Diophantine equations and cyclotomic fields

Dissertation

for the award of the degree

"Doctor of Philosophy" Ph.D. Division of Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

within the doctoral program "Mathematical Science"

of the Georg-August University School of Science (Gauss)

Submitted by

Boris Bartolomé

from Toulouse, France

Göttingen, 2015

Thesis Committee

Prof. Dr. Yuri Bilu, Institut de Mathématiques de Bordeaux, Université de Bordeaux

Prof. Dr. Preda Mihăilescu, Mathematisches Institut, Georg-August-Universität Göttingen

Members of the Examination Board

Reviewer: Prof. Dr. Jörg Brüdern, Mathematisches Institut, Georg-August-Universität Göttingen

Reviewer: Prof. Dr. Yann Bugeaud, Institut de Recherche Mathématique Avancée, Université de Strasbourg

Reviewer: Prof. Dr. Clemens Fuchs, Fachbereich Mathematik, Universität Salzburg (not member of the examination board)

Further Members of the Examination Board

Prof. Dr. Philipp Habegger, Mathematisches Institut, Universität Bäsel

Prof. Dr. Jean-François Jaulent, Institut de Mathématiques de Bordeaux, Université de Bordeaux

Prof. Dr. David Masser, Mathematisches Institut, Universität Bäsel

Ass. Prof. Dr. Fabien Pazuki, Department of Mathematical Sciences, University of Copenhagen

Date of the Oral Examination: November 26^{th}, 2015.

Acknowledgments

I would first like to thank my wife, Corinne, without whose encouragement and continuous support this work would have never happened and my life would have taken other paths. She is my stabilizer [1]. Also, a deep gratitude goes towards those who have trusted in and helped me: Yuri Bilu and Preda Mihăilescu; after months trying to find an advisor, a difficult task given my age, I found two the same day. And even though I do not believe in fate, I must admit that I have been the luckiest human being that day. I was resigned to accept any topic and any advisor, and I found two extraordinary mathematicians who would *(almost)* let me chose the topics I wanted to work on. I have understood the need to be rigourous with Yuri, always available and always open, frank and friendly, someone to admire besides his mathematical abilities. And I have witnessed an extraordinary mathematical intuition with Preda, a fighter to mathematics and to life, with whom I have developed a very deep complicity. I hope these few years are only the beginning of a long lasting frienship with both of them. Among the people who have gone beside their duty to help me out there's first Cyril Mauvillain and all the Bordeaux library of mathematics team: thank you for all you did to help me, sometimes very far from what is to be expected from you. A special thanks to Karim Belabas, maintainer of PARI, who has answered my questions when I was half world accross from him, at some indecent time at night. I am thankful also to the administrative staff at Göttingen University, especially to Stefan Halversheid, the dean of mathematics and computer science when I arrived, who hosted me in his house and helped me match the requirements of the cotutelle with the German constraints, and in the same line Max Wardetzky, current dean, who has helped shape the PhD on the German side. Also Mrs. Barann for her indefectible advice on rules I still do not understand. I am also indebted to Jörg Brüdern, Yann Bugeaud and Clemens Fuchs for having accepted to review this work and for their useful comments. Thank you also to David Masser and Jean-François Jaulent for having come out of their retirement for a day or two, as well as Philipp Habegger and Fabien Pazuki who all have travelled to Göttingen to evaluate this work. I would also like to thank my children, for the life they bring into mine, and without whom this thesis would have certainly contained more results. And my parents for having allowed me to acquire some fundamental life values. I have come accross many mathematicians during this journey, and most have been really wonderful, helpful and motivating. To all of them, thank you. Finally, to the creator of TEX, Donald Knuth, who has allowed thousands of scientists to present their results beautifully: THANX.

[1] Note to algebraists: in the human (or even chemical) sense

If you can keep your head when all about you
Are losing theirs and blaming it on you,
If you can trust yourself when all men doubt you,
But make allowance for their doubting too;
If you can wait and not be tired by waiting,
Or being lied about, don't deal in lies,
Or being hated, don't give way to hating,
And yet don't look too good, nor talk too wise:

If you can dream - and not make dreams your master;
If you can think - and not make thoughts your aim;
If you can meet with Triumph and Disaster
And treat those two impostors just the same;
If you can bear to hear the truth you've spoken
Twisted by knaves to make a trap for fools,
Or watch the things you gave your life to, broken,
And stoop and build 'em up with worn - out tools:

If you can make one heap of all your winnings
And risk it on one turn of pitch-and-toss,
And lose, and start again at your beginnings
And never breathe a word about your loss;
If you can force your heart and nerve and sinew
To serve your turn long after they are gone,
And so hold on when there is nothing in you
Except the Will which says to them: 'Hold on!'

If you can talk with crowds and keep your virtue,
Or walk with Kings - nor lose the common touch,
If neither foes nor loving friends can hurt you,
If all men count with you, but none too much;
If you can fill the unforgiving minute
With sixty seconds' worth of distance run,
Yours is the Earth and everything that's in it,
And - which is more - you'll be a Man, my son!

Rudyard Kipling, 1895

Summary

Diophantine equations and cyclotomic fields

This thesis examines some approaches to address Diophantine equations, specifically we focus on the connection between the Diophantine analysis and the theory of cyclotomic fields.

First (in Chapter 2), we propose a quick introduction to the methods of Diophantine approximation we have used in this research work. We remind the notion of height and introduce the logarithmic gcd.

Then (in Chapter 3), we address a conjecture, made by Thoralf Skolem in 1937, on an exponential Diophantine equation. For this conjecture, let $\mathbb{K}$ be a number field, $\alpha_1, \ldots, \alpha_m, \lambda_1, \ldots, \lambda_m$ non-zero elements in $\mathbb{K}$, and S a finite set of places of $\mathbb{K}$ (containing all the infinite places) such that the ring of S-integers

$$\mathcal{O}_S = \mathcal{O}_{\mathbb{K},S} = \{\alpha \in \mathbb{K} : |\alpha|_v \leq 1 \text{ for places } v \notin S\}$$

contains $\lambda_1, \ldots, \lambda_m, \alpha_1, \ldots, \alpha_m, \alpha_1^{-1}, \ldots, \alpha_m^{-1}$. For every $n \in \mathbb{Z}$, let $A(n) = \lambda_1\alpha_1^n + \cdots + \lambda_m\alpha_m^n \in \mathcal{O}_S$. Skolem suggested [Skolem 1937]:

Conjecture 0.0.1 (Exponential Local-Global Principle) *Assume that for every non zero ideal $\mathfrak{a}$ of the ring $\mathcal{O}_S$, there exists $n \in \mathbb{Z}$ such that $A(n) \equiv 0 \bmod \mathfrak{a}$. Then there exists $n \in \mathbb{Z}$ such that $A(n) = 0$.*

Let Γ be the multiplicative group generated by $\alpha_1, \ldots, \alpha_m$. Then Γ is the product of a finite abelian group and a free abelian group of finite rank. We prove that the conjecture is true when the rank of Γ is one.

This result was proved in collaboration with Florian Luca, from University of the Witwatersrand (South Africa) and Yuri Bilu. It was published in Acta Arithmetica [Bartolomé *et al.* 2013]. Shortly after its publication, Florian Luca met Andrzej Schinzel in a mathematical congress, and Schinzel told him that our result was a direct consequence of [Schinzel 1977][Theorem 6]. A quick verification proved it was true. However, this work has been done with no previous knowledge of this result and using other (subspace theorem and Baker's inequality), interesting per se, methods.

After that (in Chapter 4), we generalize a result previously published by Abouzaid ([Abouzaid 2008]). Let $F(X, Y) \in \mathbb{Q}[X, Y]$ be a $\mathbb{Q}$-irreducible polynomial. In 1929 Skolem [Skolem 1929] proved the following beautiful theorem:

Theorem 0.0.2 (Skolem) *Assume that*

$$F(0, 0) = 0.$$

Then for every non-zero integer d, the equation $F(X, Y) = 0$ has only finitely many solutions in integers $(X, Y) \in \mathbb{Z}^2$ with $\gcd(X, Y) = d$.

In 2008, Abouzaid [Abouzaid 2008] generalized this result by working with arbitrary algebraic numbers and by obtaining an asymptotic relation between the heights of the coordinates and their logarithmic gcd. He proved the following theorem:

Theorem 0.0.3 (Abouzaid) *Assume that* $(0,0)$ *is a non-singular point of the plane curve* $F(X,Y)=0$. *Let* $m=\deg_X F,\quad n=\deg_Y F,\quad M=\max\{m,n\}$. *Let* ε *satisfy* $0<\varepsilon<1$. *Then for any solution* $(\alpha,\beta)\in\bar{\mathbb{Q}}^2$ *of* $F(X,Y)=0$, *we have either*

$$\max\{\mathrm{h}(\alpha),\mathrm{h}(\beta)\}\leq 56M^8\varepsilon^{-2}\mathrm{h}_p(F)+420M^{10}\varepsilon^{-2}\log(4M),$$

or

$$\max\{|\mathrm{h}(\alpha)-n\mathrm{lgcd}(\alpha,\beta)|,|\mathrm{h}(\beta)-m\mathrm{lgcd}(\alpha,\beta)|\}\leq\varepsilon\max\{\mathrm{h}(\alpha),\mathrm{h}(\beta)\}+742M^7\varepsilon^{-1}\mathrm{h}_p(F)+5762M^9\varepsilon^{-1}\log(2m+2n).$$

However, he imposed the condition that $(0,0)$ be a non-singular point of the plane curve $F(X,Y)=0$. Using a somewhat different version of Siegel's "absolute" Lemma and of Eisenstein's Lemma, we could remove the condition and prove it in full generality. We prove the following theorem:

Theorem 0.0.4 *Let* $F(X,Y)\in\bar{\mathbb{Q}}[X,Y]$ *be an absolutely irreducible polynomial satisfying* $F(0,0)=0$. *Let* $m=\deg_X F,\quad n=\deg_Y F$ *and* $r=\min\left\{i+j:\frac{\partial^{i+j}F}{\partial^i X\partial^j Y}(0,0)\neq 0\right\}$. *Let* ε *satisfy* $0<\varepsilon<1$. *Then, for any* $\alpha,\beta\in\bar{\mathbb{Q}}$ *such that* $F(\alpha,\beta)=0$, *we have either:*

$$\mathrm{h}(\alpha)\leq 200\varepsilon^{-2}mn^6(\mathrm{h}_p(F)+5)$$

or

$$\left|\tfrac{\mathrm{lgcd}(\alpha,\beta)}{r}-\tfrac{\mathrm{h}(\alpha)}{n}\right|\leq\tfrac{1}{r}(\varepsilon\mathrm{h}(\alpha)+4000\varepsilon^{-1}n^4(\mathrm{h}_p(F)+\log(mn)+1)+30n^2m(\mathrm{h}_p(F)+\log(nm))).$$

In our proof, we closely follow Abouzaid's methods. This result was also submitted for publication in 2014, and during the review process, the referee *kindly* pointed out that this result had already been proven in Philipp Habegger's unpublished PhD Thesis, and quick check proved that to be true: the result is proved in [Habegger 2007][Appendix B, Theorem B.3] using his quasi-equivalence of heights. While we admit his solution is more "industrial" and provides a better bound, we still believe that Abouzaid's inital argument is quite enlightening and natural in certain ways. Our result has been published in [Bartolomé 2015]. Our main tool is Puiseux expansions.

Then (in Chapter 5) we give an overview of the tools we have used in cyclotomic fields. We try there to develop a systematic approach to address a certain type of Diophantine equations. We discuss on cyclotomic extensions and give some basic but useful properties, on group-ring properties and on Jacobi sums.

Finally, (in Chapter 6) we show a very interesting application of the approach developped in the previous chapter. There, we consider the Diophantine equation

$$X^n-1=BZ^n, \tag{1}$$

where $B\in\mathbb{Z}$ is understood as a parameter. Define $\varphi^*(B):=\varphi(\mathrm{rad}\,(B))$, where rad (B) is the radical of B, and assume that

$$(n,\varphi^*(B))=1. \tag{2}$$

For a fixed $B\in\mathbb{N}_{>1}$ we let

$$\mathcal{N}(B)=\{n\in\mathbb{N}_{>1}\mid\exists\ k>0\text{ such that }n|\varphi^*(B)^k\}.$$

If p is an odd prime, we shall denote by CF the combined condition requiring that

I The Vandiver Conjecture holds for p, so the class number h_p^+ of the maximal real subfield of the cyclotomic field $\mathbb{Q}[\zeta_p]$ is not divisible by p.

II We have $i_r(p) < \sqrt{p} - 1$, in other words, there is at most $\sqrt{p} - 1$ odd integers $k < p$ such that the Bernoulli number $B_k \equiv 0 \bmod p$.

Current results on Equation (1) are restricted to values of B which are built up from two small primes $p \leq 13$ [Bennett *et al.* 2006] and complete solutions for $B < 235$ ([A.Bazso *et al.* 2010]). If expecting that the equation has no solutions, – possibly with the exception of some isolated examples – it is natural to consider the case when the exponent n is a prime. Of course, the existence of solutions (X, Z) for composite n imply the existence of some solutions with n prime, by raising X, Z to a power.

The main contribution of our work has been to relate (1) in the case when n is a prime and (2) holds, to the diagonal Nagell – Ljunggren equation,

$$\frac{X^n - 1}{X - 1} = n^e Y^n, \quad e = \begin{cases} 0 & \text{if } X \not\equiv 1 \bmod n, \\ 1 & \text{otherwise.} \end{cases}$$

This way, we can apply results from [Mihăilescu 2008] and prove the following:

Theorem 0.0.5 *Let n be a prime and $B > 1$ an integer with $(\varphi^*(B), n) = 1$. Suppose that equation (1) has a non trivial integer solution different from $n = 3$ and $(X, Z; B) = (18, 7; 17)$. Let $X \equiv u \bmod n, 0 \leq u < n$ and $e = 1$ if $u = 1$ and $e = 0$ otherwise. Then:*

1. *$n > 163 \cdot 10^6$.*
2. *$X - 1 = \pm B/n^e$ and $B < n^n$.*
3. *If $u \notin \{-1, 0, 1\}$, then condition CF (II) fails for n and*

$$\begin{array}{rcl} 2^{n-1} \equiv 3^{n-1} \equiv 1 \mod n^2, & & \text{and} \\ r^{n-1} \equiv 1 \mod n^2 & & \text{for all } r | X(X^2 - 1). \end{array}$$

If $u \in \{-1, 0, 1\}$, then Condition CF (I) fails for n.

Based on this theorem, we also prove the following:

Theorem 0.0.6 *If equation (1) has a solution for a fixed B verifying the conditions (2), then either $n \in \mathcal{N}(B)$ or there is a prime p coprime to $\varphi^*(B)$ and a $m \in \mathcal{N}(B)$ such that $n = p \cdot m$. Moreover X^m, Y^m are a solution of (1) for the prime exponent p and thus verify the conditions in Theorem 0.0.5.*

This is a strong improvement of the currently known results.

As we have made heavy use of [Mihăilescu 2008], at the end of this thesis we have added an appendix to expose some new result that allows for a full justification of Theorem 3 of [Mihăilescu 2008].

Keywords

Diophantine Equations, Cyclotomic Fields, Nagell-Ljunggren Equation, Skolem, Abouzaid, Exponential Diophantine Equation, Baker's Inequality, Subspace Theorem.

Zusammenfassung

Diophantine equations and cyclotomic fields

Diese Doktorarbeit untersucht einige Verfahren zur Behandlung von Diophantischen Gleichungen. Wir behandeln insbesondere den Zusammnehang zwischen Diophantischer Analysis und der Theorie von Kreisteilungskörper.

In Kapitel 2 wird eine kurze Einführung in den Methoden der Diophantischen Approximation, die wir in dieser Arbeit verwendeten, gegeben. Insbesondere werden die Begriffe von Höhe und logarithmischen grössten geminsamen Teiler eingeführt.

Im darauffolgenden Kapitel 3, wird eine Vermutung von Thoralf Skolem aus dem Jahr 1937 behandelt, betreffend einer Diophantischen Gleichung. Sei $\mathbb{K}$ ein Zahlkörper, $\alpha_1, \ldots, \alpha_m, \lambda_1, \ldots, \lambda_m$ nicht verschwindende algebraische Zahlen aus $\mathbb{K}$ und S eine endliche Menge von Stellen aus $\mathbb{K}$, die alle unendlichen Stellen enthält und so, dass der Ring der S-ganzen Zahlen

$$\mathcal{O}_S = \mathcal{O}_{\mathbb{K},S} = \{\alpha \in \mathbb{K} : |\alpha|_v \leq 1 \text{ für Stellen } v \notin S\}$$

auch $\lambda_1, \ldots, \lambda_m, \alpha_1, \ldots, \alpha_m, \alpha_1^{-1}, \ldots, \alpha_m^{-1}$ enthält.

Für jedes $n \in \mathbb{Z}$, sei $A(n) = \lambda_1\alpha_1^n + \cdots + \lambda_m\alpha_m^n \in \mathcal{O}_S$. Skolem vermutete [Skolem 1937]:

Conjecture 0.0.7 (Exponential Local-Global Principle) *Angenommen, dass für jedes nicht triviale Ideal $\mathfrak{a}$ im Ganzheitsring $\mathcal{O}_S$, ein $n \in \mathbb{Z}$ existiert, so dass $A(n) \equiv 0 \bmod \mathfrak{a}$; dann existiert ein $n \in \mathbb{Z}$, so dass $A(n) = 0$.*

Sei Γ die durch $\alpha_1, \ldots, \alpha_m$ erzeugte multiplicative Gruppe. Dann ist Γ Produkt einer endlichen abelschen Gruppe mit einer freien abelschen Gruppe von endlichem Rang. Wir beweisen die Vermutung für den Fall in dem der freie Teil den Rang eins hat.

Das Ergebnis wurde in Zusammenarbeit mit Florian Luca, von der University of the Witwatersrand (Süd Afrika) und Yuri Bilu erhalten und wurde in Acta Arithmetica [Bartolomé *et al.* 2013] publiziert. Kurz nach der Publikation wurde Florian Luca von Andrzej Schinzel davon informiert, dass unser Ergebnis eine direkte Konsequenz von Lehrsatz 6 von [Schinzel 1977] ist, was danach leicht zu konfirmieren war. Nicht destotrotz wurde unser Ergebnis ohne Kenntnis der Arbeit von Schintzel erhalten und der Beweis verwendet wesentlich verschiedene Methoden, die in sich interessant sind.

Im Kapitel 4 wird ein früheres Ergebnis von Abouzaid ([Abouzaid 2008]) verallgemeinert. Sei $F(X, Y) \in \mathbb{Q}[X, Y]$ ein $\mathbb{Q}$-unzerlegbares Polynom. In 1929 bewies Skolem [Skolem 1929] folgenden schönen Satz:

Theorem 0.0.8 (Skolem) *Sei*

$$F(0, 0) = 0.$$

Dann ist die Menge der Lösungen $L_d = \{F(X, Y) = 0 : X, Y \in \mathbb{Z} und (X, Y) = d\}$ endlich, für jeden $d > 0$.

In 2008, verallgemeinerte Abouzaid [Abouzaid 2008] dieses Ergebnis, indem er in Zahlenkörper arbeitete. Er bewies folgenden Satz:

Theorem 0.0.9 (Abouzaid) *Sei $(0, 0)$ ein nicht - singurlärer Punkt der ebenen Kurve $F(X, Y) = 0$. Sei $m = \deg_X F$, $n = \deg_Y F$, $M = \max\{m, n\}$ und ε genüge den Ungleichungen $0 < \varepsilon < 1$. Dann gilt für jede Lösung $(\alpha, \beta) \in \bar{\mathbb{Q}}^2$ von $F(X, Y) = 0$, entweder*

$$\max\{\mathrm{h}(\alpha), \mathrm{h}(\beta)\} \leq 56M^8\varepsilon^{-2}\mathrm{h}_p(F) + 420M^{10}\varepsilon^{-2}\log(4M),$$

oder

$$\max\{|\mathrm{h}(\alpha) - n\mathrm{lgcd}(\alpha,\beta)|, |\mathrm{h}(\beta) - m\mathrm{lgcd}(\alpha,\beta)|\} \leq \varepsilon \max\{\mathrm{h}(\alpha), \mathrm{h}(\beta)\} + 742M^7\varepsilon^{-1}\mathrm{h}_p(F) + 5762M^9\varepsilon^{-1}\log(2m+2n).$$

Die Bedingung, dass $(0,0)$ ein nicht singulärer Punkt sei, ist eine Einschränkung in diesem Ergebnis. Wir konnten diese Einschränkung aufheben, in dem wir eine leicht veränderte Version des "absoluten" Lemma von Siegel und des Eisenstein-Lemmas verwendeten. Folgender Satz ergibt sich:

Theorem 0.0.10 *Sei $F(X,Y) \in \bar{\mathbb{Q}}[X,Y]$ ein total unzerlegbarer Polynom mit $F(0,0) = 0$. Sei $m = \deg_X F$, $n = \deg_Y F$ und $r = \min\left\{i+j : \frac{\partial^{i+j}F}{\partial^i X \partial^j Y}(0,0) \neq 0\right\}$. Sei $\varepsilon \in \mathbb{R}$ mit $0 < \varepsilon < 1$. Dann gilt für jede Lösung $\alpha, \beta \in \bar{\mathbb{Q}}$ von $F(X,Y) = 0$, entweder:*

$$\mathrm{h}(\alpha) \leq 200\varepsilon^{-2}mn^6(\mathrm{h}_p(F) + 5)$$

oder

$$\left|\frac{\mathrm{lgcd}(\alpha,\beta)}{r} - \frac{\mathrm{h}(\alpha)}{n}\right| \leq \frac{1}{r}(\varepsilon\mathrm{h}(\alpha) + 4000\varepsilon^{-1}n^4(\mathrm{h}_p(F) + \log(mn) + 1) + 30n^2m(\mathrm{h}_p(F) + \log(nm))).$$

Dieses Ergebnis wurde 2014 zur Publikation eingereicht; doch wies der Referee darauf hin, dass dieses Ergebnis in einer Doktorarbeit von Philipp Habegger bewiesen wurde – tatsächlich befindet sich das Ergebnis in [Habegger 2007][Appendix B, Theorem B.3] und wird bewiesen mittels der von Habegger eingeführten quasi-Äquivalenz von Höhen. Unser Ergebnis ist weniger technisch und verwendet einleuchtende Methoden, die auf Puiseux-Reihen basieren. Es wurde publiziert in [Bartolomé 2015].

Im Kapitel 5 werden einige Ergebnisse aus der Theorie der Kreisteilungskörper bewiesen, um einen systematischen Lösungsvorgang für bestimmte exponentielle Diophantische Gleichungen darzustellen. Wir besprechen auch einige Eigenschaften von Gruppenringe und von Jacobi-Summen. Darauf basierend wird in Kapitel 6 eine interessante Anwendung entwickelt. Wir betrachten die Diophantische Gleichung

$$X^n - 1 = BZ^n, \tag{3}$$

wobei $B \in \mathbb{Z}$ als Parameter zu verstehen ist. Sei $\varphi^*(B) := \varphi(\mathrm{rad}\ (B))$, mit rad (B) dem Radikal von B, und nehme an, dass

$$(n, \varphi^*(B)) = 1. \tag{4}$$

Zudem definieren wir für festen $B \in \mathbb{N}_{>1}$

$$\mathcal{N}(B) = \{n \in \mathbb{N}_{>1} \mid \exists\ k > 0 \text{ such that } n|\varphi^*(B)^k\}.$$

Falls p eine ungerade Primzahl ist, dann bezeichnen wir mit CF das Bedingungspaar

I Die Vermutung von Vandiver ist wahr für p: somit ist die Klassenzahl h_p^+ des maximalen reellen Teilkörpers des p-ten Kreisteilungskörpers $\mathbb{Q}[\zeta_p]$ nicht durch p teilbar.

II Der irregularitätsindex ist beschränkt durch $i_r(p) < \sqrt{p} - 1$; es gibt also höchstens $\sqrt{p} - 1$ ungerade $k < p$ für denen der Zähler der Bernoullizahl $B_k \equiv 0 \bmod p$.

Die besten Ergebnisse sind zur Zeit auf Parameter B eingeschränkt, die durch Primzahlen $q \leq 13$ teilbar sind [Bennett *et al.* 2006] und es sind vollständige Lösungen für $B < 235$ ([A.Bazso *et al.* 2010]) bekannt.

Wenn man von der Erwartung ausgeht, dass die Gleichung keine Lösungen besitzt, ist es natürlich vom Falle auszugehen, in dem der Exponent n eine Primzahl ist: die Existenz von Lösungen für einen zusammengesetzten Exponent n impliziert die Existenz von Lösungen für dessen Primteiler, als Exponent.

Das Hauptergebnis der Arbeit besteht darin, die Gleichung (3), unter Voraussetzung dass n prim ist und (4) gilt, auf dem besser verstandenen Diagonalfall der Gleichung von Nagell – Ljunggren zu beziehen:

$$\frac{X^n - 1}{X - 1} = n^e Y^n, \quad e = \begin{cases} 0 & \text{Falls } X \not\equiv 1 \bmod n, \\ 1 & \text{sonst.} \end{cases}$$

Damit können Ergebnisse aus [Mihăilescu 2008] verwendet werden und wir beweisen

Theorem 0.0.11 *Sei n prim und $B > 1$ eine ganze Zahl mit $(\varphi^*(B), n) = 1$. Angenommen, die Gleichung (3) habe eine nicht-triviale Lösung, die verschieden ist von $n = 3$ und $(X, Z; B) = (18, 7; 17)$, sei $X \equiv u \bmod n, 0 \leq u < n$ mit $e = 1$ falls $u = 1$ and $e = 0$ sonst. Dann gilt:*

1. *$n > 163 \cdot 10^6$.*
2. *$X - 1 = \pm B/n^e$ und $B < n^n$.*
3. *Falls $u \notin \{-1, 0, 1\}$, dann wird die Bedingung CF (II) durch n nicht erfüllt und*

$$\begin{array}{rcl} 2^{n-1} & \equiv & 3^{n-1} \equiv 1 \mod n^2, \quad \text{und} \\ r^{n-1} & \equiv & 1 \mod n^2 \quad \text{für alle } r | X(X^2 - 1). \end{array}$$

Falls $u \in \{-1, 0, 1\}$, dann ist die Bedingung CF (I) für n falsch.

Aus diesem Satz folgern wir:

Theorem 0.0.12 *Falls die Gleichung (3) für ein festes B, das die Bedingungen (4) erfüllt, eine Lösung besitzt, dann ist entweder $n \in \mathcal{N}(B)$ oder es gibt eine Primzahl p, die zu $\varphi^*(B)$ teilerfremd ist und ein $m \in \mathcal{N}(B)$, so dass $n = p \cdot m$. Zudem bilden X^m, Y^m eine Lösung von (3) für den primen Exponent p und erfüllen somit die Bedingungen in Satz 0.0.11.*

Dies verbessert die aktuelle Ergebnisse wesentlich.

Im Appendix wird eine ausführliche Beweisführung des Theorems 3 in [Mihăilescu 2008] angegeben, das im Kapitel 6 eine wesntliche Rolle spielt.

Keywords

Diophantine Equations, Cyclotomic Fields, Nagell-Ljunggren Equation, Skolem, Abouzaid, Exponential Diophantine Equation, Baker's Inequality, Subspace Theorem.

Résumé

Diophantine equations and cyclotomic fields

Cette thèse examine quelques approches aux équations diophantiennes, en particulier les connexions entre l'analyse diophantienne et la théorie des corps cyclotomiques.

Tout d'abord (au chapitre 2), nous proposons une introduction très sommaire et rapide aux méthodes d'analyse diophantienne que nous avons utilisées dans notre travail de recherche. Nous rappelons la notion de hauteur et présentons le PGCD logarithmique.

Ensuite (au chapitre 3), nous attaquons une conjecture, formulée par Skolem en 1937, sur une equation diophantienne exponentielle. Pour cette conjecture, soit $\mathbb{K}$ un corps de nombres, $\alpha_1, \ldots, \alpha_m, \lambda_1, \ldots, \lambda_m$ des éléments non-nuls de $\mathbb{K}$, et S un ensemble fini de places de $\mathbb{K}$ (qui contient toutes les places infinies), de telle sorte que l'anneau de S-entiers

$$\mathcal{O}_S = \mathcal{O}_{\mathbb{K},S} = \{\alpha \in \mathbb{K} : |\alpha|_v \leq 1 \text{ for places } v \notin S\}$$

contienne $\lambda_1, \ldots, \lambda_m, \alpha_1, \ldots, \alpha_m, \alpha_1^{-1}, \ldots, \alpha_m^{-1}$. Pour chaque $n \in \mathbb{Z}$, soit $A(n) = \lambda_1\alpha_1^n + \cdots + \lambda_m\alpha_m^n \in \mathcal{O}_S$. Skolem a suggéré [Skolem 1937]:

Conjecture 0.0.13 (Principe local-global exponentiel) *Supposons que pour chaque ideal non-nul* $\mathfrak{a}$ *de l'anneau* $\mathcal{O}_S$*, il existe* $n \in \mathbb{Z}$ *tel que* $A(n) \equiv 0 \bmod \mathfrak{a}$*. Alors il existe* $n \in \mathbb{Z}$ *tel que* $A(n) = 0$*.*

Soit Γ le groupe multiplicatif engendré par $\alpha_1, \ldots, \alpha_m$. Alors Γ est le produit d'un groupe abélien fini et d'un groupe libre de rang fini. Nous démontrons que cette conjecture est vraie lorsque le rang de Γ est un.

Ce résultat a été démontré en collaboration avec Florian Luca, de l'université de Witwatersrand (Afrique du Sud) et Yuri Bilu. Il a été publié dans Acta Arithmetica [Bartolomé *et al.* 2013]. Juste aprs sa publication, Florian Luca a rencontré Andrzej Schinzel à un congrès mathématique, et Schinzel lui a dit que notre résultat était une conséquence directe de [Schinzel 1977][Theorem 6]. Une vérification rapide a montré que c'était bien vrai. Cependant, ce travail a été mené sans aucune connaissance préalable de ce résultat et en utilisant d'autres méthodes (le théorème du sous-espace et l'inégalité de Baker).

Après cela, (au chapitre 4), nous généralisons un résultat précédent de Mourad Abouzaid ([Abouzaid 2008]). Soit $F(X, Y) \in \mathbb{Q}[X, Y]$ un $\mathbb{Q}$-polynôme irréductible. En 1929, Skolem [Skolem 1929] a démontré le beau théorème suivant:

Theorem 0.0.14 (Skolem) *Supposons que*

$$F(0, 0) = 0.$$

Alors, pour tout entier non-nul d*, l'équation n'adment qu'un nombre fini de solutions entières* $(X, Y) \in \mathbb{Z}^2$ *telles que* $\gcd(X, Y) = d$*.*

En 2008, Mourad Abouzaid [Abouzaid 2008] a généralisé ce résultat en travaillant avec des entiers algébriques arbitraires et en obtenant une relation asymptotique entre les hauteurs des coordonnées et leur PGCD logarithmique. Il a démontré le théorème suivant:

Theorem 0.0.15 (Abouzaid) *Supposons que* $(0,0)$ *soit un point non-singulier de la courbe plane* $F(X,Y)=0$. *Soit* $m=\deg_X F, \quad n=\deg_Y F, \quad M=\max\{m,n\}$. *Soit* ε *tel que* $0<\varepsilon<1$. *Alors, pour toute solution* $(\alpha,\beta)\in\bar{\mathbb{Q}}^2$ *de* $F(X,Y)=0$, *nous avons soit*

$$\max\{\mathrm{h}(\alpha),\mathrm{h}(\beta)\}\leq 56M^8\varepsilon^{-2}\mathrm{h}_p(F)+420M^{10}\varepsilon^{-2}\log(4M),$$

soit

$$\max\{|\mathrm{h}(\alpha)-n\mathrm{lgcd}(\alpha,\beta)|,|\mathrm{h}(\beta)-m\mathrm{lgcd}(\alpha,\beta)|\}\leq \varepsilon\max\{\mathrm{h}(\alpha),\mathrm{h}(\beta)\}+742M^7\varepsilon^{-1}\mathrm{h}_p(F)+5762M^9\varepsilon^{-1}\log(2m+2n).$$

Cependant, il a imposé la condition que $(0,0)$ soit un point non-singulier de la courbe plane $F(X,Y)=0$. En utilisant des versions quelque peu différentes du lemme “absolu" de Siegel et du Lemme d'Eisenstein, nous avons pu lever la condition et démontrer le théorème de façon générale. Nous démontrons le théorème suivant:

Theorem 0.0.16 *Soit* $F(X,Y)\in\bar{\mathbb{Q}}[X,Y]$ *un polynôme absolument irréductible qui satisfasse* $F(0,0)=0$. *Soit* $m=\deg_X F, \quad n=\deg_Y F$ *et* $r=\min\left\{i+j:\frac{\partial^{i+j}F}{\partial^i X\partial^j Y}(0,0)\neq 0\right\}$. *Soit* ε *tel que* $0<\varepsilon<1$. *Alors, pour tout* $\alpha,\beta\in\bar{\mathbb{Q}}$ *tel que* $F(\alpha,\beta)=0$, *nous avons soit*

$$\mathrm{h}(\alpha)\leq 200\varepsilon^{-2}mn^6(\mathrm{h}_p(F)+5)$$

ou

$$\left|\tfrac{\mathrm{lgcd}(\alpha,\beta)}{r}-\tfrac{\mathrm{h}(\alpha)}{n}\right|\leq\tfrac{1}{r}(\varepsilon\mathrm{h}(\alpha)+4000\varepsilon^{-1}n^4(\mathrm{h}_p(F) \quad + \quad \log(mn)+1)+ \quad + \quad 30n^2m(\mathrm{h}_p(F)+\log(nm))).$$

Dans notre démonstration nous suivons de près les méthodes de Mourad Abouzaid. Ce résultat a aussi été soumis pour publication en 2014, et pendant le processus de revue, l'arbitre nous a *gentiment* indiqué que ce résultat avait déjà été démontré dans la thèse de doctorat, jamais publiÃ©e, de Philipp Habegger. Une vérification rapide a aussi démontré que cela était vrai: le résultat est démontré à [Habegger 2007][Appendix B, Theorem B.3] en utilisant sa quasi-équivalence des hauteurs. Alors que nous admettons que sa solution est plus "industrielle" et donne une meilleure borne, nous croyons cependant que l'argument initial de Mourad Abouzaid est plus naturel et propose quelque éclairage supplémentaire sur ce qui se passe. Notre résultat a été publié à [Bartolomé 2015]. Notre principal outil sont les développements en séries de Puiseux.

Ensuite (au chapitre 5) nous donnons un aperçu des outils que nous avons utilisés dans les corps cyclotomiques. Nous tentons de développer une approche systématique pour un certain genre d'équations diophantiennes. Nous proposons quelques résultats sur les corps cyclotomiques, les anneaux de groupe et les sommes de Jacobi, qui nous seront utiles pour ensuite décrire l'approche.

Finalement (au chapitre 6) nous développons une application de l'approche précédemment expliquée. Nous considérerons l'équation diophantienne

$$X^n-1=BZ^n, \tag{5}$$

où $B\in\mathbb{Z}$ est un paramètre. Definissons $\varphi^*(B):=\varphi(\mathrm{rad}\,(B))$, où rad (B) est le radical de B, et supposons que

$$(n,\varphi^*(B))=1. \tag{6}$$

où rad (B) est le radical de B. Pour $B \in \mathbb{N}_{>1}$ fixé, soit

$$\mathcal{N}(B) = \{n \in \mathbb{N}_{>1} \mid \exists\ k > 0 \text{ tel que } n | \varphi^*(B)^k\}.$$

Si p est un premier impair, nous appellerons CF les conditions combinées

I La conjecture de Vandiver est vraie pour p, c'est-à-dire que le nombre de classe h_p^+ du sous-corps réel maximal du corps cyclotomique $\mathbb{Q}[\zeta_p]$, n'est pas divisible par p.

II Nous avons $i_r(p) < \sqrt{p} - 1$, en d'autre mots, il y a au plus $\sqrt{p} - 1$ entiers impairs $k < p$ tels que le nombre de Bernouilli $B_k \equiv 0 \bmod p$.

Les résultats actuels sur (5) sont restreints aux valeurs de B composées du produit de deux premiers petits $p \leq 13$ [Bennett *et al.* 2006] et de solutions complètes pour $B < 235$ ([A.Bazso *et al.* 2010]). Si nous pensons que l'équation n'a pas de solutions , – avec l'exception potentielle de quelques exemples isolés – il est naturel de considérer le cas où l'exposant n est premier. Bien sûr, l'existence de solutions (X, Z) pour n composé implique l'existence de quelques solutions pour n premier, en élevant X, Z à une puissance.

La contribution principale de notre travail a été de trouver un lien entre (5) lorsque n est premier et que (6) est vérifié, à l'équation diagonale de Nagell – Ljunggren,

$$\frac{X^n - 1}{X - 1} = n^e Y^n, \quad e = \begin{cases} 0 & \text{si } X \not\equiv 1 \bmod n, \\ 1 & \text{sinon.} \end{cases}$$

Ainsi, nous pouvons appliquer des résultats de [Mihăilescu 2008] et démontrer le théorème suivant:

Theorem 0.0.17 *Soit n un nombre premier et $B > 1$ un entier tel que $(\varphi^*(B), n) = 1$. Supposons que l'équation (5) admette une solution entière non-triviale, différente de $n = 3$ et $(X, Z; B) = (18, 7; 17)$. Soit $X \equiv u \bmod n$, $0 \leq u < n$ et $e = 1$ si $u = 1$ et $e = 0$ sinon. Alors:*

1. *$n > 163 \cdot 10^6$.*
2. *$X - 1 = \pm B/n^e$ et $B < n^n$.*
3. *Si $u \notin \{-1, 0, 1\}$, alors la condition CF (II) n'est pas vérifiée pour n et*

$$\begin{array}{rcl} 2^{n-1} & \equiv & 3^{n-1} \equiv 1 \mod n^2, \quad \text{et} \\ r^{n-1} & \equiv & 1 \mod n^2 \quad \text{pour tout } r | X(X^2 - 1). \end{array}$$

Si $u \in \{-1, 0, 1\}$, alors la condition CF (I) n'est pas vérifiée pour n.

Sur la base de ce théorème, nous démontrons ensuite:

Theorem 0.0.18 *Si l'équation (5) admet une solution pour B fixé vérifiant les conditions (6), alors, soit $n \in \mathcal{N}(B)$, ou bien il y a un nombre premier p, premier avec $\varphi^*(B)$ et un $m \in \mathcal{N}(B)$ tels que $n = p \cdot m$. De plus X^m, Y^m sont une solution de (5) pour l'exposant premier p et donc vérifient les conditions du théorème 0.0.17.*

Cela est une amélioration très considérable par rapport aux résultats actuellement connus.

Comme nous utilisons de façon intensive l'article [Mihăilescu 2008], nous avons rajouté en annexe des résultats nouveaux qui permettent de justifier pleinement les résultats annoncés en [Mihăilescu 2008][Theorem 3].

Mots clef

Equations diophantiennes, corps cyclotomiques, equations de Nagell-Ljunggren, Skolem, Abouzaid, equations diophantiennes exponentielles, inégalité de Baker, théorème du sous-espace.

Contents

Chapter 1
Introduction

1.1 Diophantine equations

A Diophantine equation (named after one of the first mathematicians to have introduced symbols into algebra, Diophantus, AD 250) is an equation in two or more variables in which only the integer solutions are sought. This thesis addresses the study of Diophantine equations. Obviously, the term *integer* depends on the field we are working on. The most famous Diophantine equation is:

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

Fermat's Last Theorem (FLT) states that equation (1.1) does not have integer solutions (in $\mathbb{Z}$), all different from zero, if $n \geq 3$. Pierre De Fermat stated this conjecture around 1660 while working on problem 8 of *Aritmetica*, Diophantus' book, where he wrote

> Cuius rei demonstrationem mirabilem sane detexi hanc marginis exiguitas non caperet
>
> (I have discovered a really marvelous proof of this statement, which this margin is too narrow to contain)

This conjecture has fueled the development of mathematics in several directions (algebraic number theory, analytic number theory, Diophantine approximation, Diophantine geometry, algebraic geometry) for 350 years. Among the mathematicians having contributed to the proof of this conjecture we can name Leonhard Euler, Carl Friedrich Gauß, Sophie Germain, Ernst Kummer, Yutaka Taniyama, Goro Shimura, Gerhard Frey, Jean-Pierre Serre, Ken Ribet, Barry Mazur, Andrew Wiles, Richard Taylor. The general steps of the proof are:

A. Work on odd prime exponents n.

B. If x, y, z is a non-trivial solution to Fermat's Last Equation, where x, y, z are relatively primes, then associate to it a Frey-Hellegouarch elliptic curve. In 1986, Ken Ribet [Ribet 1990] proved Jean-Pierre Serre's ε conjecture that the Frey-Hellegouarch curve cannot be parametrized with modular forms.

C. In 1994, Andrew Wiles [Wiles 1995] proved the Shimura-Taniyama-Weil conjecture that any elliptic curve can be parametrized with modular forms. Thus a contradiction arouse in the case of Fermat's Last Equation.

Another famous Diophantine equation is:

$$x^n - y^m = 1 \tag{1.2}$$

Eugène Catalan conjectured in 1842 [Catalan 1842] that equation (1.2) admits only one solution in non-zero integers (that is, $3^2 - 2^3 = 1$). Some mathematicians having contributed to the proof of this conjecture are Victor Amédée Lebesgue, Trygve Nagell, Sigmund Selberg, Kustaa Inkeri, Seppo Hyyrö, Ko Chao, J. W. S. Cassels, Yann Bugeaud, Guillaume Hanrot, Maurice Mignotte, Preda Mihăilescu. It was Preda Mihăilescu [Mihăilescu 2004] who finished the proof, using a cyclotomic approach, in 2001.

1.2 Structure of the thesis

In **Chapter 2**, we propose a quick introduction to the methods of Diophantine approximation we have used in this research work. We remind the notion of height and introduce the logarithmic *gcd*.

In **Chapter 3**, we address a conjecture, made by Thoralf Skolem in 1937, on an exponential Diophantine equation. For this conjecture, let $\mathbb{K}$ be a number field, $\alpha_1, \ldots, \alpha_m$ as well as $\lambda_1, \ldots, \lambda_m$ be non-zero elements in $\mathbb{K}$, and S a finite set of places of $\mathbb{K}$ (containing all the infinite places) such that the ring of S-integers

$$\mathcal{O}_S = \mathcal{O}_{\mathbb{K},S} = \{\alpha \in \mathbb{K} : |\alpha|_v \leq 1 \text{ for places } v \notin S\}$$

contains $\lambda_1, \ldots, \lambda_m, \alpha_1, \ldots, \alpha_m, \alpha_1^{-1}, \ldots, \alpha_m^{-1}$. For every $n \in \mathbb{Z}$, let $A(n) = \lambda_1\alpha_1^n + \cdots + \lambda_m\alpha_m^n \in \mathcal{O}_S$. Skolem suggested [Skolem 1937]:

Conjecture 1.2.1 (Exponential Local-Global Principle) *Assume that for every non zero ideal $\mathfrak{a}$ of the ring $\mathcal{O}_S$, there exists $n \in \mathbb{Z}$ such that $A(n) \equiv 0 \bmod \mathfrak{a}$. Then there exists $n \in \mathbb{Z}$ such that $A(n) = 0$.*

Let Γ be the multiplicative group generated by $\alpha_1, \ldots, \alpha_m$. Then Γ is the product of a finite abelian group and a free abelian group of finite rank. In chapter 3, we prove that the conjecture is true when the rank of Γ is one.

This result was proved in collaboration with Florian Luca, from University of the Witwatersrand (South Africa) and Yuri Bilu. It was published in Acta Arithmetica [Bartolomé *et al.* 2013]. Shortly after its publication, Florian Luca met Andrzej Schinzel in a mathematical congress, and Schinzel told him that our result was a direct consequence of [Schinzel 1977][Theorem 6]:

Theorem 1.2.2 *Let α_{hij}, β_{hi} be non-zero elements of a number field $\mathbb{K}$, D a positive integer. If the system of congruences*

$$\prod_{h=1}^{g_i}\left(\prod_{j=1}^{k}\alpha_{hij}^{x_j} - \beta_{hi}\right) \equiv 0 \mod \mathfrak{m}\ (i = 1, 2, \cdots, l)$$

is soluble for all moduli $\mathfrak{m}$ prime to D, then the corresponding system of equations is soluble in integers.

Andrej Schnizel used Tchebotarev's theorem to prove Theorem 1.2.2, whereas we did not use it. Our work has been done with no previous knowledge of this result and using other (subspace theorem and Baker's inequality), interesting per se, methods.

In **Chapter 4**, we generalize a previous result by Abouzaid ([Abouzaid 2008]). Let $F(X, Y) \in \mathbb{Q}[X, Y]$ be a $\mathbb{Q}$-irreducible polynomial. In 1929 Skolem [Skolem 1929] proved the following beautiful theorem:

Theorem 1.2.3 (Skolem) *Assume that $F(0, 0) = 0$. Then for every non-zero integer d, the equation $F(X, Y) = 0$ has only finitely many solutions in integers $(X, Y) \in \mathbb{Z}^2$ with $\gcd(X, Y) = d$.*

In 2008, Abouzaid [Abouzaid 2008] generalized this result by working with arbitrary algebraic numbers and by obtaining an asymptotic relation between the heights of the coordinates and their logarithmic gcd. He proved the following theorem:

Theorem 1.2.4 (Abouzaid) *Assume that $(0,0)$ is a non-singular point of the plane curve $F(X,Y)=0$. Let $m=\deg_X F,\quad n=\deg_Y F,\quad M=\max\{m,n\}$. Let ε satisfy $0<\varepsilon<1$. Then for any solution $(\alpha,\beta)\in\bar{\mathbb{Q}}^2$ of $F(X,Y)=0$, we have either*

$$\max\{\mathrm{h}(\alpha),\mathrm{h}(\beta)\}\leq 56M^8\varepsilon^{-2}\mathrm{h}_p(F)+420M^{10}\varepsilon^{-2}\log(4M),$$

or

$$\max\{|\mathrm{h}(\alpha)-n\mathrm{lgcd}(\alpha,\beta)|,|\mathrm{h}(\beta)-m\mathrm{lgcd}(\alpha,\beta)|\}\leq \varepsilon\max\{\mathrm{h}(\alpha),\mathrm{h}(\beta)\}+742M^7\varepsilon^{-1}\mathrm{h}_p(F)+5762M^9\varepsilon^{-1}\log(2m+2n).$$

However, he imposed the condition that $(0,0)$ be a non-singular point of the plane curve $F(X,Y)=0$. Using a somewhat different version of Siegel's "absolute" Lemma and of Eisenstein's Lemma, we could remove the condition and prove it in full generality. We prove the following theorem:

Theorem 1.2.5 *Let $F(X,Y)\in\bar{\mathbb{Q}}[X,Y]$ be an absolutely irreducible polynomial satisfying $F(0,0)=0$. Let $m=\deg_X F,\quad n=\deg_Y F$ and $r=\min\left\{i+j:\frac{\partial^{i+j}F}{\partial^i X\partial^j Y}(0,0)\neq 0\right\}$. Let ε satisfy $0<\varepsilon<1$. Then, for any $\alpha,\beta\in\bar{\mathbb{Q}}$ such that $F(\alpha,\beta)=0$, we have either:*

$$\mathrm{h}(\alpha)\leq 200\varepsilon^{-2}mn^6(\mathrm{h}_p(F)+5)$$

or

$$\left|\frac{\mathrm{lgcd}(\alpha,\beta)}{r}-\frac{\mathrm{h}(\alpha)}{n}\right|\leq\frac{1}{r}(\varepsilon\mathrm{h}(\alpha)+4000\varepsilon^{-1}n^4(\mathrm{h}_p(F)+\log(mn)+1)+30n^2m(\mathrm{h}_p(F)+\log(nm))).$$

In our proof, we closely follow Abouzaid's methods. This result was also submitted for publication in 2014, and during the review process, the referee *kindly* pointed out that this result had already been proven in Philipp Habegger's unpublished PhD thesis; a quick check proved that to be true: the result is proved in [Habegger 2007][Appendix B, Theorem B.3] using his quantitative version of the quasi-equivalence of heights. Philipp Habegger's theorem is:

Theorem 1.2.6 *Let $P\in\bar{\mathbb{Q}}[X,Y]$ be irreducible with $n=\deg_X P>0$, $m=\deg_Y P>0$ and $d=\deg P$. If $P(x,y)=0$ where x and y are non-zero algebraic numbers, then*

$$\max\left\{\left|\mathrm{lgcd}(x,y)-\frac{e(P)}{m}\mathrm{h}(x)\right|,\left|\mathrm{lgcd}(x,y)-\frac{e(P)}{n}\mathrm{h}(y)\right|\right\}\leq 183d\max\{d,\mathrm{h}_p(P)\}^{1/2}\max\{1,\mathrm{h}(x),\mathrm{h}(y)\}.$$

To prove this theorem, he uses his explicit version of quasi-equivalence of heights:

Theorem 1.2.7 *Let $P\in\bar{\mathbb{Q}}[X,Y]$ be irreducible with $n=\deg_X P>0$, $m=\deg_Y P>0$ and $d=\deg P$. If $P(x,y)=0$ where $x,y\in\bar{\mathbb{Q}}$, then*

$$\left|\frac{\mathrm{h}(x)}{m}-\frac{\mathrm{h}(y)}{n}\right|\leq 51\max\{n,m,\mathrm{h}_p(P)\}^{1/2}\max\{1,\mathrm{h}(x),\mathrm{h}(y)\}^{1/2}.$$

Philipp Habegger used his sharp quantitative method of the quasi-equivalence of heights to prove Theorem 1.2.6, while the main ingredient of our proof are Puiseux expansions and we closely follow Abouzaid's arguments. While we admit that Philipp Habegger's solution is more "industrial" and provides a better bound, we still believe that Abouzaid's inital argument is quite enlightening and natural in certain ways. Our result has been published in [Bartolomé 2015].

The next chapter (**Chapter 5**) describes our approach to a certain type of exponential Diophantine equations:

$$\frac{x^p - y^p}{(x-y)^f} = B.z^q \text{ with } x, y \in \mathbb{Z},\ B \in \mathbb{Z},\ f \in \{0,1\},\ (p,q) \in \mathbb{Z}^2.$$

We start by giving an overview of some of the tools we have used: we give some basic properties of cyclotomic extensions, group-rings and Jacobi sums, and of general binomial series developments. Then we describe our approach in three main steps. Finally, we show how this approach has been specialized in two specific cases: the proof of Catalan's conjecture, as well as some conditions and bounding of the potential solutions of the diagonal Nagell-Ljunggren equation:

$$\frac{x^p - 1}{x - 1} = p^e \cdot y^p \quad \text{with} \quad x, y \in \mathbb{Z} \quad e \in \{0,1\}.$$

Chapter 6 shows a very interesting application of the approach developed in the previous chapter. There, we consider the Diophantine equation

$$X^n - 1 = BZ^n, \tag{1.3}$$

where $B \in \mathbb{Z}$ is understood as a parameter. Define $\varphi^*(B) := \varphi(\text{rad }(B))$, where rad (B) is the radical of B, and assume that

$$(n, \varphi^*(B)) = 1. \tag{1.4}$$

where rad (B) is the radical of B. For a fixed $B \in \mathbb{N}_{>1}$ we let

$$\mathcal{N}(B) = \{n \in \mathbb{N}_{>1} \mid \exists\ k > 0 \text{ such that } n | \varphi^*(B)^k\}.$$

If p is an odd prime, we shall denote by CF the combined condition requiring that

I The Vandiver Conjecture holds for p, so the class number h_p^+ of the maximal real subfield of the cyclotomic field $\mathbb{Q}[\zeta_p]$ is not divisible by p.

II We have $i_r(p) < \sqrt{p} - 1$, in other words, there is at most $\sqrt{p} - 1$ odd integers $k < p$ such that the Bernoulli number $B_k \equiv 0 \bmod p$.

Current results on Equation (1.3) are restricted to values of B which are built up from two small primes $p \leq 13$ [Bennett *et al.* 2006] and complete solutions for $B < 235$ ([A.Bazso *et al.* 2010]). If expecting that the equation has no solutions, – possibly with the exception of some isolated examples – it is natural to consider the case when the exponent n is a prime. Of course, the existence of solutions (X, Z) for composite n imply the existence of some solutions with n prime, by raising X, Z to a power.

The main contribution of this chapter is to relate (1.3) in the case when n is a prime and (1.4) holds, to the diagonal Nagell – Ljunggren equation,

$$\frac{X^n-1}{X-1} = n^e Y^n, \quad e = \begin{cases} 0 & \text{if } X \not\equiv 1 \bmod n, \\ 1 & \text{otherwise.} \end{cases}$$

This way, we can apply results from [Mihăilescu 2008] and prove the following:

Theorem 1.2.8 *Let n be a prime and $B > 1$ an integer with $(\varphi^*(B), n) = 1$. Suppose that the equation (1.3) has a non trivial integer solution different from $n = 3$ and $(X, Z; B) = (18, 7; 17)$. Let $X \equiv u \bmod n, 0 \leq u < n$ and $e = 1$ if $u = 1$ and $e = 0$ otherwise. Then:*

1. *$n > 163 \cdot 10^6$.*
2. *$X - 1 = \pm B/n^e$ and $B < n^n$.*
3. *If $u \notin \{-1, 0, 1\}$, then condition CF (II) fails for n and*

$$\begin{array}{rcl} 2^{n-1} \equiv 3^{n-1} \equiv 1 \mod n^2, & & \text{and} \\ r^{n-1} \equiv 1 \mod n^2 & & \text{for all } r | X(X^2-1). \end{array}$$

If $u \in \{-1, 0, 1\}$, then Condition CF (I) fails for n.

Based on this theorem, we prove the following:

Theorem 1.2.9 *If equation (1.3) has a solution for a fixed B verifying the conditions (1.4), then either $n \in \mathcal{N}(B)$ or there is a prime p coprime to $\varphi^*(B)$ and a $m \in \mathcal{N}(B)$ such that $n = p \cdot m$. Moreover X^m, Y^m are a solution of (1.3) for the prime exponent p and thus verify the conditions in Theorem 1.2.8.*

This is a strong improvement of the currently known results.

Chapter 2
Diophantine approximation

In this chapter we introduce some notions, definitions and properties of Diophantine analysis we will use in Chapters 3 and 4. We introduce heights and logarithmic GCD.

2.1 Diophantine analysis

Diophantine analysis, in its most classical form, studies integral and rational points on algebraic varieties over number fields. One can speak on several aspects of this study:

- the finiteness aspect, or, more generally, the non-density aspect: proving that, under sufficiently general assumptions, integral or rational points are finite in number or, in higher dimension, are not Zariski dense;
- the counting aspect: when there are infinitely many integral (or rational) points, give upper bounds or even asymptotics for their counting functions;
- the existence aspect: decide whether at least one integral point exists;
- the effectiveness aspect: determine, at least in principle, all integral points (say, give an explicit upper bound for their heights);
- the algorithmic aspect: give a practical method permitting to determine integral points, using computers.

Of course, this classification is very rough and incomplete, but it gives some initial idea on the subject.

The finiteness/density and the counting aspects are most well developed. The finiteness aspect in dimension 1 is almost completely solved by the classical theorems of Siegel [Siegel 1929] and Faltings [Faltings 1983]: there are finitely many integral points on affine curves of genus at least 1 (or even of genus 0 but with at least 3 points at infinity), and finitely many rational points on projective curves of genus at least 2. In higher dimension much less is known, but some substantial progress has been made in the last decade in the work of Corvaja, Zannier, Levin and Autissier, starting from the pioneering articles of Corvaja and Zannier of 2002 [Corvaja & Zannier 2002] and 2004 [Corvaja & Zannier 2004].

The counting aspect is well advanced too, and is presented by seminal works of Tchinkel, Pila, Heath-Brown and many others.

The existence aspect is much less elaborated. The celebrated result of Matiyasevich [Matiyasevich 1970] states that on affine varieties of sufficiently high dimension the existence problem for integral points is not decidable. It is believed, however, that it is decidable in low dimensions, most notably, in dimension 1. While decidability of existence of an integral/rational point on a general affine/projective curve is still an open problem, some results in this direction are obtained, and most of them are based on so-called *effective methods* in Diophantine Analysis.

The above-mentioned general finiteness theorems of Siegel and Faltings are non-effective in the sense that none of them implies any explicit bound for the height of the points.

Partial effectivization of Siegel's theorem is obtained by Baker's method based on Baker's theory of logarithmic forms. Another effective method in Diophantine analysis is Runge's method, which is elementary, but remarkably efficient when it applies, and which was used, for instance, by Preda Mihăilescu in the course of his proof of Catalan's conjecture [Mihăilescu 2004]. Both these methods, when they apply, give explicit upper bounds for the heights of integral points on certain affine algebraic curves. In the most basic form, they explicitly bound solutions of certain polynomial Diophantine equations. In particular, these results imply that, in principle, one can determine all the solutions just by enumerating all possible integers below the bound. Unfortunately, the bound is usually too high for this to be practical, and if one wants to solve completely the equation in question, one should apply special reduction and enumeration techniques.

2.2 Heights and logarithmic gcd

In this section we recall definitions and collect various results about absolute values and heights.

We normalize the absolute values on number fields so that they extend standard absolute values on $\mathbb{Q}$: if $v \mid p$ (non-Archimedean) then $|p|_v = p^{-1}$ and if $v \mid \infty$ (Archimedean) then $|2015|_v = 2015$.

2.2.1 Heights and lgcd of algebraic numbers

Let $\mathbb{K}$ be a number field, $d = [\mathbb{K} : \mathbb{Q}]$ and $d_v = [\mathbb{K}_v : \mathbb{Q}_v]$. The *height* of an algebraic number $\alpha \in \mathbb{K}$ is defined as

$$\mathrm{h}(\alpha) = \frac{1}{d} \sum_{v \in M_{\mathbb{K}}} d_v \log^+ |\alpha|_v.$$

where $M_{\mathbb{K}}$ is the set of places (normalized absolute values) of the number field $\mathbb{K}$ and $\log^+ = \max\{\log, 0\}$. It is well-known that the height does not depend on the particular choice of $\mathbb{K}$, but only on the number α itself. It is equally well-known that $\mathrm{h}(\alpha) = \mathrm{h}(\alpha^{-1})$, so that

$$\mathrm{h}(\alpha) = \frac{1}{d} \sum_{v \in M_{\mathbb{K}}} -d_v \log^- |\alpha|_v = \sum_{v \in M_{\mathbb{K}}} \mathrm{h}_v(\alpha),$$

where $\log^- = \min\{\log, 0\}$ and

$$\mathrm{h}_v(\alpha) = -\frac{d_v}{d} \log^- |\alpha|_v.$$

The quantities $\mathrm{h}_v(\alpha)$ can be viewed as "local heights". Clearly, $\mathrm{h}_v(\alpha) \geq 0$ for any v and α.

We define the *logarithmic gcd* of two algebraic numbers α and β, not both 0, as

$$\mathrm{lgcd}(\alpha, \beta) = \sum_{v \in M_K} \min\{\mathrm{h}_v(\alpha), \mathrm{h}_v(\beta)\},$$

where $\mathbb{K}$ is a number field containing both α and β. It again depends only α and β, not on $\mathbb{K}$. A simple verification shows that for $\alpha, \beta \in \mathbb{Z}$ we have $\mathrm{lgcd}(\alpha, \beta) = \log \gcd(\alpha, \beta)$.

Now let $\mathbb{K}$ be a number field and S be a set of places of $\mathbb{K}$. We define the *S-height* by

$$\mathrm{h}_S(\alpha) = \sum_{v \in S} \mathrm{h}_v(\alpha).$$

Similarly we define lgcd_S. We shall use the inequality $\mathrm{lgcd}_S(\alpha, \beta) \leq \mathrm{h}_S(\alpha) \leq \mathrm{h}(\alpha)$ without special reference.

2.2.2 Affine and projective heights of polynomials

We define the projective and the affine height of a vector $\underline{a} = (a_1, \ldots, a_m) \in \bar{\mathbb{Q}}^m$ with algebraic entries, by

$$\mathrm{h}_p(\underline{a}) = \frac{1}{d} \sum_{v \in M_{\mathbb{K}}} d_v \log \max_{1 \leq k \leq m} |a_k|_v \qquad (\underline{a} \neq \underline{0}),$$

$$\mathrm{h}_a(\underline{a}) = \frac{1}{d} \sum_{v \in M_{\mathbb{K}}} d_v \log^+ \max_{1 \leq k \leq m} |a_k|_v,$$

where $\mathbb{K}$ is a number field containing $a_1, \ldots, a_m$. Here d, d_v are defined as in the previous subsection. We notice that the height of an algebraic number defined in the previous subsection corresponds to the affine height of a one-dimensional vector.

We define the projective and affine height of a polynomial as the corresponding heights of the vector of its non-zero coefficients. If F is a non-zero polynomial, then, for $\alpha \in \bar{\mathbb{Q}}^*$ we have $\mathrm{h}_p(\alpha F) = \mathrm{h}_p(F)$. Also, $\mathrm{h}_p(F) \leq \mathrm{h}_a(F)$, with $\mathrm{h}_p(F) = \mathrm{h}_a(F)$ if F has a coefficient equal to 1.

In [Schmidt 1990, Lemma 4], Schmidt proves the following lemma:

Lemma 2.2.1 *Let $F(X, Y) \in \bar{\mathbb{Q}}[X, Y]$ be a polynomial with algebraic coefficients, such that $m = \deg_X F$ and $n = \deg_Y F$. Let $R_F(X) = Res_Y(F, F'_Y)$ be the resultant of F and its derivative polynomial with respect to Y. Then:*

$$\mathrm{h}_p(R_F) \leq (2n-1)\mathrm{h}_p(F) + (2n-1)\log((m+1)(n+1)\sqrt{n}). \tag{2.1}$$

It is well-known that the height of a root of a polynomial is bounded in terms of the height of the polynomial itself. The following lemma can be found in the article [Bilu & Borichev 2013, Proposition 3.6]:

Lemma 2.2.2 *Let $F(X)$ be a polynomial of degree m with algebraic coefficients. Let α be a root of F. Then, $\mathrm{h}(\alpha) \leq \mathrm{h}_p(F) + \log 2$*

We want to generalize this to a system of two algebraic equations in two variables.

Lemma 2.2.3 *Let $F_1(X, Y)$ and $F_2(X, Y)$ be polynomials with algebraic coefficients, having no common factor. Put:*

$$m_i = \deg_X F_i, \ n_i = \deg_Y F_i \qquad (i = 1, 2).$$

Let α, β be algebraic numbers satisfying $F_1(\alpha, \beta) = F_2(\alpha, \beta) = 0$. Then

$$\mathrm{h}(\alpha) \leq n_1 \mathrm{h}_p(F_2) + n_2 \mathrm{h}_p(F_1) + (m_1 n_2 + m_2 n_1) + (n_1 + n_2)\log(n_1 + n_2) + \log 2.$$

Proof Since F_1 and F_2 have no common factor, their Y-resultant $R(X)$ is a non-zero polynomial, and $R(\alpha) = 0$. [Abouzaid 2008, Proposition 2.4] gives the estimate

$$\mathrm{h}_p(R) \leq n_1 \mathrm{h}_p(F_2) + n_2 \mathrm{h}_p(F_1) + (m_1 n_2 + m_2 n_1) + (n_1 + n_2)\log(n_1 + n_2).$$

Combining this with Lemma 2.2.2, the result follows. □

We will also use [Abouzaid 2008, Proposition 2.5]:

Lemma 2.2.4 *Let $F(X,Y) \in \bar{\mathbb{Q}}[X,Y]$ be a polynomial with $m = \deg_X F$ and $n = \deg_Y F$ and let α, β be two algebraic numbers. Then*

A. We have $\mathrm{h}(F(\alpha,\beta)) \leq \mathrm{h}_a(F) + m\mathrm{h}(\alpha) + n\mathrm{h}(\beta) + \log((m+1)(n+1))$.

B. If $F(\alpha,\beta) = 0$ with $F(\alpha, Y)$ not vanishing identically, then:

$$\mathrm{h}(\beta) \leq \mathrm{h}_p(F) + m\mathrm{h}(\alpha) + n + \log(m+1).$$

Proposition 2.2.5 *We let S be a set of places of the number field $\mathbb{K}$, and $\neg S$ be the complement of S in the set of all places of $\mathbb{K}$.*

A. For non-zero algebraic numbers α, β, γ we have

$$\mathrm{lgcd}(\alpha\beta, \gamma) \leq \mathrm{lgcd}(\alpha, \gamma) + \mathrm{lgcd}(\beta, \gamma),$$

and similarly for lgcd_S.

In the sequel $\mathbb{K}$ is a number field, S a set of places of $\mathbb{K}$ containing the infinite places, and α, β, γ belong to the ring $\mathcal{O}_S$ of S-integers.

B. α and β are co-prime in $\mathcal{O}_S$ if and only if $\mathrm{lgcd}_{\neg S}(\alpha, \beta) = 0$.

C. If α and β are co-prime in $\mathcal{O}_S$ then

$$\mathrm{lgcd}_{\neg S}(\alpha\beta, \gamma) = \mathrm{lgcd}_{\neg S}(\alpha, \gamma) + \mathrm{lgcd}_{\neg S}(\beta, \gamma).$$

D. We have $\mathrm{lgcd}_{\neg S}(\alpha, \beta) \leq \mathrm{h}_{\neg S}(\alpha)$, with equality exactly when α divides β in $\mathcal{O}_S$.

Chapter 3
On the Exponential Local-Global Principle

3.1 Abstract

Skolem conjectured that the "power sum" $A(n) = \lambda_1\alpha_1^n + \cdots + \lambda_m\alpha_m^n$ satisfies a certain local-global principle. We prove this conjecture in the case when the multiplicative group generated by $\alpha_1, \cdots, \alpha_m$ is of rank 1.

3.2 Introduction

Let $\mathbb{K}$ be a number field, $\alpha_1, \ldots, \alpha_m, \lambda_1, \ldots, \lambda_m$ non-zero elements in $\mathbb{K}$, and S a finite set of places of $\mathbb{K}$ (containing all the infinite places) such that the ring of S-integers

$$\mathcal{O}_S = \mathcal{O}_{\mathbb{K},S} = \{\alpha \in \mathbb{K} : |\alpha|_v \leq 1 \text{ for places } v \notin S\}$$

contains $\lambda_1, \ldots, \lambda_m, \alpha_1, \ldots, \alpha_m, \alpha_1^{-1}, \ldots, \alpha_m^{-1}$. Then, for every $n \in \mathbb{Z}$

$$A(n) = \lambda_1\alpha_1^n + \cdots + \lambda_m\alpha_m^n \in \mathcal{O}_S.$$

The expression $A(n)$ will be called *power sum*. The following conjecture was suggested by Skolem [Skolem 1937].

Conjecture 3.2.1 (Exponential Local-Global Principle) *Assume that for every non zero ideal $\mathfrak{a}$ of the ring $\mathcal{O}_S$, there exists $n \in \mathbb{Z}$ such that $A(n) \equiv 0 \bmod \mathfrak{a}$. Then there exists $n \in \mathbb{Z}$ such that $A(n) = 0$.*

Some particular cases of this conjecture all addressing the instance when $m = 2$ and $\{A(n)\}_{n\geq 0} \subseteq \mathbb{Z}$, have been dealt with in [Broughan & Luca 2010, Schinzel 1975, Schinzel 1977, Schinzel 2003]. For some results on the analogous Skolem conjecture over function fields, see [Sun 2011].

In this chapter, we prove this conjecture in a special case. Let Γ be the multiplicative group generated by $\alpha_1, \ldots, \alpha_m$. Then Γ is the product of a finite abelian group and a free abelian group of finite rank, say ρ. In this case we shall call $A(n)$ a *power sum of rank ρ*.

Theorem 3.2.2 *Conjecture 3.2.1 holds for power sums of rank one.*

Surprisingly enough, our proof makes no use of the Tchebotarev theorem, usually an indispensable ingredient in this kind of arguments. Instead, it relies on two "powerful tools" from the Diophantine Approximations. One is the celebrated Subspace Theorem of Schmidt-Schlickewei, which is used through a theorem of Corvaja and Zannier (Theorem 3.3.1). The other tool is Baker's inequality (Theorem 3.4.5).

3.3 Using the subspace theorem through a Theorem of Corvaja and Zannier

In this section we state one theorem of Corvaja and Zannier and obtain a consequence of this theorem, which will be one of our principal tools.

We remind here the result of Corvaja and Zannier [Corvaja & Zannier 2005, page 204, Corollary 1]:

Theorem 3.3.1 *Let Γ be a finitely generated subgroup of $\bar{\mathbb{Q}}^\times$, and $\varepsilon > 0$. Then for multiplicatively independent $\alpha, \beta \in \Gamma$ we have*

$$\lgcd(\alpha - 1, \beta - 1) \leq \varepsilon \max\{\mathrm{h}(\alpha), \mathrm{h}(\beta)\} + O(1),$$

where the constant implied by $O(1)$ depends on Γ and ε (but not on α and β).

We shall use it through the following statement.

Corollary 3.3.2 *Let $\mathbb{K}$ be a number field, S a finite subset of $M_{\mathbb{K}}$ containing the infinite places, $\beta, \gamma \in \mathcal{O}_S^\times$ multiplicatively independent, and $\varepsilon > 0$. Then for $k, n \in \mathbb{Z}$ we have*

$$\lgcd_{\neg S}(\gamma^k - 1, \gamma^n - \beta) \leq \varepsilon |k| + O(1),$$

where the implied constant depends on γ, β, $\mathbb{K}$, S and ε, but not on k and n.

Proof Replacing, if necessary, γ by γ^{-1}, we may assume that $k > 0$. Also, since $n \equiv n' \mod k$ implies the congruence $\gamma^n \equiv \gamma^{n'} \bmod (\gamma^k - 1)$ in the ring $\mathcal{O}_S$, we may assume that $0 \leq n < k$. Applying Theorem 3.3.1 with $\Gamma = \langle \gamma, \beta \rangle$, with γ^k as α and with $\gamma^n \beta^{-1}$ as β, we obtain

$$\begin{aligned} \lgcd_{\neg S}(\gamma^k - 1, \gamma^n - \beta) &\leq \lgcd_{\neg S}(\gamma^k - 1, \gamma^n \beta^{-1} - 1) \\ &\leq \lgcd(\gamma^k - 1, \gamma^n \beta^{-1} - 1) \\ &\leq \varepsilon\big(k\mathrm{h}(\gamma) + \mathrm{h}(\beta)\big) + O(1) \\ &= \varepsilon \mathrm{h}(\gamma) k + O(1). \end{aligned}$$

Redefining ε, we obtain the result. □

3.4 Cyclotomic polynomials

In this section we establish properties of the cyclotomic polynomials, needed for the proof. We denote by $\Phi_k(T)$ the k-th cyclotomic polynomial. Since $T^k - 1 = \prod_{d|k} \Phi_d(T)$, we have

$$\Phi_k(T) = \prod_{d|k} (T^d - 1)^{\mu(k/d)}, \tag{3.1}$$

where μ is the Möbius function. We shall systematically use this in the sequel.

3.4.1 Divisibility

All the results of this subsection are well-known, but it is easier to supply quick proofs than to find references.

Proposition 3.4.1 *Let k and ℓ be distinct positive integers. Then the resultant of $\Phi_k(T)$ and $\Phi_\ell(T)$ divides (in $\mathbb{Z}$) a power of $k\ell$.*

Proof The resultant of these polynomials is a product of factors of the type $\zeta_k - \zeta_\ell$, where ζ_k (respectively, ζ_ℓ) is a primitive k-th (respectively, ℓ-th) root of unity. The elementary theory of cyclotomic fields (see, for instance, [Washington 1997, Chapters 1 and 2]) implies that $\zeta_k - \zeta_\ell$ divides $k\ell$ in the ring $\mathbb{Z}[\zeta_{k\ell}]$. Hence the resultant divides a power of $k\ell$ in $\mathbb{Z}[\zeta_{k\ell}]$. Since $\mathbb{Q} \cap \mathbb{Z}[\zeta_{k\ell}] = \mathbb{Z}$, the resultant divides the same power of $k\ell$ in $\mathbb{Z}$. □

Corollary 3.4.2 *Let $\mathbb{K}, S$ be like in the Introduction, and k, ℓ like in Proposition 3.4.1.*

A. Assume that S contains the places dividing $k\ell$. Then for any $\gamma \in \mathcal{O}_S$ we have $\gcd\big(\Phi_k(\gamma), \Phi_\ell(\gamma)\big) = 1$ in the ring $\mathcal{O}_S$; that is, no prime ideal of $\mathcal{O}_S$ divides both $\Phi_k(\gamma)$ and $\Phi_\ell(\gamma)$.

B. Assume that S contains the places dividing $k\ell$ and $k \nmid \ell$. Then for any $\gamma \in \mathcal{O}_S$ we have $\gcd\big(\Phi_k(\gamma), \gamma^\ell - 1\big) = 1$ in the ring $\mathcal{O}_S$.

C. Assume that S contains the places dividing k. For $\gamma \in \mathcal{O}_S^\times$ let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_S$ dividing $\Phi_k(\gamma)$. Then γ is of exact order k in $(\mathcal{O}_S/\mathfrak{p})^\times$. In particular, if for some $n \in \mathbb{Z}$ we have $\gamma^n \equiv 1 \bmod \mathfrak{p}$ then $k \mid n$.

Proof Part A is immediate from Proposition 3.4.1. For part B observe that $\gamma^\ell - 1$ is a product of factors of the type $\Phi_{\ell'}(\gamma)$ with $\ell' \mid \ell$, and by the assumption none of these ℓ' is equal to k. Hence part B follows from part A. Finally, part C follows immediately from part B. □

3.4.2 Heights and cyclotomic polynomials

We need an asymptotic expression for the height of the algebraic number $\Phi_k(\gamma)$, in terms of $\mathrm{h}(\gamma)$ and k. In general, if $f(x)$ is a polynomial with algebraic coefficients, then, using basic properties of heights, it is not difficult to show that $\mathrm{h}(f(\gamma)) = \deg f \mathrm{h}(\gamma) + O(1)$ as f is fixed and γ is varying. We, however, need a result of different type: find asymptotics for $\mathrm{h}\big(\Phi_k(\gamma)\big)$ as γ is fixed, but k is growing.

For a positive integer k we denote by $\varphi(k)$ the Euler function and by $\omega(k)$ the number of distinct prime divisors of k.

Proposition 3.4.3 *Let γ be an algebraic number. Then*

$$\big|\mathrm{h}\big(\Phi_k(\gamma)\big) - \varphi(k)\mathrm{h}(\gamma)\big| \le 2^{\omega(k)}(\log k + O(1)),$$

where the constant implied by $O(1)$ depends on γ (but not on k).

The proof requires a complex analytic lemma.

Lemma 3.4.4 *For a positive integer k we have*

$$\max_{|z|\le 1} \log|\Phi_k(z)| \le 2^{\omega(k)}(\log k + O(1)),$$

the maximum being over the unit disc on the complex plane, and the implied constant being absolute.

Proof By the maximum principle, it suffices to show that

$$\log|\Phi_k(z)| \le 2^{\omega(k)}(\log k + O(1)). \tag{3.2}$$

for a complex z with $|z| = 1$. Thus, fix such z. We can write it in a unique way as $z = \zeta e^{2\pi i\theta/k}$, where ζ is a k-th root of unity (not necessarily primitive) and $-1/2 < \theta \le 1/2$. Let ℓ be the exact order of ζ; thus, ℓ is a divisor of k and ζ is a primitive ℓ-th root of unity. Let d be any other divisor of k. If $\ell \nmid d$ then $2 \ge |z^d - 1| \ge 2\sin(\pi d/2k)$, which implies that

$$\big|\log|z^d - 1|\big| \le \log k + O(1). \tag{3.3}$$

And if $\ell \mid d$ then, we have $|z^d - 1| = 2\sin(\pi\theta d/k)$. Writing $d = d'\ell$, this implies that

$$\log|z^{d'\ell} - 1| = \log d' - \log(k/\ell\theta) + O(1). \tag{3.4}$$

Identity (3.1) implies that

$$\begin{aligned}\log|\Phi_k(z)| &= \sum_{d\mid k} \mu(k/d)\log|z^d - 1| \\ &= \sum_{d\mid k,\ell\nmid d} \mu(k/d)\log|z^d - 1| + \sum_{d'\mid k/\ell} \mu\big((k/\ell)/d'\big)\log|z^{\ell d'} - 1|.\end{aligned}$$

Notice that the first sum above has at most $2^{\omega(k)} - 1$ non-zero summands. Now substituting here (3.3) and (3.4), we obtain

$$\begin{aligned}\log|\Phi_k(z)| &\le (2^{\omega(k)} - 1)(\log k + O(1)) + \sum_{d'\mid k/\ell} \mu\left(\frac{k/\ell}{d'}\right)\left(\log d' - \log\left(\frac{k}{\ell\theta}\right)\right) \\ &= (2^{\omega(k)} - 1)(\log k + O(1)) + \Lambda(k/\ell) - \delta\log(k/\ell\theta),\end{aligned}$$

where $\Lambda(\cdot)$ is the von Mangoldt function, $\delta = 0$ if $\ell < k$ and $\delta = 1$ if $\ell = k$. In any case we obtain (3.2), proving the lemma. □

Proof of Proposition 3.4.3 Fix a number field $\mathbb{K}$ containing γ. For a finite place v of $\mathbb{K}$ we, obviously, have

$$\log^+|\Phi_k(\gamma)|_v = \begin{cases}\varphi(k)\log|\gamma|_v, & |\gamma|_v > 1,\\ 1, & |\gamma|_v \le 1.\end{cases} \tag{3.5}$$

For infinite places we have similar "approximate" statements

$$\log^+|\Phi_k(\gamma)|_v \begin{cases}= \varphi(k)\log|\gamma|_v + O(2^{\omega(k)}), & |\gamma|_v > 1,\\ \le 2^{\omega(k)}(\log k + O(1)), & |\gamma|_v \le 1.\end{cases} \tag{3.6}$$

The second inequality follows from Lemma 3.4.4. To prove the first one, assume that $|\gamma|_v > 1$. Then for $n \geq 1$ we have $\log|\gamma^n - 1|_v = n \log|\gamma|_v + O(1)$. Using (3.1) we find

$$\begin{aligned}\log|\Phi_k(\gamma)|_v &= \sum_{d|k} \mu(k/d) \log|\gamma^d - 1|_v \\ &= \log|\gamma|_v \sum_{d|k} d\mu(k/d) + O(2^{\omega(k)}) \\ &= \varphi(k) \log|\gamma|_v + O(2^{\omega(k)}),\end{aligned}$$

as wanted.

The (in)equalities (3.5) and (3.6) imply that

$$\left|\log^+|\Phi_k(\gamma)|_v - \varphi(k)\log^+|\gamma|_v\right| \begin{cases} = 0, & v \text{ finite}, \\ \leq 2^{\omega(k)}(\log k + O(1)), & v \text{ infinite}. \end{cases}$$

Summing this up over $v \in M_{\mathbb{K}}$, we obtain the result. □

3.4.3 Using Baker's Inequality

Besides Theorem 3.3.1 of Corvaja and Zannier, our second principal tool is the celebrated inequality of Baker, see the first two contributions in [Wüstholz 2002].

Theorem 3.4.5 *Let $\gamma_1, \dots, \gamma_r$ be non-zero algebraic numbers, and v a place of a number field containing them. Then for any $n_1, \dots, n_r \in \mathbb{Z}$ we have either $\gamma_1^{n_1} \cdots \gamma_r^{n_r} = 1$ or*

$$\left|\gamma_1^{n_1} \cdots \gamma_r^{n_r} - 1\right|_v \geq e^{-C \log N}, \qquad N = \max\{2, n_1, \dots, n_r\},$$

where C is a positive constant depending on $\gamma_1, \dots, \gamma_r$ and v, but not on $n_1, \dots, n_r$.

We deduce from it the following property of cyclotomic polynomials, inspired by the work of Schinzel [Schinzel 1974] and Stewart [Stewart 1977].

Proposition 3.4.6 *Let $\mathbb{K}$ be a number field, S a finite set of places of $\mathbb{K}$, and $\gamma \in \mathbb{K}$ not a root of unity. Then for any integer $k > 1$ we have*

$$\mathrm{h}_S\big(\Phi_k(\gamma)\big) = O(2^{\omega(k)} \log k),$$

where the implied constant depends on $\mathbb{K}$, S and γ, but not on k.

Proof Since the set S is finite, it suffices to prove that for any $v \in M_{\mathbb{K}}$ we have

$$\mathrm{h}_v\big(\Phi_k(\gamma)\big) = O(2^{\omega(k)} \log k),$$

where here and below the constants implied by $O(\cdot)$ depend only on γ and v. Equivalently, we have to show that

$$\left|\log^-|\Phi_k(\gamma)|_v\right| = O(2^{\omega(k)} \log k). \tag{3.7}$$

If $|\gamma|_v > 1$ then $\log|\Phi_k(\gamma)|_v = \varphi(k)\log|\gamma|_v + O(2^{\omega(k)})$, see the proof of Proposition 3.4.3. It follows that $\log^-|\Phi_k(\gamma)|_v = O(2^{\omega(k)})$, better than (3.7).

Now assume that $|\gamma|_v \le 1$. Using Theorem 3.4.5 with $r = 1$, we obtain that $|\gamma^n - 1|_v \ge e^{-C \log n}$ with $C > 0$ depending on γ and v. Hence

$$\log 2 \ge \log |\gamma^n - 1|_v \ge -C \log n,$$

which implies that $\left|\log |\gamma^n - 1|_v\right| = O(\log n)$. Using (3.1), we obtain

$$\log |\Phi_k(\gamma)|_v = \sum_{d|k} \mu(k/d) \log |\gamma^d - 1|_v = O(2^{\omega(k)} \log k),$$

which proves (3.7). □

Combining Propositions 3.4.3 and 3.4.6, we obtain the following consequence.

Corollary 3.4.7 *In the set-up of Proposition 3.4.6 we have*

$$\mathrm{h}_{\neg S}(\Phi_k(\gamma)) = \varphi(k)\mathrm{h}(\gamma) + O(2^{\omega(k)} \log k).$$

3.5 Proof of Theorem 3.2.2

Let $A(n) = \lambda_1 \alpha_1^n + \cdots + \lambda_m \alpha_m^n$ be a power sum of rank 1. Assume that

(L) for every non-zero ideal $\mathfrak{a}$ of the ring $\mathcal{O}_S$ there exists $n \in \mathbb{Z}$ such that $A(n) \equiv 0 \bmod \mathfrak{a}$.

We want to prove that

(G) there exists $n \in \mathbb{Z}$ such that $A(n) = 0$.

3.5.1 General Observations

We start with some general observations, which hold true for any power sum, not just power sums of rank 1.

Extension of the set of places We may replace the set S by any bigger (finite) set of places. Indeed, condition (G) does not depend on S, and condition (L) becomes weaker when S is replaced by a bigger set. In particular, extending the set S, we may assume that

$$\lambda_1, \ldots, \lambda_m \in \mathcal{O}_S^\times. \tag{3.8}$$

Extension of the base field We may replace the field $\mathbb{K}$ by a finite extension $\mathbb{K}'$, the set S being replaced by the set of places S' of $\mathbb{K}'$ extending those from $\mathbb{K}$. Condition (G) is again not concerned, and condition (L) is replaced by an equivalent one (each ideal of $\mathcal{O}_{\mathbb{K}',S'}$ is contained in an ideal coming from $\mathcal{O}_{\mathbb{K},S}$).

The group Γ is torsion-free We may assume that the group Γ, generated by the "roots" $\alpha_1, \ldots, \alpha_m$ is torsion-free. Indeed, since it is finitely generated, its torsion subgroup is finite; denote its order by μ. Then the group $\Gamma^\mu = \{x^\mu : x \in \Gamma\}$ is torsion free. Now consider instead of $A(n)$ the power sum

$$\widetilde{A}(n) = A(\mu n)A(\mu n + 1) \cdots A(\mu n + \mu - 1) = \tilde{\lambda}_1 \tilde{\alpha}_1^n + \cdots + \tilde{\lambda}_{\tilde{m}} \tilde{\alpha}_{\tilde{m}}^n.$$

Clearly, each of the conditions (L) and (G) holds simultaneously for $A(n)$ and $\widetilde{A}(n)$, and the group generated by $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_m$ is contained in Γ^μ, a torsion-free group. Hence we may replace $A(n)$ by $\widetilde{A}(n)$ and assume in the sequel that Γ is torsion-free.

3.5.2 Using the Rank 1 Assumption

Now we use the assumption that the rank of Γ is 1. Since we may assume it is torsion-free, this means that $\Gamma = \langle\gamma\rangle$, where $\gamma \in \mathbb{K}^\times$ is not a root of unity. Write $\alpha_j = \gamma^{\nu_j}$ with $\nu_j \in \mathbb{Z}$. Assuming that $\nu_1 < \nu_2 < \cdots < \nu_m$, we write

$$A(n) = \lambda_m \gamma^{\nu_1 n} P(\gamma^n),$$

where

$$P(T) = T^{\nu_m - \nu_1} + \frac{\lambda_{m-1}}{\lambda_m} T^{\nu_{m-1} - \nu_1} + \cdots + \frac{\lambda_2}{\lambda_m} T^{\nu_2 - \nu_1} + \frac{\lambda_1}{\lambda_m} \in \mathbb{K}[T].$$

Extending the field $\mathbb{K}$, we may assume that it contains all the roots of the polynomial $P(T)$. It follows from (3.8) that

$$\text{the roots of } P(T) \text{ are } S\text{-units.} \tag{3.9}$$

Condition (G) is equivalent to saying that one of the roots of $P(T)$ belongs to Γ. Thus, we assume from now on that

$$\text{no root of } P(T) \text{ belongs to } \Gamma, \tag{3.10}$$

and we shall find a non-zero ideal $\mathfrak{a}$ of $\mathcal{O}_S$ such that $P(\gamma^n) \not\equiv 0 \bmod \mathfrak{a}$ for any $n \in \mathbb{Z}$. This will prove the theorem, since $A(n)$ is equal to $P(\gamma^n)$ times an S-unit.

3.5.3 The Ideal $\mathfrak{a}$

We are going now to define the ideal $\mathfrak{a}$. First of all, we split the polynomial $P(T)$ into two factors: $P(T) = P_{\text{ind}}(T)P_{\text{dep}}(T)$, such that each of the roots of $P_{\text{ind}}(T)$ is multiplicatively independent of γ, and those of $P_{\text{dep}}(T)$ are multiplicatively dependent with γ. Fix a positive integer q such that $\beta^q \in \Gamma$ for every root β of $P_{\text{dep}}(T)$. Then for every such β we have $\beta^q = \gamma^r$, where $r = r(\beta) \in \mathbb{Z}$. Further, fix a prime number p, not dividing q and such that

$$r(\beta) \not\equiv r(\beta') \bmod p \tag{3.11}$$

for any roots β, β' of $P_{\text{dep}}(T)$ such that $r(\beta) \neq r(\beta')$. Extending the set S we may assume that

$$\text{all places dividing } pq \text{ belong to } S. \tag{3.12}$$

Assumption (3.12) has one implication that will be crucial in the sequel.

Observation *Let ζ_μ be a primitive μ-th root of unity for some $\mu \mid pq$. Then ζ_μ is of exact order μ modulo $\mathfrak{p}$ for any prime ideal $\mathfrak{p}$ of $\mathcal{O}_S$.*

Indeed, if this is not true, then $\mathfrak{p} \mid \zeta_{\mu'} - 1$ for some $\mu' \mid \mu$, $\mu' > 1$, which implies that $\mathfrak{p} \mid \mu$, contradicting (3.12).

We let $\mathfrak{a}$ be the principal ideal generated by $a = \Phi_{p^\ell}(\gamma)\Phi_{p^\ell q}(\gamma)$, where Φ_k denotes the k-th cyclotomic polynomial and the positive integer ℓ will be specified later. We will show that both $P_{\text{ind}}(\gamma^n)$ and $P_{\text{dep}}(\gamma^n)$ have "small" common divisor with $\mathfrak{a}$. This will imply that, when ℓ is chosen suitably, $P(\gamma^n)$ cannot be divisible by $\mathfrak{a}$ for any n.

Until the end of the proof the constants implied by $O(\cdot)$ may depend on the polynomial $P(T)$, on γ, on p and q, and on the parameter ε introduced below, but they do not depend on ℓ nor on n.

We claim the following.

Claim I Fix $\varepsilon > 0$. Then for any $n \in \mathbb{Z}$ we have

$$\operatorname{lgcd}_{\neg S}\big(P_{\mathrm{ind}}(\gamma^n), a\big) \le \varepsilon p^\ell + O(1).$$

Claim D *Let n be a rational integer. Then in the ring $\mathcal{O}_S$ we have either* $\gcd\big(P_{\mathrm{dep}}(\gamma^n), \Phi_{p^\ell}(\gamma)\big) = 1$ *or* $\gcd\big(P_{\mathrm{dep}}(\gamma^n), \Phi_{p^\ell q}(\gamma)\big) = 1$.

We postpone the proof of the Claims until later, and show now how they imply the theorem.

3.5.4 Proof of the Theorem (Assuming the Claims)

Assuming the Claims, we will show now that when the parameter ℓ is chosen large enough, we have $P(\gamma^n) \not\equiv 0 \bmod \mathfrak{a}$ for any $n \in \mathbb{Z}$.

Thus, assume that for some n we have $P(\gamma^n) \equiv 0 \bmod \mathfrak{a}$. In other words, both $\Phi_{p^\ell}(\gamma)$ and $\Phi_{p^\ell q}(\gamma)$ divide $P(\gamma^n)$ in the ring $\mathcal{O}_S$. In addition to this, Corollary 3.4.2:A together with (3.12) implies that they are co-prime in $\mathcal{O}_S$. It follows that

$$\begin{aligned}\operatorname{lgcd}_{\neg S}(P(\gamma^n), a) &= \operatorname{lgcd}_{\neg S}\big(P(\gamma^n), \Phi_{p^\ell}(\gamma)\big) + \operatorname{lgcd}_{\neg S}\big(P(\gamma^n), \Phi_{p^\ell q}(\gamma)\big)\\ &= \mathrm{h}_{\neg S}\big(\Phi_{p^\ell}(\gamma)\big) + \mathrm{h}_{\neg S}\big(\Phi_{p^\ell q}(\gamma)\big)\\ &= \varphi(p^\ell)\mathrm{h}(\gamma) + \varphi(p^\ell q)\mathrm{h}(\gamma) + O(\ell), \qquad (3.13)\end{aligned}$$

see Corollary 3.4.7.

On the other hand, Claim D implies that

$$\operatorname{lgcd}_{\neg S}(P_{\mathrm{dep}}(\gamma^n), a) \le \max\big\{\mathrm{h}_{\neg S}\big(\Phi_{p^\ell}(\gamma)\big), \mathrm{h}_{\neg S}\big(\Phi_{p^\ell q}(\gamma)\big)\big\} = \varphi(p^\ell q)\mathrm{h}(\gamma) + O(\ell),$$

again by Corollary 3.4.7. Combining this with Claim I, we obtain

$$\operatorname{lgcd}_{\neg S}(P(\gamma^n), a) \le \varepsilon p^\ell + \varphi(p^\ell q)\mathrm{h}(\gamma) + O(\ell). \qquad (3.14)$$

Now select ε to have $\varepsilon < (1 - p^{-1})\mathrm{h}(\gamma)$. Then (3.13) and (3.14) become contradictory for large ℓ. This proves the theorem. □

3.5.5 Proof of Claim I

Clearly, $a \mid \gamma^{p^\ell q} - 1$. Corollary 3.3.2 implies that

$$\operatorname{lgcd}_{\neg S}(\gamma^n - \beta, a) \le \operatorname{lgcd}_{\neg S}\big(\gamma^n - \beta, \gamma^{p^\ell q} - 1\big) \le \varepsilon p^\ell q + O(1).$$

Hence

$$\operatorname{lgcd}_{\neg S}\big(P_{\mathrm{ind}}(\gamma^n), a\big) \le \varepsilon p^\ell q \deg P_{\mathrm{ind}} + O(1).$$

Redefining ε, we obtain the result. □

3.5.6 Proof of Claim D

Let us assume the contrary and let $\mathfrak{p}, \mathfrak{p}'$ be prime ideals of $\mathcal{O}_S$ such that $\mathfrak{p}$ divides $\gcd\big(P_{\mathrm{dep}}(\gamma^n), \Phi_{p^\ell}(\gamma)\big)$ and $\mathfrak{p}'$ divides $\gcd\big(P_{\mathrm{dep}}(\gamma^n), \Phi_{p^\ell q}(\gamma)\big)$. There exist (not necessarily distinct) roots β, β' of $P_{\mathrm{dep}}(T)$ such that

$$\gamma^n \equiv \beta \bmod \mathfrak{p}, \qquad \gamma^n \equiv \beta' \bmod \mathfrak{p}'.$$

Further, let $r \in \mathbb{Z}$ be such that $\beta^q = \gamma^r$, see the beginning of Subsection 3.5.3. Then $\gamma^{qn-r} \equiv 1 \bmod \mathfrak{p}$.

On the other hand, Corollary 3.4.2:C implies that for any root β of $P_{\mathrm{ind}}(T)$ we have

$$\gamma \text{ is of exact order } p^\ell \text{ in } (\mathcal{O}_S/\mathfrak{p})^\times. \tag{3.15}$$

In particular, $qn \equiv r \bmod p^\ell$. Similarly, if $r' \in \mathbb{Z}$ is such that $(\beta')^q = \gamma^{r'}$ then Corollary 3.4.2:C implies that $qn \equiv r' \bmod p^\ell q$. We obtain the congruence $r \equiv r' \bmod p$, which, by our choice of p (see (3.11)) implies that $r = r'$. Thus, we have $qn \equiv r \bmod p^\ell q$, which implies that $q \mid r$. It follows that $\beta = \zeta\gamma^\nu$ with $\nu \in \mathbb{Z}$ and ζ a q-th root of unity, not necessarily primitive.

Now it is the time to use our basic assumption (3.10). We obtain that $\beta \notin \Gamma$, which means that $\zeta \neq 1$. Thus, $\zeta = \zeta_\mu$ is a primitive μ-th root of unity with $\mu \mid q$ and $\mu > 1$.

Since $\zeta_\mu \equiv \gamma^{n-\nu} \bmod \mathfrak{p}$, the image of ζ_μ in $(\mathcal{O}_S/\mathfrak{p})^\times$ belongs to the subgroup generated by the image of γ. Hence the order of ζ_μ modulo $\mathfrak{p}$ divides the order of γ. But the order of ζ_μ is μ, see the "Observation" in Subsection 3.5.3, and the order of γ is p^ℓ, see (3.15). Thus, $\mu \mid p^\ell$, which contradicts co-primarity of p and q. This proves the claim. □

Chapter 4
Skolem-Abouzaid's theorem in the singular case

4.1 Abstract

Let $F(X,Y) \in \mathbb{Q}[X,Y]$ be a $\mathbb{Q}$-irreducible polynomial. In 1929, Skolem ([Skolem 1929]) proved a result allowing explicit bounding of the solutions of $F(X,Y) = 0$ such that $\gcd(X,Y) = d$ in terms of the coefficients of F and d. In 2008, Abouzaid [Abouzaid 2008] generalized this result by working with arbitrary algebraic numbers and by obtaining an asymptotic relation between the heights of the coordinates and their logarithmic gcd. However, he imposed the condition that $(0,0)$ be a non-singular point of the plane curve $F(X,Y) = 0$. In this chapter, this constraint is removed.

4.2 Introduction

Let $F(X,Y) \in \mathbb{Q}[X,Y]$ be a $\mathbb{Q}$-irreducible polynomial. In 1929 Skolem [Skolem 1929] proved the following beautiful theorem:

Theorem 4.2.1 (Skolem) *Assume that*

$$F(0,0) = 0. \tag{4.1}$$

Then for every non-zero integer d, the equation $F(X,Y) = 0$ has only finitely many solutions in integers $(X,Y) \in \mathbb{Z}^2$ with $\gcd(X,Y) = d$.

The same year, Siegel obtained his celebrated finiteness theorem for integral solutions of Diophantine equations: equation $F(X,Y) = 0$ has finitely many solutions in integers unless the corresponding plane curve is of genus 0 and has at most 2 points at infinity. While Siegel's result is, certainly, deeper and more powerful than Theorem 4.2.1, the latter has one important advantage. Siegel's theorem is known to be non-effective: it does not give any bound for the size of integral solutions. On the contrary, Skolem's method allows one to bound the solutions explicitly in terms of the coefficients of the polynomial F and the integer d. Indeed, such a bound was obtained by Walsh [Walsh 1992]; see also [Poulakis 2004].

In 2008, Abouzaid [Abouzaid 2008] gave a far-going generalization of Skolem's theorem. He extended it in two directions.

First, he studied solutions not only in rational integers, but in arbitrary algebraic numbers. To accomplish this, he introduced the notion of *logarithmic gcd* of two algebraic numbers α and β, which coincides with the logarithm of the usual gcd when $\alpha, \beta \in \mathbb{Z}$.

Second, he not only bounded the solution in terms of the logarithmic gcd, but obtained a sort of asymptotic relation between the heights of the coordinates and their logarithmic gcd.

Let us state Abouzaid's principal result (see [Abouzaid 2008, Theorem 1.3]). In the sequel we assume that $F(X,Y) \in \bar{\mathbb{Q}}[X,Y]$ is an absolutely irreducible polynomial, and use the notation

$$m = \deg_X F, \quad n = \deg_Y F, \quad M = \max\{m,n\}. \tag{4.2}$$

We denote by $\mathrm{h}(\alpha)$ the absolute logarithmic height of $\alpha \in \bar{\mathbb{Q}}$ and by $\mathrm{lgcd}(\alpha, \beta)$ the logarithmic gcd of $\alpha, \beta \in \mathbb{Q}$. We also denote by $\mathrm{h}_p(F)$ the projective height of the polynomial F. For all definitions, see Subsection 2.2.1.

Theorem 4.2.2 (Abouzaid) *Assume that $(0,0)$ is a non-singular point of the plane curve $F(X,Y) = 0$. Let ε satisfy $0 < \varepsilon < 1$. Then for any solution $(\alpha, \beta) \in \bar{\mathbb{Q}}^2$ of $F(X,Y) = 0$, we have either*

$$\max\{\mathrm{h}(\alpha), \mathrm{h}(\beta)\} \leq 56M^8\varepsilon^{-2}\mathrm{h}_p(F) + 420M^{10}\varepsilon^{-2}\log(4M),$$

or

$$\max\{|\mathrm{h}(\alpha) - n\mathrm{lgcd}(\alpha,\beta)|, |\mathrm{h}(\beta) - m\mathrm{lgcd}(\alpha,\beta)|\} \leq \varepsilon \max\{\mathrm{h}(\alpha), \mathrm{h}(\beta)\} + 742M^7\varepsilon^{-1}\mathrm{h}_p(F) + 5762M^9\varepsilon^{-1}\log(2m+2n).$$

Informally speaking,

$$\frac{\mathrm{h}(\alpha)}{n} \sim \frac{\mathrm{h}(\beta)}{m} \sim \mathrm{lgcd}(\alpha, \beta) \tag{4.3}$$

as $\max\{\mathrm{h}(\alpha), \mathrm{h}(\beta)\} \to \infty$.

Unfortunately, Abouzaid's assumption is slightly more restrictive than Skolem's (4.1): he assumes not only that the point $(0,0)$ belongs to the plane curve $F(X,Y) = 0$, but also that $(0,0)$ is a non-singular point on this curve.

Denote by r the "order of vanishing" of $F(X,Y)$ at the point $(0,0)$:

$$r = \min\left\{ i + j : \frac{\partial^{i+j} F}{\partial^i X \partial^j Y}(0,0) \neq 0 \right\}. \tag{4.4}$$

Clearly, $r > 0$ if and only if $F(0,0) = 0$ and $r = 1$ if and only $(0,0)$ is a non-singular point of the plane curve $F(X,Y) = 0$.

We can now state our principal result.

Theorem 4.2.3 *Let $F(X,Y) \in \bar{\mathbb{Q}}[X,Y]$ be an absolutely irreducible polynomial satisfying $F(0,0) = 0$. Let ε satisfy $0 < \varepsilon < 1$. Then, for any $\alpha, \beta \in \bar{\mathbb{Q}}$ such that $F(\alpha,\beta) = 0$, we have either:*

$$\mathrm{h}(\alpha) \leq 200\varepsilon^{-2}mn^6(\mathrm{h}_p(F) + 5)$$

or

$$\left|\tfrac{\mathrm{lgcd}(\alpha,\beta)}{r} - \tfrac{\mathrm{h}(\alpha)}{n}\right| \leq \tfrac{1}{r}(\varepsilon\mathrm{h}(\alpha) + 4000\varepsilon^{-1}n^4(\mathrm{h}_p(F) + \log(mn) + 1) + 30n^2m(\mathrm{h}_p(F) + \log(nm))).$$

By symmetry, the same kind of bound holds true for the difference $\frac{\mathrm{lgcd}(\alpha,\beta)}{r} - \frac{\mathrm{h}(\beta)}{m}$. Informally speaking,

$$\frac{\mathrm{h}(\alpha)}{n} \sim \frac{\mathrm{h}(\beta)}{m} \sim \frac{\mathrm{lgcd}(\alpha,\beta)}{r} \tag{4.5}$$

as $\max\{\mathrm{h}(\alpha), \mathrm{h}(\beta)\} \to \infty$.

Validity of (4.5) was stated without proof by Abouzaid, see the end of Section 1 in [Abouzaid 2008] (Abouzaid's definition of r looks different, but it can be easily shown that it is equivalent to ours).

As indicated above, our argument follows, in principle, Abouzaid's pattern. However, we had to substantially refine his proof at certain points, to accommodate it for the more general set-up of Theorem 4.2.3. For instance, our Proposition 4.6.1 comparing the logarithmic gcd with certain "partial height" is considerably more involved than its prototype from [Abouzaid 2008].

Plan of the chapter Section 4.4 is preliminary: we compile therein some definitions and results from different sources, which will be used in the article. In Section 4.5 we establish the "Main Lemma", which is the heart of the proof of Theorem 4.2.3. In Section 4.6 we complete the proof of Theorem 4.2.3 using the "Main Lemma".

4.3 Heights

We remind that we normalize the absolute values on number fields so that they extend standard absolute values on $\mathbb{Q}$: if $v \mid p$ (non-Archimedean) then $|p|_v = p^{-1}$ and if $v \mid \infty$ (Archimedean) then $|2015|_v = 2015$.

4.3.1 Coefficients versus roots

In this subsection we establish some simple relations between coefficients and roots of a polynomial over a field with absolute value, needed in the proof of our main result. It will be convenient to use the notion of *v-Mahler measure* of a polynomial.

Let $\mathbb{K}$ be a field with absolute value v and $f(X) \in \mathbb{K}[X]$ a polynomial of degree n. Let $\beta_1, \dots, \beta_n \in \bar{\mathbb{K}}$ be the roots of f:

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \ldots + a_0 = a_n (X - \beta_1) \dots (X - \beta_n).$$

Define the v-Mahler measure of f by

$$M_v(f) = |a_n|_v \prod_{i=1}^{n} \max\{1, |\beta_i|_v\},$$

where we extend v somehow to $\bar{\mathbb{K}}$. (Clearly, $M_v(f)$ does not depend on the particular extension of v.) It is well-known that $|f|_v = M_v(f)$ for non-archimedean v ("Gauß lemma") and $M_v(f) \le (n+1)|f|_v$ for archimedean v (Mahler).

Lemma 4.3.1 *Let $\beta_1, \dots, \beta_{\ell+1}$ be $\ell + 1$ distinct roots of $f(X)$, where $0 \le \ell \le n-1$. Then*

$$\max\{|\beta_1|_v, \dots, |\beta_{\ell+1}|_v\} \ge c_v(n) \frac{|a_\ell|_v}{|f|_v},$$

where $c_v(n) = 1$ for non-archimedean v and $c_v(n) = (n+1)^{-1} 2^{-n}$ for archimedean v.

Proof We have

$$a_\ell = \pm a_n \sum_{1 \le i_1 < \ldots < i_{n-\ell} \le n} \beta_{i_1} \dots \beta_{i_{n-\ell}}, \tag{4.6}$$

where $\beta_1, \dots, \beta_n$ are all roots of $f(X)$ in $\bar{\mathbb{K}}$ counted with multiplicities. Observe that each term in the sum above contains one of the roots $\beta_1, \dots, \beta_{\ell+1}$, and the product of the other

roots together with a_n is v-bounded by $M_v(f)$. Hence, denoting $\mu = \max\{|\beta_1|_v, \ldots, |\beta_{\ell+1}|_v\}$, we obtain $|a_\ell|_v \le \mu M_v(f)$ in the non-archimedean case and $|a_\ell|_v \le \binom{n}{\ell}\mu M_v(f)$ in the archimedean case. Since $\binom{n}{\ell} \le 2^n$, the result follows. □

4.3.2 Siegel's "Absolute" Lemma

In this section we give a version of the Absolute Siegel's Lemma due to David and Philippon [David & Philippon 1999], adapted for our purposes.

We start from a slightly modified definition of the projective height of a non-zero vector $\underline{a} = (a_1, \ldots, a_n) \in \bar{\mathbb{Q}}^n$. As before, we fix a number field $\mathbb{K}$ containing $a_1, \ldots, a_n$ and set $d = [\mathbb{K} : \mathbb{Q}]$, $d_v = [\mathbb{K}_v : \mathbb{Q}_v]$ for $v \in M_{\mathbb{K}}$.

Now we define

$$\mathrm{h}_s(\underline{a}) = \sum_{v \in M_{\mathbb{K}}} \frac{d_v}{d} \log \|\underline{a}\|_v,$$

where

$$\|\underline{a}\|_v = \begin{cases} \max\{|a_1|_v, \ldots, |a_n|_v\}, & v < \infty, \\ (|a_1|_v^2 + \ldots + |a_n|_v^2)^{1/2}, & v \mid \infty. \end{cases}$$

This definition is the same as for $\mathrm{h}_p(\underline{a})$, except that for the archimedean places the sup-norm is replaced by the euclidean norm. We have clearly $\mathrm{h}_s(\lambda\underline{a}) = \mathrm{h}_s(\underline{a})$ for $\lambda \in \bar{\mathbb{Q}}^\times$, and

$$\mathrm{h}_p(\underline{a}) \le \mathrm{h}_s(\underline{a}) \le \mathrm{h}_p(\underline{a}) + \frac{1}{2}\log n. \tag{4.7}$$

Now let us define the height of a linear subspace of $\bar{\mathbb{Q}}^n$. If W is a 1-dimensional subspace of $\bar{\mathbb{Q}}^n$ then we set

$$\mathrm{h}_s(W) := \mathrm{h}_s(\underline{w}),$$

where $\underline{w}$ is an arbitrary non-zero vector from W. Clearly, $\mathrm{h}_s(W)$ does not depend on the particular choice of the vector $\underline{w}$.

To extend this to subspaces of arbitrary dimension, we use Grassmann spaces. Recall that the mth Grassmann space $\wedge^m\bar{\mathbb{Q}}^n$ is of dimension $\binom{n}{m}$, and has a standard basis consisting of the vectors

$$e_{i_1} \wedge \ldots \wedge e_{i_m}, \qquad (1 \le i_1 < \ldots < i_m \le n),$$

where $e_1, \ldots, e_n$ is the standard basis of $\bar{\mathbb{Q}}^n$. If W is an m-dimensional subspace of $\bar{\mathbb{Q}}^n$ then $\wedge^m W$ is a 1-dimensional subspace of $\wedge^m\bar{\mathbb{Q}}^n$, and we simply define

$$\mathrm{h}_s(W) := \mathrm{h}_s(\wedge^m W).$$

Finally, we set $\mathrm{h}_s(W) = 0$ for the zero subspace $W = \{\underline{0}\}$.

To make this more explicit, pick a basis $\underline{w}_1, \ldots, \underline{w}_m$ of W. Then $\wedge^m W$ is generated by $\underline{w}_1 \wedge \ldots \wedge \underline{w}_m$, and we have

$$\mathrm{h}_s(W) = \mathrm{h}_s(\underline{w}_1 \wedge \ldots \wedge \underline{w}_m). \tag{4.8}$$

This allows one to estimate the height of a subspace generated by a finite set of vectors in terms of heights of generators.

Proposition 4.3.2 *Let W be a subspace of $\bar{\mathbb{Q}}^n$ generated by vectors $\underline{w}_1, \ldots, \underline{w}_m \in \bar{\mathbb{Q}}^n$. Then*

$$\mathrm{h}_s(W) \le \mathrm{h}_s(\underline{w}_1) + \ldots + \mathrm{h}_s(\underline{w}_m).$$

Proof Selecting among $\underline{w}_1, \dots, \underline{w}_m$ a maximal linearly independent subset, we may assume that $\underline{w}_1, \dots, \underline{w}_m$ is a basis of W. Then we have (4.8). It remains to observe that for any place v we have

$$\|\underline{w}_1 \wedge \dots \wedge \underline{w}_m\|_v \le \|\underline{w}_1\|_v \dots \|\underline{w}_m\|_v.$$

For non-archimedean v this is obvious, and for archimedean v this is the classical Hadamard's inequality. □

We denote by $(\underline{x} \cdot \underline{y})$ the standard inner product on $\bar{\mathbb{Q}}^n$:

$$(\underline{x} \cdot \underline{y}) = x_1 y_1 + \dots + x_n y_n.$$

Let $W^\perp$ denote the orthogonal complement to W with respect to this product. It is well-known that the coordinates of $\wedge^m W$ (where $m = \dim W$) in the standard basis of $\wedge^m \bar{\mathbb{Q}}^n$ are the same (up to a scalar multiple) as the coordinates of $\wedge^{n-m} W^\perp$ in the standard basis of $\wedge^{n-m} \bar{\mathbb{Q}}^n$. In particular,

$$\mathrm{h}_s(W) = \mathrm{h}_s(W^\perp). \tag{4.9}$$

We use this to estimate the height of the subpace defined by a system of linear equations.

Proposition 4.3.3 *Let $L_1, \dots, L_m$ be non-zero linear forms on $\bar{\mathbb{Q}}^n$, and let W be the subspace of $\bar{\mathbb{Q}}^n$ defined by $L_1(\underline{x}) = \dots = L_m(\underline{x}) = 0$. Then*

$$\mathrm{h}_s(W) \le \mathrm{h}_p(L_1) + \dots + \mathrm{h}_p(L_m) + \frac{m}{2} \log n. \tag{4.10}$$

Proof Let $\underline{a}_1, \dots, \underline{a}_m$ be vectors in $\bar{\mathbb{Q}}^n$ such that $L_i(\underline{x}) = (\underline{x} \cdot \underline{a}_i)$. Then

$$\mathrm{h}_p(L_i) = \mathrm{h}_p(\underline{a}_i) \qquad (i = 1, \dots, m). \tag{4.11}$$

The space $W^\perp$ is generated by $\underline{a}_1, \dots, \underline{a}_m$. Applying to it Proposition 4.3.2 and using (4.7), we obtain

$$\mathrm{h}_s(W^\perp) \le \mathrm{h}_s(\underline{a}_1) + \dots + \mathrm{h}_s(\underline{a}_m) \le \mathrm{h}_p(\underline{a}_1) + \dots + \mathrm{h}_p(\underline{a}_m) + \frac{m}{2} \log n.$$

Together with (4.9) and (4.11), this gives (4.10). □

Remark 4.3.4 *It is not difficult to slightly refine (4.10), replacing $\log n$ by $\log m$ in the right-hand side, but this would not lead to any substantial improvement of our results.*

In [Bilu & Borichev 2013, Lemma 4.7] the following version of "absolute Siegel's lemma" is given.

Proposition 4.3.5 *Let W be an ℓ-dimensional subspace of $\bar{\mathbb{Q}}^n$ and $\varepsilon > 0$. Then, there is a non-zero vector $\underline{x} \in W$, satisfying:*

$$\mathrm{h}_p(\underline{x}) \le \frac{\mathrm{h}_s(W)}{\ell} + \frac{1}{2\ell} \sum_{i=1}^{\ell-1} \sum_{k=1}^{i} \frac{1}{k} + \varepsilon.$$

Corollary 4.3.6 *Let $L_1, \dots, L_m$ be non-zero linear forms in n variables with algebraic coefficients. Then, there exists a non-zero vector $\underline{x} \in \bar{\mathbb{Q}}^n$ such that $L_1(\underline{x}) = \dots = L_m(\underline{x}) = 0$ and*

$$\mathrm{h}_p(\underline{x}) \le \frac{1}{n-m} \left(\mathrm{h}_p(L_1) + \dots + \mathrm{h}_p(L_m) \right) + \frac{1}{2} \frac{n}{n-m} \log n. \tag{4.12}$$

Proof We apply Proposition 4.3.5, where W is the vector subspace of $\bar{\mathbb{Q}}^n$ defined by $L_1(\underline{x}) = \ldots = L_m(\underline{x}) = 0$. Denoting $\ell = \dim W$, we have clearly $n - m \leq r \leq n$ and

$$\frac{1}{2\ell} \sum_{i=1}^{\ell-1} \sum_{k=1}^{i} \frac{1}{k} < \frac{1}{2} \log \ell \leq \frac{1}{2} \log n.$$

Hence there exists a non-zero $\underline{x} \in W$ satisfying

$$\mathrm{h}_p(\underline{x}) \leq \frac{1}{n-m} \mathrm{h}_s(W) + \frac{1}{2} \log n.$$

Using (4.10), we find

$$\mathrm{h}_p(\underline{x}) \leq \frac{1}{n-m} \left(\mathrm{h}_p(L_1) + \ldots + \mathrm{h}_p(L_m)\right) + \frac{1}{2} \frac{m}{n-m} \log n + \frac{1}{2} \log n,$$

which is (4.12). □

4.4 Power series

In this section we recall various results about power series, used in our proof.

4.4.1 Puiseux Expansions

Let $\mathbb{K}$ be a field of characteristic 0, and $\mathbb{K}[[x]]$ the field of formal power series over $\mathbb{K}$. It is well-known that an extension of $\mathbb{K}[[x]]$ of degree n is a subfield of a field of the form $\mathbb{L}[[x^{1/e}]]$, where e is a positive integer (the ramification index), $\mathbb{L}$ is a finite extension of $\mathbb{K}$, and

$$[\mathbb{L} : \mathbb{K}], e \leq n.$$

This fact (quoted sometimes as the "Theorem of Puiseux") has the following consequence: if we fix an algebraic closure $\bar{\mathbb{K}}$ of $\mathbb{K}$, then the algebraic closure of $\mathbb{K}[[x]]$ can be given by

$$\overline{\mathbb{K}[[x]]} = \bigcup_{e=1}^{\infty} \bigcup_{\substack{\mathbb{K} \subset \mathbb{L} \subset \bar{\mathbb{K}} \\ [\mathbb{L}:\mathbb{K}] < \infty}} \mathbb{L}[[x^{1/e}]],$$

where the interior union is over all subfields $\mathbb{L}$ of $\bar{\mathbb{K}}$ finite over $\mathbb{K}$.

Another immediate consequence of the "Theorem of Puiseux" is the following statement:

Proposition 4.4.1 *Let*

$$F(X, Y) = f_n(X)Y^n + \cdots + f_0(X) \in \mathbb{K}[X, Y]$$

be a polynomial of Y-degree n. Then there exists a finite extension $\mathbb{L}$ of $\mathbb{K}$, positive integers $e_1, \ldots, e_n$, all not exceeding n, and series $y_i \in \mathbb{L}[[x^{1/e_i}]]$ such that

$$F(x, Y) = f_n(x)(Y - y_1) \cdots (Y - y_n). \tag{4.13}$$

We write the series $y_1, \dots, y_n$ as

$$y_i = \sum_{k=\kappa_i}^{\infty} a_{ik} x^{k/e_i}$$

with $a_{i\kappa_i} \neq 0$. It is well-known and easy to show that

$$|\kappa_i| \leq \deg_X F \qquad (i = 1, \dots, n).$$

This inequality will be used throughout the article without special notice.

We want to link the numbers e_i and κ_i with the "order of vanishing" at $(0,0)$, introduced in (4.4).

Proposition 4.4.2 *Let $F(X,Y) \in \mathbb{K}[X,Y]$ and $y_1, \dots, y_n$ be as above, and assume that $F(0,Y)$ is not identically 0. Then the quantity r, introduced in (4.4), satisfies*

$$r = \sum_{\kappa_i > 0} \min\{1, \kappa_i/e_i\}, \tag{4.14}$$

where the sum extends only to those i for which $\kappa_i > 0$.

Proof We denote by ν_x the standard additive valuation on $\mathbb{K}[[x]]$, normalized to have $\nu_x(x) = 1$. This ν_x extends in a unique way to the algebraic closure $\overline{\mathbb{K}[[x]]}$; precisely, for

$$y(x) = \sum_{k=\kappa}^{\infty} a_k x^{k/e} \in \overline{\mathbb{K}[[x]]} \qquad (a_\kappa \neq 0)$$

we have $\nu_x(y) = \kappa/e$. Furthermore, for

$$G(x,Y) = g_s(x)Y^s + \cdots + g_0(x) \in \overline{\mathbb{K}[[x]]}[Y]$$

we set $\nu_x(G) = \min\{\nu_x(g_0), \dots, \nu_x(g_s)\}$. Gauß' lemma asserts that for $G_1, G_2 \in \overline{\mathbb{K}[[x]]}[Y]$, we have $\nu_x(G_1G_2) = \nu_x(G_1) + \nu_x(G_2)$.

Since $F(0,Y)$ is not identically 0, we have $\nu_x(F(x,Y)) = 0$. Applying Gauß' lemma to (4.13), we obtain

$$\nu_x(f_n(x)) + \sum \min\{0, \kappa_i/e_i\} = 0.$$

Hence, setting $\tilde{f}_n = x^{-\nu_x(f_n(x))} f_n(x)$, we may re-write (4.13) as

$$F(x,Y) = \prod_{\kappa_i > 0} (Y - y_i) \cdot \tilde{f}_n(x) \prod_{\kappa_i \leq 0} (x^{-\kappa_i/e_i} Y - x^{-\kappa_i/e_i} y_i). \tag{4.15}$$

Now set $G(x,Y) = F(x,xY)$. Then clearly $r = \nu_x(G)$. Applying Gauß' Lemma to the decomposition

$$G(x,Y) = \prod_{\kappa_i > 0} (xY - y_i) \cdot \tilde{f}_n(x) \prod_{\kappa_i \leq 0} (x^{1-\kappa_i/e_i} Y - x^{-\kappa_i/e_i} y_i),$$

we obtain (4.14). □

Here is one more useful property.

Proposition 4.4.3 *In the set-up of Proposition 4.4.2, assume that $\kappa_i > 0$ for exactly ℓ indexes $i \in \{1, \dots, n\}$. Then $f_k(0) = 0$ for $k < \ell$, but $f_\ell(0) \neq 0$.*

Proof Re-write (4.15) as

$$F(x,Y) = \prod_{\kappa_i>0}(Y-y_i)\prod_{\kappa_i=0}(Y-y_i)\cdot\tilde{f}_0(x)\prod_{\kappa_i<0}(x^{-\kappa_i/e_i}Y - x^{-\kappa_i/e_i}y_i).$$

Substituting $x=0$, every factor in the first product becomes Y, every factor in the second product becomes $Y-a_{i0}$, with $a_{i0}\neq 0$, and every factor in the third product (including $\tilde{f}_0(0)$) becomes constant. Whence the result. □.

4.4.2 Eisenstein's theorem

In this subsection, we recall the quantitative version of Eisentsein's theorem, due to work from Dwork, Robba [Dwork & Robba 1979], Schmidt [Schmidt 1990] and Van der Poorten [Dwork & van der Poorten 1992], as given in [Bilu & Borichev 2013]. It will be convenient to use the notion of $M_{\mathbb{K}}$-divisor.

An $M_{\mathbb{K}}$-divisor is an infinite vector $(A_v)_{v\in M_{\mathbb{K}}}$ of positive real numbers, each A_v being associated to one $v\in M_{\mathbb{K}}$, such that for all but finitely many $v\in M_{\mathbb{K}}$ we have $A_v=1$. An $M_{\mathbb{K}}$-divisor is effective if for all $v\in M_{\mathbb{K}}$, $A_v\geq 1$.

We define the *height* of an $M_{\mathbb{K}}$-divisor $\mathscr{A}=(A_v)_{v\in M_{\mathbb{K}}}$ as

$$\mathrm{h}(\mathscr{A}) = \sum_{v\in M_{\mathbb{K}}}\frac{d_v}{d}\log A_v. \tag{4.16}$$

The following version of Eisenstein's theorem is from [Bilu & Borichev 2013, Theorem 7.5].

Theorem 4.4.4 *Let $F(X,Y)$ be a separable polynomial of degrees $m=\deg_X F$ and $n=\deg_Y F$. Further, let $y(x)=\sum_{k=\kappa}^{\infty}a_k x^{k/e}\in\mathbb{K}[[x^{1/e}]]$ be a power series satisfying $F(x,y(x))=0$. (Here we do not assume that $a_\kappa\neq 0$.) Then there exists an effective $M_{\mathbb{K}}$-divisor $\mathscr{A}=(A_v)_{v\in M_{\mathbb{K}}}$ such that:*

$$|a_k|_v \leq \max\{1, |a_{e\lfloor\kappa/e\rfloor}|_v\}A_v^{k/e-\lfloor\kappa/e\rfloor},$$

for any $v\in M_{\mathbb{K}}$ and any $k\geq\kappa$, and such that $\mathrm{h}(\mathscr{A})\leq 4n\mathrm{h}_p(F)+3n\log(nm)+10en$.

Applying this theorem to the series of the form $a_1x^{1/e}+a_2x^{2/e}+\ldots$ (that is, with $a_k=0$ for $k\leq 0$) and setting $\kappa=0$, we obtain that:

Corollary 4.4.5 *Let $F(X,Y)$ be a separable polynomial of degrees $m=\deg_X F$ and $n=\deg_Y F$. Further, let $y(x)=\sum_{k=1}^{\infty}a_k x^{k/e}\in\mathbb{K}[[x^{1/e}]]$ be a power series satisfying $F(x,y(x))=0$. Then, there exists an effective $M_{\mathbb{K}}$-divisor $\mathscr{A}=(A_v)_{v\in M_{\mathbb{K}}}$ such that:*

$$|a_k|_v \leq A_v^{k/e} \qquad (v\in M_{\mathbb{K}},\quad k=1,2,\ldots), \tag{4.17}$$

and such that

$$\mathrm{h}(\mathscr{A}) \leq 4n\mathrm{h}_p(F)+3n\log(nm)+10en. \tag{4.18}$$

The following lemma is a slightly modified version of Proposition 2.7 of Abouzaid's article [Abouzaid 2008]:

Lemma 4.4.6 *Let $\mathbb{K}$ be a number field and let $y(x) = \sum_{k=1}^{\infty} a_k x^{k/e}$ be a series with coefficients in $\mathbb{K}$. Assume further that there exists an effective $M_{\mathbb{K}}$-divisor $\mathscr{A} = (A_v)_{v \in M_{\mathbb{K}}}$, such that for all $k \geq 1$ we have $|a_k|_v \leq A_v^{k/e}$. For $\ell \in \mathbb{N}$ write $y(x)^\ell = \sum_{k=1}^{\infty} a_k^{(\ell)} c^{k/e}$. Then, for any $v \in M_K$ and for all $k \geq 1$ we have:*

$$|a_k^{(\ell)}|_v \leq \begin{cases} 2^{\ell+k} A_v^{k/e}, & \text{if } v|\infty, \\ A_v^{k/e}, & \text{if } v < \infty. \end{cases} \tag{4.19}$$

In [Abouzaid 2008], a slightly sharper estimate, with $\binom{\ell+k-1}{k}$ instead of $2^{\ell+k}$ is given.

4.5 The "Main Lemma"

In this section we prove an auxiliary statement which is crucial for the proof of Theorem 4.2.3. It can be viewed as a version of the famous Theorem of Sprindzhuk, see [Bombieri 1983, Bilu & Masser 2006]. In fact, our argument is an adaptation of that from [Bilu & Masser 2006]. We follow [Abouzaid 2008, Sections 3.1–3.3] with some changes.

4.5.1 Statement of the Main Lemma

In this section $\mathbb{K}$ is a number field, $F(X, Y) \in \mathbb{K}[X, Y]$ an absolutely irreducible polynomial of degrees $m = \deg_X F$ and $n = \deg_Y F$, and $\alpha, \beta \in \mathbb{K}^\times$ satisfy $F(\alpha, \beta) = 0$. Furthermore, everywhere in this section except Subsection 4.5.6

$$y(x) = \sum_{k=1}^{\infty} a_k x^k \in \mathbb{K}[[x]]$$

is a power series satisfying $F(x, y(x)) = 0$; in particular, $F(0, 0) = 0$.

We consider the following finite subset of $M_{\mathbb{K}}$:

$$T = \{v \in M_{\mathbb{K}} : |\alpha|_v < 1 \text{ and } y(x) \text{ converges } v\text{-adically to } \beta \text{ at } x = \alpha\}.$$

Lemma 4.5.1 ("Main Lemma") *Let ε satisfy $0 < \varepsilon \leq 1$. Then we have either*

$$\mathrm{h}(\alpha) \leq 200\varepsilon^{-2} mn^4 (\mathrm{h}_p(F) + 5), \tag{4.20}$$

or

$$\left| \frac{\mathrm{h}(\alpha)}{n} - \mathrm{h}_T(\alpha) \right| \leq \varepsilon n \mathrm{h}(\alpha) + 200\varepsilon^{-1} n^2 (\mathrm{h}_p(F) + \log(mn) + 10). \tag{4.21}$$

4.5.2 Preparations

The proof of the "Main Lemma" requires some preparation. First of all, recall that, according to Eisenstein's Theorem as given in Corollary 4.4.5, there exists an effective $M_{\mathbb{K}}$-divisor $\mathscr{A} = (A_v)_{v \in M_{\mathbb{K}}}$ such that both (4.17) and (4.18) hold with $e = 1$:

$$\begin{aligned} |a_k|_v &\leq A_v^k \qquad (v \in M_{\mathbb{K}}, \quad k = 1, 2, \ldots), \\ \mathrm{h}(\mathscr{A}) &\leq 4n\mathrm{h}_p(F) + 3n \log(nm) + 10n. \end{aligned}$$

We fix this $\mathscr{A}$ until the end of the section.

Next, we need to construct an "auxiliary polynomial".

Proposition 4.5.2 (Auxiliary polynomial) *Let δ be a real number $0 < \delta \leq 1/2$ and let N be a positive integer. There exists a non-zero polynomial $G(X,Y) \in \bar{\mathbb{Q}}[X,Y]$ satisfying $\deg_X G \leq N$, $\deg_Y G \leq n-1$,*

$$\nu_x(G(x,y(x))) \geq (1-\delta)Nn, \tag{4.22}$$

$$\mathrm{h}_p(G) \leq \delta^{-1}nN(\mathrm{h}(\mathscr{A})+3). \tag{4.23}$$

Proof It is quite analogous to the proof of Proposition 3.1 in [Abouzaid 2008]. Condition (4.22) is equivalent to a system of $(1-\delta)Nn$ linear equations in the $n(N+1)$ coefficients of G. Each coefficient of each linear equation is a coefficient of x^k, for $k \leq Nn$, one of the series $y(x)^\ell$ for $\ell = 0,\dots,n-1$.

Using Corollary 4.4.5 and Lemma 4.4.6, we estimate the height of every equation as $nN\mathrm{h}(\mathscr{A}) + (Nn+n)\log 2$. Corollary 4.3.6 implies now that we can find a non-zero solution of our system of height at most

$$\delta^{-1}(nN\mathrm{h}(\mathscr{A}) + (Nn+n)\log 2 + \frac{1}{2}\delta^{-1}\log(nN)).$$

This is smaller than the right-hand side of (4.23). □

4.5.3 Upper Bound

Now we can obtain an upper bound for $\mathrm{h}_T(\alpha)$ in terms of $\mathrm{h}(\alpha)$.

Proposition 4.5.3 (Upper bound for $\mathrm{h}_T(\alpha)$) *Let δ satisfy $0 < \delta \leq 1/2$. Then we have either*

$$\mathrm{h}(\alpha) \leq 10\delta^{-2}mn^4(\mathrm{h}_p(F)+5), \tag{4.24}$$

or

$$n\mathrm{h}_T(\alpha) \leq (1+4\delta)\mathrm{h}(\alpha) + 8\delta^{-1}n(\mathrm{h}(\mathscr{A})+10) + \mathrm{h}_p(F). \tag{4.25}$$

Proof Fix a positive integer N, to be specified later, and let $G(X,Y)$ be the auxiliary polynomial introduced in Proposition 4.5.2. Extending the field $\mathbb{K}$, we may assume that $G(X,Y) \in \mathbb{K}[X,Y]$. We may also assume that G has a coefficient equal to 1; in particular, $|G|_v \geq 1$ for all $v \in M_{\mathbb{K}}$, where we denote by $|G|_v$ the maximum of v-adic norms of coefficients of G.

The series $z(x) = G(x,y(x)) \in \mathbb{K}[[x]]$ can be written as

$$z(x) = \sum_{k=\eta}^{\infty} b_k x^k$$

with $\eta \geq (1-\delta)Nn \geq \frac{1}{2}Nn$ (recall that $\delta \leq 1/2$). Again using (4.17) and Lemma 4.4.6, we estimate the coefficients b_k as follows: for $v < \infty$ we have $|b_k|_v \leq |G|_v A_v^k$, and for $v \mid \infty$ we have $|b_k|_v \leq n(N+1)2^{k+n-1}|G|_v A_v^k$. Since for $k \geq \eta \geq \frac{1}{2}Nn$ we have $n(N+1)2^{k+n-1} \leq 8^k$, we obtain the estimate

$$|b_k| \leq \begin{cases} |G|_v A_v^k, & v < \infty, \\ |G|_v (8A_v)^k, & v \mid \infty. \end{cases} \qquad (v \in M_k, \quad , k \geq \eta). \tag{4.26}$$

Now we distinguish two cases.

Case 1: $G(\alpha,\beta)=0$ In this case we have $F(\alpha,\beta)=G(\alpha,\beta)=0$. We want to apply Lemma 2.2.3; for this, we have to verify that polynomials F and G do not have a common factor. This is indeed the case, because F is absolutely irreducible, and $\deg_Y G<\deg_Y F$.

Lemma 2.2.3, combined with (4.23) and (4.18), gives

$$\begin{aligned}\mathrm{h}(\alpha)&\le n\mathrm{h}_p(G)+(n-1)\mathrm{h}_pF+(m(n-1)+Nn)+(2n-1)\log(2n-1)+\log 2\\&\le\delta^{-1}Nn^2(\mathrm{h}(\mathscr{A})+6)+(n-1)(\mathrm{h}_p(F)+m)\\&\le 5\delta^{-1}Nn^3(\mathrm{h}_p(F)+5)+mn.\end{aligned}\tag{4.27}$$

Below, after specifying N, we will see that this is sharper than (4.24).

Case 2: $G(\alpha,\beta)=\gamma\neq 0$ To treat this case it will be convenient to use, instead of the set T, a slightly smaller subset $\widetilde{T}$, consisting of $v\in T$ satisfying

$$|\alpha|_v<\begin{cases}A_v^{-1}, & v<\infty,\\ (16A_v)^{-1}, & v\mid\infty.\end{cases}$$

We have clearly

$$0\le\mathrm{h}_T(\alpha)-\mathrm{h}_{\widetilde{T}}(\alpha)\le\mathrm{h}(\mathscr{A})+\log 16,\tag{4.28}$$

and (4.26) implies the estimate

$$|b_k\alpha^k|_v<\begin{cases}|G|_vA_v^{\eta}|\alpha|_v^{\eta}, & v<\infty,\\ |G|_v(8A_v)^{\eta}|\alpha|_v^{\eta}\cdot(1/2)^{k-\eta}, & v\mid\infty.\end{cases}\qquad(v\in\widetilde{T},\quad k\ge\eta).\tag{4.29}$$

Recall that for $v\in T$, the series $y(x)$ converges v-adically to β at $x=\alpha$. Hence the same holds true for $v\in\widetilde{T}$. It follows that, for $v\in\widetilde{T}$, the series $z(x)=G(x,y(x))$ converges v-adically to[1] $G(\alpha,\beta)=\gamma$.

Using (4.29), we can estimate $|\gamma|_v$ for $v\in\widetilde{T}$:

$$|\gamma|_v<\begin{cases}|G|_vA_v^{\eta}|\alpha|_v^{\eta}, & v<\infty,\\ 2|G|_v(8A_v)^{\eta}|\alpha|_v^{\eta}, & v\mid\infty.\end{cases}\qquad(v\in\widetilde{T},\quad k\ge\eta).$$

Using this and remembering that $|G|_v\ge 1$ for all v, we obtain the following lower estimate for $\mathrm{h}(\gamma)$:

$$\begin{aligned}\mathrm{h}(\gamma)&\ge\mathrm{h}_{\widetilde{T}}(\gamma)\\&\ge\eta\mathrm{h}_{\widetilde{T}}(\alpha)-\mathrm{h}_p(G)-\eta\mathrm{h}(\mathscr{A})-\eta\log 16-\log 2\\&\ge Nn(1-\delta)\mathrm{h}_{\widetilde{T}}(\alpha)-2\delta^{-1}nN(\mathrm{h}(\mathscr{A})+6).\end{aligned}$$

Combining this with (4.28), we obtain

$$\mathrm{h}(\gamma)\ge Nn(1-\delta)\mathrm{h}_T(\alpha)-3\delta^{-1}nN(\mathrm{h}(\mathscr{A})+6).\tag{4.30}$$

[1]For archimedean v to make this conclusion we need absolute convergence of $y(x)$ at $x=\alpha$, which is obvious for $v\in\widetilde{T}$.

On the other hand, using Lemma 2.2.4 it is easy to bound $\mathrm{h}(\gamma)$ from above. Indeed, part B of this lemma implies that

$$\mathrm{h}(\beta) \leq \mathrm{h}_p(F) + m\mathrm{h}(\alpha) + n + \log(m+1),$$

and part A implies that

$$\mathrm{h}(\gamma) \leq \mathrm{h}_a(G) + N\mathrm{h}(\alpha) + (n-1)\mathrm{h}(\beta) + \log((N+1)n).$$

Since G has a coefficient equal to 1, we have $\mathrm{h}_a(G) = \mathrm{h}_p(G) \leq \delta^{-1}nN(\mathrm{h}(\mathscr{A}) + 3)$. Hence

$$\begin{aligned}\mathrm{h}(\gamma) &\leq \mathrm{h}_p(G) + N\mathrm{h}(\alpha) + (n-1)(\mathrm{h}_p(F) + m\mathrm{h}(\alpha) + n + \log(m+1)) + \log((N+1)n)\\ &\leq (N+mn)\mathrm{h}(\alpha) + \delta^{-1}nN(\mathrm{h}(\mathscr{A}) + 4) + n\mathrm{h}_p(F) + n^2 + n\log(m+1).\end{aligned}$$

Combining this with (4.30) and dividing by N, we obtain

$$n(1-\delta)\mathrm{h}_T(\alpha) \leq \left(1 + \frac{mn}{N}\right)\mathrm{h}(\alpha) + 4\delta^{-1}n(\mathrm{h}(\mathscr{A}) + 6) + N^{-1}(n\mathrm{h}_p(F) + n^2 + n\log(m+1)). \tag{4.31}$$

Completing the proof of Proposition 4.5.3 Now it is the time to specify N: we set $N = \lceil \delta^{-1}mn \rceil$. With this choice of N, inequality (4.27) is indeed sharper than (4.24), and inequality (4.31) implies the following:

$$n(1-\delta)\mathrm{h}_T(\alpha) \leq (1+\delta)\mathrm{h}(\alpha) + 4\delta^{-1}n(\mathrm{h}(\mathscr{A}) + 10) + \delta\mathrm{h}_p(F).$$

Since $\delta \leq 1/2$, this is sharper than (4.25). □

4.5.4 Lower Bound

Our next objective is a lower bound for $\mathrm{h}_T(\alpha)$. We will see that it easily follows from the upper bound.

Proposition 4.5.4 (Lower bound for $\mathrm{h}_T(\alpha)$) *Let δ satisfy $0 < \delta \leq 1/2$. Then we have either (4.24) or*

$$n\mathrm{h}_T(\alpha) \geq (1 - 4n\delta)\mathrm{h}(\alpha) - 9\delta^{-1}n^2(\mathrm{h}(\mathscr{A}) + 10) - n\mathrm{h}_p(F). \tag{4.32}$$

Proof Remark first of all that we may assume that the polynomial $F(\alpha, Y)$ is of degree n and separable. Indeed, if this is not the case, then $R_F(\alpha) = 0$, where $R_F(X)$ is the Y-resultant of $F(X,Y)$ and its Y-derivative $F'_Y(X,Y)$. In this case, the joint application of Lemmas 2.2.1 and 2.2.2 gives

$$\mathrm{h}(\alpha) \leq 2n\mathrm{h}_p(F) + 2n\log((m+1)(n+1)\sqrt{n}) + \log 2,$$

sharper than (4.24).

Thus, $F(\alpha, Y)$ has n distinct roots in $\bar{\mathbb{Q}}$, one of which is β; we denote them $\beta_1 = \beta, \ldots, \beta_n$. Extending the field $\mathbb{K}$, we may assume that $\beta_1, \ldots, \beta_n \in \mathbb{K}$.

Set $S = \{v \in M_{\mathbb{K}} : |\alpha|_v < 1\}$. For $i = 1, \dots, n$ we let T_i be the set of $v \in S$ such that $y(x)$ converges v-adically to β_i at $x = \alpha$; in particular, $T_1 = T$. The sets $T_1, \dots, T_n$ are clearly disjoint, and we have

$$S \supset T_1 \cup \ldots \cup T_n \supset \widetilde{S}, \tag{4.33}$$

where $\widetilde{S}$ consists of $v \in S$ for which $|\alpha|_v < A_v^{-1}$. The left inclusion in (4.33) is trivial, and to prove the right one just observes that for every $v \in \widetilde{S}$, the series $y(x)$ absolutely converges v-adically at $x = \alpha$, and, since $F(x, y(x)) = 0$, the sum must be a root of $F(\alpha, Y)$.

Clearly,

$$0 \le \mathrm{h}(\alpha) - \mathrm{h}_{\widetilde{S}}(\alpha) = \mathrm{h}_S(\alpha) - \mathrm{h}_{\widetilde{S}}(\alpha) \le \mathrm{h}(\mathscr{A}).$$

It follows that

$$\mathrm{h}_{T_1}(\alpha) + \cdots + \mathrm{h}_{T_n}(\alpha) \ge \mathrm{h}_{\widetilde{S}}(\alpha) \ge \mathrm{h}(\alpha) - \mathrm{h}(\mathscr{A}).$$

Now observe that the upper bound (4.25) holds true with T replaced by any T_i:

$$n\mathrm{h}_{T_i}(\alpha) \le (1 + 4\delta)\mathrm{h}(\alpha) + 8\delta^{-1}n(\mathrm{h}(\mathscr{A}) + 10) + \mathrm{h}_p(F) \qquad (i = 1, \dots, n).$$

The last two inequalities imply that

$$n\mathrm{h}_T(\alpha) = n\mathrm{h}_{T_1}(\alpha) \ge n(\mathrm{h}(\alpha) - \mathrm{h}(\mathscr{A})) - (n-1)((1+4\delta)\mathrm{h}(\alpha) + 8\delta^{-1}n(\mathrm{h}(\mathscr{A}) + 10) + \mathrm{h}_p(F)),$$

which easily transforms into (4.32). □

4.5.5 Proof of the "Main Lemma"

Using Propositions 4.5.3 and 4.5.4 with $\delta = \varepsilon/4$ and dividing by n, we obtain that either (4.20) holds, or

$$\left|\mathrm{h}_T(\alpha) - \frac{\mathrm{h}(\alpha)}{n}\right| \le \varepsilon\mathrm{h}(\alpha) + 40\varepsilon^{-1}n(\mathrm{h}(\mathscr{A}) + 10) + \mathrm{h}_p(F).$$

Combining this with (4.18), we obtain (4.21). □

4.5.6 "Ramified Main Lemma"

We will actually need a slightly more general statement, allowing ramification in the series $y(x)$. The set-up is as before, except that now we consider the series

$$y(x) = \sum_{k=1}^{\infty} a_k x^{k/e} \in \mathbb{K}[[x^{1/e}]]$$

satisfying $F(x, y(x)) = 0$. We fix an e-th root $\alpha^{1/e}$ and we will assume that it belongs to $\mathbb{K}$. We will now say that the series $y(x)$ converges v-adically to β at α if the series $y(x^e)$ converges v-adically to β at $\alpha^{1/e}$. (Of course, this depends on the particular choice of the root $\alpha^{1/e}$.) We again define T as the set of all $v \in S$ for which $y(x)$ converges v-adically to β at α.

Lemma 4.5.5 ("Ramified Main Lemma") *Let ε satisfy $0 < \varepsilon \le 1$. Then we have either*

$$\mathrm{h}(\alpha) \le 200\varepsilon^{-2}me^2n^4(\mathrm{h}_p(F) + 5), \tag{4.34}$$

or

$$\left|\frac{\mathrm{h}(\alpha)}{n} - \mathrm{h}_T(\alpha)\right| \le \varepsilon\mathrm{h}(\alpha) + 200\varepsilon^{-1}en^2(\mathrm{h}_p(F) + 2\log(mn) + 10). \tag{4.35}$$

Proof The proof is by reduction to the unramified case. Apply Lemma 4.5.1 to the polynomial $F(X^e, Y)$, the series $y(x^e)$ and the number $\alpha^{1/e}$. We obtain that either

$$\mathrm{h}(\alpha^{1/e}) \le 200\varepsilon^{-2}men^6(\mathrm{h}_p(F) + 5),$$

or

$$|\mathrm{h}(\alpha^{1/e}) - n\mathrm{h}_T(\alpha^{1/e})| \le \varepsilon\mathrm{h}(\alpha^{1/e}) + 200\varepsilon^{-1}n^4(\mathrm{h}_p(F) + \log(men) + 10).$$

These estimates easily transform into (4.34) and (4.35), respectively, using that

$$\mathrm{h}(\alpha^{1/e}) = e^{-1}\mathrm{h}(\alpha), \quad \mathrm{h}_T(\alpha^{1/e}) = e^{-1}\mathrm{h}_T(\alpha), \quad e \le n. \qquad \square$$

4.6 Proof of the Main Theorem

In this section we prove Theorem 4.2.3. First of all, we investigate the relation between $\mathrm{h}_T(\alpha)$ and $\mathrm{lgcd}_T(\alpha, \beta)$, where T is defined as in Section 4.5.

4.6.1 Comparing $\mathrm{h}_T(\alpha)$ and $\mathrm{lgcd}_T(\alpha, \beta)$

In this subsection we retain the set-up of Subsection 4.5.1, except that we allow ramification in the series $y(x)$, as we did in Subsection 4.5.6. Thus, in this subsection:

- $\mathbb{K}$ is a number field;
- $F(X, Y) \in \mathbb{K}[X, Y]$ is an absolutely irreducible polynomial;
- $\alpha, \beta \in \mathbb{K}$ satisfy $F(\alpha, \beta) = 0$;
- $y(x) = \sum_{k=1}^{\infty} a_k x^{k/e} \in \mathbb{K}[[x^{1/e}]]$ satisfies $F(x, y(x)) = 0$;
- $T \subset M_{\mathbb{K}}$ is the set of all $v \in M_{\mathbb{K}}$ such that $|\alpha|_v < 1$ and $y(x)$ converges v-adically at α to β.

The v-adic convergence is understood in the same sense as in Subsection 4.5.6: we fix an e-th root $\alpha^{1/e}$, assume that it belongs to $\mathbb{K}$ and and define v-adic convergence of $y(x)$ to β at α as v-adic convergence of $y(x^e)$ to β at $\alpha^{1/e}$.

Let κ be the smallest k such that $a_k \ne 0$; by the assumption, $\kappa > 0$. Then we have $\nu_x(y) = \kappa/e$ and

$$y(x) = \sum_{k=\kappa}^{\infty} a_k x^{k/e}$$

with $a_\kappa \ne 0$. In this subsection we prove that $\mathrm{lgcd}_T(\alpha, \beta)$ can be approximated by $\min\{1, \kappa/e\}\mathrm{h}_T(\alpha)$.

Proposition 4.6.1 *In the above set-up we have*

$$|\mathrm{lgcd}_T(\alpha, \beta) - \min\{\kappa/e, 1\}\mathrm{h}_T(\alpha)| \le 30n\kappa\mathrm{h}_p(F) + 30n\kappa\log(nm) + 15en. \tag{4.36}$$

This statement corresponds to Proposition 3.6 in [Abouzaid 2008]. Our proof is, however, much more involved, in particular because Abouzaid did not need the lower estimate.

Proof Let $\mathscr{A} = (A_v)_{v\in M_{\mathbb{K}}}$ be the $M_{\mathbb{K}}$-divisor from Corollary 4.4.5. For the reader's convenience, we reproduce here (4.17) and (4.18):

$$\begin{aligned} |a_k|_v &\le A_v^{k/e} \qquad (v \in M_{\mathbb{K}}, \quad k \ge 1),\\ \mathrm{h}(\mathscr{A}) &\le 4n\mathrm{h}_p(F) + 3n\log(nm) + 10en. \end{aligned}$$

As we already did several times in Section 4.5, it will be convenient to replace T by a smaller subset. Thus, let $\widetilde{T}$ consist of $v \in T$ satisfying

$$|\alpha|_v < \begin{cases} A_v^{-\kappa-1}\min\{1, |a_\kappa|_v\}^e, & v < \infty,\\ (1/4)^e A_v^{-\kappa-1}\min\{1, |a_\kappa|_v\}^e, & v < \infty. \end{cases} \tag{4.37}$$

(Attention: this is not the same $\widetilde{T}$ as in Subsection 4.5.3!) Clearly,

$$0 \le \mathrm{h}_T(\alpha) - \mathrm{h}_{\widetilde{T}}(\alpha) \le (\kappa+1)\mathrm{h}(\mathscr{A}) + e\mathrm{h}_{T\smallsetminus\widetilde{T}}(a_\kappa).$$

Using (4.17) we estimate $\mathrm{h}(a_\kappa) \le (\kappa/e)\mathrm{h}(\mathscr{A})$. We obtain

$$0 \le \mathrm{h}_T(\alpha) - \mathrm{h}_{\widetilde{T}}(\alpha) \le (\kappa+1)\mathrm{h}(\mathscr{A}) \le 3\kappa\mathrm{h}(\mathscr{A}) + e\log 4, \tag{4.38}$$

where for the latter estimate we use $\kappa \ge 1$. In particular,

$$0 \le \mathrm{lgcd}_T(\alpha,\beta) - \mathrm{lgcd}_{\widetilde{T}}(\alpha,\beta) \le 3\kappa\mathrm{h}(\mathscr{A}) + e\log 4. \tag{4.39}$$

After this preparation, we can now proceed with the proof. For every $v \in \widetilde{T}$ we want to obtain an estimate of the form $c_v|\alpha|_v^{\kappa/e} \le |\beta|_v \le c'_v|\alpha|_v^{\kappa/e}$, where c_v and c'_v are some quantities not depending on α.

Upper estimate for $|\beta|_v$. This is easy. It follows from (4.37) that

$$|\alpha|_v < \begin{cases} A_v^{-1}, & v < \infty,\\ (4^e A_v)^{-1}, & v < \infty. \end{cases}$$

From this and (4.17) we deduce that

$$|a_k\alpha^{k/e}|_v < \begin{cases} A_v^{\kappa/e}|\alpha|_v^{\kappa/e}, & v < \infty,\\ A_v^{\kappa/e}|\alpha|_v^{\kappa/e}\cdot(1/4)^{k-\kappa}, & v \mid \infty \end{cases} \qquad (k \ge \kappa). \tag{4.40}$$

Hence

$$|\beta|_v < \begin{cases} A_v^{\kappa/e}|\alpha|_v^{\kappa/e}, & v < \infty,\\ 2A_v^{\kappa/e}|\alpha|_v^{\kappa/e}, & v \mid \infty. \end{cases}$$

Lower estimate for $|\beta|_v$. The lower estimate is slightly more subtle. First, we bound the difference $\beta - a_\kappa\alpha^{\kappa/e}$ from above using (4.37).

Similarly to (4.40), we have

$$|a_k\alpha^{k/e}|_v < \begin{cases} A_v^{(\kappa+1)/e}|\alpha|_v^{(\kappa+1)/e}, & v<\infty, \\ A_v^{(\kappa+1)/e}|\alpha|_v^{(\kappa+1)/e}\cdot(1/4)^{(k-\kappa-1)/e}, & v\mid\infty \end{cases} \qquad (k\geq\kappa+1).$$

Hence, presenting $\beta - a_\kappa\alpha^{\kappa/e}$ as the v-adic sum of the series

$$y(x) - a_\kappa x^{\kappa/e} = \sum_{k=\kappa+1}^{\infty} a_k x^{k/e}$$

at $x=\alpha$, we obtain the estimate

$$|\beta - a_\kappa\alpha^{\kappa/e}|_v < \begin{cases} A_v^{(\kappa+1)/e}|\alpha|_v^{(\kappa+1)/e}, & v<\infty, \\ 2A_v^{(\kappa+1)/e}|\alpha|_v^{(\kappa+1)/e}, & v\mid\infty. \end{cases}$$

Combining this with (4.37), we find

$$|\beta - a_\kappa\alpha^{\kappa/e}|_v < \begin{cases} \min\{|a_\kappa|_v,1\}|\alpha|_v^{\kappa/e}, & v<\infty, \\ (1/2)\min\{|a_\kappa|_v,1\}|\alpha|_v^{\kappa/e}, & v\mid\infty. \end{cases}$$

Hence

$$|\beta|_v \geq \begin{cases} \min\{|a_\kappa|_v,1\}|\alpha|_v^{\kappa/e}, & v<\infty, \\ (1/2)\min\{|a_\kappa|_v,1\}|\alpha|_v^{\kappa/e}, & v\mid\infty, \end{cases}$$

the lower estimate we were seeking.

Completing the proof of Proposition 4.6.1 Thus, we proved that

$$c_v|\alpha|_v^{\kappa/e} \leq |\beta|_v \leq c_v'|\alpha|_v^{\kappa/e}, \tag{4.41}$$

with

$$c_v = \begin{cases} \min\{|a_\kappa|_v,1\}, & v<\infty, \\ (1/2)\min\{|a_\kappa|_v,1\}, & v\mid\infty, \end{cases}, \qquad c_v' = \begin{cases} A_v^{\kappa/e}, & v<\infty, \\ 2A_v^{\kappa/e}, & v\mid\infty. \end{cases}$$

From (4.41) we deduce that for $v\in\widetilde{T}$

$$c_v|\alpha|_v^{\min\{\kappa/e,1\}}\max\{|\alpha|_v,|\beta|_v\} \leq c_v'|\alpha|_v^{\min\{\kappa/e,1\}}.$$

(We use here the obvious inequality $c_v\leq 1\leq c_v'$.) Hence

$$-(\kappa/e)\mathrm{h}(\mathscr{A}) - \log 2 \leq \mathrm{lgcd}_{\widetilde{T}}(\alpha,\beta) - \min\{\kappa/e,1\}\mathrm{h}_{\widetilde{T}}(\alpha) \leq \mathrm{h}(a_\kappa) + \log 2.$$

Since $\mathrm{h}(a_\kappa)\leq(\kappa/e)\mathrm{h}(\mathscr{A})$, this implies

$$|\mathrm{lgcd}_{\widetilde{T}}(\alpha,\beta) - \min\{\kappa/e,1\}\mathrm{h}_{\widetilde{T}}(\alpha)| \leq (\kappa/e)\mathrm{h}(\mathscr{A}) + \log 2,$$

which, together with (4.38) and (4.39) gives

$$|\mathrm{lgcd}_{\widetilde{T}}(\alpha,\beta) - \min\{\kappa/e,1\}\mathrm{h}_{\widetilde{T}}(\alpha)| \leq 7\kappa\mathrm{h}(\mathscr{A}) + 4e.$$

Combining this with (4.17), we obtain (4.36). □

4.6.2 Proving Theorem 4.2.3

Now we are fully equipped for the proof of our main result. We want to show that, assuming

$$\mathrm{h}(\alpha) \geq 200\varepsilon^{-2}mn^6(\mathrm{h}_p(F)+5), \tag{4.42}$$

we have

$$\begin{aligned}\left|\tfrac{\lgcd(\alpha,\beta)}{r} - \tfrac{\mathrm{h}(\alpha)}{n}\right| \leq \tfrac{1}{r}(\varepsilon\mathrm{h}(\alpha) + 4000\varepsilon^{-1}n^4(\mathrm{h}_p(F) &+ \log(mn)+1)+\\ &+ 30n^2m(\mathrm{h}_p(F)+\log(nm))).\end{aligned} \tag{4.43}$$

Write $F(X,Y) = f_n(X)Y^n + \cdots + f_0(X)$. According to Proposition 4.4.1 we have

$$F(x,Y) = f_n(x)(Y-y_1)\cdots(Y-y_n).$$

where

$$y_i = \sum_{k=\kappa_i}^{\infty} a_{ik}x^{k/e_i} \in \mathbb{K}((x^{1/e_i})) \qquad (i=1,\ldots,n).$$

We assume that $a_{i\kappa_i} \neq 0$ for $i=1,\ldots,n$, so that $\kappa_i/e_i = \nu_x(y_i)$.

Denoting by ℓ the number of indexes i such that $\kappa_i > 0$, we may assume that $\kappa_1,\ldots,\kappa_\ell > 0$ and $\kappa_{\ell+1},\ldots,\kappa_n \leq 0$. Proposition 4.4.2 implies that

$$r = \sum_{i=1}^{\ell} \min\{1, \kappa_i/e_i\}, \tag{4.44}$$

and Proposition 4.4.3 implies that $f_\ell(0) \neq 0$. We may normalize polynomial $F(X,Y)$ to have

$$f_\ell(0) = 1.$$

In particular, $|F|_v \geq 1$ for every $v \in M_{\mathbb{K}}$, where $|F|_v$ denotes the maximum of v-adic norms of the coefficients of F, and also $\mathrm{h}_p(F) = \mathrm{h}_a(F)$.

Set $E = \operatorname{lcm}(e_1,\ldots,e_\ell)$ and fix an E-th root $\alpha^{1/E}$. This fixes uniquely the roots $\alpha^{1/e_1},\ldots,\alpha^{1/e_\ell}$. Extending the field $\mathbb{K}$ we may assume that the coefficients of the series $y_1,\ldots,y_\ell$ belong to $\mathbb{K}$, and the same is true for $\alpha^{1/E}$ (and hence for $\alpha^{1/e_1},\ldots,\alpha^{1/e_\ell}$ as well). Having fixed the root $\alpha^{1/e_i} \in \mathbb{K}$, we may define v-adic convergence of y_i at α, see Subsection 4.5.6.

Extending further the field $\mathbb{K}$, we may assume that it contains all the roots of the polynomial $F(\alpha,Y)$. Hence, if one of the series $y_1,\ldots,y_\ell$ converges v-adically at α (and if the convergence is absolute in the archimedean case), then the sum must belong to $\mathbb{K}$.

Consider the following subsets of $M_{\mathbb{K}}$:

$$\begin{aligned} S &= \{v \in M_K : |\alpha|_v < 1\},\\ T_i &= \{v \in S : \text{the series } y_i \text{ converges } v\text{-adically to } \beta \text{ at } \alpha\} \qquad (i=1,\ldots,\ell).\end{aligned}$$

(These sets are not the same T_i as in Subsection 4.5.4!)

We have clearly $\lgcd(\alpha,\beta) = \lgcd_S(\alpha,\beta)$. If we manage to show that the sets T_i are pairwise disjoint, and that $\mathrm{h}_{S\setminus(T_1\cup\cdots\cup T_\ell)}(\beta)$ is "negligible", then joint application of Lemma 4.5.5,

Proposition 4.6.1 and identity (4.44) would prove Theorem 4.2.3. We will argue like this, only with the sets T_i replaced by slightly smaller subsets.

Let $\mathscr{A}_i = (A_{iv})_{v\in M_\mathbb{K}}$ be the $M_\mathbb{K}$-divisor for the series y_i given by Corollary 4.4.5. Define the $M_\mathbb{K}$-divisor $\mathscr{A} = (A_v)_{v\in M_\mathbb{K}}$ by

$$A_v = \max\{A_{1v}, \dots, A_{\ell v}\} \qquad (v \in M_\mathbb{K}).$$

We have clearly

$$\begin{aligned} |a_{ki}|_v &\le A_v^{k/e} \qquad (v \in M_\mathbb{K}, \quad 1 \le i \le \ell, \quad k \ge \kappa_i), \\ \mathrm{h}(\mathscr{A}) &\le \mathrm{h}(\mathscr{A}_1) + \cdots + \mathrm{h}(\mathscr{A}_\ell) \\ &\le 4n^2\mathrm{h}_p(F) + 3n^2\log(nm) + 10n^3. \end{aligned} \tag{4.45}$$

Now let $\widetilde{S}$ consist of $v \in S$ satisfying

$$|\alpha|_v < \begin{cases} |F|_v^{-n} A_v^{-1}, & v < \infty, \\ ((n+1)2^{n+3}|F|_v)^{-n} A_v^{-1}, & v \mid \infty, \end{cases} \tag{4.46}$$

and set $\widetilde{T}_i = T_i \cap \widetilde{S}$. (This is not the same $\widetilde{S}$ that in Subsection 4.5.4!) Clearly,

$$\begin{aligned} 0 \le \mathrm{lgcd}(\alpha,\beta) - \mathrm{lgcd}_{\widetilde{S}}(\alpha,\beta) &\le \mathrm{h}(\alpha) - \mathrm{h}_{\widetilde{S}}(\alpha) \\ &= \mathrm{h}_{S\setminus\widetilde{S}}(\alpha) \\ &\le \mathrm{h}(\mathscr{A}) + n\mathrm{h}_p(F) + \log((n+1)2^{n+3}) \\ &\le 5n^2\mathrm{h}_p(F) + 3n^2\log(nm) + 15n^3, \end{aligned} \tag{4.47}$$

$$\begin{aligned} 0 \le \mathrm{lgcd}_{T_i\setminus\widetilde{T}_i}(\alpha,\beta) &\le \mathrm{h}_{S\setminus\widetilde{S}}(\alpha) \\ &\le 5n^2\mathrm{h}_p(F) + 3n^2\log(nm) + 15n^3 \qquad (i = 1, \dots, \ell). \end{aligned} \tag{4.48}$$

Here we used the equality $\mathrm{h}_p(F) = \mathrm{h}_a(F)$.

Mention also that for $v \in \widetilde{S}$, we have $|\alpha|_v < A_v^{-1}$, which implies that the series $y_1, \dots, y_\ell$ converge v-adically at α in the completion $\mathbb{K}_v$, the convergence being absolute when v is archimedean. Hence, as we have seen above, the sum must belong to $\mathbb{K}$.

Proposition 4.6.2 *The sets $\widetilde{T}_1, \dots, \widetilde{T}_\ell$ pairwise disjoint. Furthermore, if $v \in \widetilde{S}$ but $v \notin \widetilde{T}_1 \cup \ldots \cup \widetilde{T}_\ell$ then*

$$|\beta|_v \ge \begin{cases} |F|_v^{-1}, & v < \infty, \\ ((n+1)2^{n+2}|F|_v)^{-1}, & v \mid \infty. \end{cases} \tag{4.49}$$

Proof The polynomial

$$Q(Y) = (Y - y_1)\cdots(Y - y_\ell) \in \mathbb{K}[[x^{1/E}]][Y].$$

divides $F(x,Y)$ in the ring $\mathbb{K}((x^{1/E}))[Y]$. By Gauß' Lemma, $Q(Y)$ divides $F(x,Y)$ in the ring $\mathbb{K}[[x^{1/E}]][Y]$ as well. Moreover, writing $F(x,Y) = Q(Y)U(Y)$ with

$$U(Y) = f_n(x)Y^{n-\ell} + u_{n-\ell-1}Y^{n-\ell-1} + \cdots + u_0 \in \mathbb{K}[[x^{1/E}]](Y),$$

the coefficients $u_0, \dots, u_{n-\ell-1}$ belong to the ring[2] $\mathbb{K}[x, y_1, \dots, y_\ell]$. Recall that for $v \in \widetilde{S}$ the series $y_1, \dots, y_\ell$ converge v-adically at α in the field $\mathbb{K}$, the convergence being absolute when v is archimedean. Hence so do the coefficients of U.

Fix $v \in \widetilde{S}$ and write

$$F(\alpha, Y) = (Y - y_1(\alpha)) \cdots (Y - y_\ell(\alpha))(f_n(\alpha)Y^{n-\ell} + u_{n-\ell-1}(\alpha)Y^{n-\ell-1} + \cdots + u_0(\alpha)),$$

where $y_1(\alpha), \dots, y_\ell(\alpha) \in \mathbb{K}$ the v-adic sum of the corresponding series at α, and similarly for $u_{n-\ell-1}(\alpha), \dots, u_0(\alpha)$. We claim that $F(\alpha, Y)$ is a separable polynomial of degree n; indeed, if this is not the case, then, as we have seen in Subsection 4.5.4, our α must satisfy (4.43), which contradicts (4.42).

Now if $v \in T_i \cap T_j$ for $i \neq j$ then $\beta = y_i(\alpha) = y_j(\alpha)$, and $F(\alpha, Y)$ must have β as a double root, a contradiction. This proves disjointedness of the sets $\widetilde{T}_i$.

Now assume that $v \in \widetilde{S}$ but $v \notin \widetilde{T}_1 \cup \ldots \cup \widetilde{T}_\ell$. Then none of the sums $y_1(\alpha), \dots, y_\ell(\alpha)$ is equal to β; in other words $y_1(\alpha), \dots, y_\ell(\alpha), \beta$ are $\ell + 1$ distinct roots of the polynomial

$$P(Y) = F(\alpha, Y) = f_n(\alpha)Y^n + \cdots + f_0(\alpha).$$

We are going to use Lemma 4.3.1. Since $f_\ell(0) = 1$ and

$$|\alpha|_v < \begin{cases} |F_v|^{-1}, & v < \infty, \\ (2|F|_v)^{-1}, & v \mid \infty, \end{cases}$$

we have

$$|f_\ell(\alpha)|_v \geq \begin{cases} 1, & v < \infty, \\ 1/2, & v \mid \infty, \end{cases}, \qquad |P|_v \leq \begin{cases} |F|_v, & v < \infty, \\ 2|F|_v, & v \mid \infty. \end{cases}$$

Now Lemma 4.3.1 implies that

$$\max\{|y_1(\alpha)|_v, \dots, |y_\ell(\alpha)|_v, |\beta|_v\} \geq \begin{cases} |F|_v^{-1}, & v < \infty, \\ ((n+1)2^{n+2}|F|_v)^{-1}, & v \mid \infty. \end{cases} \tag{4.50}$$

On the other hand, we may estimate $|y_i(\alpha)|_v$ from above using (4.45) and (4.46). In what follows we repeatedly use the inequality $e_i \leq n$. Since

$$|\alpha|_v < \begin{cases} A_v^{-1}, & v < \infty, \\ (2^{e_i}A_v)^{-1}, & v \mid \infty \end{cases} \qquad (i = 1, \dots, \ell),$$

we have

$$|a_k\alpha^{k/e_i}|_v < \begin{cases} (A_v|\alpha|_v)^{1/e_i}, & v < \infty, \\ (A_v|\alpha|_v)^{1/e_i} \cdot (1/2)^{k-1}, & v \mid \infty \end{cases} \qquad (k \geq 1, \quad i = 1, \dots, \ell),$$

which implies

$$|y_i(\alpha)|_v < \begin{cases} (A_v|\alpha|_v)^{1/e_i}, & v < \infty, \\ 2(A_v|\alpha|_v)^{1/e_i}, & v \mid \infty \end{cases} \qquad (i = 1, \dots, \ell).$$

[2]This is a consequence of the general algebraic property: let R be a commutative ring, R' a subring and $Q(Y), F(Y) \in R'[Y]$, the polynomial Q being monic; assume that $Q \mid F$ in $R[Y]$; then $Q \mid F$ in $R'[Y]$. Indeed, denoting by a the leading coefficient of F, the polynomial Q divides $G = F - aY^{\deg F - \deg Q}Q$ in $R[Y]$, and $\deg G < \deg F$, so by induction $Q \mid G$ in $R'[Y]$.

Now since

$$|\alpha|_v < \begin{cases} |F|_v^{-e_i} A_v^{-1}, & v < \infty, \\ ((n+1)2^{n+3}|F|_v)^{-e_i} A_v^{-1}, & v \mid \infty \end{cases} \qquad (i = 1, \dots, \ell),$$

we obtain finally

$$|y_i(\alpha)|_v < \begin{cases} |F|_v^{-1}, & v < \infty, \\ ((n+1)2^{n+2}|F|_v)^{-1}, & v \mid \infty \end{cases} \qquad (i = 1, \dots, \ell).$$

Compared with (4.50), this implies (4.49). The proposition is proved. □

An immediate consequence of the second statement of Proposition 4.6.2 is the estimate

$$\operatorname{lgcd}_{\widetilde{S} \smallsetminus (\widetilde{T}_1 \cup \dots \cup \widetilde{T}_\ell)} \leq \mathrm{h}_{\widetilde{S} \smallsetminus (\widetilde{T}_1 \cup \dots \cup \widetilde{T}_\ell)}(\beta) \leq \mathrm{h}_p(F) + \log((n+1)2^{n+2}) \tag{4.51}$$

(we again use $\mathrm{h}_a(F) = \mathrm{h}_p(F)$).

Now we collect everything together to prove Theorem 4.2.3. According to Lemma 4.5.5, condition (4.42) implies that

$$\left| \frac{\mathrm{h}(\alpha)}{n} - \mathrm{h}_{T_i}(\alpha) \right| \leq \varepsilon \mathrm{h}(\alpha) + 200\varepsilon^{-1} n^3 (\mathrm{h}_p(F) + 2\log(mn) + 10) \qquad (i = 1, \dots, \ell).$$

Combining this with Proposition 4.6.1 and estimate (4.48), we obtain

$$\begin{aligned} \left| \min\Big\{ \frac{\kappa_i}{e_i}, 1 \Big\} \frac{\mathrm{h}(\alpha)}{n} - \operatorname{lgcd}_{\widetilde{T}_i}(\alpha, \beta) \right| &\leq \varepsilon \mathrm{h}(\alpha) + 3000\varepsilon^{-1} n^3 (\mathrm{h}_p(F) + \log(mn) + 1) \\ &\quad + 30nm\mathrm{h}_p(F) + 30nm\log(nm). \qquad (i = 1, \dots, \ell). \end{aligned}$$

Summing up, using (4.44) and the disjointedness of the sets $\widetilde{T}_i$, we obtain

$$\begin{aligned} \left| r \frac{\mathrm{h}(\alpha)}{n} - \operatorname{lgcd}_{\widetilde{T}_1 \cup \dots \cup \widetilde{T}_\ell}(\alpha, \beta) \right| &\leq \varepsilon \mathrm{h}(\alpha) + 3000\varepsilon^{-1} n^4 (\mathrm{h}_p(F) + \log(mn) + 1) \\ &\quad + 30n^2 m\mathrm{h}_p(F) + 30n^2 m\log(nm). \end{aligned}$$

Finally, combining this with (4.47) and (4.51), we obtain (4.43). □

Chapter 5
A cyclotomic approach to Diophantine equations

5.1 Introduction

In this chapter, we generalize an approach initiated by Preda Mihăilescu in his article [Mihăilescu 2004] and developed since in [Mihăilescu 2008], [Bartolomé & Mihăilescu 2015] (Chapter 6) and still ongoing, to address certain types of Diophantine equations, of general form

$$\frac{x^p - y^p}{(x-y)^f} = B.z^q \text{ with } x, y \in \mathbb{Z},\ B \in \mathbb{Z},\ f \in \{0,1\},\ (p,q) \in \mathbb{Z}^2.$$

We will discuss some prerequisites and develop step by step the approach. In Chapter 6, the equation $X^n - 1 = B.Z^n$ is treated without any further reference to this chapter; the reader will however be able to follow the approach through the chapter.

5.2 Prerequisites

In this section, we will state, sometimes without proof, the technical knowledge required for our approach. We will use the following facts on cyclotomic fields (these basic facts can be found in any introductory book to algebraic number theory, for instance Chapter IV of [Lang 1994]):

- Let $n \in \mathbb{N}_{>1}$. The discriminant of the n-th cyclotomic extension is $\mathrm{Discr}(\mathbb{Q}(\zeta_n)) = (-1)^{\varphi(n)/2}\frac{n^{\varphi(n)}}{\prod_{p|n} p^{\varphi(n)/(p-1)}}$, so that a prime ramifies in $\mathbb{Q}(\zeta_n)$ if and only if it divides n.
- Decomposition of a prime in $\mathbb{Z}$ not dividing n: if f is the smallest positive integer such that $q^f \equiv 1 \mod n$, then, $q\mathbb{Z}[\zeta_n] = \mathfrak{q}_1 \cdots \mathfrak{q}_g$, where $g = \varphi(n)/f$ and each $\mathfrak{q}_i$ has residue class degree f. This means that a prime q splits completely in the n-th cyclotomic extension if and only if $q \equiv 1 \mod n$.
- An integral basis of the cyclotomic field $\mathbb{Q}(\zeta_n)$ is $(1, \zeta_n, \cdots, \zeta_n^{\varphi(n)-1})$.
- Finally, we know that the cyclotomic field $\mathbb{Q}(\zeta_n)$ is a cyclic Galois extension with Galois group isomorphic to $(\mathbb{Z}/n\mathbb{Z})^\times$ and that its ring of integers is $\mathbb{Z}[\zeta_n]$. The same holds for its maximal real subfield $\mathbb{K}^+ = \mathbb{Q}[\zeta_n + \zeta_n^{-1}]$, with ring of integers $\mathcal{O}_\mathbb{K}^+ = \mathbb{Z}[\zeta_n + \zeta_n^{-1}]$ and Galois group G^+.

Let p be a prime and let ζ be primitive p-th root of unity. We work in the p-th cyclotomic extension $\mathbb{K} = \mathbb{Q}(\zeta)$ because it factors the right hand side of our equation. The ring of integers is thus $\mathcal{O}_\mathbb{K} = \mathbb{Z}[\zeta]$. We let σ_c be the $\mathbb{Q}$-automorphism of $\mathbb{K}$ such that $\sigma_c(\zeta) = \zeta^c$. Also, let $G = \mathrm{Gal}\,(\mathbb{K}/\mathbb{Q}) = \{\sigma_c | c = 1, \cdots, p-1\}$ (and $G^+ = \{\sigma_c' : \zeta + \zeta^{-1} \mapsto \zeta^c + \zeta^{p-c}; c \in \{1, \cdots, (p-1)/2\}\}$). Finally, let $\lambda = 1 - \zeta$ be a generator of the ramified prime above p.

We will work with group rings of the form $\mathcal{R}[G]$ (or $\mathcal{R}[G^+]$), where $\mathcal{R}$ can be $\mathbb{Z}$ or $\mathbb{F}_q$ (where q is the other prime appearing in the equation). If $\mathcal{R} = \mathbb{F}_q$, then we consider the canonical lift of $\theta = \sum_{\sigma\in G} l_\sigma \sigma \in \mathbb{F}_q[G]$ to the unique $\Theta = \sum_{\sigma\in G} n_\sigma \sigma \in \mathbb{Z}[G]$ such that $0 \leq n_\sigma < q$ and $n_\sigma \equiv l_\sigma \mod q$. We use exponential notation for the action of elements of the group ring onto elements of the cyclotomic field: the action of $\sum_{c=1}^{p-1} n_c\sigma_c \in \mathbb{Z}[G]$ on the number $\epsilon \in \mathbb{K}$ is noted as $\epsilon^{\sum_{c=1}^{p-1} n_c\sigma_c}$ and means $\prod_{c=1}^{p-1} \sigma_c(\epsilon)^{n_c}$. This is why, for instance, the norm is written $\mathbf{N}_{\mathbb{K}/\mathbb{Q}} = \sum_{c=1}^{p-1} \sigma_c$.

In group ring theory, we will especially consider annihilators. We remind that if R is a ring, M is an R-module and we write exponentially the action of an element of R on an element of this module, then we call annihilator M the ideal $ann(M) = \{a \in R|$ for all $s \in M,\ s^a = 1\}$. The Stickelberger ideal is defined as follows: let the element $\vartheta = \frac{1}{p}\sum_{c=1}^{p-1} c \cdot \sigma_c^{-1} \in \mathbb{Q}[G]$ be the *Stickelberger* element; then the Stickelberger ideal is $I = (\vartheta\mathbb{Z}[G]) \cap \mathbb{Z}[G]$. It is of dimension $(p+1)/2$. When M is the class group of an abelian extension $\mathbb{K}$ of $\mathbb{Q}$, the Stickelberger ideal I annihilates of M in $\mathbb{Z}[G]$. That is, for any element $\Theta \in I$, for any ideal $\mathfrak{a} \subset \mathbb{K}$, the ideal $\mathfrak{a}^\Theta$ is principal. The proof can be found in [Stickelberger 1890] or in a modern form in [Washington 1997].

For $\theta = \sum_{\sigma\in G} n_\sigma\sigma \in \mathbb{Z}[G]$, we define the *absolute weight* of θ, $w(\theta) = \sum_{\sigma\in G} n_\sigma$. A notion specific to elements in the Stickelberger ideal is the *relative weight*: for $\theta = \sum_{\sigma\in G} n_\sigma\sigma \in I$ we have the relation $\theta + \jmath\theta = \varsigma(\theta) \cdot \mathbf{N}_{\mathbb{K}/\mathbb{Q}}$, where $\varsigma(\theta) \in \mathbb{Z}$ is called the *relative weight* of θ.

Now we remind a series of notions and properties related specifically to the Stickelberger ideal I: the interested reader can find a deeper treatise on the Stickelberger ideal in several standard books, for instance in [Washington 1997], [Ireland. & Rosen 1990] or [Jha 1992]. Let $\left(\psi_d = \sum_{\nu=1}^{p-1} \left(\lfloor (d+1)\nu/p \rfloor - \lfloor d\nu/p \rfloor\right) \sigma_\nu^{-1}\right)_{d=1}^{\frac{p-1}{2}} \in \mathbb{Z}[G]$ be the family of *Fueter* elements. Together with the norm, it consitutes a basis of the Stickelberger ideal as a $\mathbb{Z}$-module ([Fueter 1922]). Note that $\varsigma(\psi_d) = 1$ for all d. The Fuchsian elements are

$$\Theta_k = (k - \sigma_k) \cdot \vartheta = \sum_{c=1}^{p-1} \left\lfloor \frac{kc}{p} \right\rfloor \cdot \sigma_c^{-1}, \quad 1 \leq k \leq p.$$

They also generate I as a $\mathbb{Z}$ - module. Note that Θ_p is the norm, and that we have the following relationship between the Fueter and the Fuchsian elements:

$$\psi_1 = \Theta_2 \text{ and } \psi_k = \Theta_{k+1} - \Theta_k,\ k \geq 2$$

An element $\Theta = \sum_{\sigma\in G} n_\sigma\sigma \in \mathbb{Z}[G]$ is *positive* if $n_\sigma \geq 0$ for all $\sigma \in G$. We write $I^+ \subset I$ for the set of all positive elements of the Stickelberger ideal. They form a multiplicative and an additive semigroup.

There exists an additive map φ from I into $\{0, 1, \cdots, p-1\}$ such that, for any $\Theta \in I$, $\zeta^{\varphi(\Theta)} = \zeta^\Theta$. It is given by $\varphi : \theta = \sum_{c=1}^{p-1} n_c\sigma_c \mapsto \varphi(\theta)$, where $\varphi(\theta)$ is the element of $\{0, 1, \cdots, p-1\}$ such that $\varphi(\theta) \equiv \sum_{c=1}^{p-1} cn_c \mod p$. It is called the Fermat quotient map, and it verifies, for any Fuchsian element Θ_n, $\varphi(\Theta_n) = \frac{n^p - n}{p}$. The kernel of the Fermat quotient map is $I_f = \{\theta \in I : \zeta^\theta = 1\}$ (the *Fermat module*). We will note $I_f^+ = \{\theta = \sum_{c=1}^{p-1} n_c\sigma_c \in I_f : n_c \geq 0\}$.

Proposition 5.2.1 *The Fermat quotient map enjoys the properties:*

$$\begin{aligned}\zeta^{\theta} &= \zeta^{\varphi(\theta)},\\ (1+\zeta)^{\theta} &= \zeta^{\varphi(\theta)/2},\\ (1-\zeta)^{\theta} &= \zeta^{\varphi(\theta)/2}\cdot\left(\left(\frac{-1}{p}\right)p\right)^{\varsigma(\theta)/2},\end{aligned}$$

where $\left(\frac{-1}{p}\right)$ *is the Legendre symbol.*

Note that for $\theta \in I$ with $\varsigma(\theta) = 2$ we have $(1-\zeta)^{2\theta} = \zeta^{\varphi(\theta)} \cdot p^2$.

Theorem 5.2.2 (Jacobi sums and the Stickelberger ideal) *Let l be an odd prime number such that $l \equiv 1 \mod p$, χ be a Dirichlet character of order p, and $\tau(\chi) = \sum_{c\in\mathbb{F}_l}\chi(c).\zeta_l^c \in \mathbb{Q}(\zeta,\zeta_l)$ be its Gauß sum. Let $J(\chi^a,\chi^b) = -\sum_{x\in\mathbb{Z}/l\mathbb{Z}}\chi^a(x)\cdot\chi^b(1-x) \in \mathbb{Q}(\zeta,\zeta_l)$ be the Jacobi sum $\left(J(\chi^a,\chi^b) = -\frac{\tau(\chi^a)\cdot\tau(\chi^b)}{\tau(\chi^{a+b})}\right)$. Then, the ideal $(J(\chi^a,\chi^b))$ can be expressed as the action of $\psi(a,b) = \sum_{c=1}^{p-1}\left(\left\lfloor\frac{c(a+b)}{p}\right\rfloor - \left\lfloor\frac{ca}{p}\right\rfloor - \left\lfloor\frac{cb}{b}\right\rfloor\right).\sigma_c^{-1}$ over any ideal $\mathfrak{L}$ of $\mathbb{Z}[\zeta_l]$ above the conductor l of χ.*

The proof of this theorem can be found in [Ireland. & Rosen 1990][Chapter 14, Section 4]. We define by multiplicativity the *set of Jacobi integers* to be the multiplicative semigroup generated by the Jacobi sums $\mathbf{J} \subset \mathbb{Z}[\zeta]$. Let's also define $\mathfrak{J} = \{(\mathbf{j}) : \mathbf{j} \in \mathbf{J}\}$ the subset of principal ideals generated by Jacobi integers. Let $\mathfrak{A} \subset \mathbb{Z}[\zeta]$ be an ideal with $\mathbf{N}(\mathfrak{A}) = t$, such that t factors into powers of primes $\ell \equiv 1 \mod p$. Then Stickelberger's theorem implies in particular that $\mathfrak{A}^{\Theta} \in \mathfrak{J}, \quad \forall\Theta \in I$.

The following lemma is a special case adaptation of [Jha 1992][Proposition 1.2], which relates $\mathfrak{J}$ to $\mathbf{J}$. The adaptation was provided in [Mihăilescu 2008].

Lemma 5.2.3 *Let ι be the natural map $\iota : \mathbf{J} \to \mathfrak{J}$ given by $\mathbf{j} \mapsto (\mathbf{j})$. Then ι is injective. In particular, a principal ideal can be generated by at most one Jacobi integer and if for some $\alpha \in \mathbb{Z}[\zeta]$ with $\alpha\cdot\overline{\alpha} \in \mathbb{Z}$, the equality $(\alpha) = \mathfrak{j} \in \mathfrak{J}$ holds, then there is a unique Jacobi integer $\mathbf{a}$ with:*

$$\alpha = \pm\zeta^n \cdot \mathbf{a}. \tag{5.1}$$

for some $n \in \mathbb{Z}$.

Proof. Let α generate the principal ideal $\mathfrak{j} \in \mathfrak{J}$ and let $\mathbf{a} \in \mathbf{J}$ with $(\alpha) = (\mathbf{a})$: such a Jacobi integer exists by definition of $\mathfrak{J}$. The principal ideals being equal, there is a unit $\varepsilon \in \mathbb{Z}[\zeta]$ such that $\alpha = \varepsilon\cdot\mathbf{a}$. Furthermore, $\mathbf{a}\cdot\overline{\mathbf{a}} \in \mathbb{N}$ follows by multiplicativity from the property of Gauß sums and since $\alpha\cdot\overline{\alpha} \in \mathbb{Z}$, it follows that $\varepsilon\cdot\overline{\varepsilon} = 1$. By Kronecker's unit theorem, ε is a root of unity. This proves (5.1).

We still have to prove that the Jacobi integer $\mathbf{a}$ is unique. Iwasawa shows in [Iwasawa 1975] (see also [Ireland. & Rosen 1990][p. 226, Ex. 13]) that Jacobi integers $\mathbf{j}$ verify

$$\mathbf{j} \equiv 1 \mod (1-\zeta)^2\mathbb{Z}[\zeta]. \tag{5.2}$$

This property is useful for establishing the power n in (5.1). In particular, assuming there is a second Jacobi integer **a'**, with $(\alpha) = (\mathbf{a'})$, we would have by (5.1) that $\alpha = \pm\zeta^{n'}\mathbf{a'}$ for some $n' \in \{0, \cdots, p-1\}$. By Iwasawa's relation, $\alpha \equiv s\zeta^n \equiv s'\zeta^{n'} \mod \lambda^2$, where s, s' are the implicit signs for $\mathbf{a}, \mathbf{a'}$, and $\lambda = 1-\zeta$. In particular $s\zeta^n \equiv s \equiv s'\zeta^{n'} \equiv s' \mod \lambda$ and thus $s = s'$. Also, $1-\zeta^n \equiv n\lambda \equiv 1-\zeta^{n'} \equiv n'\lambda \mod \lambda^2$, so $n \equiv n' \mod p$. Consequently, $\mathbf{a} = \mathbf{a'}$, which completes the proof. □

We will be interested in series expansions of algebraic numbers $\mu \in \mathbb{Z}[\zeta]$ which verify $\mu^q = (1+\zeta/z)^{\Theta}$, where q is again the prime appearing in the equation and $\Theta = \sum_{\sigma\in G} n_\sigma \sigma \in \mathbb{F}_q[G]$. That is, $\mu = (1+\zeta/z)^{\Theta/q}$. For that, we will use properties of the generalized binomial series $f(z) = \sum_{k\geq 0} \binom{\alpha}{k} z^k$, where $\alpha, z \in \mathbb{C}$ and $\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}$. We know that if $|z| < 1$, this series converges absolutely to $(1+z)^\alpha$. The next definition links binomial series and the action of group ring elements on numbers:

Definition 5.2.4 (Elementary q-th root power series) *Let $\mathbb{Q}(\zeta)[[T]]$ be the ring of formal power series over the $p-$*th *cyclotomic field. The* elementary $q-$th root power *series is defined as:*

$$f(T) = (1+\zeta T)^{1/q} = \sum_{k\geq 0} \binom{1/q}{k} \cdot (\zeta \cdot T)^k = \sum_{k\geq 0} a_k \cdot T^k \in \mathbb{Q}(\zeta)[[T]],$$

where $a_k = \binom{1/q}{k} \cdot \zeta^k$. Let $\sigma \in G$; then, when σ acts on $f(T)$, it acts on ζ and not upon the formal variable T.

We can thus generalize the previous definition: let $\Theta = \sum_{\sigma\in G} n_\sigma \sigma \in \mathbb{Z}[G]$. Then:

$$f[\Theta](T) = (1+\zeta T)^{\Theta/q} = \prod_{c\in P} f[n_c\sigma_c](T) = \prod_{c\in P} \left(\sum_{k\geq 0} \binom{n_c/q}{k} (\zeta^c \cdot T)^k \right). \tag{5.3}$$

Since $f[\Theta] \in \mathbb{Q}(\zeta)[[T]]$, it also has a development as a simple power series. We shall write this development as

$$\begin{aligned} f[\Theta](T) &= \sum_{k\geq 0} a_k(\Theta) \cdot T^k \quad \textit{and let} \\ \tilde{b}_k(\Theta) &= a_k(\Theta) \cdot \left(q^k \cdot k!\right). \end{aligned} \tag{5.4}$$

We next prove a general lemma, and then show two important properties of the coefficients of the $q-$th power series. These results have been used in [Mihăilescu 2004] and in [Mihăilescu 2008].

Lemma 5.2.5 *For $k \geq 0$, let $E(k) = k + v_q(k!)$. Then $E(k)$ is strictly monotonous and verifies*

$$E(k) < k \cdot \frac{q}{q-1}. \tag{5.5}$$

Furthermore

$$q^{E(k)} \cdot \binom{n/q}{k} \in \mathbb{Z} \quad \textit{and} \quad E(k) = -v_q\left(\binom{n/q}{k}\right), \tag{5.6}$$

this last equality holding only if $n \not\equiv 0 \mod q$.

Proof. Since $k+1 > k$ and $v_q((k+1)!) \geq v_q(\ k!\ )$, the function $E(k)$ is strictly monotonous. We now use Legendre's theorem, so $v_q(\ k!\ ) = \sum_{i>0} \lfloor k/q^i \rfloor < k/(q-1)$; this leads to the upper bound for $E(k)$.

If $n \equiv 0 \mod q$, then the first part of (5.6) is obvious. Let's now assume that $n \not\equiv 0 \mod q$. Developing $\binom{n/q}{k}$, we find $\binom{n/q}{k} = \frac{n(n-q)\cdots(n-(k-1)q)}{q^k k!}$. The numerator is not divisible by q (as we assumed $n \not\equiv 0 \mod q$), thus obviously $v_q\left(\binom{n/q}{k}\right) = -k - v_q(k!) = -E[k]$, which proves the second part of (5.6).

Let's compute the valuation v_ℓ of the numerator of $\binom{n/q}{k}$ for any prime $\ell \nmid q$. The pigeon hole principle shows that the number of multiples of ℓ^i in the above numerator is $\lfloor k/\ell^i \rfloor$. Adding up we find:

$$v_\ell\left(n \cdot (n-q) \ldots (n-(k-1)q)\right) \geq \sum_{i>0} \lfloor k/\ell^i \rfloor = v_\ell(k!).$$

Therefore, $k!$ divides $n(n-q)\cdots(n-(k-1)q)$ and $q^{E[k]}\binom{n/q}{k} \in \mathbb{Z}$, which proves the first part of (5.6). □

For $\Theta = \sum_{c \in P} n_c \cdot \sigma_c \in \mathbb{Z}[G]$ we define the additive map $\rho : \mathbb{Z}[G] \to \mathbb{Z}[\zeta]$

$$\rho \ : \ \Theta \mapsto -\sum_{\sigma \in G} n_c \cdot \zeta^\sigma.$$

Corollary 5.2.6 *Writing $\tilde{b}_k(n\sigma) = (q^k \cdot k!) \cdot a_k(n\sigma)$, the following congruence holds:*

$$\tilde{b}_k(n\sigma) \equiv (-n\zeta^\sigma)^k = \rho(n\sigma)^k \mod q\mathbb{Z}[\zeta]. \tag{5.7}$$

Proof. We know that $\tilde{b}_k(n\sigma) = (n \cdot (n-q) \ldots (n-(k-1)q)) \cdot (-\zeta)^{k\sigma}$. It is then obvious, using Equation (5.6), that

$$\tilde{b}_k(n\sigma) \equiv (-n\zeta^\sigma)^k \mod q\mathbb{Z}[\zeta].$$

□

The arithmetic properties of the coefficients $a_k, \tilde{b}_k$ are given by the following:

Lemma 5.2.7 *The coefficients $a_k(\Theta), \tilde{b}_k(\Theta)$ of the series $f[\Theta] \in \mathbb{Q}(\zeta)[[T]]$ have the properties:*

1. *Both $a_k, \tilde{b}_k$ commute with the Galois action, i.e.: $a_k(\sigma\Theta) = (a_k(\Theta))^\sigma$ and the same for $\tilde{b}_k$.*
2. *$q^{E(k)} \cdot a_k(\Theta) \in \mathbb{Z}[\zeta]$.*
3. *On the coefficients $\tilde{b}_k$:*

$$\tilde{b}_k(\Theta) \in \mathbb{Z}[\zeta] \text{ and } \tilde{b}_k(\Theta) \equiv \rho(\Theta)^k \mod q \cdot \mathbb{Z}[\zeta]. \tag{5.8}$$

4. If $\Theta = \jmath\Theta$, then $a_k(\Theta) \in \mathbb{R}$.

Proof. Property 1. follows from the definition of $f[\Theta]$ (5.3) and the fact that σ acts on ζ but not on the formal parameter T. Suppose that $\Theta = \Theta_1 + \Theta_2$; then the coefficients of $f[\Theta]$ are derived from the ones of $f[\Theta_i], i = 1, 2$ as follows:

$$a_k(\Theta) \;=\; a_k(\Theta_1 + \Theta_2) \;=\; \sum_{l=0}^{k} a_l(\Theta_1) \cdot a_{(k-l)}(\Theta_2) \tag{5.9}$$

$$\tilde{b}_k(\Theta) \;=\; \tilde{b}_k(\Theta_1 + \Theta_2) \;=\; \sum_{l=0}^{k} \left(\begin{array}{c} k \\ l \end{array} \right) \cdot \left(\tilde{b}_l(\Theta_1) \cdot \tilde{b}_{(k-l)}(\Theta_2) \right). \tag{5.10}$$

We have

$$v_p\left(a_l(\Theta_1) \cdot a_{(k-l)}(\Theta_2)\right) = -\left(E(l) + E(k-l)\right) = -E(k) + v_q\left(\left(\begin{array}{c} k \\ l \end{array} \right)\right) \geq -E(k);$$

thus $v_q(a_k(\Theta)) \geq -E(k)$, by induction on the canonical weights $w(\Theta_1), w(\Theta_2)$. This proves 2. By Equation (5.7), we have $\tilde{b}_k(\Theta) \equiv \rho(\Theta)^k \mod q\mathbb{Z}[\zeta]$ for $\Theta = n\sigma$, $\forall \sigma \in G$. Assume that Equation (5.8) holds for $\Theta_i, i = 1, 2$. Then by the previous identity for $\tilde{b}_k(\Theta_1 + \Theta_2)$, we have

$$\begin{aligned} \tilde{b}_k(\Theta_1 + \Theta_2) \;&=\; \sum_{l=0}^{k} \left(\begin{array}{c} k \\ l \end{array} \right) \cdot \left(\tilde{b}_l(\Theta_1) \cdot \tilde{b}_{(k-l)}(\Theta_2) \right) \\ &\equiv\; \sum_{l=0}^{k} \left(\begin{array}{c} k \\ l \end{array} \right) \rho(\Theta_1)^l \cdot \rho(\Theta_2)^{k-l} = \rho(\Theta_1 + \Theta_2)^k \mod q\mathbb{Z}[\zeta]. \end{aligned}$$

Relation (5.8) follows from this by induction on the weights of Θ_1, Θ_2. □

The coefficients $\tilde{b}_n[\Theta]$ have the advantage of being algebraic integers. Writing $a_k[\Theta] = \frac{\tilde{b}_k[\Theta]}{k! q^k}$, we see from the above, keeping the same definition for E, that the denominator and numerator have a massive common factor. Therefore we shall define

$$b_k[\Theta] \;=\; a_k[\Theta] \cdot q^{E(k)} \in \mathbb{Z}[\zeta], \quad E(k) = k + v_q(k!). \tag{5.11}$$

In particular,

$$b_k[\Theta] \cdot \frac{k!}{q^{v_q(k!)}} = a_k[\Theta] \cdot (k! \cdot q^k),$$

and for $k < q$ we have

$$b_k[\Theta] \;\equiv\; \rho[\Theta]^k / k! \bmod q\mathbb{Z}[\zeta]. \tag{5.12}$$

We now give an estimate of the error term in the evaluation of the general series $f[\Theta](1/z)$. We know that the exponent in our function is not integer. The convergence radius of $f(T)$ being one, it follows by multiplicativity, that the series $f[\Theta](T)$ also have the same domain of convergence, for all $\Theta \in \mathbb{Z}[G]$. Let

$$S_m(\Theta; T) = \sum_{k=0}^{m} a_k(\Theta) \cdot T^m$$

be the $m-$th partial sum of $f[\Theta]$ and $R_m(\Theta;T) = f[\Theta] - S_m(\Theta;T)$ the remainder term. We estimate this remainder, when T is replaced by a complex number $|z| < 1$ (inspired by [Bilu 2004][Proposition 8.2.1]).

Lemma 5.2.8 *Let $\Theta = \sum_{\sigma\in G} n_\sigma \sigma \in \mathbb{Z}[G]$ have weight $w(\Theta) = H$. If $z \in \mathbb{C}$, $|z| < 1$, then*

$$|f[\Theta](z) - S_m(\Theta;z)| \leq \binom{-H/q}{m+1} \cdot \frac{|z|^{m+1}}{(1-|z|)^{m+1+h}}. \tag{5.13}$$

Furthermore, the coefficients $b_k[\Theta]$ are bounded by:

$$|b_k[\Theta]| \leq q^{E(k)} \cdot \binom{k+t}{k}, \qquad \text{where } t = \left\lfloor \frac{H}{q} \right\rfloor. \tag{5.14}$$

Proof. For this proof we remind the notion of dominance of series. A power series $f(T) = \sum_{k=0}^{\infty} a_k T^k$ with complex coefficients is *dominated* by the series $g(T) = \sum_{k=0}^{\infty} A_k T^k$ with non-negative real coefficients if $|a_k| \leq A_k$ for $k = 0, 1, \ldots$; if this is the case, we write $f \ll g$. The relation of dominance is preserved by addition and multiplication of power series.

Let r be a real number, and s a complex number satisfying $|s| \leq 1$. Then, the binomial series $(1+sT)^r = \sum_{k=0}^{\infty} \binom{r}{k} s^k T^k$ is dominated by $(1-T)^{-|r|} = \sum_{k=0}^{\infty} (-1)^k \binom{-|r|}{k} T^k$. Indeed, the coefficients of the latter series are positive and $\left|\binom{r}{k}\right| \leq \left|\binom{-|r|}{k}\right|$.

It follows that $f[\Theta](T) \ll (1-T)^{-H/p}$. This together with the definition of $b_n[\Theta]$ and the properties of generalized binomial numbers yield (5.14).

From common remainder estimates for Taylor series, we obtain the following:

$$\begin{aligned} |f[\Theta](z) - S_m(\Theta;z)| &\leq \left|(1-|z|)^{-H/q} - S_m(|z|)\right| \\ &\leq \sup_{0\leq|\xi|\leq|z|} \left| \left(\frac{d^{m+1}(1-T)^{-H}}{dT^{m+1}}\bigg|_{T=\xi} \right) \right| \frac{|z|^{m+1}}{(m+1)!} \\ &= \binom{-H/q}{m+1} \cdot \frac{|z|^{m+1}}{(1-|z|)^{H/q+m+1}}, \end{aligned}$$

as claimed. □

We will use the Voronoi identities – see [Jha 1992][Lemma 1.0] –, which we remind here for convenience:

Lemma 5.2.9 *Let m be an even integer such that $2 \leq m \leq n-1$. Let a be an integer, coprime to n. Then*

$$a^m \sum_{j=1}^{n-1} \left\lfloor \frac{aj}{n} \right\rfloor j^{m-1} \equiv \frac{(a^{m+1}-a)B_m}{m} \mod n, \tag{5.15}$$

where B_m is the m-th Bernoulli number. In particular, for $m = n-1$, we get

$$\sum_{j=1}^{n-1} \left\lfloor \frac{aj}{n} \right\rfloor j^{n-2} \equiv \frac{a^n - a}{n} \mod n,$$

which is the Fermat quotient map of the a-th Fuchsian element, $\varphi(\Theta_a)$.

5.3 The binomial cyclotomic series approach

We remind that we work on equations of the type:

$$\frac{x^p - y^p}{(x-y)^f} = B.z^q \text{ with } x, y \in \mathbb{Z},\ B \in \mathbb{Z},\ f \in \{0,1\},\ (p,q) \in \mathbb{Z}^2. \tag{5.16}$$

This approach has been applied, for instance, to the case of Catalan's conjecture (cf. [Mihăilescu 2004]) where $f = 0,\ y = 1,\ B = 1$, to the case of Diagonal Nagell-Ljunggren (cf. [Mihăilescu 2008]), where $f = 1,\ y = 1,\ B = 1,\ q = p$ or to the case of binomial Thue (Chapter 6), where $f = 0,\ y = 1,\ q = p$.

We will be using the same notation than and the results stated in the previous section. As usual, we start by assuming that there is a solution (x, y, z, p, q, f, B) to Equation (5.16). In this section, we first describe the binomial cyclotomic series approach step by step in the general context of Equation (5.16). We will then show how this approach has been applied in three different cases: the conjecture of Catalan ([Mihăilescu 2004]), conditions for the diagonal Nagell-Ljunggren equation ([Mihăilescu 2008]) and binary Thue ([Bartolomé & Mihăilescu 2015], Chapter 6).

A very simple but useful lemma we will be using is Euler's:

Lemma 5.3.1 (Euler) *Let p be an odd prime, and $(x, y) \in \mathcal{O}_{\mathbb{K}}^2$ such that $(x, y) = 1$. Then, $\delta = \left(\frac{x^p-y^p}{x-y}, x - y\right)$ divides p, and if $\delta = p$, then p^2 does not divide $\frac{x^p-y^p}{x-y}$.*

Proof. We can rewrite $\frac{x^p-y^p}{x-y} = \frac{((x-y)+y)^p-y^p}{x-y} = \sum_{k=1}^{p-1} \binom{p}{k} (x-y)^{k-1}y^{p-k} + (x-y)^{p-1}$. Therefore, $\frac{x^p-y^p}{x-y} \equiv (x-y)^{p-1} \mod p$, and if $\frac{x^p-y^p}{x-y} \equiv 0 \mod p$, then $(x-y) \equiv 0 \mod p$. We can also rewrite $\frac{x^p-y^p}{x-y} = py^{p-1} + (x-y)\sum_{k=2}^{p} \binom{p}{k} (x-y)^{k-2}y^{p-k}$, which implies that $\delta | py^{p-1}$. As $(x, y) = 1$, $(\delta, y) = 1$ and thus $\delta | p$. Together with the previous result, this proves that $\delta \in \{1, p\}$. Finally, writing $\frac{x^p-y^p}{x-y} = py^{p-1} + (x-y)\left(\sum_{k=2}^{p-1} \binom{p}{k} (x-y)^{k-2}y^{p-k} + (x-y)^{p-2}\right)$, we see that $\frac{x^p-y^p}{x-y} \equiv py^{p-1} \mod p^2$ and the lemma is proved. □

A. Kummer was the first to thoroughly study a similar equation in cyclotomic extensions. The results he obtained were the most advanced results on Fermat's Last Equation known at that time. The splitting of the initial equation is also the point of departure of the binomial cyclotomic series approach. We work in the p-th cyclotomic extension of $\mathbb{Q}$, $\mathbb{K} = \mathbb{Q}(\zeta)$ (where ζ is a primitive p-th root of unity), with ring of integers $\mathcal{O}_{\mathbb{K}} = \mathbb{Z}[\zeta]$ and Galois group $G = \mathrm{Gal}\,(\mathbb{K}/\mathbb{Q}) = \{\sigma_c : \zeta \mapsto \zeta^c\}$. Indeed, in this extension, the right hand side of Equation (5.16) factors. The equation naturally yields the *characteristic* algebraic integer, let it be ([Lennon & McCartney 1970]) $\alpha \in \mathcal{O}_{\mathbb{K}}$. This algebraic integer (in $\mathbb{K}$) is such that our initial equation can be written $\mathbb{N}_{\mathbb{K}/\mathbb{Q}}(\alpha) = z^q$. We choose the algebraic integer α such that it is not divisible by λ (and so, neither are its conjugates). For instance, if we assumed $f = 1$ and $B = 1$ in Equation (5.16), we might choose $\alpha = (x - \zeta y)/(1-\zeta)^e$, where $e = 1$ if $(\frac{x^p-y^p}{x-y}, x - y) = p$ (cf. Lemma 5.3.1) and $e = 0$ otherwise. This characteristic algebraic integer α gives naturally rise to the *characteristic* ideal $\mathfrak{A} = (\alpha, z) \subset \mathcal{O}_{\mathbb{K}}$ such that $\mathfrak{A}^q = (\alpha)$. *To see this, with α chosen as above ($\alpha = (x - \zeta y)/(1-\zeta)^e$), we*

have, for $c \neq d$, $(1-\zeta^d)^e\alpha^{\sigma_d}-(1-\zeta^c)^e\alpha^{\sigma_c} = (\zeta^d-\zeta^c)y$ *and thus* $(y\lambda) \subset (\alpha^{\sigma_c}, \alpha^{\sigma_d})$. *We also have* $\bar{\zeta}^d(1-\zeta^d)^e\alpha^{\sigma_d} - \bar{\zeta}^c(1-\zeta^c)^e\alpha^{\sigma_c} = x(\zeta^{-d}-\zeta^{-c})$, *and thus* $(x\lambda) \subset (\alpha^{\sigma_c}, \alpha^{\sigma_d})$. *We have chosen* α *such that neither* α *nor its conjugates are divisible by* $\mathfrak{p}$, *and* $(x, y) = 1$. *However, the above linear combination shows that* $(\alpha^{\sigma_c}, \alpha^{\sigma_d})$ *divides the prime* $\mathfrak{p}$. *Thus, the* α^{σ_c} *are coprime. Given that* $\mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\alpha) = \prod_{\sigma\in G}\alpha^\sigma = z^q$, *each ideal* (α^σ) *is the* q*-th power of an ideal.* Considering the generators of $(\alpha, z)^q$ and the fact that $\left(\mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\alpha)/\alpha, \alpha\right) = 1$, we see that $\mathfrak{A} = (\alpha, z)$.

B. Let $\theta = \sum_{\sigma\in G} n_\sigma\sigma \in \mathcal{R}[G]$ be an annihilator of $\mathfrak{A}$ in the class group of $\mathbb{K}$ (remember that $\mathcal{R} = \mathbb{F}_q$ or $\mathbb{Z}$). Thus, there exists $\nu \in \mathcal{O}_\mathbb{K}$ such that $\mathfrak{A}^\theta = (\nu)$, and thus $\mathfrak{A}^{q\theta} = (\alpha^\theta) = (\nu^q)$. Therefore, there exists a unit $\varepsilon_\theta \in \mathcal{O}_\mathbb{K}^\times$ such that $\alpha^\theta = \varepsilon_\theta \cdot \nu^q$. One first objective is to have a pure q-th equation, and thus to eliminate the unit ε_θ. For that, we have two main strategies, the *plus* and the *minus* stategy, depending on the situation. We say that an element θ of $\mathcal{R}[G]$ belongs to the *plus* part of the group ring if $\theta \in \mathcal{R}[G]^+ = (1 + \jmath)\mathcal{R}[G]$, where $\jmath$ denotes complex conjugation. And an element θ of $\mathcal{R}[G]$ belongs to the *minus* part of the group ring if $\theta \in \mathcal{R}[G]^- = (1 - \jmath)\mathcal{R}[G]$.

Plus Let $\theta' \in \mathcal{R}[G]^+$ annihilate $\mathfrak{A}^{1+\jmath} \in \mathcal{O}_\mathbb{K}^+$ in the class group of $\mathbb{K}^+$, and $\theta = (1 + \jmath)\theta' \in \mathcal{R}[G]$ (where $(1 + \jmath)\theta'$ is the composition of $(1 + \jmath)$ with a lift of θ' to $\mathbb{Z}[G]$). Then again, there exists $\nu \in \mathcal{O}_\mathbb{K}^+$ and a real unit $\varepsilon_\theta \in \mathcal{O}_\mathbb{K}^{+\times}$ such that $\alpha^\theta = (\alpha.\bar{\alpha})^{\theta'} = \varepsilon_\theta \cdot \nu^q$. The approach can be completed if the properties of the solution to our initial equation allow a choice of θ such that $\varepsilon_\theta = 1$.

Minus Alternatively, one chooses an annihilator $\theta' \in \mathcal{R}[G]^-$ of $\mathfrak{A}^{1-\jmath}$. Proceeding as previously, let $\theta = (1 - \jmath)\theta' \in \mathcal{R}[G]$. Thus, $\alpha^\theta = \alpha^{(1-\jmath)\theta'} = \left(\frac{\alpha}{\bar{\alpha}}\right)^{\theta'} = \frac{\varepsilon_\theta}{\bar{\varepsilon}_\theta} \cdot \left(\frac{\nu}{\bar{\nu}}\right)^q$. We see that $\frac{\varepsilon_\theta}{\bar{\varepsilon}_\theta}$ is a root of unity in $\mathbb{K}$ (it is an algebraic integer, all of whose conjugates have absolute value 1 and thus it is a root of unity by Kronecker's theorem), and as such belongs to $< -\zeta >$.

C. If we rewrite the previous α as $\alpha = \frac{x}{(1-\zeta)^e}(1 - \zeta\left(\frac{y}{x}\right))$, and in a more generic form as $\alpha = \upsilon(1 + \varpi T)$, where $\upsilon \in \mathcal{O}_\mathbb{K}^\times, \varpi \in \mathcal{O}_\mathbb{K}$, we obtain that:

Plus $\nu = \upsilon^{\theta/q} \cdot (1 + \varpi T)^{\theta/q}$. We can rationalize the factor $\upsilon^{\theta/q}$ by selecting θ such that $|w(\theta)|$ is a multiple of q,

Minus $\frac{\nu}{\bar{\nu}} = \left(\frac{\varepsilon_\theta}{\bar{\varepsilon}_\theta}\right)^{-1/q} \cdot \left(\frac{\upsilon}{\bar{\upsilon}}\right)^{\theta/q} \cdot \left(\frac{1+\varpi T}{1+\bar{\varpi}T}\right)^{\theta/q}$. We note that $\frac{\varepsilon_\theta}{\bar{\varepsilon}_\theta}$, as well as $\frac{\upsilon}{\bar{\upsilon}}$ are both roots of unity, and since any root of unity in $\mathbb{K}$ is a q-th power, they are q-th powers in $\mathbb{K}$.

We will consider the formal series development $f(\theta, T, \varpi) \in \mathbb{K}[[T]]$ of ν in the *plus* case, and of $\nu/\bar{\nu}$ in the *minus* case: $f(\theta, T, \varpi) = (1 + \varpi T)^{\theta/q}$ in the *plus* case, and $f(\theta, T, \varpi) = (1 + \varpi T)^{\theta/q} \cdot (1 + \bar{\varpi}T)^{-\theta/q}$ in the minus case. This series converges in a local or global topology and in whatever local or global topology it converges, the limit verifies $\nu^q = f^q(\theta, T, \varpi)$ in the *plus* case, and $\left(\frac{\nu}{\bar{\nu}}\right)^q = f^q(\theta, T, \varpi)$ in the *minus* case. What happens to this equation under Galois action? The same equation holds for all the conjugates: $\nu^{q\sigma_c} = f^q(\sigma_c \cdot \theta, T, \varpi)$ in the *plus* case, and $\left(\frac{\nu}{\bar{\nu}}\right)^{q\sigma_c} = f^q(\sigma_c \cdot \theta, T, \varpi)$ in the *minus* case. When taking the q-th root on both sides of the equation, we obtain

(a) in the *plus* case

$$\nu^{\sigma_c} = \xi^{\kappa_{\theta,\sigma_c}} \cdot f(\sigma_c \cdot \theta, T, \varpi), \tag{5.17}$$

(b) in the *minus* case

$$\left(\frac{\nu}{\bar{\nu}}\right)^{\sigma_c} = \xi^{\kappa_{\theta,\sigma_c}} \cdot f(\sigma_c \cdot \theta, T, \varpi), \tag{5.18}$$

where ξ is a primitive q-th root of unity and κ_{θ,σ_c} is called the *Galois exponent*. In general the Galois exponents are unknown. They are additive and verify $\kappa_{\jmath\theta,\sigma} + \kappa_{\theta,\sigma} = 0$, where $\jmath \in G$ denotes again complex conjugation. Beyond this relation, the Galois exponents are random and in particular they do not commute with Galois action on ξ: we say that Equation (5.17) and Equation (5.18) are not Galois covariant.

The approach is different depending on which case we are in:

Plus We have seen that in the *plus* case, ν is real. When we have been able to eliminate the real root ε_θ, we can choose θ such that $|w(\theta)|$ be a mutliple of q: $|w(\theta)| = hq$. We obtain then an equation like Equation (5.17). The only real roots of unity in $\mathbb{R}$ being ± 1, the κ map constantly vanishes and we obtain an equation of the type $x^h \cdot f(\theta, T, \varpi) = polynomial(x) + remainder$. Provided that there are sufficiently large lower bounds on $|x|$, one can show that the remainder is null. This is a " Runge type " approach.

Minus In the *minus* case, ν is not real. In this case, we treat the Galois exponents as unknowns. The approach differs depending on whether $p = q$ or $p \neq q$.

$p = q$ In this case Equation (5.18) becomes

$$\left(\frac{\nu}{\bar{\nu}}\right)^{\sigma_c} = \zeta^{\kappa_{\theta,\sigma_c}} \cdot f(\sigma_c \cdot \theta, T, \varpi),$$

Let $m = |G|/2$ and $\Delta = \sum_{c=1}^{m} \lambda_c \left(\frac{\nu}{\bar{\nu}}\right)^{\sigma_c} + \overline{\lambda_c \left(\frac{\nu}{\bar{\nu}}\right)^{\sigma_c}} = \sum_{c=1}^{m} \lambda_c \zeta^{\kappa_{\theta,\sigma_c}} f(\sigma_c \cdot \theta, x, \varpi)) + \bar{\lambda}_c \zeta^{-\kappa_{\theta,\sigma_c}} \overline{f(\sigma_c \cdot \theta, x, \varpi))}$, where $\lambda_1 \cdots \lambda_m \in \mathcal{O}_{\mathbb{K}}$. Let also $\Delta_a = \sigma_a(\Delta),\ a \in \{1, \cdots, |G|\}$. The purpose of elimination in this case will be to find a set of $\lambda_1 \cdots \lambda_m$ such that $\Delta_a = O(\sqrt{y}/x),\ \forall a \in \{1, \cdots, |G|\}$. This condition is equivalent to:

$$\begin{pmatrix} \zeta^{-\kappa_{\theta,\sigma_1}/1} & \zeta^{-\kappa_{\theta,\sigma_2}/1} & \cdots & \zeta^{-\kappa_{\theta,\sigma_m}/1} \\ \zeta^{-\kappa_{\theta,\sigma_2}/2} & \zeta^{-\kappa_{\theta,\sigma_4}/2} & \cdots & \zeta^{-\kappa_{\theta,\sigma_{2m}}/2} \\ \vdots & \vdots & \vdots & \vdots \\ \zeta^{-\kappa_{\theta,\sigma_m}/m} & \zeta^{-\kappa_{\theta,\sigma_{2m}}/m} & \cdots & \zeta^{-\kappa_{\theta,\sigma_{m^2}}/m} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where the indexes of σ are mod p. Let $\mathbf{A} = \left(\zeta^{-\kappa_{\theta,\sigma_{ij}}/j}\right)_{i,j=1}^{m}$. We first assume that $\mathbf{A}$ is regular. In order to ascertan that Δ is not zero, we impose the condition

$$\begin{pmatrix} \zeta^{-\kappa_{\theta,\sigma_1}/1} & \zeta^{-\kappa_{\theta,\sigma_2}/1} & \cdots & \zeta^{-\kappa_{\theta,\sigma_m}/1} \\ \zeta^{-\kappa_{\theta,\sigma_2}/2} & \zeta^{-\kappa_{\theta,\sigma_4}/2} & \cdots & \zeta^{-\kappa_{\theta,\sigma_{2m}}/2} \\ \vdots & \vdots & \vdots & \vdots \\ \zeta^{-\kappa_{\theta,\sigma_k}/k} & \zeta^{-\kappa_{\theta,\sigma_{2k}}/k} & \cdots & \zeta^{-\kappa_{\theta,\sigma_{km}}/k} \\ \vdots & \vdots & \vdots & \vdots \\ \zeta^{-\kappa_{\theta,\sigma_m}/m} & \zeta^{-\kappa_{\theta,\sigma_{2m}}/m} & \cdots & \zeta^{-\kappa_{\theta,\sigma_{m^2}}/m} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ C \neq 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Let $\vec{d} = (C\delta_{c,k})_{c=1}^{m}$ (here, δ is the Kronecker symbol). By Cramer's rule, the solutions of our system are $\lambda_c = A_c/A$, where $A = \det(\mathbf{A})$ and A_c is the determinant of a minor of $\mathbf{A}$ obtained by replacing the c-th column by the column vector $\vec{d}$. Using Hadamard's inequality, we find that $|A_c| \leq \left(\frac{p-3}{2}\right)^{(p-3)/4}$ and $|A| \leq \left(\frac{p-1}{2}\right)^{(p-1)/4}$. Then, obviously $A\Delta \in \mathcal{O}_{\mathbb{K}}$. Lemma 5.2.8 yields an upper bound for $\mathbf{N}(A\Delta)$. And since $\mathbf{N}(A\Delta)$ is an integer different from zero, it must be greater or equal to 1, which leads to an upper bound for $|x|$.
When $\mathbf{A}$ is not regular, we restraint the system to its largest regular subsystem. Then, using Lemma A.1.1 from Appendix A, we can add an additional equation and the system remains regular. Finally, proceeding now as in the case when $\mathbf{A}$ is regular, we again find upper bounds for $|x|$. These upper bounds are slightly better than in the regular case. We can thus use in both cases the bounds from the regular case.

$p \neq q$ The idea of the approach in this case is the same as in the case $p = q$, always building linear combinations and trying to cancel enough leading coefficients in the power series developments. The same distinction between the case when $\mathbf{A}$ is regular or irregular will be made. Depending on the annihilators and the characteristics of the system, p/q will play an essential role. Also, given that G does not act on the q-th root of unity ξ, there will be no division of the exponent κ_{θ,σ_c} by an integer in the matrix $\mathbf{A}$. This case and some of its possible ramifications are illustrated in [Mihăilescu 2007].

This approach was initiated with the proof of Catalan's conjecture by Preda Mihăilescu in [Mihăilescu 2004]. It was pursued with the objective of bounding potential solutions of the Diagonal Nagell-Ljunggren equation in [Mihăilescu 2008] and has been further pursued in [Bartolomé & Mihăilescu 2015] to study the binary Thue equation $X^n - B.Y^n = 1$. We will now see how this approach can be found in some works.

5.3.1 Catalan's conjecture

In 1842, the Belgian mathematician Eugène Catalan conjectured that two consecutive numbers, other than 8 and 9, could not be exact powers [Catalan 1842]. Catalan held a position as "répétiteur" (he was in the process of obtaining his teaching degree, and this training process included helping individually students with issues "repeat" their lesson after class) at Ecole Polytechnique, near Paris, at that time. He sent a letter to August Leopold Crelle, to be published in his journal [Catalan 1844]:

192

13.
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'admet qu'une seule „solution."

160 years after Catalan published his conjecture in [Catalan 1842], Preda Mihăilescu proved it to be a theorem. We will be interested in part of the proof (which we will not reproduce here, it has been discussed abundantly in [Mihăilescu 2004], [Bilu 2004], [Schoof 2008], [Bilu *et al.* 2014]). We will show how our approach was applied in the case of Catalan.

The facts that were known before Mihăilescu's last part of the proof ([Mihăilescu 2004]) were:

- $p \not\equiv 1 \mod q$,
- $|x| \geq q(2p+1)(2q^{p-1}+1)$,
- There exist a non-zero integer a and a positive integer v such that $x - 1 = p^{q-1}a^q$, $y = pav$ and $\frac{x^p-1}{x-1} = pv^q$, and there exist a non-zero integer b and a positive integer u such that $y + 1 = q^{p-1}b^p$, $x = qub$ and $\frac{y^q+1}{y+1} = qu^p$,
- $q^2|x$, and $p^{q-1} \equiv 1 \mod q^2$,

Catalan's conjecture was reduced to prove that for any two odd primes p and q, there do not exist two integers x and y (other than the known solution $2^3 - 3^2 = 1$) such that $x^p - y^q = 1$. We find Equation (5.16) with $f = 0$, $y = 1$, $B = 1$.

Our general approach specializes to the case of Catalan's conjecture as follows:

A. The characteristic algebraic integer is $\alpha = (x - \zeta)/(1 - \zeta) \in \mathbb{Z}[\zeta]$in this case, and it verifies $\mathbf{N}(\alpha) = v^q$. The characteristic ideal associated to the characteristic algebraic integer is $\mathfrak{A} = (\alpha, v)$. It verifies $\mathfrak{A}^q = (\alpha)$.

B. In the case of Catalan's equation, we find ourselves in the *plus* situation. However, Mihăilescu finds in this case a special set of annihilators ("Mihăilescu's ideal"), that is the set of $\Theta \in \mathbb{Z}[G]$ such that $(x - \zeta)^\Theta$ is a q-th power.

C. Mihăilescu then proves that the plus part of $\mathcal{I}_M$ contains at least one element Θ that is non-trivial. As we are in the plus case, ν is real and the Galois exponents are cancelled.

The series and bounds are used to prove that for the real annihilators chosen, the rest of the series development is null. This can be seen as an application of Runge's method, even though Mihăilescu had not heard about this method until after his proof.

5.3.2 Diagonal Nagell-Ljunggren

We are interested in equation:

$$\frac{x^n-1}{x-1} = y^q, \text{ in integers } x>1,\ y>1,\ n>2,\ q\geq 2. \tag{5.19}$$

The study of this equation started in [Nagell 1920a] and [Nagell 1920b] with Nagell, and later Ljunggren brought some precisions to Nagell's proof in [Ljunggren 1943]. They proved that

Theorem 5.3.2 (Nagell-Ljunggren) *Apart from the solutions*

$$\frac{3^5-1}{3-1} = 11^2, \quad \frac{7^4-1}{7-1} = 20^2 \quad \textit{and} \quad \frac{18^3-1}{18-1} = 7^3,$$

equation (5.19) has no other solution (x,y,n,q) *if either one of the following conditions is satisfied:*

(i) q*=2,*

(ii) *3 divides* n*,*

(iii) *4 divides* n*,*

(iv) $q=3$ *and* $n\not\equiv 5 \mod 6$.

The diagonal case of the Nagell-Ljunggren equation is

$$\frac{x^p-1}{x-1} = p^e\cdot y^p \quad \text{with} \quad x,y\in\mathbb{Z} \quad e\in\{0,1\}, \tag{5.20}$$

and p an odd prime. The only known non - trivial solution is

$$\frac{18^3-1}{18-1} = 7^3,$$

and it is conjectured to be also the only such solution. However, it is not even proved that (5.20) has only finitely many solutions. In [Mihăilescu 2008], Preda Mihăilescu uses again this approach to find conditions and upper bounds on the potential solutions to this equation. We find Equation (5.16) with $f=1,\ y=1,\ B=1,\ q=p$.

A. The characteristic algebraic integer is $\alpha = (x-\zeta)/(1-\zeta)\in\mathbb{Z}[\zeta]$ and it verifies $\mathbf{N}(\alpha)=y^p$. The characteristic ideal associated to the characteristic algebraic integer is $\mathfrak{A}=(\alpha,y)$. It verifies $\mathfrak{A}^p=(\alpha)$.

B. Because $q=p$, it is necessary to work in Fermat's ideal to cancel the unit, and we can even work in the plus part of Fermat's ideal. We are thus in the *plus* situation.

C. However, the Galois exponents cannot not be cancelled. The linear system Δ is thus created and its norm bounded.

D. A non-vanishing linear combination of the values of $\nu^{\sigma.\Theta_1}$ and $\nu^{\sigma.\Theta_2}$, where Θ_1 and Θ_2 have been selected in the positive part of the Fermat ideal, is used. The series development and its conjugates converge to the conjugate of the limit in the complex topology. The bounds computed on the determinant of the linear system as well as equation-specific properties allow then to find the upper bounds:

$$|x| < \begin{cases} 4.(p-3/2)^{(p+2)/2} & \text{if } x \not\equiv \pm 1, 0 \mod p \\ (4p)^{(p-1)/2} & \text{if } x \equiv 0 \mod p \\ 4.(p-2)^p & \text{otherwise} \end{cases}$$

.

5.3.3 Binary Thue

We call binomial Thue equation:

$$A.X^n - B.Y^n = C,$$

where $n \geq 3$ and A, B and C are non-zero integers. Thue proved in 1909 in [Thue 1909] that, for a fixed n, this equation has at most a finite number of solutions in integers (x, y). Currently, even the best numerical bounds on the solutions are too large for numerical resolution. This equation has been totally solved in some particular cases, always with $C = \pm 1$. We study equation:

$$X^n - B.Y^n = 1, \tag{5.21}$$

which we call binary Thue. The detailed study of this function is the object of Chapter 6.

5.4 Conclusion

In this chapter, we have shown an approach to certain exponential Diophantine equations, using several techniques from cyclotomy and series develomment, for which we have given some prerequisite properties and definitions. However, one can notice that it has so far been applied only to binomial equations (that is, with two unknowns). It is an interesting follow-up of this work to apply this approach to ternary (that is, with three unknowns) equations.

Chapter 6

On the equation $X^n - 1 = B \cdot Z^n$

6.1 Abstract

We consider the Diophantine equation $X^n - 1 = BZ^n$, where $B \in \mathbb{Z}$ is understood as a parameter. We prove that if this equation has a solution, then either the Euler totient of the radical, $\varphi(\text{rad }(B))$, has a common divisor with the exponent n, or the exponent is a prime and the solution stems from a solution to the diagonal case of the Nagell–Ljunggren equation: $\frac{X^n-1}{X-1} = n^e Y^n$, $e \in \{0, 1\}$. This allows us to apply recent results on this equation to the binary Thue equation in question. In particular, we can then display parametrized families for which the Thue equation has no solution. The first such family was proved by Bennett in his seminal paper on binary Thue equations [Bennet 2001].

6.2 Introduction

Let $B \in \mathbb{Z}, n \in \mathbb{N}_{>1}$, define $\varphi^*(B) := \varphi(\text{rad }(B))$, where rad (B) is the radical of B, and assume that:

$$(n, \varphi^*(B)) = 1. \tag{6.1}$$

This condition implies that B has no prime factors $t \equiv 1 \bmod n$. In particular, none of its prime factors splits completely in the $n-$th cyclotomic field.

More generally, for a fixed $B \in \mathbb{Z}$ we let

$$\mathcal{N}(B) = \{n \in \mathbb{N}_{>1} \mid \exists\ k > 0 \text{ such that } n | \varphi^*(B)^k\}. \tag{6.2}$$

If p is an odd prime, we shall denote by CF the combined condition requiring that

I The Vandiver Conjecture holds for p, so the class number h_p^+ of the maximal real subfield of the cyclotomic field $\mathbb{Q}[\zeta_p]$ is not divisible by p.

II The index of irregularity of p is small, namely $i_r(p) < \sqrt{p} - 1$, so there are $i_r(p)$ odd integers $k < p$ such that the Bernoulli number $B_k \equiv 0 \bmod p$.

The second condition was discovered by Eichler, as a sufficient condition for the first case of Fermat's Last Theorem (FLT) to be true. It is known from recent computations of Buhler and Harvey [Buhler & Harvey 2011] that the condition CF is satisfied for primes up to $163 \cdot 10^6$.

We consider the Binary Thue equation

$$X^n - 1 \quad = \quad B \cdot Z^n, \tag{6.3}$$

where solutions with $Z \in \{-1, 0, 1\}$ are considered to be trivial. The assertion that equation (6.3) has finitely many solutions other than the trivial ones is a special case of the general Pillai conjecture (cf. [Bilu *et al.* 2014][Conjecture 13.17]):

Conjecture 6.2.1 *Let $a,\ b,\ c$ be nonzero integers. Then the equation $ax^m + by^n = c$ has finitely many solutions in integers $x,\ y$ and positive integers $m,\ n$ such that $x,\ y \neq 0, \pm 1$; $m,\ n > 1$; $(m, n) \neq (2, 2)$.*

One has to exclude the case $m = n = 2$, because equation $ax^2 + by^2 = c$ may have infinitely many solutions.

The Binary Thue equation is encountered as a particular case of binomial Thue equations of the type

$$aX^n - bY^n = c, \tag{6.4}$$

see [Bennett *et al.* 2006]. In a seminal paper [Bennet 2001], Michael Bennett proves that in the case when $c = \pm 1$, there is at most one solution for fixed $(a, b; n)$ and deduces that the parametric family $(a + 1, a; n)$ has the only solution $(1, 1)$ for all n. Equation (6.3) inserts naturally in the family of equations (6.4), with $a = c = \pm 1$. Current results on Equation (1.3) are restricted to values of B which are built up from two small primes $p \leq 13$ [Bennett *et al.* 2006] and complete solutions for $B < 235$ ([A.Bazso *et al.* 2010]).

A conjecture directly related to Equation (6.3) states that

Conjecture 6.2.2 *Let n be a prime and $B > 1$ an integer, verifying condition (6.1). Then, Equation (6.3) has no other non-trivial solution than $(X, Y; B, n) = (18, 7; 17, 3)$.*

The main contribution of this paper is to relate Equation (6.3) to the diagonal Nagell – Ljunggren equation,

$$\frac{X^n - 1}{X - 1} = n^e Y^n, \quad e = \begin{cases} 0 & \text{if } X \not\equiv 1 \bmod n, \\ 1 & \text{otherwise,} \end{cases} \tag{6.5}$$

in the case when n is a prime and condition (6.1) holds. We can then apply results from [Mihăilescu 2008] and prove the following:

Theorem 6.2.3 *Let n be a prime and $B > 1$ an integer with $(\varphi^*(B), n) = 1$. Suppose that Equation (6.3) has a non trivial integer solution different from $n = 3$ and $(X, Z; B) = (18, 7; 17)$. Let $X \equiv u \bmod n, 0 \leq u < n$ and $e = 1$ if $u = 1$ and $e = 0$ otherwise. Then:*

1. *$n > 163 \cdot 10^6$.*
2. *$X - 1 = \pm B/n^e$ and $B < n^n$.*
3. *If $u \notin \{n - 1, 0, 1\}$, then condition CF (II) fails for n and*

$$\begin{array}{rcl} 2^{n-1} & \equiv & 3^{n-1} \equiv 1 \mod n^2, \quad \text{and} \\ r^{n-1} & \equiv & 1 \mod n^2 \quad \text{for all } r | X(X^2 - 1). \end{array}$$

If $u \in \{n - 1, 0, 1\}$, then Condition CF (I) fails for n.

The particular solution $n = 3$ and $(X, Z; B) = (18, 7; 17)$ is reminiscent of a solution of the diagonal Nagell equation; it is commonly accepted that the existence of non trivial solutions tends to render Diophantine equations more difficult to solve. Based on Theorem 6.2.3, we prove the following

Theorem 6.2.4 *If Equation (6.3) has a solution for a fixed B verifying condition (6.1), then either $n \in \mathcal{N}(B)$ or there is a prime p coprime to $\varphi^*(B)$ and an $m \in \mathcal{N}(B)$ such that $n = p \cdot m$. Moreover X^m, Y^m is a solution of (6.3) for the prime exponent p and thus verifies the conditions of Theorem 6.2.3.*

Remark 6.2.5 *Theorem 6.2.3 uses criteria from the diagonal case of the Nagell-Ljunggren equation, the relation being established by point (2.) of the theorem. The criteria were proved in [Mihăilescu 2008] and are in part reminiscent from classical cyclotomic results on Fermat's Last Theorem. Thus, the criteria for the First Case, which are ennounced in point (3.) are the Eichler criterion CF (II) and the criteria of Wieferich and Furtwängler (cf. [Mihăilescu 2008][Theorem 2]). For the Second Case of Diagonal Nagell-Ljunggren, in point (3.), it was possible to restrict the two conditions proved by Kummer for the second case of FLT to the single condition CF (I), namely Vandiver's conjecture (cf. Theorem 4 of [Mihăilescu 2008]). This is a consequence of the fact that unlike FLT, Nagell-Ljunggren is a binary equation, a fact which allowed also to prove upper bounds for the solutions, which are given in Theorem 6.4.2. The fact that the Nagell-Ljunggren equation is not homogenous in X makes it difficult to prove lower bounds, thus leaving a gap on the way to a complete proof of Conjecture 6.2.2.*

The plan of the chapter is as follows: in Section 2 we drop the condition that n be a prime and use Theorem 6.2.3 to deduce the results on Equation (6.3) for arbitrary exponents n which are stated in Theorem 6.2.4. In Section 3 we establish the connection between equations (6.3) and (6.5), review some basic properties of Stickelberger ideals and prove auxiliary technical lemmata concerning coefficients of binomial series development.

With these prerequisites, we complete the proof of Theorem 6.2.3. Given the reduction to the Nagell-Ljunggren Diagonal Case, the proof focuses on point (2.) of Theorem 6.2.3.

6.3 Proof of Theorem 6.2.4 assuming Theorem 6.2.3

In this section we derive Theorem 6.2.4 assuming Theorem 6.2.3. For this we assume that Equation (6.3) has a solution with $(\varphi^*(B), n) = 1$, since our results only hold in this case, a fact which is reflected also in the formulation of Theorem 6.2.4.

Consider the case when $n = p \cdot q$ is the product of two distinct primes. If $(n, B) = 1$, then Theorem 6.2.3 holds for both p and q with the value $e = 0$. If X, Z is a solution, then Theorem 6.2.3 . (2.) implies that $X^p = \pm B + 1$ and $X^q = \pm B + 1$. Consequently either $X^p + X^q = 2$ or $X^p - X^q = 2$. This is impossible for $|X| > 2$ and a simple case distinction implies that there are no solutions. As a consequence,

Corollary 6.3.1 *Consider Equation (6.3) for fixed B and suppose that n is an integer which has two distinct prime divisors $q > p > 2$ with $(p, B) = (q, B) = 1$. Then Equation (6.3) has no solutions for which condition (6.1) holds.*

If all divisors of n are among the primes dividing B, we are led to the following equation: $p(X^q - 1) = q(X^p - 1)$, which has no solutions in the integers other than 1. Indeed, assume $X \neq 1$ to be a solution of the previous equation, and $q = p + t$, $t \geq 0$. The real function $f(t) = p(X^{p+t} - 1) - (p + t)(X^p - 1)$ is strictly monotonous and $f(0) = 0$. Therefore, the equation $p(X^q - 1) = q(X^p - 1)$ has no solutions. There is only the case left in which n is built from two primes, one dividing B and one not. In this case, one obtains that equation $p(X^q - 1) = X^p - 1$ which can also be shown not to have non trivial solutions, using the above remark, this time with $f(t) = p(X^{p+t} - 1) - (X^p - 1)$. Hence:

Corollary 6.3.2 *Equation (6.3) has no solutions for exponents n which are divisible by more than one prime such that condition (6.1) holds.*

We are left to consider the case of prime powers $n = p^c$ with $c > 1$. If $p \nmid B$, we obtain $X^{n/p} - 1 = B/p^e$, so in particular $B/p^e + 1 \geq 2^{p^{c-1}}$ is a $p^{c-1}-$th power. Since in this case, Equation (6.3) has in particular a solution for the exponent p, Theorem 6.2.3 implies that $B < p^p$; when $c > 2$, combining this with the previous lower bound implies that there are no solutions. For $c = 2$, we deduce that $|X| < p$ and, after applying Theorem 6.2.3 again and letting $\xi = \zeta^{1/p}$ be a primitive p^2-th root of unity, we obtain the following equation

$$Y^{p^2} = \frac{X^{p^2} - 1}{p^e(X^p - 1)} = \mathbf{N}_{\mathbb{Q}[\xi]/\mathbb{Q}}(\alpha) \qquad \alpha = \frac{X - \xi}{(1 - \xi)^e}.$$

As usual, the conjugates of the ideal (α) are pairwise coprime. We let $\mathfrak{A} = (Y, \alpha)$ be an ideal with $N(\mathfrak{A}) = (Y)$; moreover, if $\mathfrak{L}|\mathfrak{A}$ is a prime ideal and $N(\mathfrak{L}) = (\ell)$, then the rational prime ℓ is totally split in $\mathbb{Q}[\xi]$, the factors being the primes $(\ell, \sigma_c(\alpha))$. Being totally split, it follows in particular that $\ell \equiv 1 \bmod p^2$ so $Y \geq \ell > 2p^2$, in contradiction with $Y < X < p + 1$. This shows that there are no solutions for $n = p^2$.

Corollary 6.3.3 *If Equation (6.3) in which $n = p^c$ is a prime power has non trivial solutions for which condition (6.1) holds, then $c = 1$.*

□

The primes dividing the exponent n used in the above corollaries are by definition coprime to $\varphi^*(B)$. As a consequence, if n is an exponent for which Equation (6.3) has a solution and $m|n$ is the largest factor of n with $m \in \mathcal{N}(B)$ – as defined in (6.2) – then the corollaries imply that there is at most one prime dividing n/m and the exponent of this prime in the prime decomposition of n must be one. This is the first statement of Theorem 6.2.4, which thus follows from these corollaries and Theorem 6.2.3.

6.4 Proof of Theorem 6.2.3

6.4.1 Preliminary results

It should be noted that we followed the tradition when addressing this equation, and noted n for a prime. This is rather disturbing (even for us), but we felt we should not change the tradition.

The proof of Theorem 6.2.3 emerges by relating Equation (6.3) to the Diagonal Case of the Nagell – Ljunggren conjecture. In this section we shall recall several technical tools used for reducing one conjecture to the other. The reduction is performed in the next section.

6.4.1.1 Link of (6.3) with the diagonal Nagell – Ljunggren equation

We note that $\delta = \left(\frac{X^n-1}{X-1}, X - 1\right)$ divides n and $\delta = n$ exactly when $X \equiv 1 \bmod n$. The first part is Euler's Lemma 5.3.1. If $X \equiv 1 \bmod n$, then $\delta = n$ and thus $n|(X - 1)$ must hold. Conversely, inserting $X \equiv 1 \bmod n$ in the previous expression shows that in this case $\delta = n$.

We first show that any solution of the Binary Thue equation (6.3) leads to a solution of the Diagonal Nagell-Ljunggren equation (6.5). Let $\zeta \in \mathbb{C}$ be a primitive $n-$th root of

unity. Then the numbers $\alpha_c = \frac{X-\zeta^c}{(1-\zeta^c)^e} \in \mathbb{Z}[\zeta]$ by definition of e, and $(\alpha_c, n) = 1$. Since for distinct $c, d \not\equiv 0 \bmod n$ we have $(1-\zeta^d)^e \cdot \alpha_d - (1-\zeta^c)^e \cdot \alpha_c = \zeta^c - \zeta^d$, it follows that $(\alpha_c, \alpha_d) \mid (1-\zeta)$ and in view of $(\alpha_c, n) = 1$, it follows that the α_c are coprime.

Let $F = \prod_{c=1}^{n-1} \alpha_c = \frac{X^n-1}{n^e(X-1)}$ and $q \mid F$ be a rational prime. In the ring $\mathbb{Z}[\zeta]$, it splits completely as the product of prime ideals $\mathfrak{Q}_c = (q, \alpha_c)$, $c = 1, 2, \ldots, n-1$: these ideals are coprime, as a consequence of the coprimality of the α_c. Therefore $q \equiv 1 \bmod n$ and it follows from Equation (6.1) that $(q, B) = 1$, so $q|Z$. Furthermore, Equation (6.3) implies that there exists $j_q > 0$ such that $q^{j_q n}||Z^n$ and thus $q^{j_q n}||F$. This holds for all primes $q \mid \text{rad}\,(F)$. It follows that Equation (6.5) is verified for $Y = \prod_{q|F} q^{j_q}$ and $Y \mid Z$. We have thus proved that if (X, Z) is a solution of Equation (6.3) for the prime n, then there exists $C \in \mathbb{Z}$ such that $Z = C \cdot Y$ with Y as above, and:

$$\frac{X^n - 1}{n^e(X-1)} = Y^n \quad \text{and} \tag{6.6}$$

$$X - 1 = B \cdot C^n / n^e. \tag{6.7}$$

From now on, we shall write $D = X - 1$.

From the above, we conclude that any integer solution of Equation (6.3) (binary Thue) induces one of Equation (6.5) (Diagonal Nagell-Ljunggren). Conversely, if (X, Y) is a solution of Equation (6.5), then $(X, Y; n^e(X-1))$ is a solution of Equation (6.3). For instance, the particular solution $(X, Y; B) = (18, 7; 17)$ of Equation (6.3) stems from

$$\frac{18^3 - 1}{18 - 1} = 7^3,$$

which is supposed to be the only non trivial solution of Equation (6.5).

Remark 6.4.1 *Note that if (X, Z) verify Equation (6.3), then $(-X, Z)$ is a solution of $\frac{X^n+1}{X+1} = BZ^n$, so the results apply also to the equation:*

$$X^n + 1 = BZ^n.$$

6.4.1.2 Bounds to the solutions of Equation (6.5)

We shall use [Mihăilescu 2008][Theorem 2]:

Theorem 6.4.2 *Suppose that X, Y are integers verifying Equation (6.5) (diagonal Nagell-Ljunggren) with $n \geq 17$ being a prime. Let $u = (X \bmod n)$. Then there is an $E \in \mathbb{R}_+$ such that $|X| < E$. The values of E in the various cases of the equation are the following:*

$$E = \begin{cases} 4 \cdot \left(\frac{n-3}{2}\right)^{\frac{n+2}{2}} & \text{if } u \notin \{-1, 0, 1\} \\ (4n)^{\frac{n-1}{2}} & \text{if } u = 0\ , \\ 4 \cdot (n-2)^n & \text{otherwise.} \end{cases} \tag{6.8}$$

Comparing the bounds (6.8) with Equation (6.7), it follows that $|C| < 2n-1$. In particular, any prime dividing C would not split completely in $\mathbb{Q}[\zeta_n]$ – since a prime splitting in this field has the form $r = 2kn + 1 > 2n$.

Remark 6.4.3 *Note that $|C| < 2n-1$ implies a fortiori that for all primes $r|C$, $r^2 \not\equiv 1 \bmod n$. If $d(r) \subset Gal(\mathbb{Q}[\zeta]/\mathbb{Q})$ is the decomposition group of the unramified prime r, it follows that $|d(r)| \geq 3$; moreover, either $d(r)$ contains a subcycle $d' \subset d(r)$ of odd order $|d'| \geq 3$ or it is a cyclic 2-group with at least 4 elements.*

6.4.1.3 A combinatorial lemma

Lemma 6.4.4 *Let p be an odd prime, $k \in \mathbb{N}$ with $1 < k < \log_2(p)$ and $P = \{1, 2, \ldots, p-1\}$. If $S = \{a_1, a_2, \ldots, a_k\} \subset P$ be a set of numbers coprime to p and such that $a_i \not\equiv \pm a_j \bmod p$ for $i \neq j$. We set the bound $A = 2\lceil p^{1/k} \rceil$; then there are k numbers $b_i \in \mathbb{Z}$, $i = 1, 2, \ldots, k$, not all zero, with $0 \leq |b_i| \leq A$ and such that*

$$\sum_{i=1}^{k} a_i b_i \equiv 0 \bmod p.$$

For $k = 2$, we can choose the b_i such that the additional condition

$$\sum_{i=1}^{2} b_i/a_i \not\equiv 0 \bmod p.$$

holds.

Proof. Let $T = \{1, 2, \ldots, A\} \subset P$. Consider the functional $f : T^k \to \mathbb{Z}/(p \cdot \mathbb{Z})$ given by

$$f(\vec{t}) \equiv \sum_{i=1}^{k} t_i a_i \mod p, \quad \text{with} \quad \vec{t} = (t_1, t_2, \ldots, t_k) \in T^k.$$

Since $|T^k| > p$, by the pigeon hole principle there are two vectors $\vec{t} \neq \vec{t'}$ such that $f(\vec{t}) \equiv f(\vec{t'}) \bmod p$. Let $b_i = t_i - t'_i$; by construction, $0 \leq |b_i| \leq A$ and not all b_i are zero, since $\vec{t} \neq \vec{t'}$. The choice of these vectors implies $\sum_{i=1}^{k} a_i b_i \equiv 0 \bmod p$, as claimed.

We now turn to the second claim. If the claim were false, then

$$a_1 b_1 + a_2 b_2 \equiv 0 \mod p \text{ and } b_1/a_1 + b_2/a_2 \equiv 0 \mod p,$$

a homogenous linear system S with determinant $\det(S) = \frac{a_1^2 - a_2^2}{a_1 a_2}$, which is non vanishing under the premise of the lemma. This would imply that the solution b_1, b_2 is trivial, in contradiction with our construction. This completes the proof. □

6.4.1.4 Some notation

We assume that n is prime and let ζ be a primitive $n-$th root of unity, $\mathbb{K} = \mathbb{Q}(\zeta)$ the $n-$th cyclotomic field and $G = \mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ its Galois group. The automorphisms $\sigma_a \in G$ are given by $\zeta \mapsto \zeta^a$, $a = 1, 2, \ldots, n-1$; complex conjugation is denoted by $\jmath \in \mathbb{Z}[G]$. In the ring of integers $\mathbb{Z}[\zeta]$, one has finite λ-*adic* expansions: for any $\alpha \in \mathbb{Z}[\zeta]$, there are some $N_\alpha > 0$ and $a_j \in \{-(p-1)/2, \cdots, 0, 1, \cdots, (p-3)/2\}, j = 0, 1, \ldots N$ such that:

$$\alpha = \sum_{j=0}^{N_\alpha} a_j (1 - \zeta)^j. \tag{6.9}$$

We shall use the algebraic $O(\cdot)$-notation, to suggest the remainder of a power series. This occurs explicitly in the following four contexts

(i) In a λ-adic development of the type (6.9), we write $\alpha = x + O(\lambda^m)$ to mean that there is some $y \in \mathbb{Z}[\zeta]$ such that $\alpha - x = \lambda^m y$. Since $(n) = (\lambda^{p-1})$, powers of n can occur as well as powers of λ in this notation.

(ii) We also use formal power series, often written $f = f(D) \in \mathbb{K}[[D]]$. For $f = \sum_{k=0}^{\infty} f_k D^k$ with partial sum $S_m(f) = \sum_{k=0}^{m} f_k D^k$ we may also use the $O(\cdot)$-notation and denote the remainder by $f(D) = S_m(D) + O(D^{m+1})$.

(iii) Suppose that D is an integer and the formal power series converges in the completion $\mathbb{K}_{\mathfrak{P}}$ at some prime $\mathfrak{P} \subset \mathcal{O}(\mathbb{K})$ dividing D. Suppose also that in this case all coefficients of f are integral: then the remainder $f(D) - S_m(D)$ is by definition divisible by $\mathfrak{P}^{m+1}$, so $O(D^{m+1})$ means in this context that the remainder is divisible by $\mathfrak{P}^{m+1}$.

(iv) If $f(D)$ converges at all the prime ideals dividing some integer $a|D$, then $O(D^{m+1})$ will denote a number divisible by a^{m+1}. In this paper we shall use this fact in the context in which $a = p$ is an integer prime dividing D and such that $f(D)$ converges at all prime ideals of $\mathbb{K}$ above p.

6.4.2 Auxiliary facts on the Stickelberger module

The following results are mostly deduced in [Mihăilescu 2008][§4], with some being deduced in [Mihăilescu 2004][§2.1-2.3 and 4.1]. The results shall only be mentioned here without proof. As usual, we note the Stickelberger ideal I and the Fermat's module I_f (cf Section 5.2). Also, I^+ is the set of all positive elements of the Stickelberger ideal I.

We shall want to consider the action of elements of $\theta \in \mathbb{F}_n[G]$ on explicit algebraic numbers $\beta \in \mathbb{K}$. Unless otherwise specified, an element $\theta = \sum_{c=1}^{n-1} m_c \sigma_c \in \mathbb{F}_n[G]$ is lifted to $\sum_{c=1}^{n-1} n_c \sigma_c$, where $n_c \in \mathbb{Z}$ are the unique integers with $0 \leq n_c < p$ and $n_c \equiv m_c \mod p$. In particular, lifts are always positive, of bounded weight $w(\theta) \leq (p-1)^2$. Rather than introducing an additional notation for the lift defined herewith, we shall always assume, unless otherwise specified, that $\theta \in \mathbb{F}_n[G]$ acts upon $\beta \in \mathbb{K}$ via this lift.

Using this lift, we define the following additive maps:

$$\rho_0 : \mathbb{F}_n[G] \to \mathbb{Q}(\zeta) \quad \theta = \sum_{c=1}^{n-1} n_c \sigma_c \mapsto \sum_{c \in P} \frac{n_c}{1 - \zeta^c},$$

and

$$\rho : \mathbb{F}_n[G] \to \mathbb{Z}[\zeta] \quad \theta \mapsto (1 - \zeta) \cdot \rho_0[\theta].$$

The $i-$th moment of an element $\theta = \sum_{c=1}^{n-1} n_c \sigma_c$ of $\mathbb{Z}[G]$ is defined as:

$$\phi^{(i)}(\theta) = \sum_{c=1}^{n-1} n_c c^i \mod n.$$

Note that $\phi^{(1)}$ is the *Fermat quotient map*: $\phi^{(1)} = \varphi$. Moments are linear maps of $\mathbb{F}_n$-vector spaces and homomorphisms of algebras, verifying:

$$\begin{array}{lcll} \phi^{(i)}(a\theta_1 + b\theta_2) & = & a\phi^{(i)}(\theta_1) + b\phi^{(i)}(\theta_2), & \text{and} \\ \phi^{(i)}(\theta_1\theta_2) & = & \phi^{(i)}(\theta_1)\phi^{(i)}(\theta_2), & \text{with} \quad \theta_j \in \mathbb{F}_n[G]; a, b \in \mathbb{F}_n. \end{array} \tag{6.10}$$

The linearity in the first identity is a straight-forward verification from the definition. For the second, note that for $\theta = \sum_c n_c \sigma_c$ we have

$$\phi^{(i)}(\sigma_a \theta) = \phi^{(i)}\left(\sum_c n_c \sigma_{ac}\right) = \sum n_c \cdot (ac)^i = a^i \cdot \phi^{(i)}(\theta).$$

Using the already established linearity, one deduces the multiplicativity of $\phi^{(i)}$ as a ring homomorphism.

Let $\alpha = \frac{X-\zeta}{(1-\zeta)^e} \in \mathbb{Z}[\zeta]$, as before, and define $c_X \in \{0, 1, \cdots, n-1\}$ such that $c_X \equiv 1/(X-1) \bmod n$ if $e = 0$ and $c_X = 0$ if $e = 1$. For any $\theta \in I^+$, there is a *Jacobi integer* $\beta[\theta] \in \mathbb{Z}[\zeta]$ such that $\beta[\theta]^n = (\zeta^{c_X}\alpha)^\theta$, normed by $\beta[\theta] \equiv 1 \bmod (1-\zeta)^2$ (Lemma 5.2.3). The definition of the relative weight $\varsigma(\theta)$ (cf. Section 5.2) implies that

$$\beta[\theta] \cdot \overline{\beta[\theta]} = \mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\alpha)^{\varsigma(\theta)} = Y^{\varsigma(\theta)}. \tag{6.11}$$

We have for any $\theta \in I^+$,

$$\beta[\theta]^n = (\zeta^{c_X}\alpha)^\theta = (\zeta^{c_X}(1-\zeta)^{1-e})^\theta \cdot \left(1 + \frac{X-1}{1-\zeta}\right)^\theta \tag{6.12}$$

Lemma 6.4.5 *We remind that $D = X - 1$. For any $\theta \in 2 \cdot I_f^+$, for any prime ideal $\mathfrak{P} \mid D$, there is a $\kappa = \kappa_{\mathfrak{P}}(\theta) \in \mathbb{Z}/(n \cdot \mathbb{Z})$ such that*

$$\beta[\theta] \equiv \zeta^\kappa \cdot Y^{\frac{\varsigma(\theta)}{2}} \bmod \mathfrak{P}.$$

Proof. Let θ_0 be an element of I_f^+, and let $\theta = 2\theta_0$. Note that from (6.11) we have $Y^{\varsigma(\theta_0)n} = \beta[\theta_0]^n \cdot \overline{\beta}[\theta_0]^n$. Thus $\beta[\theta]^n = \beta[\theta_0]^{2n} = Y^{\varsigma(\theta_0)n} \cdot \left(\beta[\theta_0]/\overline{\beta}[\theta_0]\right)^n$. Using (6.12) and the previous observations, we find:

$$\begin{aligned}
\beta[\theta]^n &= Y^{\varsigma(\theta_0)n} \cdot \left(\zeta^{c_X} \cdot (1-\zeta)^{1-e}\right)^{(\theta_0 - \jmath\theta_0)} \cdot \left(1 + \frac{X-1}{1-\zeta}\right)^{(\theta_0 - \jmath\theta_0)} \\
&= Y^{\varsigma(\theta_0)n} \cdot \zeta^{(2c_X+1)\varphi(\theta_0)} \cdot (1 + D/(1-\zeta))^{(\theta_0 - \jmath\theta_0)} \\
\beta[\theta]^n &= Y^{\varsigma(\theta_0)n} \cdot \left(1 + \frac{D}{1-\zeta}\right)^{(\theta_0 - \jmath\theta_0)}.
\end{aligned} \tag{6.13}$$

Thus for any prime ideal $\mathfrak{P} \mid D$ there is a $\kappa = \kappa_{\mathfrak{P}}(\theta) \in \mathbb{Z}/(n \cdot \mathbb{Z})$ such that

$$\beta[\theta] \equiv \zeta^\kappa \cdot Y^{\varsigma(\theta_0)} \bmod \mathfrak{P}. \tag{6.14}$$

□

In the sequel, we indicate how to choose θ such that $\kappa = 0$ in Lemma 6.4.5. In this case, Equation (6.12) leads to a $\mathfrak{P}$-adic binomial series expansion for $\beta[\theta]$.

Lemma 6.4.6 *Let ψ_k denote the k-th Fueter element. Then, there exists a linear combination $\theta = \sigma\psi_k + \tau\psi_l \in I$ with $\sigma, \tau \in G$ and $1 \le k, l < n$, such that $\phi^{(1)}(\theta) = 0$ and $\phi^{(-1)}(\theta) \neq 0$.*

The proof of this Lemma is elementary, using the Voronoi relations (5.15); since the details are rather lengthy, they will be given at the end of the chapter.

The following two lemmata contain computational information for the binomial series developments that we shall use below. First, we remind that ρ_0 is the following additive map:

$$\rho_0 : \mathbb{F}_n[G] \to \mathbb{Q}(\zeta) \quad \theta = \sum_{c=1}^{n-1} n_c \sigma_c \mapsto \sum_{c \in P} \frac{n_c}{1-\zeta^c}$$

Lemma 6.4.7 *Let D be an indeterminate. Let $\theta = \sum_{c=1}^{n-1} n_c \sigma_c \in \mathbb{Z}[G]$ and $f[\theta] = \left(1 + \frac{D}{1-\zeta}\right)^{\theta/n} \in \mathbb{K}[[D]]$. Let $0 < N < n$ be a fixed integer. Then,*

$$f[\theta] = 1 + \sum_{k=1}^{N} \frac{a_k[\theta]}{k! n^k} D^k + O(D^{N+1}),$$

where, for $1 \leq k \leq N$, we have

$$a_k[\theta] = \rho_0^k[\theta] + O\left(\frac{n}{(1-\zeta)^k}\right).$$

In the above identity, $a_k[\theta], \rho_0^k[\theta] \in \mathbb{Z}[\zeta, \frac{1}{n}]$ are not integral, but their difference is an algebraic integer $a_k[\theta] - \rho_0^k[\theta] \in \frac{n}{(1-\zeta)^k} \cdot \mathbb{Z}[\zeta]$.

Proof. Let $\theta = \sum_c n_c \sigma_c$ and $m = m(\theta) = \mid \{c \ : \ n_c \neq 0\} \mid$ be the number of non vanishing coefficients of θ. We prove this result by induction on m. First, note that

$$\binom{n_c/n}{k} = \frac{1}{k!} \cdot \frac{n_c^k}{n^k} \cdot (1 + O(n)).$$

Thus, if $\theta = n_c \sigma_c$ and $m = 1$, then:

$$f[\theta] = 1 + \sum_{k=1}^{n-1} \frac{1}{k!} \cdot \frac{n_c^k}{n^k} \cdot (1 + O(n)) \cdot \frac{D^k}{(1-\zeta)^k} = 1 + \sum_{k=1}^{N} \frac{a_k[\theta]}{k! n^k} D^k + O(D^{N+1}),$$

where, for $1 \leq k \leq N$,

$$a_k[\theta] = \rho_0^k[\theta] + O\left(\frac{n}{(1-\zeta)^k}\right),$$

which confirms the claim for $m = 1$. Suppose the claim holds for all $j \leq m$ and let $\theta = \theta_1 + \theta_2$ with $m(\theta_i) < m$ and $m(\theta) = m$. Then,

$$\begin{array}{rcl} f[\theta] & = & \left(1 + \frac{D}{1-\zeta}\right)^{\theta_1/n} \cdot \left(1 + \frac{D}{1-\zeta}\right)^{\theta_2/n} \\ & = & 1 + \sum_{k=1}^{N} \alpha_k[\theta] D^k + O(D^{N+1}), \end{array}$$

where for $k < n-1$ we have

$$\begin{aligned}\alpha_k[\theta] &= \sum_{j=1}^{k} \frac{a_j[\theta_1]}{n^j j!(1-\zeta)^l} \cdot \frac{a_{k-j}[\theta_2]}{n^{k-j}(k-j)!(1-\zeta)^{k-l}} \cdot (1+O(n)) \\ &= \frac{1}{k!n^k}\left(\rho_0[\theta_1]+\rho_0[\theta_2]\right)^k + O\left(\frac{n}{k!n^k(1-\zeta)^k}\right) \\ &= \frac{1}{k!n^k}\cdot\rho_0^k[\theta] + O\left(\frac{n}{k!n^k(1-\zeta)^k}\right) = \frac{1}{k!n^k}\cdot\left(\rho_0^k[\theta] + O(n/(1-\zeta)^k\right)\end{aligned}$$

This proves the claim by complete induction. □

Lemma 6.4.8 *By proceeding like in [Mihăilescu 2004][Lemma 8], we notice that* $\frac{a_k[\theta]}{k!} \in \mathbb{Z}[\zeta]$ *(notation is different between both articles).*

As a consequence, we will deduce that matrices built from the first coefficients occurring in some binary series developments are regular.

Lemma 6.4.9 *Let* $\theta = \sum_{c=1}^{n-1} n_c\sigma_c \in \mathbb{Z}[G]$ *such that* $\phi^{(-1)}(\theta) \not\equiv 0 \mod n$*, let* $f[\theta] = \left(1+\frac{D}{1-\zeta}\right)^{\theta/n}$ *and* $0 < N < n-1$ *be a fixed integer. Then,*

$$f[\theta] = 1 + \sum_{k=1}^{N} \frac{b_k[\theta]}{k!n^k(1-\zeta)^k} D^k + O(D^{N+1}), \quad \textit{with} \quad \frac{b_k[\theta]}{k!} \in \mathbb{Z}[\zeta].$$

Moreover, if $J \subset Gal(\mathbb{Q}[\zeta]/\mathbb{Q})$ *is a subset with* $|J| = N$*, then the matrix*[1]

$$A_N = (b_k[\sigma_c\theta])_{k=0;\sigma_c\in J}^{N-1} \in GL(\mathbb{K}, N)$$

.

Proof. Let $\lambda = 1-\zeta$; we show that the determinant of A_N is not zero modulo λ. Using Lemma 6.4.7, we know that we have a development of symbolic power series

$$f[\theta] = 1 + \sum_{k=1}^{N} \frac{a_k[\theta]}{k!n^k} D^k + O(D^{N+1}),$$

where

$$a_k[\theta] = \rho_0^k[\theta] + O\left(\frac{n}{(1-\zeta)^k}\right).$$

By definition, $(1-\zeta)^k \cdot a_k[\sigma_c\theta] \in \mathbb{Z}[\zeta]$ for all $\sigma_c \in G$. Let $b_k[\theta] = (1-\zeta)^k \cdot a_k[\theta] \in \mathbb{Z}[\zeta]$. Then, according to Lemma 6.4.7,

$$\begin{aligned} b_k[\sigma_c\theta] &= (1-\zeta)^k \cdot \left(\rho_0^k[\sigma_c\theta] + O\left(\tfrac{n}{(1-\zeta)^k}\right)\right) \\ &= \rho^k[\sigma_c\theta] + O(n) = \left(\sum_{l=1}^{n-1} n_l \cdot \tfrac{1-\zeta}{1-\zeta^{lc}}\right)^k + O(n) \\ &\equiv \left(\sum_{l=1}^{n-1} \tfrac{n_l}{lc}\right)^k \mod \lambda \equiv \left(\tfrac{\phi^{(-1)}[\theta]}{c}\right)^k \mod \lambda. \end{aligned}$$

[1] We shall apply this Lemma below, in a context in which J satisfies the additional condition that $i+j \neq n$ for any i, j with $\sigma_i \in J$ and $\sigma_j \in J$.

Thus, $\det A_N \equiv \left| \left(\left(\frac{\phi^{(-1)}[\theta]}{c} \right)^k \right)_{k=0,\sigma_c \in J}^{N-1} \right| \mod \lambda$. We have obtained a Vandermonde determinant:

$$\det A_N \equiv \left(\phi^{(-1)}[\theta]\right)^{N(N-1)/2} \cdot \prod_{i \neq j; \sigma_i, \sigma_j \in J} \left(\frac{1}{i} - \frac{1}{j} \right) \mod \lambda.$$

We have assumed that $\phi^{(-1)}[\theta] \not\equiv 0 \mod n$, and $1/i \not\equiv 1/j \bmod n$ for $\sigma_i, \sigma_j \in J$; this implies finally that $\prod_{\sigma_i, \sigma_j \in J} \left(\frac{1}{i} - \frac{1}{j} \right) \not\equiv 0 \mod n$, which confirms our claim. □

6.4.3 Proof of Theorem 6.2.3

[Mihăilescu 2008][Theorem 4] proves that if CF holds, then Equation (6.5) (diagonal Nagell-Ljunggren) has no solution except for (6.8). The computations in [Buhler & Harvey 2011] prove that CF holds for $n \leq 163.10^6$. This proves Theorem 6.2.3:1.. Theorem 6.2.3:3. is also proved in [Mihăilescu 2008][Theorem 4]. In the sequel we shall show that the only possible solutions are $X = \pm B/n^e + 1$. We may assume in particular that $n > 163 \cdot 10^6$.

We have already proved that $X - 1 = B \cdot C^n/n^e$ in Section 6.4.1.1. If $C = \pm 1$, then $X - 1 = \pm B/n^e$, as stated in point (2.) of Theorem 6.2.3 and X is a solution of Equation (6.5). The bounds on $|X|$ in (6.8) imply $|B| < n^n$, the second claim of Theorem 6.2.3:2..

Consequently, Theorem 6.2.3 will follow if we prove that $C = \pm 1$; we do this in this section. Assume that there is a prime $p|C$ with $p^i||C$. Let $\mathfrak{P} \subset \mathbb{Z}[\zeta]$ be a prime ideal lying above p and let $d(p) \subset G$ be its decomposition group. We shall use Remark 6.4.3 in order to derive some group ring elements which cancel the exponents κ occurring in Lemma 6.4.5.

Recall that $D = B \cdot C^n/n^e = X - 1$, with C defined by Equation (6.7). Note that Equation (6.7) implies that either $(n, D) = 1$, or $n^2|B$ and $(n, C) = 1$. Indeed, if $(n, D) \neq 1$, then $e = 1$ and $n^2|n^e(X-1) = BC^n$ and since $(C, n) = 1$, it follows that $n^2|B$; the last relation follows from the bounds $C^n \leq E < 4(n-2)^n$, hence $|C| < n$. In both cases $1/(1-\zeta)$ is congruent to an algebraic integer modulo $D/n^{v_n(D)} \cdot \mathbb{Z}[\zeta]$.

According to Remark 6.4.3, we know that there are at least two elements, $\sigma_1', \sigma_2' \in d(p)$ such that $\sigma_1' \neq \jmath \cdot \sigma_2'$. Let $\sigma_i'(\zeta) = \zeta^{c_i}$, $c_i \in (\mathbb{Z}/n\mathbb{Z})^\times$. It follows from Lemma 6.4.4 that, for $c_i \neq c_j$ when $i \neq j$, there are $h_1', h_2' \in \mathbb{Z}$ with $|h_i'| \leq \sqrt{n}$ and $\sum_{i=1}^2 h_i' c_i \equiv 0 \bmod n$ while $\sum_{i=1}^2 h_i'/c_i \not\equiv 0 \bmod n$.

We define

$$\begin{array}{rcl} \mu & = & \sum_{i=1}^2 h_i \sigma_i \in \mathbb{Z}[d(p)] \subset \mathbb{Z}[G], \quad \text{with} \\ h_i & = & \begin{cases} h_i' & \text{if } h_i' > 0 \text{ and} \\ -h_i' & \text{otherwise,} \end{cases} \quad \text{and} \\ \sigma_i & = & \begin{cases} \sigma_i' & \text{if } h_i' > 0 \text{ and} \\ \jmath\sigma_i' & \text{otherwise.} \end{cases} \end{array} \tag{6.15}$$

By construction, μ is a positive element, i.e. the coefficients $h_i \geq 0$. Let $\widehat{\cdot}: G \to (\mathbb{Z}/n\mathbb{Z})^\times$ denote the cyclotomic character and note that $h_i'\widehat{\sigma} = h_i\widehat{\sigma'}$ for $h_i' < 0$ and thus $\phi^{(1)}(\mu) = 0$. In view of Lemma 6.4.4, we also know that we can choose the h_i' and thus μ, such that

$$\phi^{(1)}(\mu) = 0, \quad \text{but} \quad \phi^{(-1)}(\mu) \neq 0.$$

Since $\mathbb{K}/\mathbb{Q}$ is abelian, all the primes $\mathfrak{P}|(p)$ have the same decomposition group $d(p)$ and μ enjoys the following stronger property: let $\mathfrak{P}|(p)$ and $S \subset G$ be a set of representatives of $G/d(p)$; let $\gamma \in \mathbb{Z}[\zeta]$ be such that $\gamma \equiv \zeta^{c_\sigma} \bmod \sigma(\mathfrak{P})$ for all $\sigma \in S$; then $\gamma^\mu \equiv 1 \bmod p\mathbb{Z}[\zeta]$, as follows directly from $\zeta^\mu \equiv 1 \bmod \sigma(\mathfrak{P})$, for all $\sigma \in S$.

In view of Lemma 6.4.6 and the fact that Fueter elements are positive, we also know that there is a $\theta_0 \in I_f^+$ such that $\varsigma(\theta_0) = 2$ and $\phi^{(-1)}(\theta_0) \neq 0$.

Let

$$\Theta = 2 \cdot \mu \cdot \theta_0.$$

In view of the properties (6.10) of moments and since for both μ and θ_0, the Fermat quotient vanishes, while $\phi^{(-1)}$ is non-null, it follows that the same must hold for Θ, so $\Theta \in 2 \cdot I_f^+$ and $\phi^{(-1)}(\Theta) \neq 0$. Let

$$h = 2 \cdot \sum_{i=1}^{l} |h_i| = 2.w(\mu),$$

where we defined the *absolute weight* $w(\sum_c n_c\sigma_c) = \sum_c |n_c|$. From subsection 6.4.2, we know that there exists a Jacobi integer $\beta[2\theta_0] \in \mathbb{Z}[\zeta]$ such that $\beta[2\theta_0]^n = (\zeta^{cx}(1-\zeta)^{1-e})^{\theta_0} \cdot \left(1 + \frac{X-1}{1-\zeta}\right)^{\theta_0}$ (see Equation (6.12)). It follows from Lemma 6.4.5, that we have $\beta[2\theta_0] \equiv \zeta^{\kappa(\theta_0)} \cdot Y^4 \bmod \mathfrak{P}$. We have chosen μ as a linear combination of two elements from the decompostion group $D(\mathfrak{P}) \subset G$, so μ acts on $\zeta \bmod \mathfrak{P}$ by $\zeta \bmod \mathfrak{P} \mapsto \zeta^\mu \equiv 1 \bmod \mathfrak{P}$. Therefore, from $\beta[\Theta] = \beta[2\theta_0]^\mu$ and thus, by the choice of μ, we have

$$\beta[\Theta] \equiv Y^h \bmod p\mathbb{Z}[\zeta]. \tag{6.16}$$

Let $\Theta = 2\sum_{c=1}^{n-1} n_c\sigma_c$; for any prime $\mathfrak{P}|(p)$, the binomial series of the $n-$th root of the right hand side in Equation (6.13) converges in the $\mathfrak{P}$ - adic valuation and its sum is equal to $\beta[\Theta]$ up to a possible $n-$th root of unity ζ^c. Here we make again use of the choice of Θ: comparing (6.16) with the product above, it follows that $\zeta^c = 1$ for all primes $\mathfrak{P} \mid (p)$. For any $N > 0$, we have $p^{inN}||D^N$ and thus

$$\beta[\Theta] \equiv Y^h \prod_{c=1}^{n-1} \left(\sum_{k=0}^{N-1} \binom{n_c/n}{k} \left(\frac{D}{1-\zeta^c} \right)^k \right) \mod p^{inN}. \tag{6.17}$$

We develop the product in a series, obtaining an expansion which converges uniformly at primes above p and is Galois covariant; for $N < n-1$ and $\sigma \in G$, we have:

$$\beta[\sigma\Theta] = Y^h \left(1 + \sum_{k=1}^{N-1} \frac{b_k[\sigma\Theta]}{(1-\zeta)^k n^k k!} \cdot D^k \right) + O(p^{inN}),$$

with $b_k[\Theta] \in \mathbb{Z}[\zeta]$. Let $P \subset \{1, 2, \cdots, n-1\}$ be a set of cardinal $1 < N < (n-1)/2$ such that if $c \in P$ then $n - c \notin P$, and $J \subset \mathbb{Z}[G]$ be the Galois automorphisms of $\mathbb{K}$ indexed by P: $J = \{\sigma_c\}_{c \in P}$. Consider the linear combination $\Delta = \sum_{\sigma \in J} \lambda_\sigma \cdot \beta[\sigma \cdot \Theta]$ where $\lambda_\sigma \in \mathbb{Q}[\zeta]$ verify the linear system:

$$\begin{aligned} \sum_{\sigma \in J} \quad & \lambda_\sigma \cdot b_k[\sigma \cdot \Theta] = 0, \text{ for } k = 0, \ldots, N-1, \ k \neq \lceil N/2 \rceil \quad \text{and} \\ \sum_{\sigma \in J} \quad & \lambda_\sigma \cdot b_{\lceil N/2 \rceil}[\sigma \cdot \Theta] = (1-\zeta)^{\lceil N/2 \rceil} n^{\lceil N/2 \rceil} \lceil N/2 \rceil !. \end{aligned} \tag{6.18}$$

Applying Lemma 6.4.9 we observe that this system is regular for any $N < n-1$. There exists therefore a unique solution in λ_σ which is not null.

We recall that a power series $\sum_{k=0}^{\infty} a_k X^k \in \mathbb{C}[[X]]$ is dominated by the series $\sum_{k=0}^{\infty} b_k X^k \in \mathbb{R}[[X]]$ with non-negative coefficients, if for all $k \geq 0$, we have $|a_k| \leq b_k$. The dominance relation is preserved by addition and multiplication of power series.

As in Lemma 5.2.8, one shows that if $r \in \mathbb{R}_{>0}$ and $\chi \in \mathbb{C}$, with $|\chi| \leq K$ with $K \in \mathbb{R}_{>0}$, then the binomial series $(1+\chi T)^r$ is dominated by $(1-KT)^{-r}$. From this, we obtain that $(1+\chi T)^{\Theta/n}$ is dominated by $(1-KT)^{-w(\Theta)/n}$. In our case of congruence (6.17), $T = D$, $\chi = \frac{1}{1-\zeta^c}$ and

$$K = \max_{1 \leq c < n} |1/(1-\zeta^c)| = 1/\sin(\pi/n) \leq n/\pi \cos(\pi/3) = 2n/\pi < n.$$

Applying this to our selected Θ, whose absolute weight is bounded by $w \leq 4n\sqrt{n}$, we find after some computations that $|b_k[\sigma \cdot \Theta]| < n^k \cdot \binom{-w/n}{k} \cdot k! < n^{3k}$ for $N < n/2$.

Let $A = \det\left(b_k[\sigma_c \cdot \Theta]\right)_{k=0;c\in I}^{N-1} \neq 0$ be the determinant of the matrix of the system (6.18), which is non vanishing, as noticed above: note that the division by $k!$ along a complete row does not modify the regularity of the matrix.

Let $\vec{d} = (1-\zeta)^{\lceil N/2 \rceil} n^{\lceil N/2 \rceil} \lceil N/2 \rceil! \left(\delta_{k,\lceil N/2 \rceil}\right)_{k=0}^{N-1}$, where $\delta_{i,j}$ is Kronecker's symbol. The solution to our system is $\lambda_\sigma = A_\sigma / A$, where $A_\sigma \in \mathbb{Z}[\zeta]$ are the determinants of some minors of $(b_k[\sigma_c \cdot \Theta])_{k=0;c\in I}^{N-1}$ obtained by replacing the respective column by $\vec{d}$.

Noticing that $|(1-\zeta)^{\lceil N/2 \rceil} n^{\lceil N/2 \rceil} \lceil N/2 \rceil !| < n^{3(N-1)}$, Hadamard's inequality implies that

$$\begin{aligned} |A_\sigma| &\leq n^{3(N-1)(N-2)/2} \cdot (N-1)^{(N-1)/2} \leq n^{3N^2/2} \cdot N^{N/2} \quad \text{and} \\ |A| &\leq n^{3N^2/2} \cdot N^{N/2} \end{aligned}$$

Let $\delta = A \cdot \Delta \in \mathbb{Z}[\zeta]$,

$$\delta = \sum_{\sigma \in J} A_\sigma \cdot \beta[\sigma \cdot \Theta] \in \mathbb{Z}[\zeta].$$

We set $N = \lceil n^{3/4} \rceil$ and claim that for such N, $\delta \neq 0$. By choice of the λ's, we have $\delta = A.p^{in\lceil N/2 \rceil}.u + p^{inN} z$ for some $z \in \mathbb{Z}[\zeta]$, where $u = \frac{D^{\lceil N/2 \rceil}}{p^{in\lceil N/2 \rceil}}.Y^h$ is a unit in $(\mathbb{Z}/p\mathbb{Z})^\times$. Therefore, if we assume that $\delta = 0$, then necessarily $p^{in\lceil N/2 \rceil}$ divides A. However, $v_p(A) < n\lceil N/2 \rceil$. Indeed, the upper bound for $|A|$ implies a fortiori that $v_p(A) \leq \lceil N/2 \cdot \log N + \frac{3N^2}{2} \log n \rceil$. Then, the assumption $\delta = 0$ would imply $n \leq 3\left\lceil n^{3/4} + \frac{1}{4} \right\rceil \log n$, which is false for $n \geq 4,5.10^6$. This contradicts thus our initial assumption. Therefore $\delta \neq 0$.

Given the bounds on A_σ, we obtain $|\delta| \leq N Y^h n^{3N^2/2} \cdot N^{N/2}$ and using the fact that $h < 4n^{1/2}$, $Y < n^n$ (Theorem 6.2.3:2.) and $N = \lceil n^{3/4} \rceil$, we find

$$|\mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\delta)| < \left(n^{\frac{11}{2} n^{3/2} + \frac{3}{8} n^{3/4} + \frac{3}{4}}\right)^{n-1}. \tag{6.19}$$

The initial homogenous conditions of the system (6.18) imply $\delta \equiv 0 \bmod p^{in\lceil N/2 \rceil}$, therefore $|\mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\delta)| \geq p^{in(n-1)N/2}$. Combining this inequality with inequality (6.19) and $n \geq 163 \cdot 10^6$, one finds that $\log p < 1.64$. This shows that $p = 2, 3$ or 5.

We consider the case $p \leq 5$ separately as follows. Note that in this case $p \not\equiv \pm 1 \bmod n$ and the decomposition group $D(p)$ contains the automorphism σ_p. We choose thus $\mu = 1 + p\jmath\sigma_p^{-1}$ and verify that $\varphi(\mu) = 0$, while $\phi^{-1}(\mu) = 1 - p^2 \not\equiv 0 \bmod n$. Consequently $\varsigma(\Theta) = 4(p+1)$ and the norm of δ is thus bounded by

$$p^{n(n-1)N/2} \leq |\mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\delta)| < \left(n^{4(p+1)+3N^2/2} \cdot N^{N/2+1}\right)^{n-1}.$$

Letting $N = 48$, we obtain the inequality

$$2^n \leq n^{73} \cdot 48^{25/24} < 64n^{73} \quad \Rightarrow \quad \frac{n-6}{73} \leq \log(n)/\log(2),$$

which is false for for $n > 695$, and a fortiori for $n > 163.10^6$. We obtain a contradiction in this case too, and thus $C = \pm 1$, which completes the proof of Theorem 6.2.3. □

6.5 Proof of Lemma 6.4.6

Proof. Let $\theta = \sigma_w\psi_u + \sigma_z\psi_v$. The conditions required by the lemma lead to the following linear system of equations over $\mathbb{F}_n$:

$$\left\{ \begin{array}{lcll} w \cdot \varphi(\psi_u) & + & z \cdot \varphi(\psi_v) & = 0 \\ 1/w \cdot \phi^{(-1)}(\psi_u) & + & 1/z \cdot \phi^{(-1)}(\psi_v) & \neq 0 \end{array} \right. \tag{6.20}$$

Considered as a linear system in the unknowns $w, z \in \mathbb{F}_n$, the above system has the matrix $M = \begin{pmatrix} \varphi(\psi_u) & \varphi(\psi_v) \\ \phi^{(-1)}(\psi_v) & \phi^{(-1)}(\psi_u) \end{pmatrix}$. Assume that the product $P(t) = \varphi(\psi_t) \cdot \phi^{(-1)}(\psi_t)$ is not constant for all $t \in (\mathbb{Z}/n \cdot \mathbb{Z})^\times$. Then there are two elements $u, v \in (\mathbb{Z}/n \cdot \mathbb{Z})^\times$ such that $P(u) \neq P(v)$; for such values u, v, the matrix M is regular over $\mathbb{F}_p$ and for any non vanishing right hand side in the second equation, the system has a unique solution (w, z). For this choice of $u, v; w, z$, the element $\theta = \sigma_w\psi_u + \sigma_z\psi_v$ satisfies the condition of the lemma.

We now show that $P(t) : (\mathbb{Z}/n \cdot \mathbb{Z})^\times \to \mathbb{F}_p$ is not a constant function. The proof uses explicit computations which include divisions by several constants which must be assumed to be non - null. Therefore we suppose that $n \notin E := \{3, 7\}$ and shall verify independently that the claim of the lemma holds for this exceptional set.

Let φ be the Fermat quotient map and Θ_k be the $k-$th Fuchsian. For any integer $1 < k < n-1$, we have:

$$\begin{array}{rcll} (n-k)^n - (n-k) & \equiv & -k^n - n + k & \bmod n^2 \\ & \equiv & -n\left(\frac{k^n-k}{n} + 1\right) & \bmod n^2. \end{array}$$

Dividing both terms by n and recalling from Lemma 5.2.9 that $\varphi(\Theta_k) = \varphi(k) \equiv \frac{k^n-k}{n} \bmod n$, we find:

$$\varphi(\Theta_{n-k}) = n - (1 + \varphi(\Theta_k)). \tag{6.21}$$

Using now (5.15) from Lemma 5.2.9, with $m = 2$, we find that:

$$\phi^{(-1)}(\Theta_k) \equiv \frac{k^3 - k}{2k^2} B_2 \equiv \frac{1}{12} \cdot \left(k - \frac{1}{k}\right) \quad \bmod n,$$

where we used the fact that $B_2 = 1/6$. Finally, using that $\psi_k = \Theta_{k+1} - \Theta_k$ for $k > 1$ while $\psi_1 = \Theta_2$, we obtain the following expressions for the moments of interest:

$$\begin{aligned} \varphi(\psi_k) &= \varphi(k+1) - \varphi(k), \\ \phi^{(-1)}(\psi_k) &\equiv \frac{1}{12} \cdot \left(1 + \frac{1}{k(k+1)}\right). \end{aligned}$$

Note that $\phi^{(-1)}(\psi_k) = 0$ iff $k^2 + k + 1 = 0$; if $n \equiv 1 \bmod 6$, the equation has two solutions in $\mathbb{F}_n$, otherwise it has none. In the latter case $\phi^{(-1)}(\psi_k) \neq 0$ for all k.

We shall assume that P is the constant function and shall show that this assumption fully determines the Fermat quotient of integers in dependence of $\varphi(2)$, and this determination is in contradiction with (6.21); the contradiction implies that P cannot be constant, thus completing the proof.

Let thus $C = \varphi(2) \cdot \phi^{(-1)}(\Theta_2) = \varphi(2) \cdot \frac{1}{8}$. Assume first that $\varphi(2) = 0$ and recall from (6.21) that $\varphi(k) + \varphi(n-k) + 1 = 0$. Therefore at least $\frac{n-1}{2}$ of the values of φ are non-vanishing. Since $\phi^{(-1)}(k) \cdot (\varphi(k+1) - \varphi(k)) = 0$ for all k we see that if $n \not\equiv 1 \bmod 6$, then φ is constantly vanishing, which is impossible.

If $n \equiv 1 \bmod 6$, let $l, m \in \mathbb{F}_n$ be the non trivial third roots of unity, so $\phi^{(-1)}(\psi_l) = \phi^{(-1)}(\psi_m) = 0$, while for all $k \notin \{l, m\}$ we must have $\varphi(k+1) = \varphi(k)$. In particular, if $l < m$, there are two integers a, b such that

$$\varphi(2) = 0 = \ldots = \varphi(l); \quad \varphi(l+1) = a = \ldots = \varphi(m); \quad \varphi(m+1) = b = \ldots = \varphi(n-1).$$

But $\varphi(n-1) = -1$ while $\varphi(n-2) = -1 - \varphi(2) = -1$, so $b = -1$. For symmetry reasons induced by (6.21), we must have $a = -1/2$ and $m = n - l$. This is absurd since $m^3 \equiv 1 \bmod n$ implies $l^3 = (n-m)^3 \equiv -m^3 \equiv -1 \bmod n$, so $n = 2 \not\equiv 1 \bmod 6$. Thus $\varphi(2) \neq 0$ in this case too. Since $\phi^{(-1)}(l) = 0$, it follows however that $C = \varphi(l) \cdot \phi^{(-1)}(l) = 0$ and thus $C = 0 = \varphi(2)/8$ and we should have $\varphi(2) = 0$, in contradiction with the facts established above. Consequently, if $n \equiv 1 \bmod 6$, then P cannot be constant.

We consider now the case $n \not\equiv 1 \bmod 6$, in which we know that $C \neq 0$. By expressing $C = P(2) = P(k)$ we obtain the following induction formula

$$\begin{aligned} C = \frac{1}{12} \cdot \frac{3\varphi(2)}{2} &= \frac{1}{12}(\varphi(k+1) - \varphi(k)) \cdot \frac{k^2+k+1}{k(k+1)}, \quad \text{hence} \\ \varphi(k+1) - \varphi(k) &= \frac{3\varphi(2)}{2} \cdot \frac{k(k+1)}{k^2+k+1}, \\ \varphi(3) - \varphi(2) &= \frac{9}{7}\varphi(2) \quad \Rightarrow \quad \varphi(3) = \frac{16}{7}\varphi(2). \end{aligned}$$

By eliminating $\varphi(2)$ from the above identity for two successive values of k one finds

$$\varphi(k+1) = \frac{2k^3}{k^3-1} \cdot \varphi(k) + \frac{k^3+1}{k^3-1} \cdot \varphi(k-1).$$

We shall use the reflexion formula (6.21) between the last and the first values in the

sequence $1, 2, \ldots, n-2, n-1$. Letting $k = n-2$ in the above induction, we find

$$\begin{aligned} -1 &\equiv \varphi(n-1) \equiv \frac{16}{9} \cdot \varphi(n-2) + \frac{7}{9} \cdot \varphi(n-3) \\ &\equiv \frac{16}{9} \cdot (-1 - \varphi(2)) + \frac{7}{9} \cdot (-1 - \varphi(3)) \bmod n, \\ 9 &\equiv 16 + 16\varphi(2) + 7 + 7\varphi(3) \equiv 23 + (16 + 7 \cdot \frac{16}{7})\varphi(2) \bmod n, \quad \text{hence} \\ -7 &\equiv 16 \cdot \varphi(2) \bmod n. \end{aligned}$$

Consequently $\varphi(2) \equiv -\frac{7}{16} \bmod n$ and thus $\varphi(3) \equiv \frac{16}{7}\varphi(2) \equiv -1 \bmod n$. But then the reflexion formula (6.21) implies that $\varphi(n-3) = -1 - \varphi(3) = 0$, and thus $C = 0$, in contradiction with the previously obtained non vanishing fact. This confirms that $P(t)$ is non constant in this case too.

It remains to verify the claim for the exceptional primes in E. For $n = 3$ the Stickelberger ideal is trivial, so there is nothing to prove. For $n = 7$ one can repeat the proof of the case $n \equiv 1 \bmod 6$, which requires no division by 7; this completes the proof of the Lemma. □

Appendix A
Addendum to [Mihăilescu 2008][Theorem 3]

The proof of Theorem 6.2.3 is based on results from [Mihăilescu 2008]. It has been pointed out that the proof of [Mihăilescu 2008][Theorem 3] may require some more detailed explanation in the case of a singular system of equations in the proof of [Mihăilescu 2008][Lemma 14]. Since the statements of [Mihăilescu 2008] are correct and can even be slightly improved, while the explanations may have seemed insufficient, we provide here for the readers interested to understand the technicalities of the proofs in [Mihăilescu 2008] some additional details and explanation, confirming those claims and results.

A.1 Clarification on the singular case of the Theorem 3 of [Mihăilescu 2008]

Let $m \in \mathbb{Z}_{>0}$ be a positive integer, $\mathbb{K}$ a field, $V = \mathbb{K}^m$ as a $\mathbb{K}$-vector space and let $L \subsetneq V$ be a proper subspace of V of dimension r. We assume that there exists at least one vector $w_1 \in L$ which is free of 0-coefficients over the canonical base $\mathcal{E}$. For $(x, y) \in V^2$, $x = (x_1, \cdots, x_m)$ and $y = (y_1, \cdots, y_m)$, the Hadamard product is defined by $[x, y] = (x_1 y_1, \cdots, x_m y_m)$. For any subspace $W \subset V$ we define the *W-bouquet* of L by

$$L_W = \langle \ \{ \ [w, x] \ : \ w \in W, \ x \in L \ \} \ \rangle_{\mathbb{K}},$$

the $\mathbb{K}$-span of all the Hadamard products of elements in W by vectors from L.

Lemma A.1.1 *Let $a_1 = (1, 1, \ldots, 1)$ over $\mathcal{E}$, and $a_2 \in V$ such that its coordinates be pairwise distinct over $\mathcal{E}$. Let $A_2 = \langle \{a_1, a_2\} \rangle_{\mathbb{K}}$ be the subspace generated by a_1, a_2. We assume that $w_1 \in W$, and we let L_{A_2} be the resulting A_2-bouquet. Then $dim(L_{A_2}) > dim(L)$.*

Proof. Obviously, $L \subset L_{A_2}$ (as $a_1 \in A$). We would like to show that $L_{A_2} \neq L$. We know that the system $(w_1, [w_1, a_2], [w_1, a_2^2], \cdots, [w_1, a_2^{m-1}])$ (the notion of power of a vector here is to be understood as an "Hadamard power") is free (as it induces a Vandermonde matrix over $\mathcal{E}$, w_1 does not have any zero among its coordinates and all coordinates of a_2 are pairwise distinct). We know that $w_1 \in L$; let us assume that $[w_1, a_2^i] \in L$ for $i \leq j$ (we know that $j \leq m - r < m$). Then, $[w_1, a_2^j] \in L$ and $[w_1, a_2^{j+1}] \notin L$. However, the Hadamard product of $[w_1, a_2^j] \in L$ by a_2, that is $[w_1, a_2^{j+1}]$, belongs to L_{A_2}. Thus, $\dim L_{A_2} > \dim L$. □

A.1.1 Application of Lemma A.1.1 to the proof of the singular case in the argument on pages 266 – 270 of [Mihăilescu 2008]

We apply here the lemma in the first case (that is $x \not\equiv s \mod p$, where $s \in \{-1, 0, 1\}$), the application to the second case being similar.

Let all notation be like in Lemma 14 in [Mihăilescu 2008]. As in [Mihăilescu 2008], we will assume that $\mathbf{A} = \left(\zeta^{-\kappa_{ac}/a}\right)_{a,c=1}^{(p-1)/2}$ (where κ_{ac} are the *Galois exponents*) is singular. Let $m = (p-1)/2$, $\mathbb{K} = \mathbb{Q}(\zeta_p)$ and $r = \text{rank }(\mathbf{A}) < (p-1)/2$. Without loss of generality, we assume that a regular r-submatrix of $\mathbf{A}$ is built with the first r rows and the first r columns. Therefore, the first r rows of $\mathbf{A}$ are independent, and we denote by W the sub-space of $V = \mathbb{K}^m$ generated by the first r row vectors $w_1, \cdots, w_r$ of $\mathbf{A}$. For $a_1 = (1, 1, \ldots, 1)$, we let a_2 be the vector of V whose components are [1] $(\eta(\sigma_c\theta))_{c=1}^{(p-1)/2}$ and $A_2 = \{a_1, a_2\}$. Then, according to Lemma A.1.1, there exists at least one vector $\vec{v} \in L_{A_2}$ which is independent on the first r vectors of $\mathbf{A}$.

Let $\mathbf{S}$ be the $(r+1) \times (r+1)$ submatrix of $\mathbf{A}$ comprising the first r rows and $r+1$ columns of $\mathbf{A}$, to which we have added an additional row: the first $r+1$ components of $\vec{v}$. Let $\vec{\lambda}'$ be the vector solution of $\mathbf{A}\vec{\lambda}' = \vec{d'}$, where $\vec{d'} = (\delta_{c,r+1})_{c=1}^{r+1}$. We know that $\vec{\lambda}' \neq \vec{0}$, as $\mathbf{S}$ is regular and $\vec{d'}$ is not the null vector. For $1 \leq c \leq r+1$, by Cramer's rule, $\lambda_c = \frac{S_c}{S}$, where S_c are the determinants of some minors of $\mathbf{S}$ obtained by replacing the c-column by $\vec{d'}$, and $S = \det \mathbf{S}$.

Let $\vec{\lambda} \in V$ be a vector whose first $r+1$ coordinates are those of $\vec{\lambda}'$ and the others are 0. Let $(\delta_{c,r+1})_{c=1}^{m}$. Then, $\vec{\lambda}$ verifies: $\mathbf{A}\vec{\lambda} = \vec{d}$.

Let $\delta = \sum_{c=1}^{r+1} \left(\lambda_c \cdot \beta_c + \overline{\lambda_c \cdot \beta_c}\right)$. Using Hadamard's inequality, we bound $|S_c| \leq \left(\frac{p-3}{2}\right)^{\frac{p-3}{4}} = D_1$ and $|S| \leq \left(\frac{p-1}{2}\right)^{\frac{p-1}{4}} = D_0$. Then, using the fact that the choice of λ_c eliminates the first term in the expansion of f_c, we find that $|S| \cdot |\delta| \leq 2x^{(p-1)/2p} \cdot \sum_{c=1}^{r+1} |S_c||R_{c,0}(x)|$, where $R_{c,0}(x) = f_c(x) - x^{(p-1)/2p}$. With the same arguments as in [Mihăilescu 2008], we deduce:

$$|S\delta| < 2(p-1)D_1 \cdot \frac{1}{|x|^{(p+1)/2p}}.$$

This inequality holds for all conjugates $\sigma_c(\delta)$, thus leading to:

$$|\mathbf{N}(S\delta)| < (2(p-1)D_1)^{(p-1)/2} \cdot \frac{1}{|x|^{\frac{(p-1)(p+1)}{4p}}}.$$

If $\delta \neq 0$, then $|\mathbf{N}(S\delta)| \geq 1$ and thus $|x| \leq 2^{5-p}\left(\frac{p}{2}\right)^{\frac{p}{2}}$. If $\delta = 0$, then $0 = S\delta = S \cdot |x|^{(p-1)/2} - \sum_{c=1}^{(p-1)/2} S_c R_{0,c}$, and thus:

$$|x| \leq \sum_c |S_c|/|S| < (p-1)D_1 < 3\left(\frac{p-3}{2}\right)^{(p-3)/2}.$$

These bounds are better than the ones in [Mihăilescu 2008], and this concludes the clarification.

[1] In the context of [Mihăilescu 2008], η corresponds to $b_1[\theta]$ in our context

Bibliography

[A.Bazso *et al.* 2010] A.Bazso, A.Bérczes, K.György and A.Pintér. *On the resolution of equations $Ax^n - By^n = C$ in integers x , y and $n \geq 3$, II.* Publicationes Mathematicae Debrecen, vol. 76, pages 227 – 250, 2010. (Cited on pages xi, xv, xix, 4 and 56.)

[Abouzaid 2008] M. Abouzaid. *Heights and logarithmic gcd on algebraic curves.* International Journal of Number Theory, no. 4, pages 177 – 197, 2008. (Cited on pages ix, xiii, xvii, 2, 10, 21, 22, 23, 28, 29, 30 and 34.)

[Bartolomé & Mihăilescu 2015] B. Bartolomé and P. Mihăilescu. *On the equation $x^n - 1 = B.z^n$.* Submitted to International Journal of Number Theory, 2015. (Cited on pages 41, 48 and 51.)

[Bartolomé *et al.* 2013] B. Bartolomé, Y. Bilu and F. Luca. *On the exponential local-global principle.* Acta Arithmetica, vol. 159, pages 101 – 111, 2013. (Cited on pages ix, xiii, xvii and 2.)

[Bartolomé 2015] B. Bartolomé. *The Skolem-Abouzaid theorem in the singular case.* Rendiconti Lincei - Matematica e Applicazioni, vol. 26, pages 263 – 289, 2015. (Cited on pages x, xiv, xviii and 4.)

[Bennet 2001] M. A. Bennet. *Rational Approximation To Algebraic Numbers Of Small Height: The Diophantine Equation $|ax^n - by^n| = 1$.* Journal für die reine und angewandte Mathematik, vol. 535, pages 1 – 49, 2001. (Cited on pages 55 and 56.)

[Bennett *et al.* 2006] M. A. Bennett, K. Győry, M. Mignotte and Á. Pintér. *Binomial Thue equations and polynomial powers.* Compositio Mathematica, vol. 142, pages 1103 – 1121, 2006. (Cited on pages xi, xv, xix, 4 and 56.)

[Bilu & Borichev 2013] Y. Bilu and A. Borichev. *Remarks on Eisenstein.* Journal of the Australian Mathematical Society, no. 94, pages 158 – 180, 2013. (Cited on pages 9, 25 and 28.)

[Bilu & Masser 2006] Y. Bilu and D. Masser. More sets, graphs and numbers, Chapter: A quick proof of Sprindzhuk's decomposition theorem, pages 25 – 32. Springer, Berlin, 2006. (Cited on page 29.)

[Bilu *et al.* 2014] Y. Bilu, Y. Bugeaud and M. Mignotte. The problem of Catalan. Springer, 2014. (Cited on pages 52 and 55.)

[Bilu 2004] Y. Bilu. *Catalan's conjecture (after Mihăilescu).* In Séminaire Bourbaki, editor, Exposé 909, pages 1 – 26, 2004. 55ème année (2002-2003); Astérisque 294. (Cited on pages 47 and 52.)

[Bombieri 1983] E. Bombieri. *On Weil's "Théorème de Décomposition".* American Journal of Mathematics, no. 105, pages 295 – 308, 1983. (Cited on page 29.)

[Broughan & Luca 2010] K. A. Broughan and F. Luca. *On the Fürstenberg closure of a class of binary recurrences.* Journal of Number Theory, no. 130, pages 696 – 706, 2010. (Cited on page 11.)

[Buhler & Harvey 2011] J. P. Buhler and D. Harvey. *Irregular primes to 163 million.* Mathematics of computation, vol. 80, no. 276, pages 2435 – 2444, 2011. (Cited on pages 55 and 65.)

[Catalan 1842] E. Catalan. *Théorèmes et problèmes.* Nouvelles Annales de Mathématiques, vol. 1, pages 519 – 521, 1842. (Cited on pages 1, 51 and 52.)

[Catalan 1844] E. Catalan. *Note extraite d'une lettre adressée à l'éditeur.* Journal für die reine und angewandte Mathematik, vol. 27, page 192, 1844. (Cited on page 51.)

[Corvaja & Zannier 2002] P. Corvaja and U. Zannier. *A subspace theorem approach to integral points on curves.* C. R., Math., Acad. Sci. Paris, vol. 334, no. 4, pages 267–271, 2002. (Cited on page 7.)

[Corvaja & Zannier 2004] P. Corvaja and U. Zannier. *On integral points on surfaces.* Annals of Mathematics, vol. 160, no. 2, pages 705 – 726, 2004. (Cited on page 7.)

[Corvaja & Zannier 2005] P. Corvaja and U. Zannier. *A lower bound for the height of a rational function at S-unit points.* Monatshefte für Mathematik, no. 144, pages 203 – 224, 2005. (Cited on page 12.)

[David & Philippon 1999] S. David and P. Philippon. *Minorations des hauteurs normalisées des sous-variétés des tores.* Annali della Scuola Normale Superiore di Pisa, Classe di Scienze, vol. 28, no. 3, pages 489 – 543, 1999. (Cited on page 24.)

[Dwork & Robba 1979] B. Dwork and P. Robba. *On natural radii of p-adic convergence.* Transactions of the American Mathematical Society, no. 256, pages 199 – 213, 1979. (Cited on page 28.)

[Dwork & van der Poorten 1992] B. M. Dwork and A. J. van der Poorten. *The Eisenstein Constant.* Duke Mathematical Journal, no. 65(1), pages 23 – 43, 1992. (Cited on page 28.)

[Faltings 1983] G. Faltings. *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern.* Inventiones mathematicae, vol. 73, no. 3, pages 349–366, 1983. (Cited on page 7.)

[Fueter 1922] R. Fueter. *Kummers Kriterium zum letzten Theorem von Fermat.* Mathematische Annalen, vol. 85, no. 1, pages 11–20, 1922. (Cited on page 42.)

[Habegger 2007] P. Habegger. *Heights and multiplicative relations on algebraic varieties.* PhD thesis, Universität Basel, 2007. (Cited on pages x, xiv, xviii and 3.)

[Ireland. & Rosen 1990] K. Ireland. and M.I. Rosen. A classical introduction to modern number theory. Graduate Texts in Mathematics. Springer, 1990. (Cited on pages 42 and 43.)

[Iwasawa 1975] K. Iwasawa. *A note on Jacobi sums.* Symposia Mathematica, vol. 15, pages 447 – 459, 1975. (Cited on page 43.)

[Jha 1992] V. Jha. *The Stickelberger Ideal in the Spirit of Kummer with Application to the First Case of Fermat's Last Theorem.* PhD thesis, Panjab University, 1992. (Cited on pages 42, 43 and 47.)

[Krantz 1994] S.G. Krantz. *Glossary of common math terms.* The Mathematical Intelligencer, vol. 16, no. 1, pages 36–36, 1994. (Not cited.)

[Lang 1994] S. Lang. Algebraic number theory. Graduate Texts in Mathematics. Springer New York, 1994. (Cited on page 41.)

[Lennon & McCartney 1970] John Lennon and Paul McCartney. *Let it be.* Abbey Road Studios, 1970. (Cited on page 48.)

[Ljunggren 1943] W. Ljunggren. *Noen Setninger om ubestemte likninger av formen* $(x^n - 1)/(x - 1) = y^q$. Norsk Mat. Tidsskr., vol. 25, pages 17 – 20, 1943. (Cited on page 53.)

[Matiyasevich 1970] Y. Matiyasevich. *Enumerable sets are Diophantine.* Doklady Akademii Nauk SSSR, no. 191, pages 279 – 282, 1970. (Cited on page 7.)

[Mihăilescu 2004] P. Mihăilescu. *Primary cyclotomic units and a proof of Catalan's conjecture.* Journal für die reine und angewandte Mathematik, vol. 572, pages 167 – 195, 2004. (Cited on pages 1, 8, 41, 44, 48, 51, 52, 61 and 64.)

[Mihăilescu 2007] Preda Mihăilescu. *New bounds and conditions for the equation of Nagell-Ljunggren.* Journal of Number Theory, vol. 124, no. 2, pages 380 – 395, 2007. (Cited on page 51.)

[Mihăilescu 2008] P. Mihăilescu. Diophantine approximation, Chapter: Class Number Conditions for the Diagonal Case of the Equation of Nagell and Ljunggren, pages 245 – 273. Springer Verlag, Development in Mathematics, 2008. (Cited on pages xi, xv, xix, xxiv, 5, 41, 43, 44, 48, 51, 53, 56, 57, 59, 61, 65, 71 and 72.)

[Nagell 1920a] T. Nagell. *Des équations indeterminées* $x^2 + x + 1 = y^n$ *et* $x^2 + x + 1 = 3y^n$. Norsk Mat. Forenings Skr., vol. 1, no. 2, 1920. (Cited on page 53.)

[Nagell 1920b] T. Nagell. *Note sur l'équation indeterminée* $(x^n - 1)/(x - 1) = y^q$. Norsk Mat. Tidsskr., vol. 1, no. 2, pages 75 – 78, 1920. (Cited on page 53.)

[Poulakis 2004] D. Poulakis. *Integer points on rational curves with fixed gcd.* Publicationes Mathematicae Debrecen, vol. 64, no. 3 – 4, pages 369 – 379, 2004. (Cited on page 21.)

[Ribet 1990] K.A. Ribet. *On modular representations of* $Gal(\bar{\mathbb{Q}}/\mathbb{Q})$ *arising from modular forms.* Inventiones mathematicae, vol. 100, pages 431–476, 1990. (Cited on page 1.)

[Schinzel 1974] A. Schinzel. *Primitive divisors of the expression* $A^n - B^n$ *in algebraic number fields.* Journal für die reine und angewandte Mathematik, no. 268/269, pages 27 – 33, 1974. (Cited on page 15.)

[Schinzel 1975] A. Schinzel. *Power residues and exponential congruences.* Acta Arithmetica, no. 27, pages 397 – 420, 1975. (Cited on page 11.)

[Schinzel 1977] A. Schinzel. *Abelian binomials, power residues and exponential congruences.* Acta Arithmetica, no. 32, pages 245 – 274, 1977. (Cited on pages ix, xiii, xvii, 2 and 11.)

[Schinzel 2003] A. Schinzel. *On the congruence $u_n \equiv c \pmod p$, where u_n is a recurring sequence of the second order.* Acta Academiae Paedagogicae Agriensis, Sectio Mathematicae, no. 30, pages 147 – 165, 2003. (Cited on page 11.)

[Schmidt 1990] W. M. Schmidt. *Eisenstein's theorem on power series expansions of algebraic functions.* Acta Arithmetica, vol. 56, no. 2, pages 161 – 179, 1990. (Cited on pages 9 and 28.)

[Schoof 2008] R. Schoof. Catalan's conjecture. Springer London, 2008. (Cited on page 52.)

[Siegel 1929] C. L. Siegel. über einige Anwendungen Diophantischer Approximationen, pages 41 – 69. Abh. Preuss. Akad. Wiss. Phys. Math. Kl., 1929. (Cited on page 7.)

[Skolem 1929] T. Skolem. Lösung gewisser Gleichungssysteme in ganzen Zahlen oder ganzzahligen Polynomen mit beschrnktem gemeinschaftlichen Teiler. Oslo Vid. Akar. Skr. I 12, 1929. (Cited on pages ix, xiii, xvii, 2 and 21.)

[Skolem 1937] T. Skolem. *Anwendung exponentieller Kongruenzen zum Beweis der Unlösbarkeit gewisser diophantischer Gleichungen.* Avhdl. Norske Vid. Akad. Oslo I, no. 12, pages 1 – 16, 1937. (Cited on pages ix, xiii, xvii, 2 and 11.)

[Stewart 1977] C. L. Stewart. Transcendence theory: Advances and applications, Chapter: Primitive divisors of Lucas and Lehmer numbers, pages 79 – 92. Academic Press, 1977. (Cited on page 15.)

[Stickelberger 1890] L. Stickelberger. *Über eine Verallgemeinerung der Kreistheilung.* Mathematische Annalen, vol. 37, no. 3, pages 321–367, 1890. (Cited on page 42.)

[Sun 2011] C. L. Sun. Hyperplanes inside an algebraic torus and a conjecture of Skolem over a global function field. arXiv:1101.3045, 2011. (Cited on page 11.)

[Thue 1909] A. Thue. *Über Annäherungswerte Algebraischer Zahlen.* Journal für die reine und angewandte Mathematik, vol. 135, pages 284 – 305, 1909. (Cited on page 54.)

[Walsh 1992] P. G. Walsh. *A quantitative version of Runge's theorem on diophantine equations.* Acta Arithmetica, vol. LXII, no. 2, pages 157 – 172, 1992. (Cited on page 21.)

[Washington 1997] L. Washington. Introduction to cyclotomic fields. Springer, New York, 2 édition, 1997. (Cited on pages 13 and 42.)

[Wiles 1995] A.J. Wiles. *Modular elliptic curves and Fermat's Last Theorem.* Annals of mathematics, vol. 141, pages 443 – 551, 1995. (Cited on page 1.)

[Wüstholz 2002] G. Wüstholz, editor. A panorama of number theory or the view from Baker's garden, Cambridge books online. Cambridge University Press, 2002. (Cited on page 15.)

BORIS BARTOLOMÉ

Forward address :	La cour 31320 Aureville, France	**Birth date :**	February 26th, 1969
		Birth location :	Toulouse, France
Mobile phone :	+33 672 62 40 59	**Citizenship :**	French
Email address :	Boris.Bartolome@stud.uni-goettingen.de	**Marital status :**	Married

POSITIONS HELD

2010 - Present	Enteleia (consulting)	*Founder and CEO*	Toulouse, France
2007 - 2010	Oliver Wyman	*Principal*	Paris, France
2006 - 2007	Independent consultant	*Qualcomm as Product manager*	London, UK
2003 - 2006	Altran Technologies	*Senior Consultant*	Toulouse, France
2002 - 2007	Ecole Nationale Aviation Civile	*Lecturer in mathematics*	Toulouse, France
2000 - 2002	Hughes Network Systems (HNS)	*Project Manager/ Business developer*	San Diego, USA
2000 - 2002	UC San Diego/CSU Hayward	*Lecturer/Assistant professor in Electrical Engineering*	CA, USA
1997 - 1999	Alcatel Space Industries	*R&D Engineer*	Paris, France

EDUCATION

2015	Göttingen Universität PhD Division of Mathematics and Natural Sciences "Diophantine equations and cyclotomic fields"	Göttingen, Germany
2015	Université de Bordeaux PhD in Pure Mathematics "Diophantine equations and cyclotomic fields"	Bordeaux, France
2004	INSEAD *MBA* with special emphasis in finance, strategy and entrepreneurship	Fontainebleau, France
1999	Ecole Nationale Supérieure des Télécommunications *PhD* in electrical engineering with the highest French distinction "Use of turbo codes for multimedia transmissions over satellite"	Paris, France
1999	Institut National Polytechnique *MS* in Fundamental Computer Processes and Parallelism	Toulouse, France
1998	Ecole Nationale Supérieure de l'Aéronautique et de l'Espace *MS* in Applied Mathematics	Toulouse, France
1997	Ecole Nationale Supérieure des Télécommunications *MS* in Signal Processing, Digital Communications and Images	Paris, France
1997	Ecole Nationale Supérieure des Télécommunications de Bretagne *Electrical Engineer* (equivalent to Master of Science)	Brest, France
1994	Ecole Nationale de l'Aviation Civile *ATPL* (fully licensed commercial pilot)	Toulouse, France

PUBLICATIONS

2015	On the equation $X^n-1 = B.Z^n$, submitted to International Journal of Number Theory • Authors: Boris Bartolome, Preda Mihailescu.
2014	The Skolem-Abouzai'd's theorem in the singular case, Rendiconti Lincei 06/2015; 26:1-27. DOI: 10.4171/RLM • Author: Boris Bartolome.
2013	On the exponential local-global principle, Acta Arithmetica • Authors: Boris Bartolome, Yuri Bilu, Florian Luca.
2000	Parallel Turbo Codes' Interleaver Optimization, Second International Symposium on Turbo Codes & Related Topics • Authors: Boris Bartolome, N. Durand, J.M. Alliot, ML. Boucheret.
1999	Turbo Codes Optimization Using Genetic Algorithms, Congress on Evolutionary Computation • Authors: Boris Bartolome, • N. Durant, J.M. Alliot.
1999	Parallel Turbo Codes Optimization, COST Spring'99 Workshop • Authors: Boris Bartolome, N. Durand, J.M. Alliot, M.L. Boucheret.
1998	Digital Filter Banks For Satellite With On Board Processing, 6th International Workshop on Digital Signal Processing Techniques for Space Applications • Authors: Boris Bartolome, ML. Boucheret, I. Mortensen, H. Favaro.