_q(\alpha))^q \equiv \beta^q \pmod{q^2\mathbb{Z}[\zeta]}$ for some $\beta \in \mathbb{Z}[\zeta]$. Therefore,

$$1+\zeta^q \equiv \beta^q \pmod{q^2\mathbb{Z}[\zeta]}$$

for some $\beta \in \mathbb{Z}[\zeta]$.

Since $\sigma_q(1+\zeta) = 1+\zeta^q \equiv \beta^q \pmod{q^2\mathbb{Z}[\zeta]}$, we have $\sigma_q(1+\zeta) \equiv \beta^q \equiv \sigma_q(\beta) \pmod{q\mathbb{Z}[\zeta]}$. Applying σ_q^{-1}, we find $1+\zeta \equiv \beta \pmod{q\mathbb{Z}[\zeta]}$, so $1+\zeta = \beta + qa$ with $a \in \mathbb{Z}[\zeta]$. Therefore,

$$(1+\zeta)^q = (\beta+qa)^q \equiv \beta^q \pmod{q^2\mathbb{Z}[\zeta]}.$$

It follows that $(1+\zeta)^q \equiv 1+\zeta^q \pmod{q^2\mathbb{Z}[\zeta]}$, which implies

$$(1+\zeta)^q = \sum_{i=0}^{q} \binom{q}{i}\zeta^i \equiv 1+\zeta^q \pmod{q^2\mathbb{Z}[\zeta]}.$$

It follows that we have $\sum_{i=1}^{q-1}\binom{q}{i}\zeta^i = \sum_{i=0}^{p-2} q^2 b_i \zeta^i$ for integers b_i. Unique representation of elements of $\mathbb{Z}[\zeta]$ as linear combinations of $1, \zeta, \zeta^2, \ldots, \zeta^{p-2}$ over $\mathbb{Z}$ implies that q^2 divides $\binom{q}{i}$ for all $i = 1, \ldots, q-1$, which is a contradiction. This completes the proof of Catalan's conjecture.

Bibliography

[1] Yuri F. Bilu. Catalan's conjecture (after Mihăilescu). *Séminaire Bourbaki*, 909, 2002–2003.

[2] J.W.S. Cassels. On the equation $a^x - b^y = 1$. *American Journal of Mathematics*, 75:159–162, 1953.

[3] J.W.S. Cassels. On the equation $a^x - b^y = 1$ II. *Proceedings of the Cambridge Philosophical Society*, 56:97–103, 1960.

[4] E. Catalan. Note extraite d'une lettre adressée à l'éditeur. *Journal für die reine und angewandte Mathematik*, 27:192, 1844.

[5] Eugène-Charles Catalan. Quelques théorèmes empiriques. (1842–43). In *Mélanges Mathématiques*, Mémoires de la Société Royale des Sciences de Liège, deuxième série, volume 12, pages 42–43. 1885.

[6] E.Z. Chein. A note on the equation $x^2 = y^q + 1$. *Proceedings of the American Mathematical Society*, 56:83–84, 1976.

[7] Leonard Eugene Dickson. *History of the Theory of Numbers*, volume II. Chelsea Publishing Company, New York, 1952.

[8] L. Euler. Commentationes Arithmeticae I. In *Opera Omnia*, Series I, volume II, pages 56–58. B.G. Teubner, Basel, 1915.

[9] Chao Ko. On the diophantine equation $x^2 = y^n + 1$, $xy \neq 0$. *Scientia Sinica*, 14:457–460, 1965.

[10] V.A. Lebesgue. Sur l'impossibilité, en nombres entiers, de l'équation $x^m = y^2 + 1$. *Nouvelles annales de mathématiques*, 9:178–181, 1850.

[11] Wm. J. LeVeque. On the equation $a^x - b^y = 1$. *American Journal of Mathematics*, 74:325–331, 1952.

[12] A. Mąkowski. Three consecutive integers cannot be powers. *Colloquium Mathematicum*, IX:297, 1962.

[13] Preda Mihăilescu. A Class Number Free Criterion for Catalan's Conjecture.

[14] Paulo Ribenboim. *Catalan's Conjecture*. Academic Press, Boston, 1994.

[15] Joseph H. Silverman and John Tate. *Rational Points on Elliptic Curves*. Springer-Verlag, New York, 1992.

[16] Francisco Thaine. On the ideal class groups of real abelian number fields. *Annals of Mathematics*, 128:1–18, 1988.

[17] R. Tijdeman. On the equation of Catalan. *Acta Arithmetica*, 29:197–209, 1976.

[18] Lawrence C. Washington. *Introduction to Cyclotomic Fields*. Springer-Verlag, New York, 1982.

[19] Yann Bugeaud and Guillaume Hanrot. Un nouveau critère pour l'équation de Catalan. *Mathematika*, 47:63–73, 2000.