THÈSE

PRÉSENTÉE A

L'UNIVERSITÉ BORDEAUX 1

ÉCOLE DOCTORALE DES SCIENCES Mathématiques et Informatique

Par Min SHA

POUR OBTENIR LE GRADE DE DOCTEUR

SPÉCIALITÉ : Mathématiques Pures

Problèmes autour des courbes elliptiques et modulaires

(Topics in Elliptic and Modular Curves)

Directeur de recherche : Yuri BILU

Soutenue le 27 Septembre 2013 Devant la commission d'examen formée de :

M. Denis BENOIS	Université Bordeaux 1	examinateur
M. Yuri BILU	Université Bordeaux 1	directeur de thèse
M. Yann BUGEAUD	Université de Strasbourg	rapporteur, président du jury
M. Loïc MEREL	Université Paris Diderot	rapporteur
M. Pierre PARENT	Université Bordeaux 1	examinateur
M. Tarlok SHOREY	Indian Institute of Technology Bombay	examinateur

Acknowledgements

First of all, I would like to thank my advisor, Yuri Bilu, for all his guidance, supports and encouragements during the work of this thesis. He is very capable in identifying the essential difficulties in an efficient way, and in providing practical and thought-provoking viewpoints that leads to solutions of problems. His kindness and generosity provide me a comfortable circumstance for my journey in mathematics. He also strongly supports me on postdoctoral applications and provides me many opportunities. I certainly cannot feel more grateful to him.

I also would like to thank Igor Shparlinski for introducing me the subject of pairing-friendly elliptic curves, proposing valuable suggestions, and supporting my postdoctoral applications. Especially, I will do postdoctoral research with him after my PhD.

I want to express my very sincere gratitude to Yann Bugeaud and Loïc Merel for agreeing to serve as the reporters on my thesis and for reading this thesis very carefully and giving valuable comments. I would also like to thank the rest of the thesis jury: Denis Benois, Pierre Parent and Tarlok Shorey, for their stimulating comments.

My special thank goes to Linsheng Yin for his supports and encouragements. He was the advisor of my master's thesis. I started my journey in number theory under his supervision. I also want to thank Min Wu for her supports and inspirations. She was the advisor of my bachelor's thesis, and exactly at that time I chose mathematics as my whole life's career with her encouragements.

I am indebted to Qing Liu. Thanks to his recommendation, I contacted Yuri Bilu successfully and then got financial support from the Chinese government to launch my PhD study in Bordeaux. I also deeply appreciate Jean-Marc Couveignes. I learned a lot from his course "Algorithmic Number Theory" and the mini-course "Algorithms for algebraic curves". I am thankful to John Boxall for his valuable comments, which are crucial to improve my work on pairing-friendly elliptic curves, and also for his recommendation letters on my postdoctoral applications. I am also grateful to Pierre Parent for his continuous supports on searching for a postdoctoral position. I want to thank

Jilong Tong for valuable suggestions concerning my research career. I also would like to express my deep gratitude to Christine Parison for helping me deal with those affairs concerning the defence.

Many colleagues at Institut de Mathématiques de Bordeaux (IMB) have helped and encouraged me in various ways. Louis Merlin helped me deal with some administrative affairs at the beginning. Aurélien Bajolet allowed me to work with him on his work which yielded the first paper of my PhD. Nicolas Mascot helped me to obtain an account of PlaFRIM. Nicolas Mascot and Aurel Page taught me how to use PlaFRIM for computations, they also helped me a lot to rectify the French abstract. Andrea Siviero invited me to give a talk in the PhD Students Seminar in Number Theory. Exactly in that talk, Samuel Le Fourn pointed out a serious mistake in my work. Nicolas Delfosse and Andrea Siviero offered me help on editing the latex. Velibor Bojkovic helped me prepare the presentation. I would like to thank all of them heartily.

I must thank all the staff, faculty and students at IMB for making it enjoyable. The active and high-level scientific activities concerning about number theory at IMB benefit me a lot, including the famous Séminaire de Théorie des Nombres de Bordeaux, and the Séminaire de Théorie Algorithmique des Nombres organized by Andreas Enge.

Thanks to the friendship with my Chinese friends, I live a happy life in Bordeaux. I would like to thank all of them. My special thanks go to Huihui Ding, Zhuowei Du, Ziying Fan, Nianlong Guo, Yuan He, Xin Jin, Na Li, Chunhui Liu, Junjiang Liu, Na Liu, Huajun Lv, Dong Wang, Yanlan Wang, Yuxia Wang, Yongliang Xiong, Wenqing Xu, Feiying Yang, Di Zhang, Pengxiang Zhao, and Na Zhou.

Last but by no means least, I thank all my friends, teachers, and family members, who have encouraged and supported me in various stages of my life. I am also grateful to the China Scholarship Council for financial support.

Résumé

Cette thèse se divise en deux parties. La première est consacrée aux points entiers sur les courbes modulaires, et l'autre se concentre sur les courbes elliptiques à couplages sur corps finis.

1. Majorations effectives des points entiers sur les courbes modulaires

La première partie est la partie principale. Dans cette partie, nous donnons quelques majorations effectives de la hauteur des j-invariants des points entiers sur les courbes modulaires quelconques associées aux sous-groupes de congruence sur les corps de nombres quelconques en supposant que le nombre des pointes est au moins 3. De plus, dans le cas d'un groupe de Cartan non-déployé nous fournissons de meilleures bornes. Comme application, nous obtenons des résultats similaires pour certaines courbes modulaires avec moins de 3 pointes.

Soit $\mathcal{H}$ le demi-plan de Poincaré: $\mathcal{H} = \{\tau \in \mathbb{C} : \mathrm{Im}\tau > 0\}$. De plus on notera $\bar{\mathcal{H}} = \mathcal{H} \cup \mathbb{Q} \cup \{i\infty\}$ le demi-plan de Poincaré étendu. Le groupe modulaire $SL_2(\mathbb{Z})$, agit par homographie sur $\mathcal{H}$, via l'action:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (\tau) = \frac{a\tau + b}{c\tau + d}.$$

Soit $\Gamma(N)$ le sous-groupe principal de congruence de niveau N qui est le sous-groupe de $SL_2(\mathbb{Z})$ formé des classes de matrices congrues modulo N à la matrice identité. En particulier, $\Gamma(1) = SL_2(\mathbb{Z})$. On appelle sous-groupe de congruence un sous-groupe de $SL_2(\mathbb{Z})$ qui contient le sous-groupe $\Gamma(N)$ pour un entier N.

Soit Γ un sous-groupe de congruence de $SL_2(\mathbb{Z})$, on définit alors la courbe modulaire associée à Γ, par

$$X_\Gamma = \Gamma \backslash \bar{\mathcal{H}}.$$

Soit j l'invariant modulaire qui définit sur $\mathcal{H}$ par le développement familier suivant

$$j(\tau) = q_\tau^{-1} + 744 + 196844q_\tau + \cdots ,$$

où $q_\tau = e^{2\pi i\tau}$.

Le corps de fonctions de $X(1)$ est $\mathbb{Q}(j)$. Le corps de définition de $X(N)$ est $\mathbb{Q}(\zeta_N)$. On note par $\mathbb{Q}(X(N))$ le corps de fonctions de $X(N)$. Alors, $\mathbb{Q}(X(N))/\mathbb{Q}(j)$ est une extension de Galois, est le groupe de Galois est isomorphe au groupe $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\pm 1$.

Soit G un sous-groupe de $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ contenant -1. Par la théorie de Galois, G correpond uniquement à un corps intermédiaires de $\mathbb{Q}(X(N))/\mathbb{Q}(j)$. Donc, G correpond uniquement à une courbe notée par X_G. X_G est une courbe modulaire de niveau N. On note par $\nu_\infty(G)$ le nombre de pointes de X_G.

Suppose que X_G est défini sur un corps de nombres K. Soit S un ensemble de valeurs absolues normalisées de K contenant les valeurs absolues archimédiennes. Soit $\mathcal{O}_S$ l'anneau de S-entiers de K. Soit P un K-point rationnel sur X_G. Si $j(P) \in \mathcal{O}_S$, on dit que P est un point S-entier sur X_G. En particulier, P s'appelle point entier sur X_G si $j(P) \in \mathcal{O}_K$, où $\mathcal{O}_K$ est l'anneau d'entiers de K. On définit l'ensemble

$$X_G(\mathcal{O}_S) = \{P \in X(K) : j(P) \in \mathcal{O}_S\}.$$

Le théroème de Siegel implique le théroème suivant.

Théorème [Siegel] *$X_G(\mathcal{O}_S)$ est fini si le genre de X_G est plus que zéro ou j a plus que deux pôles.*

En 1995, Bilu a obtenu un théroème suivant.

Théorème [Bilu] *Le théroème de Siegel est effectif pour (X_G, j) si X_G a plus que deux pointes.*

Mais, il n'a pas donné des résultats quantitatifs. Dans cette partie, l'objectif principal est de obtenir des résultats quantitatifs sur le théroème de Bilu par utiliser la méthode de Baker.

Soit α un élément de K. On note par $\mathrm{h}(\alpha)$ la hauteur logarithmique absolue.

Soit $d = [K : \mathbb{Q}]$ et $s = |S|$. On définit

$$\Delta_0 = d^{-d}\sqrt{|D_K|}(\log |D_K|)^d \prod_{\substack{v\in S\\ v\nmid\infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v),$$

$$\Delta = d^{-d}\sqrt{N^{d_N}|D_K|^{\varphi(N)}}\left(\log(N^{dN}|D_K|^{\varphi(N)})\right)^{d\varphi(N)}\left(\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v)\right)^{\varphi(N)},$$

où D_K le discriminant de K. On note par p le premier maximal au-dessous de S. Si S contient juste les valeurs absolues archimédiennes, alors $p = 1$. On définit $\mathrm{h}(P) = \mathrm{h}(j(P))$ quand P est un point S-entier sur X_G.

Théorème [Sha] *Soit N pas une puissance d'un premier. Soit $\nu_\infty(G) \geq 3$. Soit P un point S-entier sur X_G. Alors,*

$$\mathrm{h}(P) \leq \left(CdsN^2\right)^{2sN}(\log(dN))^{3sN}p^{dN}\Delta,$$

où C est une constante absolue effective.

Théorème [Sha] *Suppose que $K \subseteq \mathbb{Q}(\zeta_N)$, et S contient juste les valeurs absolues archimédiennes. Soit N pas une puissance d'un premier. Soit $\nu_\infty(G) \geq 3$. Soit P un point S-entier sur X_G. Alors,*

$$\mathrm{h}(P) \leq C^{\varphi(N)}N^{\frac{3}{2}\varphi(N)+10}(\log N)^{\frac{5}{2}\varphi(N)-2},$$

où C est une constante absolue effective, et $\varphi(N)$ est la fonction d'Euler.

Théorème [Sha] *Suppose que $\mathbb{Q}(\zeta_N) \subseteq K$. Soit N pas une puissance d'un premier. Soit $\nu_\infty(G) \geq 3$. Soit P un point S-entier sur X_G. Alors,*

$$\mathrm{h}(P) \leq (Cds)^{2s}(\log d)^{3s}N^8p^d\Delta_0\log p,$$

où C est une constante absolue effective.

La situation est différente quand N est une pusissance d'un premier. Dans ce cas, on définit

$$M = \begin{cases} 2N & \text{si } N \text{ n'est pas une puissance de } 2, \\[2em] 3N & \text{si } N \text{ est une puissance de } 2. \end{cases}$$

Théorème [Sha] *Soit N une puissance d'un premier. Soit $\nu_\infty(G) \geq 3$. Soit P un point S-entier sur X_G. Alors, on peut obtenir trois majorations effectives pour $\mathrm{h}(P)$ par replacer N par M dans les trois Théorèmes derniers.*

A partir de maintenant, soit p un premier. Le normalisateur d'un sous-groupes de Cartan non déployé est défini par

$$\mathcal{C}_{\text{ns}}^{+}(p)=\left\{\begin{pmatrix}\alpha & \Xi\beta\\ \beta & \alpha\end{pmatrix},\begin{pmatrix}\alpha & \Xi\beta\\ -\beta & -\alpha\end{pmatrix}:\alpha,\beta\in\mathbb{F}_p,(\alpha,\beta)\neq(0,0)\right\},$$

où Ξ est un non-résidu quadratique modulo p. On note par $X_{\text{ns}}^{+}(p)$ la courbe modulaire correspondant à le groupe $\mathcal{C}_{\text{ns}}^{+}(p)$.

Théorème [Bajolet et Sha] *Soit p un premier plus que 5. Soit d un diviseur de $\frac{p-1}{2}$ plus que 2. Soit P un point entier sur $X_{\text{ns}}^{+}(p)$. Alors,*

$$\mathrm{h}(P)=\log|j(P)|<C(d)p^{6d+5}(\log p)^2,$$

où $C(d)=30^{d+5}\cdot d^{-2d+4.5}$.

Théorème [Bajolet et Sha] *Soit p un premier plus que 5. Soit P un point entier sur $X_{\text{ns}}^{+}(p)$. Alors,*

$$\log|j(P)|<41993\cdot 13^{p}\cdot p^{2p+7.5}(\log p)^2.$$

Théorème [Bajolet et Sha] *Suppose que un premier $p\geq 7$ et $p\equiv 1$ (mod 3). Soit P un point entier sur $X_{\text{ns}}^{+}(p)$. Alors,*

$$\log|j(P)|<30^{8}\cdot p^{23}(\log p)^2.$$

2. Analyses heuristiques sur les courbes elliptiques à couplages

En 1985 Miller et Koblitz ont introduit, indépendamment l'un de l'autre, la cryptographie fondée sur les groupes des points rationnels d'une courbe elliptique définie sur un corps fini. Ils proposent de généraliser des protocoles tels que léchange de clés Diffie-Hellman ou la signature d'El Gamal.

En 2000, Joux met à profit les couplages sur les courbes elliptiques en expliquant qu'il est possible, avec les propriétés de bilinéarité du couplage de Weil, de faire un échange type Diffie-Hellman entre trois personnes en un tour seulement. Lors de la conférence Crypto 2001, Boneh et Franklin proposent à leur tour un schéma de chiffrement basé sur l'identité utilisant ce couplage. La cryptographie basée sur les couplages connaît depuis un véritable engouement.

Dans cette partie, nous donnons une nouvelle majoration du nombre de classes d'isogénie de courbes elliptiques ordinaires à couplages. Nous analysons également la méthode de Cocks-Pinch pour confirmer certaines de ses propriétés communément conjecturées. Par ailleurs, nous présentons la première analyse heuristique connue qui suggère que toute construction efficace de courbes elliptiques à couplages peut engendrer efficacement de telles courbes sur tout corps à couplages. Enfin, quelques données numériques allant dans ce sens sont données.

Mots-clefs

courbe modulaire, point entier, j-invariant, méthode de Baker, courbe elliptique à couplages, méthode de Cocks-Pinch, corps à couplages.

Abstract

Abstract

This thesis is divided into two parts. One is devoted to integral points on modular curves, and the other concerns pairing-friendly elliptic curves.

In the first part, we give some effective upper bounds for the j-invariant of integral points on arbitrary modular curves corresponding to congruence subgroups over arbitrary number fields assuming that the number of cusps is not less than 3. Especially, in the non-split Cartan case we provide much better bounds. As an application, we get similar results for certain modular curves with less than three cusps.

In the second part, a new heuristic upper bound for the number of isogeny classes of ordinary pairing-friendly elliptic curves is given. We also heuristically analyze the Cocks-Pinch method to confirm some of its general consensuses. Especially, we present the first known heuristic which suggests that any efficient construction of pairing-friendly elliptic curves can efficiently generate such curves over pairing-friendly fields. Finally, some numerical evidence is given.

Keywords

modular curve, integral point, j-invariant, Baker's method, pairing-friendly elliptic curve, Cocks-Pinch method, pairing-friendly field.

Contents

I Effective Bounds for Integral Points on Modular Curves **1**

1 Introduction **2**

1.1 Modular Curves . . . 2
1.2 Siegel's Theorem . . . 4
1.3 Structure of Part I . . . 7

2 Bounding the j-invariant of integral points on modular curves **9**

2.1 Main results . . . 9
2.2 Notation and conventions . . . 11
2.3 Preparations . . . 12
2.3.1 Siegel functions . . . 12
2.3.2 Modular units on $X(N)$ . . . 14
2.3.3 X_G and X_{G_1} . . . 15
2.3.4 Modular units on X_{G_1} . . . 17
2.4 Siegel's theory of convenient units . . . 19
2.5 Baker's inequality . . . 20
2.6 The case of mixed level . . . 23
2.6.1 Preparation for Baker's inequality . . . 24
2.6.2 Using Baker's inequality . . . 26
2.6.3 Proof of Theorem 2.1 . . . 28
2.6.4 Proof of Theorem 2.2 . . . 29
2.6.5 Proof of Theorem 2.3 . . . 31
2.7 The case of prime power level . . . 31
2.8 The case $\Lambda = 1$ . . . 32

3 Bounding the j-invariant of integral points on $X_{\mathrm{ns}}^{+}(p)$ **36**

3.1 Background . . . 36
3.2 Main Results . . . 37
3.3 Notation and conventions . . . 38
3.4 Preparations . . . 39
3.4.1 Siegel functions and modular units . . . 39
3.4.2 $X_{\mathrm{ns}}^{+}(p)$ and X_H . . . 40
3.4.3 Modular units on X_H . . . 40
3.5 Baker's method on X_H . . . 42
3.5.1 Baker's inequality . . . 42
3.5.2 Cyclotomic units . . . 43
3.5.3 More about modular units on X_H . . . 43

3.5.4 Upper bound for B . . . 45
3.5.5 Preparation for Baker's inequality . . . 46
3.5.6 Using Baker's inequality . . . 47
3.6 Computations . . . 48
3.6.1 Upper Bound for m . . . 48
3.6.2 Height of η_{k-1} for $k = 2, \ldots, (p-1)/2$ . . . 49
3.6.3 Height of η_0 . . . 49
3.6.4 Height of $|\Upsilon_{c,\sigma}|$ . . . 50
3.6.5 Calculation of Ω . . . 51
3.6.6 Calculation of B_0 . . . 51
3.6.7 Final results . . . 52
3.7 The case $\Lambda = 0$ . . . 52

4 Bounding the j-invariant of integral points on certain modular curves 55
4.1 Main Results . . . 55
4.2 Quantitative Riemann existence theorem for $X_{\widetilde{\Gamma}}$ . . . 57
4.3 Quantitative Chevalley-Weil theorem for $\phi : X_{\widetilde{\Gamma}} \to X_{\Gamma}$ . . . 58
4.4 Proof of Theorems . . . 59

II Heuristics of Pairing-friendly Elliptic Curves 61

5 Introduction 62
5.1 Motivation . . . 62
5.2 Structure of Part II . . . 64
5.3 Preliminary and Notation . . . 64

6 Upper bound for isogeny classes of ordinary pairing-friendly elliptic curves 66

7 Heuristics of the Cocks-Pinch method 69
7.1 Background on the Cocks-Pinch method . . . 69
7.2 Heuristics from algebraic number theory . . . 70
7.3 Heuristics from the Bateman-Horn conjecture . . . 74

8 Involving Pairing-friendly Fields 78

9 Numerical Evidence 82

Bibliography 87

Part I

Effective Bounds for Integral Points on Modular Curves

Chapter 1

Introduction

1.1 Modular Curves

We briefly recall some basic definitions and notation concerning modular curves. For all missing details one may consult, for instance, [40, 63, 82].

Recall that for every positive integer N, the *principal congruence subgroup* $\Gamma(N)$ of $\mathrm{SL}_2(\mathbb{Z})$ is the kernel of the reduction map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. By convention we define $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$. We say that a subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is a *congruence subgroup of level* N if it contains $\Gamma(N)$. The minimal N with this property will be called the *exact level* of Γ. For every positive integer N, there are two classical congruence subgroups of level N:

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod N \right\},$$
$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod N \right\}.$$

Let $\mathcal{H}$ denote the Poincaré upper half-plane: $\mathcal{H} = \{\tau \in \mathbb{C} : \mathrm{Im}\tau > 0\}$. We also put $\bar{\mathcal{H}} = \mathcal{H} \cup \mathbb{Q} \cup \{i\infty\}$. The modular group $\mathrm{SL}_2(\mathbb{Z})$ acts on $\bar{\mathcal{H}}$ from the left as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}(\tau) = \frac{a\tau + b}{c\tau + d}.$$

As a set, the quotient space $\mathrm{SL}_2(\mathbb{Z})\backslash\bar{\mathcal{H}}$ can be identified in a natural way with the set

$$D = \{\tau \in \mathcal{H} : |\tau| \geq 1, -\frac{1}{2} \leq \mathrm{Re}(\tau) \leq 0\} \cup \{\tau \in \mathcal{H} : |\tau| > 1, 0 < \mathrm{Re}(\tau) < \frac{1}{2}\},$$

and we call D the standard fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$. Notice that in Part II, we will denote by D the CM discriminant by convention.

Under the group action above, for every congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$, the quotient space $\Gamma\backslash\bar{\mathcal{H}}$, supplied with the properly defined topology and analytic structure, gives a Riemann surface X_Γ. By Riemann existence theorem, X_Γ is a complex algebraic curve, known as *modular curve*. We call X_Γ a modular curve of level N if Γ is a congruence subgroup of level N. By convention, we denote $Y_\Gamma = \Gamma\backslash\mathcal{H}$, the finite part of X_Γ.

We defined the *cusps* of X_Γ as the Γ-equivalence classes of $\mathbb{Q}\cup\{i\infty\}$ and denote by $\nu_\infty(\Gamma)$ the number of cusps. A non-cuspidal point $P \in X_\Gamma$ is called *elliptic* if for some $\tau \in \mathcal{H}$ representing P the stabilizer $\Gamma_\tau \neq \{\pm 1\}$. It is well-known that the curve X_Γ has only finitely many elliptic points.

Since every finite subgroup of $\mathrm{SL}_2(\mathbb{Z})/\{\pm1\}$ is cyclic of order 2 or 3, we say an element of $\mathrm{SL}_2(\mathbb{Z})$ is *elliptic* if its image in $\mathrm{SL}_2(\mathbb{Z})/\{\pm1\}$ is of order 2 or 3. It is easy to see that X_Γ has elliptic points if and only if Γ has elliptic elements.

The modular curves corresponding to $\Gamma(N), \Gamma_1(N)$ and $\Gamma_0(N)$ are usually denoted by $X(N), X_1(N)$ and $X_0(N)$, respectively.

The classical j-invariant function is defined on $\mathcal{H}$ by the familiar relation

$$j(\tau) = q_\tau^{-1} + 744 + 196844 q_\tau + \cdots ,$$

where $q_\tau = e^{2\pi i\tau}$. Since j is $\mathrm{SL}_2(\mathbb{Z})$-automorphic, it defines a function on X_Γ for every congruence subgroup Γ. Moreover, it is meromorphic with poles exactly at the cusps.

In fact, everything above concerning modular curves is true for any finite index subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$. Note that there exist infinitely many finite index subgroups of $\mathrm{SL}_2(\mathbb{Z})$ which are not congruence subgroups.

For every positive integer N, the field of definition of $X(N)$ is $\mathbb{Q}(\zeta_N)$, where $\zeta_N = e^{2\pi i/N}$. Each function field $\mathbb{Q}(X(N)) = \mathbb{Q}(\zeta_N)(X(N))$ is a Galois extension of the function field $\mathbb{Q}(X(1)) = \mathbb{Q}(j)$. For the Galois group, we have

$$\mathrm{Gal}(\mathbb{Q}(X(N))/\mathbb{Q}(j)) \cong \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\pm 1,$$

which is defined up to an inner automorphism; once it is fixed, we have the following well-defined isomorphisms

$$\mathrm{Gal}(\mathbb{Q}(X(N))/\mathbb{Q}(\zeta_N, j)) \cong \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\pm 1, \quad \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \cong (\mathbb{Z}/N\mathbb{Z})^*.$$

Let G be a subgroup of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ containing -1. By Galois theory, G corresponds uniquely to an immediate field of the extension $\mathbb{Q}(X(N))/\mathbb{Q}(j)$. This gives an algebraic curve denoted by X_G. We denote by $\det G$ the image of G under the determinant map $\det : \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) \to (\mathbb{Z}/N\mathbb{Z})^*$. The curve X_G is defined over $\mathbb{Q}(\zeta_N)^{\det G}$. So in particular it is defined over $\mathbb{Q}$ if $\det G = (\mathbb{Z}/N\mathbb{Z})^*$.

If Γ is the pullback of $G \cap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ to $\mathrm{SL}_2(\mathbb{Z})$, then the set $X_G(\mathbb{C})$ of complex points is analytically isomorphic to the modular curve X_Γ. Consequently, we also call X_G a modular curve of level N. Its finite part is denoted by Y_G (that is, X_G deprived of the cusps). In this case, we use the common notation $\nu_\infty(G)$ for the number of cusps of X_G.

Here, we want to mention two special subgroups of $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$, p is a prime. The normalizer of a split Cartan subgroup is given by

$$\mathcal{C}_{\mathrm{s}}^+(p) = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}, \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} : \alpha, \beta \in \mathbb{F}_p^* \right\},$$

and the normalizer of a non-split Cartan subgroup is defined by

$$\mathcal{C}_{\mathrm{ns}}^+(p) = \left\{ \begin{pmatrix} \alpha & \Xi\beta \\ \beta & \alpha \end{pmatrix}, \begin{pmatrix} \alpha & \Xi\beta \\ -\beta & -\alpha \end{pmatrix} : \alpha, \beta \in \mathbb{F}_p, (\alpha, \beta) \neq (0,0) \right\},$$

where Ξ is a quadratic non-residue modulo p. In particular, one can choose $\Xi = -1$ if $p \equiv 3 \mod 4$. Moreover, $|\mathcal{C}_{\mathrm{ns}}^+(p)| = 2(p^2-1)$ by [17, Formula (2.3)] and $\det \mathcal{C}_{\mathrm{ns}}^+(p) = \mathbb{F}_p^*$.

The modular curves corresponding to $\mathcal{C}_{\mathrm{s}}^+(p)$ and $\mathcal{C}_{\mathrm{ns}}^+(p)$ are denoted by $X_{\mathrm{split}}^+(p)$ and $X_{\mathrm{ns}}^+(p)$, respectively. Both of them are defined over $\mathbb{Q}$ and of level p.

1.2 Siegel's Theorem

Let X be a smooth projective curve over a number field K of genus g and $f \in K(X)$ a non-constant rational function. Let S be a finite set of places of K, containing all Archimedean places. Denote by $\mathcal{O}_S$ the ring of S-integers of K.

We denote by $X(K)$ the set of K-rational points and by $X(\mathcal{O}_S, f)$ the set of S-*integral points* with respect to f:

$$X(\mathcal{O}_S, f) = \{P \in X(K) : f(P) \in \mathcal{O}_S\}.$$

Theorem 1.1 (Siegel [83]). *Assume that either $g \geq 1$ or f has at least three distinct poles. Then for any K and S as above, the set $X(\mathcal{O}_S, f)$ is finite.*

Furthermore, in 1983 Faltings [46] proved the Mordell's conjecture, which says that the set $X(K)$ is finte if $g \geq 2$.

However, the results of Siegel and Faltings are both ineffective in the sense that they imply no effective or explicit bounds for the size of S-integral or rational points. In spite of multiple efforts of many mathematicians, no effective approach to the study of rational point is known. On the other hand, there is a general method for effective analysis of integral points, developed by Alan Baker ([7–13]). Using Baker's method, we have known effective versions of Siegel's theorem for curves of genus 0 and 1 and for certain curves of higher genus.

Theorem 1.2. *Siegel's theorem is effective for (X, f) if*

1. (folklore) *g=0 and f has at least 3 poles, or*

2. (Baker and Coates [14]) *g=1, or*

3. (Bilu [32], Dvornicich and Zannier [43]) *$g \geq 1$ and $\bar{K}(X)/\bar{K}(f)$ is a Galois extension.*

Since 1995, Bilu and his collaborators have succeeded in getting effective Siegel's theorem for various classes of modular curves when choosing $f = j$. In 1995, Bilu [23] showed that Siegel's theorem is effective for modular curve X if X has at least three distinct cusps. In other words, the j-invariant of integral points of X can be effectively bounded. But there was no quantitative version therein.

Theorem 1.3 (Bilu [23]). *Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Then Siegel's theorem is effective for (X_Γ, j) if*

1. *Γ is a congruence subgroup and $\nu_\infty(\Gamma) \geq 3$, or*

2. *Γ has no elliptic elements.*

Theorem 1.3 is a fundamental criterion on effective Siegel's theorem for modular curves. For example, by Theorem 1.3, Siegel's theorem is effective for $(X(N), j)$ when $N \geq 2$, and for $(X_1(N), j)$ when $N \geq 4$. Afterwards, Bilu [24] gave the following refinement of Theorem 1.3.

Theorem 1.4 (Bilu [24]). *Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Assume that Γ has a congruence subgroup Γ' with $\nu_\infty(\Gamma') \geq 3$, and Γ' contains all elliptic elements of Γ. Then Siegel's theorem is effective for (X_Γ, j).*

Applying Theorem 1.4, Bilu [24] proved that Siegel's theorem is effective for $X_0(N)$ when $N \notin \{1, 2, 3, 5, 7, 13\}$. Furthermore, Bilu and Illengo [26] obtained effective Siegel's theorem for "almost every" modular curve. But they still gave no quantitative results.

Theorem 1.5 (Bilu and Illengo [26]). *Let Γ be a congruence subgroup of level not dividing the number $2^{20} \cdot 3^7 \cdot 5^3 \cdot 7^2 \cdot 11^{13}$. Then Siegel's theorem is effective for (X_Γ, j).*

By using Runge's method, the first explicit bound for the j-invariant of the S-integral points of X_G was given in [27, Theorem 1.2] when X_G satisfies "Runge condition" which roughly says that all the cusps are not conjugate. Especially, when $G = \mathcal{C}_{\mathrm{s}}^+(p)$, this bound can be sharply reduced, see [27, Theorem 6.1] and [28, Theorem 1.1].

Let G be a subgroup of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ containing -1 such that the corresponding modular curve X_G is defined over K. Let $\mathcal{C}(G, K)$ be the set of $\mathrm{Gal}(\bar{K}/K)$-orbits of the cusps. We denote by $\mathrm{h}(\cdot)$ the usual absolute logarithmic height. For $P \in X_G(\bar{\mathbb{Q}})$, we write $\mathrm{h}(P) = \mathrm{h}(j(P))$.

Theorem 1.6 (Bilu and Parent [27]). *Assume that $|\mathcal{C}(G, K)| > |S|$ (the "Runge condition"). Then for any $P \in Y_G(\mathcal{O}_S)$, we have*

$$\mathrm{h}(P) \leq 36 s^{s/2+1} (N^2|G|/2)^s \log(2N),$$

where $s = |S|$. If S only contains Archimedean places, we even have

$$\mathrm{h}(P) \leq 24 s^{s/2+1} (N^2|G|/2)^s \log(3N).$$

Theorem 1.7 (Bilu and Parent [28]). *There exists an absolute effective constant C such that for any prime number p and any $P \in Y_{\mathrm{split}}^+(p)(\mathbb{Z})$, we have*

$$\log |j(P)| \leq 2\pi p^{1/2} + 6 \log p + C.$$

The main task of Part I is to give effective or explicit bounds for the j-invariant of integral points on modular curves without Runge condition and by using Baker's method. More precisely, we will try to give quantitative versions for Theorems 1.3 and 1.4.

The problem of computing effective or explicit bounds for integral points on modular curves is of obvious importance, with the recent work of Bilu and Parent [28] serving as a prime example. In [28], the authors first obtained an effective upper bound for the j-invariant of integral points on the modular curve $X_{\mathrm{split}}^+(p)$. Then applying this bound, they showed that the $\mathbb{Q}$-rational points on $X_{\mathrm{split}}^+(p)$ are exactly the CM points and cusps for p greater than an absolute constant. Subsequently, they solved Serre's uniformity problem in the split Cartan case and finally left this problem with the non-split Cartan case. Moreover, Bilu, Parent and Rebolledo [29] showed that the $\mathbb{Q}$-rational points on $X^+_{\text{split}}(p^r)$ are exactly the CM points and cusps for all prime numbers $p \geq 11$, $p \neq 13$, and all integers $r \geq 1$.

Actually, the interest in integral points on the modular curves corresponding to normalizers of Cartan subgroups is motivated by their relation to imaginary quadratic fields of low class number. See Appendix A in Serre's book [78] for a nice historical account and further explanations. In particular, integral points on the curves $X^+_{\text{ns}}(24)$ and $X^+_{\text{ns}}(15)$ were studied by Heegner [55] and Siegel [84] in their classical work on the class number 1 problem. Kenku [57] determined all integral points on $X^+_{\text{ns}}(7)$, and Baran [16, 17] did this for $X^+_{\text{ns}}(9)$ and $X^+_{\text{ns}}(15)$. Most recently, a general method for computing integral points on $X^+_{\text{ns}}(p)$ has been developed by Bajolet and Bilu in [5].

In addition, the following Belyĭ's theorem tells us that effective Siegel's theorem on modular curves is crucial to obtain effective Siegel's theorem on general smooth projective curves.

Theorem 1.8 (Belyĭ [21]). *A smooth projective curve X is defined over $\bar{\mathbb{Q}}$ if and only if there exists a finite index subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ such that X is isomorphic to X_Γ.*

Here, we also would like to indicate that Surroca [86, 87] showed that the *abc* conjecture of Masser-Oesterlé implies an effective version of Siegel's theorem, and the converse is also true. In fact, this work was motivated by Elkies [44], who proved that the *abc* conjecture implies an effective version of Mordell's conjecture. It is interesting to think about whether the effective versions of Siegel's theorem on modular curves can induce some effective results towards the truth of the *abc* conjecture.

1.3 Structure of Part I

In Chapter 2, we will give some effective bounds for the j-invariant of integral points on arbitrary modular curves over arbitrary number fields assuming that the number of cusps is not less than 3. This will be based on the article [81].

In Chapter 3, for the special modular curve $X^+_{\text{ns}}(p)$, we will give explicit bounds for the j-invariant of integral points on $X^+_{\text{ns}}(p)$, which are much better than those given in Chapter 2. This is the joint work with Aurélien Bajolet [6]. Here, we want to indicate that Runge condition fails for $X^+_{\text{ns}}(p)$.

In Chapter 4, applying the results in Chapter 2, Quantitative Riemann existence theorem and Quantitative Chevalley-Weil theorem, we will give effective bounds for the j-invariant of integral points on certain modular curves which have positive genus and less than three cusps. For example, modular curves with no elliptic points. This will be based on the manuscript [79].

Chapter 2

Bounding the j-invariant of integral points on modular curves

2.1 Main results

Let G be a subgroup of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ containing -1 ($N \geq 2$), and let X_G be the corresponding modular curve. Let K_0 be a number field containing $\mathbb{Q}(\zeta_N)^{\det G}$. Then X_G is defined over K_0. Let S_0 be a finite set of absolute values of K_0, containing all the Archimedean (or infinite) places and normalized with respect to $\mathbb{Q}$. Recall that a K_0-rational point $P \in X_G(K_0)$ is an S_0-integral point if $j(P) \in \mathcal{O}_{S_0}$, where $\mathcal{O}_{S_0}$ is the ring of S_0-integers in K_0.

In this chapter, we apply Baker's method, based on Matveev [68] and Yu [94], to obtain some effective bounds for the j-invariant of S_0-integral points on X_G assuming that $\nu_\infty(G) \geq 3$.

Theorem 2.1. *Assume that $K_0 \subseteq \mathbb{Q}(\zeta_N)$, N is not a power of any prime, $\nu_\infty(G) \geq 3$, and S_0 only consists of Archimedean places. Then for any S_0-integral point P on X_G, we have*

$$\mathrm{h}(P) \leq C^{\varphi(N)} N^{\frac{3}{2}\varphi(N)+10} (\log N)^{\frac{5}{2}\varphi(N)-2},$$

where C is an absolute effective constant and $\varphi(N)$ is the Euler's totient function.

Actually, we obtain a more general Theorem 2.2, which applies to any number field and any ring of S_0-integers in it.

Put $d_0 = [K_0 : \mathbb{Q}]$ and $s_0 = |S_0|$. We define the following quantities

$$\Delta_0 = d_0^{-d_0} \sqrt{|D_0|} (\log |D_0|)^{d_0} \prod_{\substack{v \in S_0 \\ v \nmid \infty}} \log \mathcal{N}_{K_0/\mathbb{Q}}(v), \tag{2.1}$$

$$\Delta = d_0^{-d_0} \sqrt{N^{d_0 N} |D_0|^{\varphi(N)}} \left(\log(N^{d_0 N} |D_0|^{\varphi(N)}) \right)^{d_0 \varphi(N)} \left(\prod_{\substack{v \in S_0 \\ v \nmid \infty}} \log \mathcal{N}_{K_0/\mathbb{Q}}(v) \right)^{\varphi(N)}, \tag{2.2}$$

where D_0 is the absolute discriminant of K_0, and the norm of a non-Archimedean (or finite) place is, by definition, the absolute norm of the corresponding prime ideal. We denote by p the maximal rational prime below S_0, with the convention $p = 1$ if S_0 consists only of the Archimedean places.

Theorem 2.2. *Assume that N is not a power of any prime and $\nu_\infty(G) \geq 3$. Then for any S_0-integral point P on X_G, we have*

$$\mathrm{h}(P) \leq \left(C d_0 s_0 N^2 \right)^{2 s_0 N} (\log(d_0 N))^{3 s_0 N} p^{d_0 N} \Delta,$$

where C is an absolute effective constant.

In particular, if $\mathbb{Q}(\zeta_N) \subseteq K_0$, we have the following theorem.

Theorem 2.3. *Assume that $\mathbb{Q}(\zeta_N) \subseteq K_0$, N is not a power of any prime and $\nu_\infty(G) \geq 3$. Then for any S_0-integral point P on X_G, we have*

$$\mathrm{h}(P) \leq (C d_0 s_0)^{2 s_0} (\log d_0)^{3 s_0} N^8 p^{d_0} \Delta_0 \log p,$$

where C is an absolute effective constant.

The situation is different when N is a prime power, see Section 2.7. In this case we define

$$M = \begin{cases} 2N & \text{if } N \text{ is not a power of } 2, \\ 3N & \text{if } N \text{ is a power of } 2. \end{cases}$$

Notice that X_G is also a modular curve of level M.

Theorem 2.4. *Assume that N is a power of some prime and $\nu_\infty(G) \geq 3$. Then for any S_0-integral point P on X_G, we can obtain two upper bounds for $\mathrm{h}(P)$ by replacing N with M in Theorems 2.1, 2.2 and 2.3.*

2.2 Notation and conventions

Throughout this chapter, log stands for two different objects without confusion according to the context. One is the principal branch of the complex logarithm, in this case we will use the following estimate without special reference

$$|\log(1+z)| \le \frac{|\log(1-r)|}{r}|z| \quad \text{for } |z| \le r < 1,$$

see [27, Formula (4)]. The other one is the p-adic logarithm function, for example see [58, Chapter IV Section 2].

For $\mathbf{a} = (a_1, a_2) \in \mathbb{Q}^2$, we put $\ell_{\mathbf{a}} = B_2(a_1 - \lfloor a_1 \rfloor)/2$, where $B_2(T) = T^2 - T + \frac{1}{6}$ is the second Bernoulli polynomial and $\lfloor a_1 \rfloor$ is the largest integer not greater than a_1. Obviously $|\ell_{\mathbf{a}}| \le 1/12$, this will be used without special reference.

Let $\mathcal{A}_N$ be the subset of the abelian group $(N^{-1}\mathbb{Z}/\mathbb{Z})^2$ consisting of the elements with exact order N. Obviously,

$$|\mathcal{A}_N| = N^2 \prod_{p|N}(1-p^{-2}) < N^2,$$

the product runs through all primes dividing N. Moreover, we always choose a representative element of $\mathbf{a} = (a_1, a_2) \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$ satisfying $0 \le a_1, a_2 < 1$. So in the sequel, for every $\mathbf{a} \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$, we have $\ell_{\mathbf{a}} = B_2(a_1)/2$.

Throughout this chapter, we fix an algebraic closure $\bar{\mathbb{Q}}$ of $\mathbb{Q}$, which is assumed to be a subfield of $\mathbb{C}$. In particular, for every $a \in \mathbb{Q}$ we have the well-defined root of unity $e^{2\pi i a} \in \bar{\mathbb{Q}}$. Every number field used in this chapter is presumed to be a subfield of $\bar{\mathbb{Q}}$. If K is such a number field and v is a valuation on K, then we tacitly assume than v is somehow extended to $\bar{\mathbb{Q}} = \bar{K}$; equivalently, we fix an algebraic closure $\bar{K}_v$ and an embedding $\bar{\mathbb{Q}} \hookrightarrow \bar{K}_v$. In particular, the roots of unity $e^{2\pi i a}$ are well-defined elements of $\bar{K}_v$.

For a number field K, we denote by M_K the set of all valuations (or places) of K extending the standard infinite and p-adic valuations of $\mathbb{Q}$: $|2|_v = 2$ if $v \in M_K$ is infinite, and $|p|_v = p^{-1}$ if v extends the p-adic valuation of $\mathbb{Q}$. We denote by M_K^∞ and M_K^0 the subsets of M_K consisting of the infinite (or Archimedean) and the finite (or non-Archimedean) valuations, respectively.

Given a number field K of degree d, for any $v \in M_K$, K_v is the completion of K with respect to the valuation v and $\bar{K}_v$ its algebraic closure. We still denote by v the unique extension of v in $\bar{K}_v$. Let $d_v = [K_v : \mathbb{Q}_v]$ be the local degree of v.

For a number field K of degree d, the absolute logarithmic height of an algebraic number $\alpha \in K$ is defined by $\mathrm{h}(\alpha) = d^{-1}\sum_{v\in M_K} d_v \log^+|\alpha|_v$, where $\log^+|\alpha|_v = \log\max\{|\alpha|_v, 1\}$.

Throughout the chapter, the symbol $\ll$ implies an absolute effective constant. We also use the notation $O_v(\cdot)$. Precisely, $A = O_v(B)$ means that $|A|_v \le B$.

2.3 Preparations

In this section, we assume that $N \ge 2$.

2.3.1 Siegel functions

Let $\mathbf{a} = (a_1, a_2) \in \mathbb{Q}^2$ be such that $\mathbf{a} \notin \mathbb{Z}^2$, and let $g_{\mathbf{a}} : \mathcal{H} \to \mathbb{C}$ be the corresponding *Siegel function*, see [62, Section 2.1]. We have the following infinite product presentation for $g_{\mathbf{a}}$, see [27, Formula (7)],

$$g_{\mathbf{a}}(q_\tau) = -q_\tau^{B_2(a_1)/2} e^{\pi i a_2(a_1-1)} \prod_{n=0}^{\infty}(1 - q_\tau^{n+a_1}e^{2\pi i a_2})(1 - q_\tau^{n+1-a_1}e^{-2\pi i a_2}).$$

For the elementary properties of $g_{\mathbf{a}}$, see [62, Pages 27-31]. Especially, the order of vanishing of $g_{\mathbf{a}}$ at $i\infty$ (i.e., the only rational number ℓ such that the limit $\lim_{\tau\to i\infty} q_\tau^{-\ell} g_{\mathbf{a}}$ exists and is nonzero) is equal to $\ell_{\mathbf{a}}$.

For a number field K and $v \in M_K$, we define $g_{\mathbf{a}}(q)$ as the above, where $q \in \bar{K}_v$ satisfies $|q|_v < 1$. Notice that here we should fix $q^{1/(12N^2)} \in \bar{K}_v$, then everything is well defined.

Given two positive integers k and ℓ, we denote by P_k the set of partitions of k into positive summands, and let $p_\ell(k)$ be the number of partitions of k into exactly ℓ positive summands. By [3, Theorem 14.5], we easily get

$$|P_k| < e^{k/2}, \quad k \ge 64.$$

Then according to the table of partitions or computer calculations, we can obtain

$$|P_k| < e^{k/2}, \quad k \ge 1.$$

Proposition 2.5. *Let* $\mathbf{a} \in \mathcal{A}_N$*. If* $q \in \bar{K}_v$ *satisfies* $|q|_v < 1$*, then we have*

$$-q^{-\ell_{\mathbf{a}}}\gamma_{\mathbf{a}}^{-1}g_{\mathbf{a}}(q) = 1 + \sum_{k=1}^{\infty}\phi_{\mathbf{a}}(k)q^{k/N},$$

where

$$\gamma_{\mathbf{a}} = \begin{cases} e^{\pi i a_2(a_1-1)} & \text{if } a_1 \neq 0, \\ e^{-\pi i a_2}(1-e^{2\pi i a_2}) & \text{if } a_1 = 0; \end{cases}$$

and

$$\phi_{\mathbf{a}}(k) = \sum_{\ell \in S^1_{\mathbf{a}k}} m_\ell(-e^{2\pi i a_2})^\ell + \sum_{\ell \in S^2_{\mathbf{a}k}} m'_\ell(-e^{-2\pi i a_2})^\ell + \sum_{\ell \in S^3_{\mathbf{a}k}} \sum_{(\ell_1,\ell_2)\in T^\ell_{\mathbf{a}k}} m_{\ell_1\ell_2}(-e^{2\pi i a_2})^{\ell_1}(-e^{-2\pi i a_2})^{\ell_2},$$

where $S^1_{\mathbf{a}k}$*,* $S^2_{\mathbf{a}k}$ *and* $S^3_{\mathbf{a}k}$ *are three subsets of* $\{1,2,\cdots,\lfloor k/N\rfloor+1\}$*,* $T^\ell_{\mathbf{a}k}$ *is a subset of* $\{(\ell_1,\ell_2): 1 \leq \ell_1,\ell_2 \leq \lfloor k/N\rfloor + 1, \ell_1+\ell_2 = \ell\}$*, and* m_ℓ, m'_ℓ*, and* $m_{\ell_1\ell_2}$ *are some positive integers. In particular, we have*

$$|\phi_{\mathbf{a}}(k)|_v \leq e^k.$$

Proof. In this proof, we fix an integer $k \geq 1$.

Suppose that $a_1 = k_1/N$ with $0 \leq k_1 \leq N-1$. Let $S_1 = \{nN + k_1 : 0 \leq n \leq \lfloor k/N\rfloor, 0 < nN + k_1 \leq k\}$ and $S_2 = \{nN + N - k_1 : 0 \leq n \leq \lfloor k/N\rfloor, nN + N - k_1 \leq k\}$. It is easy to see that if $k_1 = 0$ or $N/2$, then $S_1 = S_2$; otherwise $S_1 \cap S_2 = \emptyset$.

Notice that the coefficient $\phi_{\mathbf{a}}(k)$ of $q^{k/N}$ equals to the coefficient of q^k in the expansion of the following finite product,

$$\prod_{n\in S_1}(1-q^n e^{2\pi i a_2})\prod_{n\in S_2}(1-q^n e^{-2\pi i a_2}). \tag{2.3}$$

If S_1 and S_2 are both empty, then the coefficient $\phi_{\mathbf{a}}(k) = 0$.

We say $\ell \in S^1_{\mathbf{a}k}$ if and only if there exist ℓ positive integers in S_1 such that the sum of them equals to k, and let m_ℓ count the number of different ways of such summations. Similarly for the definitions of $S^2_{\mathbf{a}k}$ and m'_ℓ.

We say $\ell \in S^3_{\mathbf{a}k}$ if and only if there exist ℓ_1 positive integers in S_1 and ℓ_2 positive integers in S_2 such that the sum of them equals to k, then $(\ell_1,\ell_2) \in T^\ell_{\mathbf{a}k}$ and let $m_{\ell_1\ell_2}$ count the number of different ways of such summations.

Then the desired expression of $\phi_{\mathbf{a}}(k)$ follows easily from the definitions.

For each element $x \in P_k$, let m_x be the number of the times of x appearing in the expansion of (2.3). Then we obtain

$$\begin{aligned}|\phi_{\mathbf{a}}(k)|_v &\leq \sum_{\ell \in S^1_{\mathbf{a}k}} m_\ell + \sum_{\ell \in S^2_{\mathbf{a}k}} m'_\ell + \sum_{\ell \in S^3_{\mathbf{a}k}} \sum_{(\ell_1,\ell_2)\in T^\ell_{\mathbf{a}k}} m_{\ell_1\ell_2} \\ &= \sum_{x \in P_k} m_x.\end{aligned}$$

If $k_1 \neq 0$ and $N/2$, then $S_1 \cap S_2 = \emptyset$. So for each $x \in P_k$, we have $m_x = 0$ or 1. Hence,

$$\sum_{x\in P_k} m_x \leq |P_k| < e^{k/2}.$$

If $k_1 = 0$ or $N/2$, then $S_1 = S_2$. Suppose that $\lfloor k/N \rfloor \geq 3$. Given $x \in P_k$ with ℓ entries, if $\ell \leq \lfloor k/N \rfloor$, then we have $m_x \leq 2^\ell$; otherwise we have $m_x = 0$. Hence,

$$\sum_{x\in P_k} m_x \leq \sum_{\ell \leq \lfloor k/N \rfloor} 2^\ell p_\ell(k) \leq 2^{\lfloor k/N \rfloor}|P_k| < e^k.$$

If $\lfloor k/N \rfloor \leq 2$, one can verify the inequality by explicit computations. $\square$

2.3.2 Modular units on $X(N)$

Recall that by a modular unit on a modular curve we mean that a rational function having poles and zeros only at the cusps.

For $\mathbf{a} \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$, we denote $g^{12N}_{\mathbf{a}}$ by $u_{\mathbf{a}}$, which is a modular unit on $X(N)$. Moreover, we have $u_{\mathbf{a}} = u_{\mathbf{a}'}$ when $\mathbf{a} \equiv \mathbf{a}' \bmod \mathbb{Z}^2$. Hence, $u_{\mathbf{a}}$ is well-defined when $\mathbf{a}$ is a nonzero element of the abelian group $(N^{-1}\mathbb{Z}/\mathbb{Z})^2$. Moreover, $u_{\mathbf{a}}$ is integral over $\mathbb{Z}[j]$. For more details, see [27, Section 4.2].

Furthermore, the Galois action on the set $\{u_{\mathbf{a}}\}$ is compatible with the right linear action of $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ on it. That is, for any $\sigma \in \mathrm{Gal}(\mathbb{Q}(X(N))/\mathbb{Q}(j)) = \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\pm 1$ and any $\mathbf{a} \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$, we have

$$u^\sigma_{\mathbf{a}} = u_{\mathbf{a}\sigma}.$$

Here, we borrow a result and its proof from [5] for subsequent applications and for the conveniences of readers.

Proposition 2.6 ([5]). *We have*

$$\prod_{\boldsymbol{a}\in\mathcal{A}_N} u_{\boldsymbol{a}} = \pm\Phi_N(1)^{12N} = \begin{cases} \pm\ell^{12N} & \text{if } N \text{ is a power of a prime } \ell, \\ \pm 1 & \text{if } N \text{ has at least two distinct prime factors,} \end{cases}$$

where Φ_N is the N-th cyclotomic polynomial.

Proof. We denote by u the left-hand side of the equality. Since the set $\mathcal{A}_N$ is stable with respect to $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, u is stable with respect to the Galois action over the field $\mathbb{Q}(X(1)) = \mathbb{Q}(j)$. So $u \in \mathbb{Q}(j)$. Moreover, since u is integral over $\mathbb{Z}[j]$, $u \in \mathbb{Z}[j]$. Notice that $X(1)$ has only one cusp and u has no zeros and poles outside the cusps, so we must have that u is a constant and $u \in \mathbb{Z}$.

Furthermore, we have

$$\begin{aligned} u &= \prod_{(a_1,a_2)\in\mathcal{A}_N} q^{6NB_2(a_1)} e^{12N\pi i a_2(a_1-1)} \prod_{n=0}^{\infty} (1-q^{n+a_1}e^{2\pi i a_2})^{12N}(1-q^{n+1-a_1}e^{-2\pi i a_2})^{12N} \\ &\overset{q=0}{=\!=} \pm \prod_{\substack{(a_1,a_2)\in\mathcal{A}_N \\ a_1=0}} (1-e^{2\pi i a_2})^{12N} \\ &= \pm \prod_{\substack{1\le k<N \\ \gcd(k,N)=1}} (1-e^{2k\pi i/N})^{12N} \\ &= \pm\Phi_N(1)^{12N}. \end{aligned}$$

□

2.3.3 X_G and X_{G_1}

Let $G_1 = G \cap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ and X_{G_1} be the modular curve corresponding to G_1. In this subsection, we assume that X_{G_1} is defined over a number field K. Then X_G is also defined over K. Since X_G and X_{G_1} have the same geometrically integral model, every K-rational point of X_G is also a K-rational point of X_{G_1}.

For each cusp c of X_{G_1}, let t_c be its local parameter constructed in [27, Section 3]. Put $q_c = t_c^{e_c}$, where e_c is the ramification index of the natural covering $X_{G_1} \to X(1)$ at c. Notice that $e_c|N$. Furthermore, the familiar expansion $j = q_c^{-1} + 744 + 196884q_c + \cdots$ holds in a v-adic neighborhood of c, the right-hand side converging v-adically, where $v \in M_K$ such that $c \in X_{G_1}(\bar{K}_v)$.

For any $v \in M_K$, let $\Omega_{c,v}$ be the set constructed in [27, Section 3] on which t_c and q_c are defined and analytic. Recall that D is the standard fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$.

Actually, when v is Archimedean, define

$$\widetilde{D} = D \cup \{i\infty\} \setminus \{\text{the arc connecting } i \text{ and } e^{2\pi i/3}\},$$

then $\Omega_{c,v} = \Gamma\backslash\sigma(\widetilde{D}+\mathbb{Z})$, where Γ is the pullback of G_1 to $\mathrm{SL}_2(\mathbb{Z})$, and $\sigma \in \mathrm{SL}_2(\mathbb{Z})$ is chosen such that $\sigma(i\infty)$ represents the cusp c. If v is non-Archimedean, then $\Omega_{c,v} = \{P \in X_{G_1}(\bar{K}_v) : |q_c(P)|_v < 1\}$.

Here, we quote [27, Proposition 3.1] as follows.

Proposition 2.7 ([27]). *Put*

$$X_{G_1}(\bar{K}_v)^+ = \begin{cases} \{P \in X_{G_1}(\bar{K}_v) : |j(P)|_v > 3500\} & \text{if } v \in M_K^\infty, \\ \{P \in X_{G_1}(\bar{K}_v) : |j(P)|_v > 1\} & \text{if } v \in M_K^0. \end{cases}$$

Then

$$X_{G_1}(\bar{K}_v)^+ \subseteq \bigcup_c \Omega_{c,v}$$

with equality for the non-Archimedean v, where the union runs through all the cusps of X_{G_1}. Moreover, for $P \in \Omega_{c,v}$ we have

$$\frac{1}{2}|j(P)|_v \le |q_c(P)^{-1}|_v \le \frac{3}{2}|j(P)|_v \tag{2.4}$$

if v is Archimedean, and $|j(P)|_v = |q_c(P)^{-1}|_v$ if v is non-Archimedean.

We will use the above proposition several times without special reference. Moreover, this proposition implies that for every $P \in X_{G_1}(\bar{K}_v)^+$ there exists a cusp c such that $P \in \Omega_{c,v}$. We call c a v-nearby cusp of P.

We directly obtain the following corollary from Proposition 2.5.

Corollary 2.8. *Let c be a cusp of X_{G_1}, $v \in M_K$ and $P \in \Omega_{c,v}$. Assume that $|q_c(P)|_v \le 10^{-N}$. For $\mathbf{a} \in \mathcal{A}_N$, we have*

$$-q_c^{-\ell_{\mathbf{a}}}\gamma_{\mathbf{a}}^{-1}g_{\mathbf{a}}(q_c(P)) = 1 + O_v(4|q_c(P)|_v^{1/N}).$$

The following proposition follows directly from [27, Propositions 2.3 and 2.5].

Proposition 2.9. *Let c be a cusp of X_{G_1}, $v \in M_K$ and $P \in \Omega_{c,v}$. For every $\mathbf{a} \in \mathcal{A}_N$, we have*

$$\left|\log|g_{\mathbf{a}}(q_c(P))|_v - \ell_{\mathbf{a}}\log|q_c(P)|_v\right| \begin{cases} \le \log N & \text{if } v \in M_K^\infty, \\ = 0 & \text{if } |N|_v = 1, \\ \le \frac{\log \ell}{\ell-1} & \text{if } v|\ell|N, \end{cases}$$

where ℓ is some prime factor of N.

2.3.4 Modular units on X_{G_1}

We apply the notation in the above subsection.

We denote by $\mathcal{M}_N$ the set of elements of exact order N in $(\mathbb{Z}/N\mathbb{Z})^2$. Let us consider the natural right group action of G_1 on $\mathcal{M}_N$. Following the proof of [26, Lemma 2.3], we see that the number of the orbits of $\mathcal{M}_N/G_1$ is equal to $\nu_\infty(G)$, this also explain why we transfer our problems on X_G to those on X_{G_1}.

Obviously, when we consider the natural right group action $\mathcal{A}_N/G_1$, there are also $\nu_\infty(G)$ orbits of this group action. So

$$\nu_\infty(G) \leq |\mathcal{A}_N| < N^2.$$

Let T be any subset of $\mathcal{A}_N$, we define

$$u_T = \prod_{\mathbf{a}\in T} u_{\mathbf{a}}.$$

Let $\mathcal{O}$ be an orbit of the right group action $\mathcal{A}_N/G_1$, we have

$$u_{\mathcal{O}} = \prod_{\mathbf{a}\in \mathcal{O}} u_{\mathbf{a}}. \tag{2.5}$$

By [27, Proposition 4.2 (ii)], $u_{\mathcal{O}}$ is a rational function on the modular curve X_{G_1}. In fact, $u_{\mathcal{O}}$ is a modular unit on X_{G_1}.

For any cusp c, we denote by $\mathrm{Ord}_c(u_{\mathcal{O}})$ the vanishing order of $u_{\mathcal{O}}$ at c. For $v \in M_K$, define

$$\rho_v = \begin{cases} 12N^3 \log N & \text{if } v \in M_K^\infty, \\ 0 & \text{if } v \in M_K^0 \text{ and } |N|_v = 1, \\ \frac{12N^3\log \ell}{\ell-1} & \text{if } v \in M_K^0 \text{ and } v|\ell|N, \end{cases}$$

where ℓ is some prime factor of N.

Then $u_{\mathcal{O}}$ has the following properties:

Proposition 2.10. (i) *Put $\lambda = (1-\zeta_N)^{12N^3}$. Then the functions $u_{\mathcal{O}}$ and $\lambda u_{\mathcal{O}}^{-1}$ are integral over $\mathbb{Z}[j]$.*

(ii) *For the cusp c_∞ at infinity, we have*

$$\mathrm{Ord}_{c_\infty}(u_{\mathcal{O}}) = 12Ne_{c_\infty}\sum_{\mathbf{a}\in\mathcal{O}}\ell_{\mathbf{a}}.$$

For any cusp c, we have $|\mathrm{Ord}_c(u_{\mathcal{O}})| < N^4$.

(iii) *Let c be a cusp of X_{G_1}, $v \in M_K$, and $P \in \Omega_{c,v}$. Assume that $|q_c(P)|_v \le 10^{-N}$. Then we have*

$$q_c(P)^{-\mathrm{Ord}_c(u_{\mathcal{O}})/e_c}\gamma_{\mathcal{O},c}^{-1}u_{\mathcal{O}}(P) = 1 + O_v(4^{12N^3}|q_c(P)|_v^{1/N}),$$

where $\gamma_{\mathcal{O},c} \in \mathbb{Q}(\zeta_N)$ and $\mathrm{h}(\gamma_{\mathcal{O},c}) \le 12N^3\log 2$.

(iv) *Let c be a cusp of X_{G_1} and $v \in M_K$. For $P \in \Omega_{c,v}$, we have*

$$\left|\log|u_{\mathcal{O}}(P)|_v - \frac{\mathrm{Ord}_c(u_{\mathcal{O}})}{e_c}\log|q_c(P)|_v\right| \le \rho_v.$$

(v) *For $v \in M_K^\infty$ and $P \in X_{G_1}(K_v)$, we have*

$$|\log|u_{\mathcal{O}}(P)|_v| \le N^3\log(|j(P)|_v + 2400) + \rho_v.$$

(vi) *The group generated by the principal divisor $(u_{\mathcal{O}})$, where $\mathcal{O}$ runs over the orbits of $\mathcal{A}_N/G_1$, is of rank $\nu_\infty(G) - 1$.*

Proof. (i) See [27, Proposition 4.2 (i)].

(ii) Similar to the proof of [27, Proposition 4.2 (iii)]. The q-order of vanishing of $u_{\mathcal{O}}$ at $i\infty$ is $12N\sum_{\mathbf{a}\in\mathcal{O}}\ell_{\mathbf{a}}$. Then

$$\mathrm{Ord}_{c_\infty}(u_{\mathcal{O}}) = 12Ne_{c_\infty}\sum_{\mathbf{a}\in\mathcal{O}}\ell_{\mathbf{a}}.$$

Since $|\ell_{\mathbf{a}}| \le \frac{1}{12}$, we have $|\mathrm{Ord}_{c_\infty}(u_{\mathcal{O}})| \le Ne_{c_\infty}|\mathcal{O}| < N^4$. The case of arbitrary c reduces to the case $c = c_\infty$ by replacing $\mathcal{O}$ by $\mathcal{O}\sigma$ where $\sigma \in \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ is such that $\sigma(c) = c_\infty$.

(iii) Similar to the proof of [27, Proposition 4.4] by using Corollary 2.8 except for the height of $\gamma_{\mathcal{O}}$. In fact, if $c = c_\infty$, we have $\gamma_{\mathcal{O},c} = \prod_{\mathbf{a}\in\mathcal{O}}\gamma_{\mathbf{a}}^{12N}$. Then $\mathrm{h}(\gamma_{\mathcal{O},c}) \le 12N\sum_{\mathbf{a}\in\mathcal{O}}\mathrm{h}(\gamma_{\mathbf{a}}) \le 12N|\mathcal{O}|\log 2 < 12N^3\log 2$. The general case reduces to the case $c = c_\infty$ by applying a suitable Galois automorphism.

(iv) and (v) They follow from [27, Proposition 4.4].

(vi) By Proposition 2.6, the rank of the free abelian group $(u_{\mathcal{O}})$ is at most $\nu_\infty(G)-1$. Then Manin-Drinfeld theorem, as stated in [62], tells us that this rank is maximal possible. ☐

2.4 Siegel's theory of convenient units

We recall here Siegel's construction [85] of convenient units in a number field K of degree d, in the form adapted to the needs of the present work. The results of this section are well-known, but not always in the set-up we wish them to have.

Let S be a finite set of absolute values of K, containing all the Archimedean valuations and normalized with respect to $\mathbb{Q}$. Fix a valuation $v_0 \in S$, we put

$$S' = S \setminus \{v_0\}, \quad s = |S| \geq 2, \quad r = s-1, \quad d' = \max\{d, 3\}, \quad \zeta = 1201\left(\frac{\log d'}{\log\log d'}\right)^3.$$

Let $\xi_1, \cdots, \xi_r$ be a fundamental system of S-units. The S-regulator $R(S)$ is the absolute value of the determinant of the $r \times r$ matrix

$$(d_v \log|\xi_k|_v)_{\substack{v\in S' \\ 1\leq k\leq r}} \tag{2.6}$$

(we fix some ordering for the set S'), where $d_v = [K_v : \mathbb{Q}_v]$ is the local degree of v. It is well-defined and is equal to the usual regulator R_K when S is the set of infinite places.

Proposition 2.11. *There exists a fundamental system of S-units $\eta_1, \cdots, \eta_r$ satisfying*

$$\begin{aligned}
&\mathrm{h}(\eta_1)\cdots\mathrm{h}(\eta_r) \leq d^{-r} r^{2r} R(S),\\
&(\zeta d)^{-1} \leq \mathrm{h}(\eta_k) \leq d^{-1} r^{2r} \zeta^{r-1} R(S) \quad (k = 1, \cdots, r).
\end{aligned}$$

Furthermore, the entries of the inverse matrix of (2.6) are bounded in absolute value by $r^{2r}\zeta$.

Proof. See [37, Lemma 1]. Notice that the left-hand inequality in the second inequality is a well-known result of Dobrowolski [41]. ☐

Corollary 2.12. *For the unit $\eta = \eta_1^{b_1}\cdots\eta_r^{b_r}$, where $\eta_1, \cdots, \eta_r$ are from Proposition 2.11 and $b_1, \cdots, b_r \in \mathbb{Z}$, put $B^* = \max\{|b_1|, \cdots, |b_r|\}$, then we have*

$$\begin{aligned}
&\mathrm{h}(\eta) \leq d^{-1} r^{2r+1} \zeta^{r-1} B^* R(S),\\
&B^* \leq 2dr^{2r}\zeta \mathrm{h}(\eta).
\end{aligned}$$

Proof. The first inequality follows from Proposition 2.11 and standard height estimates.

Write

$$d_v \log|\eta|_v = \sum_{k=1}^{r} d_v b_k \log|\eta_k|_v, \quad v \in S'.$$

Resolving this in terms of $b_1, \cdots, b_r$ and using the final statement of Proposition 2.11, we obtain

$$B^* \leq r^{2r}\zeta \sum_{v\in S'} d_v|\log|\eta|_v| \leq r^{2r}\zeta \sum_{v\in S} d_v|\log|\eta|_v|.$$

Since η is an S-unit,

$$\sum_{v\in S} d_v|\log|\eta|_v| = d(\mathrm{h}(\eta) + \mathrm{h}(\eta^{-1})) = 2d\mathrm{h}(\eta).$$

Then the corollary is proved. □

Finally, we quote two estimates of the S-regulator in terms of the usual regulator R_K, the class number h_K, the degree d, and the discriminant D_K of the field K.

Proposition 2.13. *We have*

$$0.1 \leq R(S) \leq h_K R_K \prod_{\substack{v\in S\\ v\nmid\infty}} \log\mathcal{N}(v),$$

$$R(S) \ll d^{-d}\sqrt{|D_K|}(\log|D_K|)^{d-1} \prod_{\substack{v\in S\\ v\nmid\infty}} \log\mathcal{N}(v).$$

For the first inequality see [37, Lemma 3]; one may remark that the lower bound $R(S) \geq 0.1$ follows from Friedman's famous lower bound [50] for the usual regulator $R_K \geq 0.2$. The second one follows from Siegel's estimate [85, Satz 1]

$$h_K R_K \ll d^{-d}\sqrt{|D_K|}(\log|D_K|)^{d-1};$$

in fact there is an explicit bound for $h_K R_K$ therein.

2.5 Baker's inequality

In this section we state Baker's inequality, which is the main technical tool of the proofs. It is actually an adaptation of a result in [1]. For the convenience of readers, we also quote its proof with slight change.

For a number field K and $v \in M_K$, we denote by p_v the underlying prime of v when v is non-Archimedean. Next, we let

- $\theta_0, \theta_1, \cdots, \theta_r$ be nonzero algebraic numbers, belonging to K;
- $\Theta_0, \Theta_1, \cdots, \Theta_r$ be real numbers satisfying

$$\Theta_k \geq \max\{d\mathrm{h}(\theta_k), 1\} \quad (k = 0, 1, \cdots, r);$$

- $b_1, \ldots, b_r$ be rational integers, $\Lambda = \theta_0\theta_1^{b_1}\cdots\theta_r^{b_r}$, $B^* = \max\{|b_1|, |b_2|, \cdots, |b_r|\}$.

Theorem 2.14 ([1])**.** *There exists an absolute constant C that can be determined explicitly such that the following holds. Assume that $\Lambda \neq 1$. Then for any real number B satisfying $B \geq B^*$ and $B \geq \max\{3, \Theta_1, \cdots, \Theta_r\}$, we have*

$$|\Lambda - 1|_v \geq e^{-\Upsilon\Theta_0\Theta_1\cdots\Theta_r \log B},$$

where

$$\Upsilon = \begin{cases} C^r d^2 \log(2d), & v \mid \infty, \\ (Cd)^{2r+6} p_v^d, & v | p_v < \infty. \end{cases}$$

Proof. The Archimedean case is due to Matveev, see Corollary 2.3 from [68]. We use this result with $n = r+1$, with $1, b_1, \ldots, b_r$ as Matveev's $b_n, b_1, \ldots, b_{n-1}$, respectively, $\Theta_0, \Theta_1, \ldots, \Theta_r$ as Matveev's $A_n, A_1, \ldots, A_{n-1}$, respectively, and B as Matveev's B.

Notice that Matveev assumes (in our notation) that

$$\Theta_k \geq |\log \theta_k|, \tag{2.7}$$

with some choice of the complex value of the logarithm. However, if we pick the principal value of the logarithm, then

$$|\log \theta_k| \leq |\log|\theta_k|| + \pi \leq d\mathrm{h}(\theta_k) + \pi \leq (1+\pi)\Theta_k.$$

Hence we may disregard (2.7) at the cost of increasing the absolute constant C in the definition of Υ.

In the case of non-Archimedean v we employ the result of Yu [94]. Precisely, we use the second consequence of his "Main Theorem" on page 190 (see the bottom of page 190 and the top of page 191), which asserts that, assuming (1.19) of [94], but without assuming (1.5) and (1.15), the first displayed equation on the top of page 191 of [94] holds.

In our notation, taking, as in the Archimedean case, $n = r+1$, using $1, b_1, \ldots, b_r$ as Yu's $b_n, b_1, \ldots, b_{n-1}$, noticing that Yu's parameters $h_n, h_1, \ldots, h_{n-1}$ do not exceed our $d^{-1}\Theta_0, d^{-1}\Theta_1, \ldots, d^{-1}\Theta_r$, and setting Yu's B_n to be 1, we re-state Yu's result as follows. Let $\mathfrak{p}$ be the prime ideal corresponding to v and δ a real number satisfying $0 < \delta \le 1/2$; then

$$\begin{aligned}&\mathrm{Ord}_{\mathfrak{p}}(\Lambda - 1) < (Cd)^{2r+5}\frac{p_v^d}{(\log p_v)^2}\max\left\{\Theta_0\Theta_1\cdots\Theta_r\log Q, \delta B\right\},\\ &Q = \delta^{-1}e^{6r^2}d^{2r}p_v^{rd}\Theta_1\cdots\Theta_r.\end{aligned}$$

Here, we replace Yu's c_0 by d^{r+1}, Yu's c_1 by $e^{6r^2}d^{3r}$, and Yu's C_0 by $(Cd)^{3r+6}p_v^d(\log p_v)^{-2}$, the constant C being absolute. Observing that

$$\log Q = \log\left(\delta^{-1}\Theta_1\cdots\Theta_r\right) + O(r^2 d\log p_v),$$

and modifying the absolute constant C, we obtain

$$\mathrm{Ord}_{\mathfrak{p}}(\Lambda - 1) < (Cd)^{2r+6}\frac{p_v^d}{\log p_v}\max\left\{\Theta_0\Theta_1\cdots\Theta_r\log\left(\delta^{-1}\Theta_1\cdots\Theta_r\right), \delta B\right\}. \tag{2.8}$$

Notice that $B \ge 3$, then $\log B > 1$. Set now

$$\delta = \min\left\{\Theta_1\cdots\Theta_r\frac{\log B}{B}, \frac{1}{2}\right\}.$$

If $\delta < 1/2$ then the maximum in (2.8) does not exceed $\Theta_0\Theta_1\cdots\Theta_r\log B$. And if $\delta = 1/2$, then

$$\frac{B}{\log B} \le 2\Theta_1\cdots\Theta_r,$$

which, by [25, Lemma 2.3.3], implies that

$$B \le 4\Theta_1\cdots\Theta_r\log\left(2\Theta_1\cdots\Theta_r\right) \le 4(r+1)\Theta_1\cdots\Theta_r\log B,$$

and the maximum in (2.8) is at most $2(r+1)\Theta_0\Theta_1\cdots\Theta_r\log B$. So in any case we obtain (again slightly adjusting the absolute constant C) the estimate

$$\mathrm{Ord}_{\mathfrak{p}}(\Lambda - 1) < (Cd)^{2r+6}\frac{p_v^d}{\log p_v}\Theta_0\Theta_1\cdots\Theta_r\log B. \tag{2.9}$$

Finally, since $|\Lambda - 1|_v = e^{-\frac{\log p_v}{e_{\mathfrak{p}}}\mathrm{Ord}_{\mathfrak{p}}(\Lambda-1)}$, where $e_{\mathfrak{p}}$ is the absolute ramification index of $\mathfrak{p}$, we obtain the result in the non-Archimedean case as well. □

Remark 2.15. We choose the form of Baker's inequality in Theorem 2.14 because of its convenience for our computations, although it is effective but not explicit. If one wants to get an explicit bound for $\mathrm{h}(P)$, one can apply Matveev [68] and Yu [94] respectively, like [73], and one can also apply [22, Theorem C] to handle uniformly with the Archimedean and non-Archimedean cases.

2.6 The case of mixed level

In this section, we assume that N has at least two distinct prime factors. Then we will apply Baker's inequality to prove Theorems 2.1 and 2.2.

In the sequel, we assume that P is an S_0-integral point of X_G and $\nu_\infty(G) \geq 3$. What we want to do is to obtain some bounds for $\mathrm{h}(P)$.

From now on we let $K = K_0 \cdot \mathbb{Q}(\zeta_N) = K_0(\zeta_N)$. Let S be the set consisting of the extensions of the places from S_0 to K, that is,

$$S = \{v \in M_K : v|v_0 \in S_0\}.$$

Then P is also an S-integral point of X_G.

Put $d = [K : \mathbb{Q}]$, $s = |S|$ and $r = s - 1$. Since $j(P) \in \mathcal{O}_S$, we have

$$\mathrm{h}(P) = d^{-1} \sum_{v \in S} d_v \log^+ |j(P)|_v \leq \sum_{v \in S} \log^+ |j(P)|_v.$$

Then there exists some $w \in S$ such that

$$\mathrm{h}(P) \leq s \log |j(P)|_w.$$

We fix this valuation w from now on. Therefore, we only need to bound $\log |j(P)|_w$.

As the discussion in Section 2.3.3, P is also an S-integral point of X_{G_1}. Hence for our purposes, we only need to focus on the modular curve X_{G_1}.

We partition the set S into three pairwise disjoint subsets: $S = S_1 \cup S_2 \cup S_3$, where S_1 consists of places $v \in S$ such that $P \in X_{G_1}(\bar{K}_v)^+$, $S_2 = M_K^\infty \setminus S_1$, and $S_3 = S \setminus (S_1 \cup S_2)$.

From now on, for $v \in S_1$ let c_v be a v-nearby cusp of P, and we write q_v for q_{c_v} and e_v for e_{c_v}. Notice that for any $v \in S_3$, it is non-Archimedean with $|j(P)|_v \leq 1$.

In the sequel, we can assume that $|j(P)|_w > 3500$, otherwise we can get a better bound than those given in Section 2.1. Then we have $w \in S_1$ and $P \in \Omega_{c_w,w}$ for some cusp c_w. Therefore, by (2.4) we only need to bound $\log |q_w(P)^{-1}|_w$.

From now on we assume that $|q_w(P)|_w \le 10^{-N}$. Indeed, applying (2.4) the inequality $|q_w(P)|_w > 10^{-N}$ yields $\mathrm{h}(P) < 3sN$, which is a much better estimate for $\mathrm{h}(P)$ than those given in Section 2.1.

Notice that under our assumptions, we see that $N \ge 2$. Moreover, in this section we assume that $s \ge 2$. In fact, if $s = 1$, then we can add another valuation to S such that $s = 2$, and then the final results of this section also hold.

2.6.1 Preparation for Baker's inequality

We fix an orbit $\mathcal{O}$ of the group action $\mathcal{A}_N/G_1$ as follows. Put $U = u_{\mathcal{O}}$, where $u_{\mathcal{O}}$ is defined in (2.5).

If $\mathrm{Ord}_{c_w}U \ne 0$, we choose $\mathcal{O}$ such that $\mathrm{Ord}_{c_w}U < 0$ according to Proposition 2.6. Noticing $v_\infty(G) \ge 3$ and combining with Proposition 2.10 (vi), we can choose another orbit $\mathcal{O}'$ such that U and V are multiplicatively independent modulo constants with $\mathrm{Ord}_{c_w}V > 0$, where $V = u_{\mathcal{O}'}$.

Define the following function

$$W = \begin{cases} U & \text{if } \mathrm{Ord}_{c_w}U = 0, \\ \\ U^{\mathrm{Ord}_{c_w}V}V^{-\mathrm{Ord}_{c_w}U} & \text{if } \mathrm{Ord}_{c_w}U \ne 0. \end{cases}$$

So we always have $\mathrm{Ord}_{c_w}W = 0$ and $W(P) \in \mathcal{O}_S$. In particular, W is integral over $\mathbb{Z}[j]$. Moreover, W is not a constant by Proposition 2.10 (vi).

By Proposition 2.10 (ii) and (iii), we have

$$\gamma_w^{-1}W(P) = 1 + O_w(4^{24N^7}|q_w(P)|_w^{1/N}), \tag{2.10}$$

where

$$\gamma_w = \begin{cases} \gamma_{\mathcal{O},c_w} & \text{if } \mathrm{Ord}_{c_w}U = 0, \\ \\ \gamma_{\mathcal{O},c_w}^{\mathrm{Ord}_{c_w}V}\gamma_{\mathcal{O}',c_w}^{-\mathrm{Ord}_{c_w}U} & \text{if } \mathrm{Ord}_{c_w}U \ne 0; \end{cases}$$

and

$$\mathrm{h}(\gamma_w) \le 24N^7 \log 2.$$

By Proposition 2.6, we know that $W(P)$ is a unit of $\mathcal{O}_S$. So there exist some integers $b_1, \cdots, b_r \in \mathbb{Z}$ such that $W(P) = \omega\eta_1^{b_1}\cdots\eta_r^{b_r}$, where ω is a root of unity and $\eta_1, \cdots, \eta_r$

are from Proposition 2.11. Let $\eta_0 = \omega\gamma_w^{-1}$. Then we set

$$\Lambda = \gamma_w^{-1} W(P) = \eta_0 \eta_1^{b_1} \cdots \eta_r^{b_r}. \tag{2.11}$$

Notice that $\eta_0, \cdots, \eta_r \in K$ and

$$|\Lambda - 1|_w \leq 4^{24N^7} |q_w(P)|_w^{1/N}. \tag{2.12}$$

For subsequent deductions, we need to bound $\mathrm{h}(W(P))$.

Proposition 2.16. *We have*

$$\mathrm{h}(W(P)) \leq 2sN^8 \log |q_w^{-1}(P)|_w + 94sN^8 \log N.$$

Proof. First suppose that $\mathrm{Ord}_{c_w} U = 0$. Then $W = U$. For $v \in S_3$, $j(P)$ is a v-adic integer. Hence, so is the number $W(P)$. In addition, it is easy to see that

$$\sum_{v \in M_K^\infty} d_v \rho_v = 12dN^3 \log N, \qquad \sum_{v \in M_K^0} d_v \rho_v \leq 12dN^3 \log N.$$

Notice that for $v \in S_1$, $|\mathrm{Ord}_{c_v}(W)| \leq N^4$. Applying Proposition 2.10 (iv) and (2.4), we have

$$\begin{aligned} d^{-1} \sum_{v \in S_1} d_v \log^+ |W(P)|_v &\leq N^4 d^{-1} \sum_{v \in S_1} d_v \log |q_v(P)^{-1}|_v + d^{-1} \sum_{v \in S_1} d_v \rho_v \\ &\leq N^4 d^{-1} \sum_{v \in S_1} d_v \log |j(P)|_v + sN^4 \log \frac{3}{2} + 24N^3 \log N \\ &\leq N^4 \mathrm{h}(P) + sN^4 \log \frac{3}{2} + 24N^3 \log N \\ &\leq sN^4 \log |j(P)|_w + sN^4 \log \frac{3}{2} + 24N^3 \log N \\ &\leq sN^4 \log |q_w(P)^{-1}|_w + sN^4 \log 3 + 24N^3 \log N. \end{aligned}$$

It follows from Proposition 2.10 (v) that

$$d^{-1} \sum_{v \in S_2} d_v \log^+ |W(P)|_v \leq N^3 \log 5900 + 12N^3 \log N.$$

Hence, we get

$$\begin{aligned}\mathrm{h}(W(P)) &= d^{-1}\sum_{v\in S_1\cup S_2} d_v \log^+|W(P)|_v \\ &\le sN^4\log|q_w(P)^{-1}|_w + sN^4\log 3 + 36N^3\log N + N^3\log 5900.\end{aligned}$$

Now suppose that $\mathrm{Ord}_{c_w}U \neq 0$. For any $v\in S_1$, we have

$$|\log|W(P)|_v| \le \frac{|\mathrm{Ord}_{c_v}(W)|}{e_v}\log|q_v(P)^{-1}|_v + 2N^4\rho_v.$$

Here note that $|\mathrm{Ord}_{c_v}(W)| \le 2N^8$. For any $v\in M_K^\infty$, we have

$$|\log|W(P)|_v| \le 2N^7\log(|j(P)|_v + 2400) + 2N^4\rho_v.$$

Apply the same argument as the above, we obtain

$$\mathrm{h}(W(P)) \le 2sN^8\log|q_w(P)^{-1}|_w + 2sN^8\log 3 + 72N^7\log N + 2N^7\log 5900.$$

Now it is easy to get the desired result. □

2.6.2 Using Baker's inequality

If $\Lambda = 1$, we can get better bounds for $\mathrm{h}(P)$ than those given in Section 2.1, see Section 2.8. So in the rest of this section we assume that $\Lambda \neq 1$.

Let $B^* = \max\{|b_1|,\cdots,|b_r|\}$, and let $\Theta_0,\Theta_1,\cdots,\Theta_r$ be real numbers satisfying

$$\Theta_k \ge \max\{d\mathrm{h}(\eta_k),1\},\quad k=0,\cdots,r.$$

By Theorem 2.14, there exists an absolute constant C which can be determined explicitly such that the following holds. Choosing $B\ge B^*$ and $B\ge\max\{3,\Theta_1,\cdots,\Theta_r\}$, we have

$$|\Lambda-1|_w \ge e^{-\Upsilon\Theta_0\Theta_1\cdots\Theta_r\log B}, \tag{2.13}$$

where

$$\Upsilon = \begin{cases} C^r d^2\log(2d), & w\mid\infty,\\ (Cd)^{2r+6}p^d, & \text{otherwise.}\end{cases}$$

Recall that p is the maximal rational prime below S_0, with the convention $p=1$ if S_0 consists only of the Archimedean places.

Applying (2.12), we have

$$e^{-\Upsilon\Theta_0\Theta_1\cdots\Theta_r\log B} \leq 4^{24N^7}|q_w(P)|_w^{1/N}.$$

Hence, we obtain

$$\log|q_w(P)^{-1}|_w \leq N\Upsilon\Theta_0\Theta_1\cdots\Theta_r\log B + 48N^8\log 2. \tag{2.14}$$

According to Proposition 2.11, we can choose

$$\Theta_k = d\zeta\mathrm{h}(\eta_k), \quad k = 1, \cdots, r.$$

So we have

$$\Theta_1\cdots\Theta_r \leq r^{2r}\zeta^r R(S).$$

Since

$$d\mathrm{h}(\eta_0) = d\mathrm{h}(\gamma_w) \leq 24dN^7\log 2,$$

we can choose

$$\Theta_0 = 24dN^7\log 2.$$

Corollary 2.12 tells us that

$$B^* \leq 2dr^{2r}\zeta\mathrm{h}(W(P)).$$

Notice that we also need $B \geq \max\{3, \Theta_1, \cdots, \Theta_r\}$, by Proposition 2.11 and Proposition 2.16 we can choose

$$B = r^{2r}\zeta^r R(S) + 2dr^{2r}\zeta\left(2sN^8\log|q_w(P)^{-1}|_w + 94sN^8\log N\right).$$

Again, we write $B = \alpha\log|q_w(P)^{-1}|_w + \beta$, where

$$\begin{aligned}
\alpha &= 4dsr^{2r}\zeta N^8,\\
\beta &= r^{2r}\zeta^r R(S) + 188dsr^{2r}\zeta N^8\log N.
\end{aligned}$$

Hence, (2.14) yields

$$\alpha\log|q_w(P)^{-1}|_w + \beta \leq \alpha N\Upsilon\Theta_0\Theta_1\cdots\Theta_r\log(\alpha\log|q_w(P)^{-1}|_w + \beta) + 48\alpha N^8\log 2 + \beta.$$

Here we put $C_1 = \alpha N \Upsilon \Theta_0 \Theta_1 \cdots \Theta_r$ and $C_2 = 48\alpha N^8 \log 2 + \beta$, then

$$\alpha \log |q_w(P)^{-1}|_w + \beta \leq C_1 \log(\alpha \log |q_w(P)^{-1}|_w + \beta) + C_2.$$

Therefore, by [25, Lemma 2.3.3] we obtain

$$\alpha \log |q_w(P)^{-1}|_w + \beta \leq 2(C_1 \log C_1 + C_2).$$

Hence

$$\log |q_w(P)^{-1}|_w \leq 2\alpha^{-1} C_1 \log C_1 + \alpha^{-1}(2C_2 - \beta).$$

That is

$$\log |j(P)|_w \leq 2\alpha^{-1} C_1 \log C_1 + \alpha^{-1}(2C_2 - \beta) + \log 2.$$

So we have

$$\mathrm{h}(P) \leq 2s\alpha^{-1} C_1 \log C_1 + s\alpha^{-1}(2C_2 - \beta) + s \log 2.$$

Finally we get

$$\mathrm{h}(P) \ll dsr^{2r}\zeta^r N^8 \Upsilon R(S) \log(d^2 sr^{4r}\zeta^{r+1} N^{16} \Upsilon R(S)). \tag{2.15}$$

To get a bound for $\mathrm{h}(P)$, we only need to calculate the quantities in the above inequality.

2.6.3 Proof of Theorem 2.1

Under the assumptions of Theorem 2.1, we have $K = \mathbb{Q}(\zeta_N)$ and $S = M_K^\infty$. Since we have assumed that $s \geq 2$, we have $\varphi(N) \geq 4$.

Then $|D| \leq N^{\varphi(N)}$ according to [93, Proposition 2.7]. It follows from Proposition 2.13 that

$$R(S) \ll \varphi(N)^{-1} N^{\varphi(N)/2} (\log N)^{\varphi(N)-1}.$$

Notice that

$$\begin{aligned}
&s = \varphi(N)/2,\\
&\zeta \ll (\log \varphi(N))^3,\\
&\Upsilon = C^{\frac{\varphi(N)}{2}-1}\varphi(N)^2\log(2\varphi(N)),\\
&\log(d^2 s r^{4r}\zeta^{r+1}N^{16}\Upsilon R(S)) \ll \varphi(N)\log N.
\end{aligned}$$

Applying (2.15) we obtain

$$\begin{aligned}
\mathrm{h}(P) &\leq C^{\varphi(N)}(\varphi(N))^{\varphi(N)+2}(\log\varphi(N))^{\frac{3}{2}\varphi(N)-2}N^{\frac{1}{2}\varphi(N)+8}(\log N)^{\varphi(N)},\\
&\leq C^{\varphi(N)}N^{\frac{3}{2}\varphi(N)+10}(\log N)^{\frac{5}{2}\varphi(N)-2},
\end{aligned}$$

the constant C being modified. Hence we prove Theorem 2.1.

2.6.4 Proof of Theorem 2.2

Now we need to give a bound for $\mathrm{h}(P)$ based on the parameters of K_0 with the assumptions of Theorem 2.2.

First, notice that

$$\begin{aligned}
&s \leq s_0\varphi(N),\\
&r = s-1 \leq s_0\varphi(N)-1,\\
&d \leq d_0\varphi(N),\\
&\zeta \ll (\log d)^3 \leq (\log(d_0\varphi(N)))^3.
\end{aligned}$$

Using Proposition 2.13, we estimate $R(S)$ as follows:

$$R(S) \ll d^{-d}\sqrt{|D_K|}(\log|D_K|)^{d-1}\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v).$$

Since $\mathcal{N}_{K/\mathbb{Q}}(v) \leq p^{[K:\mathbb{Q}]} = p^d$, this implies the upper bound

$$\log R(S) \ll \frac{1}{2}\log|D_K| + d\log\log|D_K| + s\log(dp). \tag{2.16}$$

Let D_{K/K_0} be the relative discriminant of K/K_0. Recall that D_0 is the absolute discriminant of K_0. We have

$$D_K = \mathcal{N}_{K_0/\mathbb{Q}}(D_{K/K_0}) D_0^{[K:K_0]}.$$

We denote by $\mathcal{O}_{K_0}$ and $\mathcal{O}_K$ the ring of integers of K_0 and K, respectively. Since $K = K_0(\zeta_N)$, we have

$$\mathcal{O}_{K_0} \subseteq \mathcal{O}_{K_0}[\zeta_N] \subseteq \mathcal{O}_K.$$

By [51, III (2.20) (b)] and note that the absolute value of the discriminant of the polynomial $x^N - 1$ is N^N, we get

$$D_{K/K_0} | N^N.$$

So

$$|\mathcal{N}_{K_0/\mathbb{Q}}(D_{K/K_0})| \le N^{d_0 N}.$$

Hence

$$|D_K| \le N^{d_0 N} |D_0|^{\varphi(N)}.$$

Now let v_0 be a non-Archimedean place of K_0, and $v_1, \cdots, v_m$ all its extensions to K, their residue degrees over K_0 being $f_1, \cdots, f_m$, respectively. Then $f_1 + \cdots + f_m \le [K : K_0] \le \varphi(N)$, which implies that $f_1 \cdots f_m \le 2^{\varphi(N)}$. Notice that we always have $2 \log \mathcal{N}_{K_0/\mathbb{Q}}(v_0) > 1$. Since $\mathcal{N}_{K/\mathbb{Q}}(v_k) = \mathcal{N}_{K_0/\mathbb{Q}}(v_0)^{f_k}$ for $1 \le k \le m$ and $m \le \varphi(N)$, we have

$$\begin{aligned}\prod_{k=1}^{m} \log \mathcal{N}_{K/\mathbb{Q}}(v_k) &\le 2^{\varphi(N)} (\log \mathcal{N}_{K_0/\mathbb{Q}}(v_0))^m \\ &\le 2^{\varphi(N)} (2 \log \mathcal{N}_{K_0/\mathbb{Q}}(v_0))^m \\ &\le 4^{\varphi(N)} (\log \mathcal{N}_{K_0/\mathbb{Q}}(v_0))^{\varphi(N)}.\end{aligned}$$

Hence

$$\prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v) \le 4^{s_0 \varphi(N)} \left(\prod_{\substack{v \in S_0 \\ v \nmid \infty}} \log \mathcal{N}_{K_0/\mathbb{Q}}(v) \right)^{\varphi(N)}. \tag{2.17}$$

If we now denote by Δ the quantity defined in (2.2), then using (2.16) and (2.17), we obtain the following estimates:

$$\begin{aligned}
&R(S) \ll 4^{s_0\varphi(N)}\Delta,\\
&R(S)\log R(S) \ll 4^{s_0\varphi(N)} s_0\Delta\log p,\\
&R(S)\log(d^2 sr^{4r}\zeta^{r+1}N^{16}\Upsilon R(S)) \ll 4^{s_0\varphi(N)} s_0\Delta\log(ps_0).
\end{aligned}$$

Here we always choose $\Upsilon = (Cd)^{2r+6}p^d$.

Finally, using (2.15) and noticing that $d_0 \le 2s_0$, we get

$$\begin{aligned}
\mathrm{h}(P) &\le \left(Cd_0s_0\varphi(N)^2\right)^{2s_0\varphi(N)}\left(\log(d_0\varphi(N))\right)^{3s_0\varphi(N)} N^8 p^{d_0\varphi(N)}\Delta\log p\\
&\le \left(Cd_0s_0N^2\right)^{2s_0N}\left(\log(d_0N)\right)^{3s_0N} p^{d_0N}\Delta.
\end{aligned}$$

the constant C being modified.

Therefore, Theorem 2.2 is proved.

2.6.5 Proof of Theorem 2.3

Under the assumptions of Theorem 2.3, we have $K = K_0, d = d_0, s = s_0$, and $r = s_0 - 1$.

Similar to the proof of Theorem 2.2, we get

$$\begin{aligned}
&R(S) \ll \Delta_0,\\
&R(S)\log R(S) \ll s_0\Delta_0\log p,\\
&R(S)\log(d^2 sr^{4r}\zeta^{r+1}N^{16}\Upsilon R(S)) \ll s_0\Delta_0\log(ps_0).
\end{aligned}$$

Then using (2.15) and noticing that $d_0 \le 2s_0$, we obtain

$$\mathrm{h}(P) \le (Cd_0s_0)^{2s_0}(\log d_0)^{3s_0}N^8p^{d_0}\Delta_0\log p,$$

where C is an absolute effective constant.

Therefore, Theorem 2.3 is proved.

2.7 The case of prime power level

In this section, we assume that N is a prime power.

As Section 2.6, we can define a similar function W. But in this case $W(P)$ is not a unit of $\mathcal{O}_S$ by Proposition 2.6. So we need to raise the level. Put

$$M = \begin{cases} 2N & \text{if } N \text{ is not a power of } 2, \\ \\ 3N & \text{if } N \text{ is a power of } 2. \end{cases}$$

Notice that X_G is also a modular curve of level M and $\nu_\infty(G) \geq 3$, since we have the following natural sequence of morphisms

$$X(M) \to X(N) \to X_G \to X(1).$$

Since $\mathrm{Gal}(\mathbb{Q}(X(M))/\mathbb{Q}(j)) = \mathrm{GL}_2(\mathbb{Z}/M\mathbb{Z})/\pm 1$, X_G corresponds to a subgroup $\widetilde{G}$ of $\mathrm{GL}_2(\mathbb{Z}/M\mathbb{Z})$ containing ± 1. In fact, The restriction of $\widetilde{G}$ on $X(N)$ is G. The modular curve $X_{\widetilde{G}}$ has the same integral geometric model as X_G. In particular, P is also an S_0-integral point of $X_{\widetilde{G}}$.

Therefore, from Theorems 2.1, 2.2 and 2.3, we can get two upper bounds for $\mathrm{h}(P)$ by replacing N by M, which proves Theorem 2.4.

2.8 The case $\Lambda = 1$

In this section, we suppose that N is not a prime power without loss of generality. Under the assumption $\Lambda = 1$ we can obtain better bounds for $\mathrm{h}(P)$ than those given in Section 2.1.

Let c be a cusp of X_{G_1} and $v \in M_K$. We also denote by v the unique extension of v to $\bar{K}_v$. Recall $\Omega_{c,v}$ and the q-parameter q_c mentioned in Section 2.3.3, for the modular function U defined in Section 2.6.1, we get the following lemma.

Lemma 2.17. *There exist an integer-valued function $f(\cdot)$ with respect to q_c and $\lambda_1^c, \lambda_2^c, \lambda_3^c \cdots \in \mathbb{Q}(\zeta_N)$ such that the following identity holds in v-adic sense,*

$$\log \frac{U(q_c)}{\gamma_{\mathcal{O},c} q_c^{\frac{\mathrm{Ord}_c U}{e_c}}} = 2\pi f(q_c) i + \sum_{k=1}^{\infty} \lambda_k^c q_c^{k/N}, \tag{2.18}$$

and

$$|\lambda_k^c|_v \leq \begin{cases} |k|_v^{-1} & \text{if } v \text{ is finite,} \\ 24N^2(k+N) & \text{if } v \text{ is infinite.} \end{cases}$$

In particular, for every $k \geq 1$, we have

$$\mathrm{h}(\lambda_k^c) \leq \log(24N^3 + 24kN^2) + \log k.$$

Proof. By definition, we have

$$\frac{U(q_c)}{\gamma_{\mathcal{O},c} q_c^{\frac{\mathrm{Ord}_c U}{e_c}}} = \prod_{\mathbf{a}\in\mathcal{O}} \prod_{\substack{n=0 \\ n+a_1\neq 0}}^{\infty} (1 - q_c^{n+a_1} e^{2\pi i a_2})^{12N} \prod_{n=0}^{\infty} (1 - q_c^{n+1-a_1} e^{-2\pi i a_2})^{12N}. \tag{2.19}$$

Since

$$\sum_{\mathbf{a}\in\mathcal{O}} \left(\sum_{\substack{n=0 \\ n+a_1\neq 0}}^{\infty} 12N|q_c|_v^{n+a_1} + \sum_{n=0}^{\infty} 12N|q_c|_v^{n+1-a_1} \right)$$

is convergent, it follows from [2, Chapter 5 Section 2.2 Theorem 6] that the right-hand side of (2.19) is absolutely convergent (v is infinite). It is also true when v is finite. Then we can write (2.19) as the form $\prod_{n=1}^{\infty}(1+d_n)$ such that $\prod_{n=1}^{\infty}(1+d_n)$ is absolutely convergent. Hence, [2, Chapter 5 Section 2.2 Theorem 5] (v is infinite) and [58, Chapter IV Section 2] (v is finite) give

$$\log \frac{U(q_c)}{\gamma_{\mathcal{O},c} q_c^{\frac{\mathrm{Ord}_c U}{e_c}}}$$

$$= 2\pi f(q_c) i + \sum_{\mathbf{a}\in\mathcal{O}} \left(\sum_{\substack{n=0 \\ n+a_1\neq 0}}^{\infty} 12N \log(1 - q_c^{n+a_1} e^{2\pi i a_2}) + \sum_{n=0}^{\infty} 12N \log(1 - q_c^{n+1-a_1} e^{-2\pi i a_2}) \right),$$

where by default $f(q_c)$ is always equal to 0 if v is finite. Applying the Taylor expansion of the logarithm function to the right-hand side of the above formula, we get the desired formula for $\log \frac{U(q_c)}{\gamma_{\mathcal{O},c} q_c^{\frac{\mathrm{Ord}_c U}{e_c}}}$.

For a fixed non-negative integer n (where we assume $n > 0$ if $a_1 = 0$), write

$$\log(1 - q_c^{n+a_1} e^{2\pi i a_2}) = \sum_{k=1}^{\infty} \alpha_k q^{k/N}.$$

An immediate verification shows that

$$|\alpha_k|_v \leq \begin{cases} |k|_v^{-1} & \text{if } v \text{ is finite,} \\ 1 & \text{if } v \text{ is infinite.} \end{cases}$$

Same estimates hold true for the coefficients of the q-series for $\log(1 - q_c^{n+1-a_1} e^{-2\pi i a_2})$.

For each $\mathbf{a} \in \mathcal{O}$, the number of coefficients in the q-series for $\log(1 - q_c^{n+a_1} e^{2\pi i a_2})$ which may contribute to λ_k^c (those with $0 \leq n \leq k/N$) is at most $k/N + 1$, and the same is true for the q-series for $\log(1 - q_c^{n+1-a_1} e^{-2\pi i a_2})$. The bound for $|\lambda_k^c|_v$ now follows by summation. $\square$

Corollary 2.18. *Assume that* $\mathrm{Ord}_c U = 0$. *Then* $\lambda_k^c \neq 0$ *for some* $k \leq N^6$.

Proof. Since U is not a constant, there must exist some $\lambda_k^c \neq 0$. Under the assumption $\mathrm{Ord}_c U = 0$, we have $U(c) = \gamma_{\mathcal{O},c}$, and then $f(q_c(c)) = 0$ by (2.18). We extend the additive valuation Ord_c from the field $K(X_{G_1})$ to the field of formal power series $K((q_c^{1/e_c}))$. Then $\mathrm{Ord}_c q_c^{1/e_c} = 1$ and $\mathrm{Ord}_c\,(-2\pi f(q_c) i + \log(U/\gamma_{\mathcal{O},c})) \leq \mathrm{Ord}_c \log(U/\gamma_{\mathcal{O},c}) = \mathrm{Ord}_c(U/\gamma_{\mathcal{O},c} - 1)$. The latter quantity is bounded by the degree of $U/\gamma_{\mathcal{O},c} - 1$, which is equal to the degree of U.

The degree of U is equal to $\frac{1}{2} \sum_{c_0} |\mathrm{Ord}_{c_0} U|$, here the sum runs through all the cusps of X_{G_1}. Then the result follows from Proposition 2.10 (ii). $\square$

Now we can prove a general result.

Proposition 2.19. *Assume that* $\mathrm{Ord}_c U = 0$. *Then for* $P \in \Omega_{c,v}$ *such that* $U(P) = \gamma_{\mathcal{O},c}$, *we have*

$$\log |q_c(P)^{-1}|_v \leq N\varphi(N) \log(24N^{14} + 24N^9) + N \log(48N^2(N^6 + N + 1)).$$

Proof. Let n be the smallest k such that $\lambda_k^c \neq 0$. Then $n \leq N^6$. We assume that $|q_c(P)|_v \leq 10^{-N}$, otherwise there is nothing to prove. Since $\mathrm{Ord}_c U = 0$ and $U(P) = \gamma_{\mathcal{O},c}$, it follows from Lemma 2.17 that $2\pi f(q_c(P)) i + \sum_{k=n}^{\infty} \lambda_k^c q_c(P)^{k/N} = 0$.

Suppose that $f(q_c(P)) = 0$. Then $|\lambda_n^c q_c(P)^{n/N}|_v = |\sum_{k=n+1}^{\infty} \lambda_k^c q_c(P)^{k/N}|_v$. On the one hand, we have

$$\begin{aligned}
|\sum_{k=n+1}^{\infty} \lambda_k^c q_c(P)^{k/N}|_v &\leq \sum_{k=n+1}^{\infty} |\lambda_k^c|_v |q_c(P)|_v^{k/N} \\
&\leq \sum_{k=n+1}^{\infty} 24N^2(k+N)|q_c(P)|_v^{k/N} \\
&= 48N^2(n+N+1)|q_c(P)|_v^{(n+1)/N}.
\end{aligned}$$

On the other hand, using Liouville's inequality (see [92, Formula (3.13)]), we get

$$|\lambda_n^c|_v \geq e^{-[\mathbb{Q}(\zeta_N):\mathbb{Q}]\mathrm{h}(\lambda_n^c)} \geq (24nN^3 + 24n^2N^2)^{-\varphi(N)}.$$

Then the desired result follows easily.

Suppose that $f(q_c(P)) \neq 0$. Then $2\pi \leq |\sum_{k=n}^{\infty} \lambda_k^c q_c(P)^{k/N}|_v \leq 48N^2(n+N)|q_c(P)|_v^{n/N}$. Then we get

$$\log|q_c(P)^{-1}|_v \leq N\log(48N^2(N^6+N)).$$

$\square$

Now we assume that $\mathrm{Ord}_{c_w}U = 0$. Then we have $W = U$. Since $\Lambda = 1$, $W(P) = \gamma_{\mathcal{O},c_w}$. For the S-integral point P of X_{G_1} fixed in Section 2.6, applying the above proposition to W, we obtain

$$\begin{aligned}
\mathrm{h}(P) &\leq s(\log|q_w(P)^{-1}|_w + \log 2)\\
&\leq s_0 N\left(N\varphi(N)\log(24N^{14}+24N^9) + N\log(48N^2(N^6+N+1)) + \log 2\right).
\end{aligned}$$

Now we assume that $\mathrm{Ord}_{c_w}U \neq 0$. Then $W = U^{\mathrm{Ord}_{c_w}V}V^{-\mathrm{Ord}_{c_w}U}$ with $\mathrm{Ord}_{c_w}W = 0$. Proposition 2.10 (vi) guarantees that W is not a constant. Applying the same method as the above without difficulties, we can also get a better bound than Theorems 2.1 and 2.2. We omit the details here.

In conclusion, if assuming $\Lambda = 1$, we can get polynomial bounds for $\mathrm{h}(P)$ in terms of s_0 and N, which are obviously better than those in Theorems 2.1-2.4.

Chapter 3

Bounding the j-invariant of integral points on $X_{\mathrm{ns}}^{+}(p)$

3.1 Background

In 1972, Serre [77] proved that for any elliptic curve E over $\mathbb{Q}$ without complex multiplication, there exists a constant $C(E) > 0$ with respect to E such that for every prime $p > C(E)$, the natural Galois representation

$$\rho_{E,p} : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(E[p]) \cong \mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$$

is surjective, where $E[p]$ is the p-torsion subgroup of E and $\mathrm{GL}(E[p])$ is its automorphism group.

Serre asked whether there exist an absolute constant C such that for any elliptic curve E without complex multiplication over $\mathbb{Q}$ and any prime $p > C$, $\rho_{E,p}$ is surjective, which now is called "Serre's uniformity problem".

As is well-known, the group $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ has the following types of maximal proper subgroups: Borel subgroups, exceptional subgroups, and normalizers of (split and non-split) Cartan subgroups. To solve Serre's uniformity problem, one has to show that for sufficiently large p, the image of $\rho_{E,p}$ is not contained in any of the above listed maximal subgroups. The cases of exceptional subgroups and Borel subgroups have been solved by Serre and Mazur, respectively. For the case of normalizers of Cartan subgroups, it is equivalent to prove that for sufficiently large p, the only $\mathbb{Q}$-rational points of the modular curves $X_{\mathrm{split}}^{+}(p)$ and $X_{\mathrm{ns}}^{+}(p)$ are the cusps and CM points.

Bilu and Parent [28] first obtained an effective upper bound for the j-invariant of integral points on the modular curve $X_{\mathrm{split}}^{+}(p)$. Then applying this bound, they showed that the $\mathbb{Q}$-rational points on $X_{\mathrm{split}}^{+}(p)$ are exactly the cusps and CM points for p greater than an absolute constant. Subsequently, they solved Serre's uniformity problem in the split Cartan case and finally left this problem with the non-split Cartan case.

In this chapter, we will obtain some effective upper bounds for the j-invariant of integral points on $X_{\mathrm{ns}}^{+}(p)$.

3.2 Main Results

Throughout this chapter we fix a prime number $p \geq 7$. We call a rational point $P \in X_{\mathrm{ns}}^{+}(p)(\mathbb{Q})$ an integral point with respect to j if $j(P) \in \mathbb{Z}$.

The modular curve $X_{\mathrm{ns}}^{+}(p)$ has $\frac{p-1}{2}$ cusps, and all its cusps are conjugate over $\mathbb{Q}$. Hence, by Siegel's theorem, the curve $X_{\mathrm{ns}}^{+}(p)$ has only finitely many integral points. Moreover, as follows from [23, Proposition 5.1(a)], their size can be bounded effectively in terms of p.

In this chapter we use Baker's method, more precisely Baker's inequality in the form due to Matveev [68, Corollary 2.3], to obtain two explicit bounds in terms of p for the j-invariant of integral points on $X_{\mathrm{ns}}^{+}(p)$.

Theorem 3.1. *Assume that $p \geq 7$ and let $d \geq 3$ be a divisor of $(p-1)/2$. Then for any integral point P on $X_{\mathrm{ns}}^{+}(p)$ we have*

$$\log |j(P)| < C(d) p^{6d+5} (\log p)^2,$$

where $C(d) = 30^{d+5} \cdot d^{-2d+4.5}$.

In particular, if we choose $d = \frac{p-1}{2}$ in Theorem 3.1, we obtain a bound which is explicit in p.

Theorem 3.2. *Assume that $p \geq 7$. Then for any integral point P on $X_{\mathrm{ns}}^{+}(p)$ we have*

$$\log |j(P)| < 41993 \cdot 13^{p} \cdot p^{2p+7.5} (\log p)^2.$$

By comparing these two theorems, the bound in Theorem 3.3 can be drastically reduced if $\frac{p-1}{2}$ has a small divisor. For example, if $p \equiv 1 \pmod 3$, we have the following theorem.

Theorem 3.3. *Assume that $p \geq 7$ and $p \equiv 1 \pmod 3$. Then for any integral point P on $X_{\mathrm{ns}}^{+}(p)$ we have*

$$\log |j(P)| < 30^8 \cdot p^{23} (\log p)^2.$$

3.3 Notation and conventions

Throughout this chapter, log stands for the principal branch of the complex logarithm, and let $G = \mathcal{C}_{\mathrm{ns}}^{+}(p)$.

In the sequel, we fix a subgroup H of $\mathbb{F}_p^{\times}$ such that $-1 \in H$ and $[\mathbb{F}_p^{\times} : H] \geq 3$. Put $d = [\mathbb{F}_p^{\times} : H]$, then we have

$$d \left| \frac{p-1}{2} \right. \qquad \text{and} \qquad d = [K : \mathbb{Q}],$$

where $K = \mathbb{Q}(\zeta_p)^H$ and $\zeta_p = e^{\frac{2\pi i}{p}}$. We can identify the Galois group $\mathrm{Gal}(K/\mathbb{Q})$ with $\mathbb{F}_p^{\times}/H$, we also identify $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/K)$ with H. In particular, $K \subseteq \mathbb{Q}(\zeta_p)^+$, where $\mathbb{Q}(\zeta_p)^+ = \mathbb{Q}(\zeta_p + \bar{\zeta_p})$.

Put

$$G_H = \{g \in G : \det g \in H\}.$$

Then the determinant map induces an isomorphism: $G/G_H \cong \mathbb{F}_p^{\times}/H$. We denote by X_H the modular curve corresponding to G_H, which is defined over K. Here X_H and $X_{\mathrm{ns}}^{+}(p)$ have the same geometrically integral model, and the function field of X_H is $K(X_{\mathrm{ns}}^{+}(p))$. The curve X_H also has the same cusps as $X_{\mathrm{ns}}^{+}(p)$. In particular, $\mathrm{Gal}(K(X_H)/\mathbb{Q}(X_{\mathrm{ns}}^{+}(p))) \cong \mathrm{Gal}(K/\mathbb{Q})$.

Hence, in this chapter we identify the following four groups: $\mathrm{Gal}(K(X_H)/\mathbb{Q}(X_{\mathrm{ns}}^{+}(p)))$, $\mathrm{Gal}(K/\mathbb{Q})$, $\mathbb{F}_p^{\times}/H$ and G/G_H. The readers should interpret the exact meaning based on the context.

For $\mathbf{a} = (a_1, a_2) \in \mathbb{Q}^2$, we put $\ell_{\mathbf{a}} = B_2(a_1 - \lfloor a_1 \rfloor)/2$, where $B_2(T) = T^2 - T + \frac{1}{6}$ is the second Bernoulli polynomial. Obviously $|\ell_{\mathbf{a}}| \leq 1/12$, this will be used without special reference.

We put $\mathbb{A} = \left(p^{-1}\mathbb{Z}/\mathbb{Z}\right)^2 \setminus \{(0,0)\}$. In this chapter, we also identify $p^{-1}\mathbb{Z}/\mathbb{Z}$ with $p^{-1}\mathbb{F}_p$. Moreover we always choose a representative element of $\mathbf{a} = (a_1, a_2) \in (p^{-1}\mathbb{Z}/\mathbb{Z})^2$ satisfying $0 \leq a_1, a_2 < 1$. So in the sequel for every $\mathbf{a} \in (p^{-1}\mathbb{Z}/\mathbb{Z})^2$, we have $\ell_{\mathbf{a}} = B_2(a_1)/2$.

In this chapter, we use the notation $O_1(\cdot)$. Precisely, $A = O_1(B)$ means that $|A| \le B$.

3.4 Preparations

3.4.1 Siegel functions and modular units

Recall the definition of Siegel function in Section 2.3.1. From the proof of [27, Proposition 2.3] and replacing $3|q_\tau|$ by $2.03|q_\tau|$ in [27, Formula (11)], we directly get the following lemma. Note that D is the standard fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$.

Lemma 3.4. *Let $\mathbf{a} \in \mathbb{Q}^2 \setminus \mathbb{Z}^2$. Then for $\tau \in D$, we have*

$$\log|g_{\mathbf{a}}(\tau)| = \ell_{\mathbf{a}} \log|q_\tau| + \log|1 - q_\tau^{a_1} e^{2\pi i a_2}| + \log|1 - q_\tau^{1-a_1} e^{-2\pi i a_2}| + O_1(2.03|q_\tau|).$$

For $\mathbf{a} \in (p^{-1}\mathbb{Z})^2 \setminus \mathbb{Z}^2$, we denote $g_{\mathbf{a}}^{12p}$ by $u_{\mathbf{a}}$, which is a modular unit on the principal modular curve $X(p)$ of level p. Moreover, we have $u_{\mathbf{a}} = u_{\mathbf{a}'}$ when $\mathbf{a} \equiv \mathbf{a}' \bmod \mathbb{Z}^2$. Hence, $u_{\mathbf{a}}$ is well-defined when $\mathbf{a} \in \mathbb{A}$. In addition, every $u_{\mathbf{a}}$ is integral over $\mathbb{Z}[j]$. For more details, see [27, Section 4.2].

Furthermore, the Galois action on the set $\{u_{\mathbf{a}}\}$ is compatible with the right linear action of $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$ on it. That is, for any $\sigma \in \mathrm{Gal}(\mathbb{Q}(X(p))/\mathbb{Q}(X(1))) \cong \mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})/\pm 1$ and any $\mathbf{a} \in \mathbb{A}$, we have

$$u_{\mathbf{a}}^{\sigma} = u_{\mathbf{a}\sigma}.$$

Here we borrow a result and its proof from [5] for the conveniences of readers. In fact, it is a refinement of Proposition 2.6 in the present case.

Lemma 3.5 ([5]). *We have*

$$\prod_{\mathbf{a} \in \mathbb{A}} u_{\mathbf{a}} = p^{12p}.$$

Proof. We denote by u the left-hand side of the equality. Since the set $\mathbb{A}$ is stable with respect to $\mathrm{GL}_2(\mathbb{Z}/p\mathbb{Z})$, u is stable with respect to the Galois action over the field $\mathbb{Q}(X(1)) = \mathbb{Q}(j)$. So $u \in \mathbb{Q}(j)$. Moreover, since u is integral over $\mathbb{Z}[j]$, $u \in \mathbb{Z}[j]$. Notice that $X(1)$ has only one cusp and u has no zeros and poles outside the cusps, so u must be a constant and $u \in \mathbb{Z}$. Since

$$\sum_{(a_1,a_2) \in \mathbb{A}} B_2(a_1) = 0 \qquad \text{and} \qquad \sum_{(a_1,a_2) \in \mathbb{A}} a_2(1-a_1) = \frac{p^2-1}{4},$$

Taking $q = 0$, we have

$$u = \prod_{(a_1,a_2)\in\mathbb{A}, a_1=0} (1 - e^{2\pi i a_2})^{12p} = \prod_{1\le k<p} (1 - e^{2k\pi i/p})^{12p} = p^{12p}.$$

□

3.4.2 $X_{\mathrm{ns}}^+(p)$ and X_H

It is known that the cusps of $X_{\mathrm{ns}}^+(p)$ correspond to the orbits of the (left) action of $G \cap \mathrm{SL}_2(\mathbb{Z}/p\mathbb{Z})$ on the set $\mathbb{F}_p^2 \setminus \{\binom{0}{0}\}$, see [26, Lemma 2.3]. By definition, these orbits are the sets $\mathcal{L}_a$, defined by $x^2 - \Xi y^2 = \pm a$, where a runs through $\mathbb{F}_p^\times/\{\pm 1\}$, the cusp at infinity corresponds to $a = 1$.

From now on, we fix an integral point P of $X_{\mathrm{ns}}^+(p)$ and assume that $|j(P)| > 3500$. Since every integral point of $X_{\mathrm{ns}}^+(p)$ is also an integral point of X_H, P is also an integral point of X_H. Hence for our purposes, we only need to focus on the modular curve X_H.

Notice that since all the cusps have ramification index p in the natural covering $X_{\mathrm{ns}}^+(p) \to X(1)$, so as the natural covering $X_H \to X(1)$.

We fix a uniformization $X_H(\mathbb{C}) = \bar{\mathcal{H}}/\Gamma$, and let $\tau_0 \in \bar{\mathcal{H}}$ be a lift of P. Pick $\sigma_c \in \mathrm{SL}_2(\mathbb{Z})$ such that $\tau = \sigma_c^{-1}(\tau_0) \in D$. As in the proof of [27, Proposition 3.1] and with the notations therein, we can choose the cusp $c = \sigma_c(i\infty)$ and construct a certain set Ω_c as Section 2.3.3. Recall that for the cusp c, t_c is its local parameter and $q_c = t_c^p$, both of them are defined and analytic on Ω_c. Moreover, $q_c(P) = q_\tau$.

According to [27, Proposition 3.1], we have

$$\frac{1}{2}|j(P)| \le |q_c(P)^{-1}| \le \frac{3}{2}|j(P)|. \tag{3.1}$$

We will use (3.1) several times without special reference.

In the sequel we can assume that $|q_c(P)| \le 10^{-p}$. Indeed, the inequality $|q_c(P)| > 10^{-p}$ yields a much better estimate for $\log|j(P)|$ than those given in Theorems 3.1 and 3.3.

3.4.3 Modular units on X_H

The group $\mathrm{GL}_2(\mathbb{F}_p)$ acts naturally (on the right) on the set $\mathbb{A}$. Since $G_H \subset \mathrm{GL}_2(\mathbb{F}_p)$, let us consider the natural right group action of G_H on $\mathbb{A}$. There are d orbits of this group action. These orbits are the sets $\mathcal{O}_a$, defined by $\{(x/p, y/p) : x^2 - \Xi^{-1}y^2 \in aH\}$, where

a runs through $\mathbb{F}_p^{\times}/H$. In fact, if $(x, y) \in \mathcal{O}_a$, then for any $g \in G_H$, noticing the two possible representations of g, it is straightforward to show that $(x, y) \cdot g \in \mathcal{O}_a$.

Based on our conventions in Section 3.3, we consider the natural right group action of $\mathrm{Gal}(K/\mathbb{Q})$ on the set of orbits of the group action $\mathbb{A}/G_H$. Moreover, for any $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ and any orbit $\mathcal{O}_a$, we have

$$\mathcal{O}_a\sigma = \mathcal{O}_{a\sigma}.$$

It is easy to see that this group action is transitive. So we obtain the following lemma.

Lemma 3.6. *We have* $|\mathcal{O}_a| = (p^2 - 1)/d$.

Let $\mathcal{O}$ be an orbit of $\mathbb{A}/G_H$. As (2.5), we define

$$u_{\mathcal{O}} = \prod_{\mathbf{a}\in\mathcal{O}} u_{\mathbf{a}}. \tag{3.2}$$

By [27, Proposition 4.2 (ii)], $u_{\mathcal{O}}$ is a rational function on the modular curve X_H. Furthermore, $u_{\mathcal{O}}$ is a modular unit on X_H.

We denote by $\mathrm{Ord}_c(u_{\mathcal{O}})$ the vanishing order of $u_{\mathcal{O}}$ at c. The following lemma is derived directly from Lemma 3.4 and [27, Proposition 4.2 (iii)].

Lemma 3.7. *We have*

$$\log|u_{\mathcal{O}}(P)| = \frac{\mathrm{Ord}_c(u_{\mathcal{O}})}{p}\log|q_c(P)| + \log|\gamma_c| + O_1(17p^3|q_c(P)|^{1/p}) \tag{3.3}$$

where

$$\mathrm{Ord}_c(u_{\mathcal{O}}) = 12p^2 \sum_{\mathbf{a}\in\mathcal{O}\sigma_c} \ell_{\mathbf{a}} \quad \textit{and} \quad \gamma_c = \prod_{\substack{(a_1,a_2)\in\mathcal{O}\sigma_c \\ a_1=0}} (1 - e^{2\pi i a_2})^{12p}.$$

Proof. Here we use the following identity:

$$u_{\mathcal{O}}(P) = u_{\mathcal{O}}(\tau_0) = u_{\mathcal{O}}(\sigma_c(\sigma_c^{-1}(\tau_0))) = u_{\mathcal{O}\sigma_c}(\tau).$$

Notice that for $|z| \le r < 1$, we have

$$|\log|1+z|| \le \frac{-\log(1-r)}{r}|z|,$$

see [27, Formula (4)]. Taking $r = 0.1$ and combining Lemma 3.4 with Lemma 3.6, we have

$$\begin{aligned}\log|u_{\mathcal{O}}(P)| =& \frac{\mathrm{Ord}_c(u_{\mathcal{O}})}{p}\log|q_c(P)| + \log|\gamma_c| \\ &+O_1\left(26p\frac{p^2-1}{d}|q_c(P)|^{1/p} + 25p\frac{p^2-1}{d}|q_c(P)|\right).\end{aligned}$$

Then this lemma follows from $d \geq 3$. □

We want to indicate that γ_c is a real algebraic number. Because if $(0, a_2) \in \mathcal{O}\sigma_c$, then we have $(0, -a_2) \in \mathcal{O}\sigma_c$ based on the fact that if $(x, y) \in \mathcal{O}$, then $(-x, -y) \in \mathcal{O}$.

Lemma 3.8. *The group generated by the principal divisor $(u_{\mathcal{O}})$, where $\mathcal{O}$ runs over the orbits of $\mathbb{A}/G_H$, is of rank $d-1$.*

Proof. By Lemma 3.5, the rank of the free abelian group $(u_{\mathcal{O}})$ is at most $d-1$. Then Manin-Drinfeld theorem, as stated in [62], tells us that this rank is maximal possible. □

3.5 Baker's method on X_H

In this section we obtain a bound for $\log|j(P)|$, involving various parameters. Recall that P is the integral point of $X_{\mathrm{ns}}^{+}(p)$ fixed in Section 3.4.2.

3.5.1 Baker's inequality

Here, we recall Baker's inequality in the Archimedean case due to Matveev, see [68, Corollary 2.3].

Let F be a number field of degree d over $\mathbb{Q}$ and embedded in $\mathbb{C}$. If $F \subseteq \mathbb{R}$, we put $\delta = 1$, and otherwise $\delta = 2$. We let

- $\alpha_1, \cdots, \alpha_n$ be nonzero algebraic numbers, belonging to F;
- $A_1, \cdots, A_n$ be real numbers satisfying

$$A_k \geq \max\{d\mathrm{h}(\alpha_k), |\log\alpha_k|\} \quad (k = 1, \cdots, n);$$

- $b_1, \ldots, b_n$ be rational integers, $\Lambda = b_1\log\alpha_1 + \cdots + b_n\log\alpha_n$, $B = \{|b_1|, |b_2|, \cdots, |b_n|\}$.

Theorem 3.9 (Matveev). *Suppose that $\Lambda \neq 0$. Then we have*

$$\log|\Lambda| > -C_1(n)d^2 A_1 \cdots A_n \log(ed)\log(eB),$$

where $C_1(n) = \min\{\frac{1}{\delta}(\frac{1}{2}en)^{\delta} 30^{n+3} n^{3.5}, 2^{6n+20}\}$.

3.5.2 Cyclotomic units

We introduce a set of independent cyclotomic units of $\mathbb{Q}(\zeta_p)^+$ as follows,

$$\xi_{k-1} = \zeta_p^{(1-k)/2} \cdot \frac{1-\zeta_p^k}{1-\zeta_p} = \frac{\bar{\zeta}_p^{\,k/2} - \zeta_p^{k/2}}{\bar{\zeta}_p^{\,1/2} - \zeta_p^{1/2}}, \qquad k = 2, \ldots, \frac{p-1}{2},$$

for details see [93, Lemma 8.1]. In particular, $\{-1, \xi_1, \cdots, \xi_{\frac{p-3}{2}}\}$ is a set of independent generators for the full group of cyclotomic units of $\mathbb{Q}(\zeta_p)^+$. Let m' be the index of $\langle \xi_1, \cdots, \xi_{\frac{p-3}{2}} \rangle$ in the full unit group of $\mathbb{Q}(\zeta_p)^+$ modulo roots of unity, which is equal to the class number of $\mathbb{Q}(\zeta_p)^+$.

We put

$$\eta_k = \mathcal{N}_{\mathbb{Q}(\zeta_p)^+/K}(\xi_k) = \prod_{\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_p)^+/K)} \xi_k^{\sigma}, \qquad k = 1, \ldots, \frac{p-3}{2}.$$

Let m be the exponent of $\langle \eta_1, \cdots, \eta_{\frac{p-3}{2}} \rangle$ in the full unit group of K modulo roots of unity. Since $[\mathbb{Q}(\zeta_p)^+ : K] = |H|/2 = \frac{p-1}{2d}$, we have

$$m \,\Big|\, \frac{m'(p-1)}{2d}. \tag{3.4}$$

Since m is finite and the rank of the full unit group of K is $d-1$, the group $\langle \eta_1, \cdots, \eta_{\frac{p-3}{2}} \rangle$ modulo roots of unity has rank $d-1$. In particular, in the sequel we assume that $\eta_1, \cdots, \eta_{d-1}$ are multiplicatively independent without loss of generality.

3.5.3 More about modular units on X_H

We fix an orbit $\mathcal{O}$ of the group action $\mathbb{A}/G_H$. Put $U = u_{\mathcal{O}}$, where $u_{\mathcal{O}}$ is defined in (3.2).

Based on our conventions in Section 3.3, for any $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, we can define U^{σ} as the natural Galois action. Indeed, we can view σ as an element of $\mathrm{Gal}(K(X_H)/\mathbb{Q}(X_{\mathrm{ns}}^{+}(p)))$ and $U \in K(X_H)$. Moreover, we have $U^{\sigma} = u_{\mathcal{O}^{\sigma}}$ and $U(P)^{\sigma} = U^{\sigma}(P)$.

Since the Galois group $\mathrm{Gal}(K/\mathbb{Q})$ acts transitively on the set of orbits of $\mathbb{A}/G_H$, we can rewrite Lemma 3.5 as follows.

Lemma 3.10. *We have*

$$\prod_{\sigma \in \mathrm{Gal}(K/\mathbb{Q})} U^{\sigma} = p^{12p}.$$

By Lemma 3.6 and the formula for $\mathrm{Ord}_c u_{\mathcal{O}}$ appearing in Lemma 3.7 we obtain a bound for the vanishing order of U at c.

Lemma 3.11. *We have*

$$|\mathrm{Ord}_c U| \leq \frac{p^2(p^2-1)}{d}.$$

For $1-\zeta_p$, we take the $\mathbb{Q}(\zeta_p)/K$-norm, setting $\mu = \mathcal{N}_{\mathbb{Q}(\zeta_p)/K}(1-\zeta_p)$.

Lemma 3.12. *We have* $(U(P)) = \left(\mu^{12p}\right)$.

Proof. Since P is an integral point of X_H, by [27, Proposition 4.2 (i)] and Lemma 3.10, the principal ideal $(U(P))$ is an integral ideal of the field K of the form $\mathfrak{p}^n$, where $\mathfrak{p} = (\mu)$ and n is a positive integer.

In addition, since $\mathfrak{p}$ is stable under the Galois action over $\mathbb{Q}$, we have $(U^{\sigma}(P)) = \mathfrak{p}^n$ for every $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$. Noticing that $\mathfrak{p}^d = (p)$, it follows from Lemma 3.10 that $n = 12p$. □

So Dirichlet's unit theorem gives

$$U(P)^m = \pm \eta_0^m \eta_1^{b_1} \dots \eta_{d-1}^{b_{d-1}},$$

where $\eta_0 = \mu^{12p}$ and $b_1, \cdots, b_{d-1}$ are some rational integers.

Let

$$V = U/\eta_0,$$

then we have

$$V(P)^m = \pm \eta_1^{b_1} \dots \eta_{d-1}^{b_{d-1}},$$

and $\mathrm{Ord}_c V = \mathrm{Ord}_c U$. For every $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, we have

$$V^{\sigma}(P)^m = \pm (\eta_1^{\sigma})^{b_1} \dots (\eta_{d-1}^{\sigma})^{b_{d-1}}, \tag{3.5}$$

where $V^{\sigma} = U^{\sigma}/\eta_0^{\sigma}$. Furthermore, by (3.3), we have

$$\log |V^{\sigma}(P)| = \frac{\mathrm{Ord}_c V^{\sigma}}{p} \log |q_c(P)| + \log |\Upsilon_{c,\sigma}| + O_1\left(17p^3 |q_c(P)|^{1/p}\right), \tag{3.6}$$

where $\Upsilon_{c,\sigma} = \gamma_{c,\sigma}/\eta_0^{\sigma}$ and

$$\gamma_{c,\sigma} = \prod_{\substack{(a_1,a_2)\in\mathcal{O}\sigma\sigma_c \\ a_1=0}} (1-e^{2\pi i a_2})^{12p}.$$

Notice that $\gamma_{c,\sigma} = \gamma_c$ when σ is the identity. So $\Upsilon_{c,1} = \gamma_c/\eta_0$.

Finally we put

$$B = \max\{|b_1|, \cdots, |b_{d-1}|, m\}.$$

3.5.4 Upper bound for B

We fix an order on the elements of the Galois group by supposing

$$\mathrm{Gal}(K/\mathbb{Q}) = \{\sigma_0 = 1, \sigma_1, \cdots, \sigma_{d-1}\}.$$

Since the real algebraic numbers $\eta_1, \cdots, \eta_{d-1}$ are multiplicatively independent, the $(d-1)\times(d-1)$ real matrix $A = \left(\log|\eta_\ell^{\sigma_k}|\right)_{1\le k,\ell\le d-1}$ is non-singular. Let $(\alpha_{k\ell})_{1\le k,\ell\le d-1}$ be the inverse matrix. Then by (3.5) we have

$$b_k = m\sum_{\ell=1}^{d-1}\alpha_{k\ell}\log|V^{\sigma_\ell}(P)|, \quad 1\le k\le d-1.$$

Define the following quantities:

$$\begin{aligned}
\delta_{c,k} &= \frac{m}{p}\sum_{\ell=1}^{d-1}\alpha_{k\ell}\mathrm{Ord}_c V^{\sigma_\ell},\\
\beta_{c,k} &= m\sum_{\ell=1}^{d-1}\alpha_{k\ell}\log|\Upsilon_{c,\sigma_\ell}|,\\
\kappa &= \max\{\max_k\sum_{\ell=1}^{d-1}|\alpha_{k\ell}|, 1\}.
\end{aligned}$$

According to (3.6), we have

$$b_k = \delta_{c,k}\log|q_c(P)| + \beta_{c,k} + O_1\left(17p^3 m\kappa|q_c(P)|^{1/p}\right).$$

Let $\delta = \max_k|\delta_{c,k}|$ and $\beta = \max_k|\beta_{c,k}|$. Then we have

$$B \le \delta\log|q_c(P)^{-1}| + \beta + 2p^3 m\kappa. \tag{3.7}$$

3.5.5 Preparation for Baker's inequality

We define the following function

$$W = \begin{cases} V & \text{if } \mathrm{Ord}_c V = 0, \\ \\ V^{\mathrm{Ord}_c V^{\sigma}} (V^{\sigma})^{-\mathrm{Ord}_c V} & \text{if } \mathrm{Ord}_c V \neq 0, \end{cases}$$

where $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ and $\sigma \neq 1$. So we always have $\mathrm{Ord}_c W = 0$. Moreover, W is not a constant by Lemma 3.8. In Section 2.8 we will choose special U (i.e. V) and σ to deal with an exceptional case that will occur.

Define

$$\alpha_d = \begin{cases} |\Upsilon_{c,1}|^{-1} & \text{if } \mathrm{Ord}_c V = 0, \\ \\ \left| \frac{\Upsilon_{c,1}^{\mathrm{Ord}_c V^{\sigma}}}{\Upsilon_{c,\sigma}^{\mathrm{Ord}_c V}} \right|^{-1} & \text{if } \mathrm{Ord}_c V \neq 0. \end{cases}$$

Then by (3.6) and Lemma 3.11 we obtain

$$\log |W(P)| = -\log \alpha_d + O_1\left(12p^7 |q_c(P)|^{1/p}\right). \tag{3.8}$$

Put

$$\Lambda = m \log |W(P)| + m \log \alpha_d.$$

If $\mathrm{Ord}_c V = 0$, by (3.5), we have

$$\Lambda = b_1 \log |\eta_1| + \cdots + b_{d-1} \log |\eta_{d-1}| + m \log \alpha_d.$$

In this case, we put $\alpha_k = |\eta_k|$ for $1 \leq k \leq d-1$.

If $\mathrm{Ord}_c V \neq 0$, by (3.5), we have

$$\Lambda = b_1 \log \left| \frac{\eta_1^{\mathrm{Ord}_c V^{\sigma}}}{(\eta_1^{\sigma})^{\mathrm{Ord}_c V}} \right| + \cdots + b_{d-1} \log \left| \frac{\eta_{d-1}^{\mathrm{Ord}_c V^{\sigma}}}{(\eta_{d-1}^{\sigma})^{\mathrm{Ord}_c V}} \right| + m \log \alpha_d.$$

In this case, we put $\alpha_k = \left| \frac{\eta_k^{\mathrm{Ord}_c V^{\sigma}}}{(\eta_k^{\sigma})^{\mathrm{Ord}_c V}} \right|$ for $1 \leq k \leq d-1$.

Hence, in both two cases we have

$$\Lambda = b_1 \log \alpha_1 + \cdots + b_{d-1} \log \alpha_{d-1} + m \log \alpha_d. \tag{3.9}$$

Notice that all α_k, $1 \leq k \leq d$, are contained in $\mathbb{Q}(\zeta_p)^+$.

3.5.6 Using Baker's inequality

If $\Lambda = 0$, we can get a better bound for $\log|j(P)|$, see Section 2.8. So here we assume that $\Lambda \neq 0$.

Using Theorem 3.9 and combining (3.7) and (3.8), we have

$$\begin{cases} |\Lambda| > \exp\left(-C_1(d)\Omega(\frac{p-1}{2})^2(1+\log\frac{p-1}{2})(1+\log B)\right), \\ |\Lambda| \leq \lambda |q_c(P)|^{1/p} \leq \lambda \exp\left(\frac{-B+\beta+2p^3 m\kappa}{\delta p}\right), \end{cases} \tag{3.10}$$

where

$$\begin{aligned} &C_1(d) = \min\left\{\frac{e}{2}d^{4.5}30^{d+3}, 2^{6d+20}\right\}, \\ &A_k \geq \max\{\frac{p-1}{2}\mathrm{h}(\alpha_k), |\log\alpha_k|, 0.16\}, 1 \leq k \leq d, \\ &\Omega = A_1 \cdots A_d, \quad \lambda = 12p^7 m, \end{aligned}$$

and $\mathrm{h}(\cdot)$ is the usual absolute logarithmic height.

We obtain $B \leq K_1 \log B + K_2$, where

$$\begin{aligned} K_1 &= \delta p C_1(d)\Omega(\frac{p-1}{2})^2(1+\log\frac{p-1}{2}), \\ K_2 &= \delta p C_1(d)\Omega(\frac{p-1}{2})^2(1+\log\frac{p-1}{2}) + \beta + 2p^3 m\kappa + \delta p \log\lambda. \end{aligned}$$

By [25, Lemma 2.3.3], we obtain

$$B \leq B_0 = 2(K_1 \log K_1 + K_2).$$

Then by (3.10), we have

$$|q_c(P)^{-1}| < \lambda^p \exp(pC_1(d)\Omega(\frac{p-1}{2})^2(1+\log\frac{p-1}{2})(1+\log B_0)).$$

Finally we have

$$\log|j(P)| < pC_1(d)\Omega(\frac{p-1}{2})^2(1+\log\frac{p-1}{2})(1+\log B_0) + p\log\lambda + \log 2. \tag{3.11}$$

Hence, to get a bound for $\log|j(P)|$, we only need to calculate the quantities in the above inequality, and we will do this in the next section.

It is easy to see that

$$C_1(d) = \min\left\{\frac{e}{2}d^{4.5}30^{d+3}, 2^{6d+20}\right\} < 2d^{4.5}30^{d+3}.$$

3.6 Computations

3.6.1 Upper Bound for m

Let h^+, R^+ and D^+ be the class number, regulator and discriminant of $\mathbb{Q}(\zeta_p)^+$, respectively.

By [93, Lemma 8.1 and Theorem 8.2], we have $m' = h^+$. By [93, Proposition 2.1 and Lemma 4.19], we have $|D^+| = p^{\frac{p-3}{2}}$. Then the class number formula (see [93, Page 37]) gives

$$h^+ = \left(\frac{p}{4}\right)^{\frac{p-3}{4}} \cdot \frac{1}{R^+} \prod_{\chi \neq 1} L(1,\chi).$$

Using [39, Theorem 2] to the field extension $\mathbb{Q}(\zeta_p)^+/\mathbb{Q}$, we have $R^+ > 0.32$. Applying [65, Theorem 1] to the field extension $\mathbb{Q}(\zeta_p)^+/\mathbb{Q}$ and noticing the constant $\mu_{\mathbb{Q}}$ below Formula (6) of [65], we get

$$|L(1,\chi)| < \frac{1}{2}\log p + 0.03 < \log p, \quad \text{if } \chi \neq 1.$$

Hence we have

$$h^+ < p^{\frac{p-3}{4}} (\log p)^{\frac{p-3}{2}}.$$

Finally by (3.4), we obtain

$$m \leq \frac{h^+(p-1)}{2d} < p^{\frac{p+1}{4}} (\log p)^{\frac{p-3}{2}}. \tag{3.12}$$

In the sequel we use the following formulas. For any $n \in \mathbb{Z}$ and $a_1, \cdots, a_k, \alpha \in \bar{\mathbb{Q}}$, we have

$$\begin{aligned}
&\mathrm{h}(a_1 + \cdots + a_k) \leq \mathrm{h}(a_1) + \cdots + \mathrm{h}(a_k) + \log k,\\
&\mathrm{h}(a_1 \cdots a_k) \leq \mathrm{h}(a_1) + \cdots + \mathrm{h}(a_k),\\
&\mathrm{h}(\alpha^n) = |n|\mathrm{h}(\alpha),\\
&\mathrm{h}(\zeta) = 0 \quad \text{for any root of unity } \zeta \in \mathbb{C}.
\end{aligned}$$

3.6.2 Height of η_{k-1} for $k = 2, \ldots, (p-1)/2$

Let $a \in \mathbb{F}_p^{\times}$ and $\sigma_a \in \mathrm{Gal}\,(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ induced by the automorphism of $\mathbb{Q}(\zeta_p) : \zeta_p \to \zeta_p^a$.

Since $\xi_{k-1}^{\sigma_a} = \frac{\bar{\zeta_p}^{ak/2} - \zeta_p^{ak/2}}{\bar{\zeta_p}^{a/2} - \zeta_p^{a/2}}$, we have $\mathrm{h}(\xi_{k-1}^{\sigma_a}) \leq 2\log 2$. So

$$\mathrm{h}(\eta_{k-1}^{\sigma_a}) \leq \frac{(p-1)\log 2}{d}.$$

Notice that if $-\frac{\pi}{2} < x < \frac{\pi}{2}$, then $\frac{\sin x}{x} > \frac{2}{\pi}$. Since $\xi_{k-1}^{\sigma_a} = \frac{\sin(\pi ak/p)}{\sin(\pi a/p)}$, we have

$$|\xi_{k-1}^{\sigma_a}| \leq \frac{1}{|\sin(\pi a/p)|} \leq \frac{1}{\sin(\pi/p)} < \frac{p}{2},$$

and

$$|\xi_{k-1}^{\sigma_a}| \geq |\sin(\pi ak/p)| \geq \sin(\pi/p) > \frac{2}{p}.$$

So we have $|\log|\xi_{k-1}^{\sigma_a}|| < \log\frac{p}{2}$. Hence

$$|\log|\eta_{k-1}^{\sigma_a}|| < \frac{(p-1)\log\frac{p}{2}}{2d}.$$

Since we can view $\mathrm{Gal}(K/\mathbb{Q})$ as a quotient group of $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$, for any $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, we have

$$\mathrm{h}(\eta_{k-1}^{\sigma}) \leq \frac{(p-1)\log 2}{d} \quad \text{and} \quad |\log|\eta_{k-1}^{\sigma}|| < \frac{(p-1)\log\frac{p}{2}}{2d}. \tag{3.13}$$

3.6.3 Height of η_0

Following the method in Section 3.6.2, we have $\mathrm{h}(1-\zeta_p^{\sigma_a}) \leq \log 2$. So

$$\mathrm{h}(\eta_0^{\sigma_a}) \leq \frac{12p(p-1)\log 2}{d}.$$

First we have $|1-\zeta_p^{\sigma_a}| \leq 2$. Second we have

$$|1-\zeta_p^{\sigma_a}|^2 \geq 2 - 2\cos\frac{\pi}{p} = 4\left(\sin\frac{\pi}{2p}\right)^2 > \left(\frac{2}{p}\right)^2.$$

So we have $|\log|1-\zeta_p^{\sigma_a}|| < \log\frac{p}{2}$. Hence

$$|\log|\eta_0^{\sigma_a}|| < \frac{12p(p-1)\log\frac{p}{2}}{d}.$$

Since we can view $\mathrm{Gal}(K/\mathbb{Q})$ as a quotient group of $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$, for any $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, we obtain

$$\mathrm{h}(\eta_0^\sigma) \leq \frac{12p(p-1)\log 2}{d} \quad \text{and} \quad |\log|\eta_0^\sigma|| < \frac{12p(p-1)\log\frac{p}{2}}{d}. \tag{3.14}$$

3.6.4 Height of $|\Upsilon_{c,\sigma}|$

Recall that $\Upsilon_{c,\sigma} = \gamma_{c,\sigma}/\eta_0^\sigma$, $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ and

$$\gamma_{c,\sigma} = \prod_{\substack{(a_1,a_2)\in\mathcal{O}\sigma\sigma_c \\ a_1=0}} (1-e^{2i\pi a_2})^{12p}.$$

Notice the description of $\mathcal{O}$ in Section 3.4.3, we have $|\{(a_1,a_2) \in \mathcal{O}\sigma\sigma_c : a_1 = 0\}| \leq 2|H| = \frac{2(p-1)}{d}$. Following the method in Section 3.6.2, we get

$$\mathrm{h}(\gamma_{c,\sigma}) \leq \frac{24p(p-1)\log 2}{d}.$$

Since $\Upsilon_{c,\sigma} = \gamma_{c,\sigma}/\eta_0^\sigma$, we have, by (3.14),

$$\mathrm{h}(\Upsilon_{c,\sigma}) \leq \mathrm{h}(\gamma_{c,\sigma}) + \mathrm{h}(\eta_0^\sigma) \leq \frac{36p(p-1)\log 2}{d}.$$

Noticing that $|\Upsilon_{c,\sigma}|^2 = \Upsilon_{c,\sigma}\bar{\Upsilon}_{c,\sigma}$, we get

$$\mathrm{h}(|\Upsilon_{c,\sigma}|) \leq \frac{36p(p-1)\log 2}{d}. \tag{3.15}$$

Since $a_1 = 0$, we have $a_2 \in \{\frac{1}{p}, \cdots, \frac{p-1}{p}\}$. First we have $|1-e^{2i\pi a_2}| \leq 2$. Second

$$|1-e^{2i\pi a_2}|^2 = 2(1-\cos 2\pi a_2) \geq 2(1-\cos\pi/p) = 4\sin^2\frac{\pi}{2p} \geq \frac{4}{p^2}.$$

So we have $|\log|1-e^{2i\pi a_2}|| \leq \log\frac{p}{2}$, and then

$$|\log|\gamma_{c,\sigma}|| \leq \frac{24p(p-1)\log\frac{p}{2}}{d}.$$

Hence we have, by (3.14),

$$|\log|\Upsilon_{c,\sigma}|| \leq \frac{36p(p-1)\log\frac{p}{2}}{d}. \tag{3.16}$$

3.6.5 Calculation of Ω

Recall that $\Omega = A_1 \cdots A_d$, where

$$A_k \geq \max\{\frac{p-1}{2}\mathrm{h}(\alpha_k), |\log \alpha_k|, 0.16\}, \quad 1 \leq k \leq d.$$

If $\mathrm{Ord}_c V = 0$, then $\alpha_k = |\eta_k| = \pm\eta_k$, $1 \leq k \leq d-1$, and $\alpha_d = |\Upsilon_{c,1}|^{-1}$. Then, by (3.13), for $1 \leq k \leq d-1$, we can choose $A_k = p^2/d$. For A_d, we can choose $A_d = 36p^3/d$, by (3.15) and (3.16).

If $\mathrm{Ord}_c V \neq 0$, then $\alpha_k = \left|\frac{\eta_k^{\mathrm{Ord}_c V^\sigma}}{(\eta_k^\sigma)^{\mathrm{Ord}_c V}}\right|$, $1 \leq k \leq d-1$, and $\alpha_d = \left|\frac{\Upsilon_{c,1}^{\mathrm{Ord}_c V^\sigma}}{\Upsilon_{c,\sigma}^{\mathrm{Ord}_c V}}\right|^{-1}$. For $1 \leq k \leq d-1$, combining Lemma 3.11 with (3.13) we can choose $A_k = p^6/d^2$. For A_d, we can choose $A_d = 36p^7/d^2$.

Therefore, we can choose

$$\Omega = 36p^{6d+1}/d^{2d}. \tag{3.17}$$

3.6.6 Calculation of B_0

For our purpose we need to calculate δ, β and κ. In fact, all we want to do is to get a bound for $|\alpha_{k\ell}|$, $1 \leq k, \ell \leq d-1$.

Let R_K be the regulator of K. By [93, Lemma 4.15], we have $|\det A| \geq mR_K$. Applying [39, Theorem 2] to the field extension $K/\mathbb{Q}$, we have $R_K > 0.32$. So we get $|\det A| > 0.32m$.

Notice that $\alpha_{k\ell} = \frac{1}{\det A}A_{\ell k}$, where A_{lk} is the relative cofactor. The reader should not confuse the matrix A, the constants A_k introduced in Section 3.5.6 and the cofactors A_{lk}.

By Hadamard's inequality and (3.13), we have

$$|A_{\ell k}| \leq \left[\frac{(p-1)\sqrt{d-2}\log\frac{p}{2}}{2d}\right]^{d-2}.$$

Then we have

$$\begin{aligned} |\alpha_{k\ell}| &< \left[\frac{(p-1)\sqrt{d-2}\log\frac{p}{2}}{2d}\right]^{d-2} \cdot \frac{1}{0.32m} \\ &< (p\sqrt{p}\log p)^{\frac{p-1}{2}-2}/m \\ &= p^{\frac{3p-15}{4}}(\log p)^{\frac{p-5}{2}}/m. \end{aligned}$$

Hence, we obtain

$$\begin{aligned}
&\delta < p^{\frac{3p-3}{4}}(\log p)^{\frac{p-5}{2}},\\
&\beta < 36p^{\frac{3p-7}{4}}(\log p)^{\frac{p-3}{2}},\\
&\kappa < p^{\frac{3p-11}{4}}(\log p)^{\frac{p-5}{2}}/m.
\end{aligned}$$

Notice that $d \le (p-1)/2$ and $p \ge 7$, so we get $C_1(d) \le p^{p+8}$. Therefore, we have

$$K_1 < p^{5p+9}(\log p)^{p-1}, \qquad K_2 < 4p^{5p+9}(\log p)^{p-1},$$

and then

$$B_0 < 16p^{5p+10}(\log p)^p, \qquad 1 + \log B_0 < 8p \log p.$$

3.6.7 Final results

Finally, by (3.11) we get an explicit bound for $\log|j(P)|$ as follows

$$\begin{aligned}
\log|j(P)| &< 2pC_1(d)\Omega(\frac{p-1}{2})^2(1+\log\frac{p-1}{2})(1+\log B_0)\\
&< C(d)p^{6d+5}(\log p)^2,
\end{aligned}$$

where $C(d) = 30^{d+5} \cdot d^{-2d+4.5}$. Hence we obtain Theorem 3.1.

If we choose $d = (p-1)/2$, applying the bound $p - 1 \ge 6p/7$ and a few numerical computations, we get Theorem 3.3.

3.7 The case $\Lambda = 0$

In this section, we suppose that $\Lambda = 0$. Using the method in Section 2.8, we will obtain a better bound for $\log|j(P)|$ than Theorem 3.1.

First we assume that $\mathrm{Ord}_c V = 0$, i.e. $\mathrm{Ord}_c U = 0$. Then we have $|U(P)| = |\gamma_c|$. Since $U(P)$ and γ_c are real, we have $U(P)^2 = \gamma_c^2$, i.e. $U^2(P) = \gamma_c^2$.

Recall Ω_c and the q-parameter q_c mentioned in Section 3.4.2. Let v be an absolute value of $\mathbb{Q}(\zeta_p)$ normalized to extend a standard absolute value on $\mathbb{Q}$. For the modular function U^2, by Lemma 2.17 we get the following lemma.

Lemma 3.13. *There exist an integer function $f(\cdot)$ with respect to q_c and $\lambda_1^c, \lambda_2^c, \lambda_3^c \cdots \in \mathbb{Q}(\zeta_p)$ such that the following identity holds in Ω_c,*

$$\log \frac{U^2(q_c)}{\gamma_c^2 q_c^{\frac{2\mathrm{Ord}_c U}{p}}} = 2\pi f(q_c) i + \sum_{k=1}^{\infty} \lambda_k^c q_c^{k/p}, \tag{3.18}$$

and

$$|\lambda_k^c|_v \leq \begin{cases} |k|_v^{-1} & \text{if } v \text{ is finite,} \\ 48p^2(k+p) & \text{if } v \text{ is infinite.} \end{cases}$$

In particular, for every $k \geq 1$ we have

$$\mathrm{h}(\lambda_k^c) \leq \log(48p^3 + 48kp^2) + \log k.$$

Corollary 3.14. *With the assumption $\mathrm{Ord}_c U = 0$, we have $\lambda_k^c \neq 0$ for some $k \leq p^5$.*

Proof. Since $\mathrm{Ord}_c U = 0$ and U is not a constant, there must exist some $\lambda_k^c \neq 0$. Under the assumption $\mathrm{Ord}_c U = 0$, we have $U(c) = \gamma_c$, and then $f(q_c(c)) = 0$ by (3.18). We extend the additive valuation Ord_c from the field $K(X_H)$ to the field of formal power series $K((q_c^{1/p}))$. Then $\mathrm{Ord}_c q_c^{1/p} = 1$ and $\mathrm{Ord}_c \left(-2\pi f(q_c) i + \log(U^2/\gamma_c^2)\right) \leq \mathrm{Ord}_c \log(U^2/\gamma_c^2) = \mathrm{Ord}_c(U^2/\gamma_c^2 - 1)$. The latter quantity is bounded by the degree of $U^2/\gamma_c^2 - 1$, which is equal to the degree of U^2.

The degree of U^2 is equal to $\sum\limits_{c_0} |\mathrm{Ord}_{c_0} U|$, here the sum runs through all the cusps of X_H. Then the result follows from Lemma 3.11. □

Now we can get a bound for $\log|j(P)|$.

Proposition 3.15. *Under the assumptions $\Lambda = 0$ and $\mathrm{Ord}_c U = 0$, we have*

$$\log|j(P)| \leq p^2 \log(48p^{12} + 48p^8) + p \log(96p^2(p^5 + p + 1)) + \log 2.$$

Proof. Let n be the smallest k such that $\lambda_k^c \neq 0$. Then $n \leq p^5$. We assume that $|q_c(P)| \leq 10^{-p}$, otherwise there is nothing to prove. Since $\mathrm{Ord}_c U = 0$ and $U^2(P) = \gamma_c^2$, it follows from (3.18) that $2\pi f(q_c(P)) i + \sum\limits_{k=n}^{\infty} \lambda_k^c q_c(P)^{k/p} = 0$.

Suppose that $f(q_c(P)) = 0$. Then $|\lambda_n^c q_c(P)^{n/p}| = |\sum\limits_{k=n+1}^{\infty} \lambda_k^c q_c(P)^{k/p}|$. On one side, we have

$$\begin{aligned} |\sum_{k=n+1}^{\infty} \lambda_k^c q_c(P)^{k/p}| \leq \sum_{k=n+1}^{\infty} |\lambda_k^c| |q_c(P)|^{k/p} &\leq \sum_{k=n+1}^{\infty} 48p^2(k+p)|q_c(P)|^{k/p} \\ &\leq 96p^2(n+p+1)|q_c(P)|^{(n+1)/p}. \end{aligned}$$

On the other side, using Liouville's inequality (see [92, Formula (3.13)]), we get

$$|\lambda_n^c| \geq e^{-[\mathbb{Q}(\zeta_p):\mathbb{Q}]\mathrm{h}(\lambda_n^c)} \geq (48np^3 + 48n^2p^2)^{-p+1}.$$

Then we obtain

$$\log|q_c(P)^{-1}| \leq p^2 \log(48p^{12} + 48p^8) + p\log(96p^2(p^5 + p + 1)).$$

Finally, the desired result follows from (3.1).

Suppose that $f(q_c(P)) \neq 0$. Then $2\pi \leq |\sum_{k=n}^{\infty} \lambda_k^c q_c(P)^{k/p}| \leq 96p^2(n+p)|q_c(P)|^{n/p}$. Then we get $\log|q_c(P)^{-1}| \leq p\log(96p^2(p^5+p))$. So we have

$$\log|j(P)| \leq p\log(96p^2(p^5+p)) + \log 2.$$

□

Now we assume that $\mathrm{Ord}_c V \neq 0$, i.e. $\mathrm{Ord}_c U \neq 0$. By Lemma 3.10, we can choose a U such that $\mathrm{Ord}_c U < 0$. Then we choose a σ such that $\mathrm{Ord}_c U^\sigma > 0$. Put $n_1 = -\mathrm{Ord}_c U$ and $n_2 = \mathrm{Ord}_c U^\sigma$. Since $U(P)$ and γ_c are real, we have $U(P)^{2n_2} U^\sigma(P)^{2n_1} = \gamma_c^{2n_2}\gamma_{c,\sigma}^{2n_1}$, i.e. $U^{2n_2}(U^\sigma)^{2n_1}(P) = \gamma_c^{2n_2}\gamma_{c,\sigma}^{2n_1}$. Lemma 3.8 guarantees that $U^{2n_2}(U^\sigma)^{2n_1}$ is not a constant.

Applying the same method as above without difficulties, we can also get a better bound than Theorem 3.1. We omit the details here.

Chapter 4

Bounding the j-invariant of integral points on certain modular curves

4.1 Main Results

Let Γ be a congruence subgroup of level N ($N \geq 2$) and X_Γ its corresponding modular curve. Assume that X_Γ is defined over a number field K. Let S be a finite set of absolute values of K, containing all the Archimedean valuations and normalized with respect to $\mathbb{Q}$.

In this chapter, we will give quantitative version for Theorem 1.4. As an application, it is also a quantitative version for Theorem 1.3 when Γ has no elliptic elements, as well as for certain modular curves which have positive genus and less than three cusps. For example, the classical modular curve $X_0(p)$ for a prime $p > 13$, it has positive genus and two cusps.

Recall that a non-cuspidal point $P \in X_\Gamma$ is called elliptic if for some $\tau \in \mathcal{H}$ representing P the stabilizer $\Gamma_z \neq \{\pm 1\}$. Notice that the curve X_Γ has finitely many elliptic points. We assume that the set of its elliptic points is $\{P_1, P_2, \cdots, P_n\}$. For each elliptic point P_i, we fix a pre-image z_i in $\mathcal{H}$. We denote by Γ_{z_i} the stabilizer of z_i in Γ. It is well-known that each Γ_{z_i} is cyclic of order 3, 4, or 6.

Let $\widetilde{\Gamma}$ be the congruence subgroup generated by $\Gamma(N)$ and $\{\Gamma_{z_1}, \cdots, \Gamma_{z_n}\}$. Consider the natural finite covering $\phi : X_{\widetilde{\Gamma}} \to X_\Gamma$. For any point $\tilde{P} \in X_{\widetilde{\Gamma}}$, fix a pre-image $z \in \mathcal{H}$, the ramification index of $\tilde{P}$ over X_Γ is equal to the index $[\pm\Gamma_z : \pm\widetilde{\Gamma}_z]$ which does not depend on the choice of z. Therefore, ϕ is unramified outside the cusps.

Assume that Γ has a congruence subgroup Γ' such that $X_{\Gamma'}$ has at least three cusps and the finite covering $X_{\Gamma'} \to X_\Gamma$ is unramified outside the cusps. Then we must have $\widetilde{\Gamma} \subseteq \Gamma'$, subsequently $X_{\widetilde{\Gamma}}$ also has at least three cusps. Under this assumption, by the results in Chapter 2 we can get effective Siegel's theorem for $X_{\widetilde{\Gamma}}$. Then the effective Siegel's theorem for X_Γ follows from quantitative Riemann existence theorem [30] and quantitative Chevalley-Weil theorem [31].

First we fix some notation. Put

$$d_N = \begin{cases} \frac{1}{2}N^3\prod_{\ell|N}(1-1/\ell^2) & \text{if } N>2, \\ \\ 6 & \text{if } N=2, \end{cases}$$

where ℓ runs through all primes dividing N. Let $d=[K:\mathbb{Q}]$, $s=|S|$, and

$$D^* = D_K^{d_N} e^{(\mathrm{h}(S)+(1+\log 1728)\Lambda)dd_N},$$

where D_K is the absolute discriminant of K,

$$\Lambda = \left(\left(\frac{d_N(N-6)}{12N}+2\right)d_N \right)^{25(\frac{d_N(N-6)}{12N}+2)d_N},$$

and

$$\mathrm{h}(S) = \frac{1}{d}\sum_{\substack{v\in S\\ v\nmid\infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v).$$

Next we define

$$\Delta_1 = d^{-d}\sqrt{N^{Ndd_N}|D^*|^{\varphi(N)}}\left(\log(N^{Ndd_N}|D^*|^{\varphi(N)})\right)^{\varphi(N)dd_N}\left(\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v)\right)^{\varphi(N)d_N}. \tag{4.1}$$

In addition, we denote by p the maximal rational prime below S, with the convention $p=1$ if S consists only of the infinite places. Now we are ready to state the main results.

Theorem 4.1. *Assume that Γ has a congruence subgroup Γ' with $\nu_\infty(\Gamma')\geq 3$, and Γ' contains all elliptic elements of Γ. Furthermore, suppose that N is not a power of any prime. Then for any S-integral point P on X_Γ, we have*

$$\mathrm{h}(P) \leq \left(Cdsd_N^2N^2\right)^{2sNd_N}(\log(dNd_N))^{3sNd_N}p^{dNd_N}\Delta_1,$$

where C is an absolute effective constant.

When N is a prime power, we define

$$M = \begin{cases} 2N & \text{if } N \text{ is not a power of } 2, \\ \\ 3N & \text{if } N \text{ is a power of } 2. \end{cases}$$

Theorem 4.2. *Assume that Γ has a congruence subgroup Γ' with $\nu_\infty(\Gamma') \geq 3$, and Γ' contains all elliptic elements of Γ. Furthermore, suppose that N is a power of some prime. Then for any S-integral point P on X_Γ, we can get an upper bound for $\mathrm{h}(P)$ by replacing N with M in Theorem 4.1.*

Here, we would like to give some examples satisfying the assumptions in Theorems 4.1 and 4.2.

Example 4.3. Assume that Γ has no elliptic elements. Then the principal congruence subgroup $\Gamma(N)$ is such a subgroup of Γ when $N \geq 2$.

Example 4.4. For a prime $p > 13$, the classical modular curve $X_0(p)$ has positive genus and two cusps. By [24, Proof of Theorem 10], it has a congruence subgroup Γ' with $\nu_\infty(\Gamma') \geq 3$, and Γ' contains all elliptic elements of $\Gamma_0(p)$.

Example 4.5. Assume that $\Gamma_{z_1}, \cdots, \Gamma_{z_n}$ generate a finite subgroup G and $|G| < \frac{1}{4}N^2 \prod\limits_{\ell | N}(1 - \ell^{-2})$, where the product being taken over all primes ℓ dividing N. By [26, Corollary 2.4], $X_{\widetilde{\Gamma}}$ has at least three cusps. Then $\widetilde{\Gamma}$ is such a subgroup of Γ.

4.2 Quantitative Riemann existence theorem for $X_{\widetilde{\Gamma}}$

The Riemann Existence Theorem asserts that every compact Riemann surface is (analytically isomorphic to) a complex algebraic curve. Bilu and Strambi [30, Theorem 1.2] gave a quantitative version of Riemann Existence Theorem, which is a key tool in this chapter.

Notice that the j-invariant induces naturally two coverings $X_\Gamma \to \mathbb{P}^1(\mathbb{C})$ and $X_{\widetilde{\Gamma}} \to \mathbb{P}^1(\mathbb{C})$, respectively. We use the same notation j to denote both of them without confusions. In addition, the j-invariant also defines an isomorphism $X(1) \cong \mathbb{P}^1(\mathbb{C})$.

For the covering $j : X_{\widetilde{\Gamma}} \to \mathbb{P}^1(\mathbb{C})$, we assume that its degree is $\tilde{n}$ and the genus of the curve $X_{\widetilde{\Gamma}}$ is $\tilde{g}$. Then there exists a rational function $y \in \bar{K}(X_{\widetilde{\Gamma}})$ such that $\bar{K}(X_{\widetilde{\Gamma}}) =$

$\bar{K}(j,y)$ and the rational functions $j,y \in \bar{K}(X_{\widetilde{\Gamma}})$ satisfy the equation $\tilde{f}(j,y)=0$, where $\tilde{f}(X,Y) \in \bar{K}[X,Y]$ is an absolutely irreducible polynomial satisfying

$$\deg_X \tilde{f} = \tilde{g}+1, \qquad \deg_Y \tilde{f} = \tilde{n}. \tag{4.2}$$

Consider the natural sequence of coverings $X(N) \to X_{\widetilde{\Gamma}} \to \mathbb{P}^1(\mathbb{C})$. Applying the formula in the bottom of [40, Page 101], we know that the degree of the covering $X(N) \to \mathbb{P}^1(\mathbb{C})$ is d_N. Combining with the genus formula of $X(N)$ (see [40, Figure 3.4]), we have

$$\tilde{n} \le d_N, \qquad \tilde{g} \le 1 + \frac{d_N(N-6)}{12N}. \tag{4.3}$$

4.3 Quantitative Chevalley-Weil theorem for $\phi : X_{\widetilde{\Gamma}} \to X_\Gamma$

The Chevalley-Weil theorem asserts that for an étale covering of projective varieties over a number field F, the discriminant of the field of definition of the fiber over an F-rational point is uniformly bounded. Bilu, Strambi and Surroca [31] got a fully explicit version of this theorem in dimension one, which is another key tool of this chapter.

For the covering $j : X_\Gamma \to \mathbb{P}^1(\mathbb{C})$, since there are only two elliptic points $\mathrm{SL}_2(\mathbb{Z})i$ and $\mathrm{SL}_2(\mathbb{Z})e^{2\pi i/3}$ of $X(1)$, it is unramified outside the two points $j(i) = 1728$ and $j(e^{2\pi i/3}) = 0$, and the point at infinity. For the covering $\phi : X_{\widetilde{\Gamma}} \to X_\Gamma$, it is unramified outside the cusps. Notice that the poles of j-invariant are exactly the cusps. Then by [31, Theorem 1.6], for every $P \in X_\Gamma(K)$ and $\tilde{P} \in X_{\widetilde{\Gamma}}(\bar{K})$ such that $\phi(\tilde{P}) = P$, we have

$$\mathcal{N}_{K/\mathbb{Q}}(D_{K(\tilde{P})/K}) \le e^{[K(\tilde{P}):\mathbb{Q}]\cdot(\mathrm{h}(S)+(1+\log 1728)\tilde{\Lambda})}, \tag{4.4}$$

where $D_{K(\tilde{P})/K}$ is the relative discriminant of $K(\tilde{P})/K$, and $\tilde{\Lambda} = ((\tilde{g}+1)\tilde{n})^{25(\tilde{g}+1)\tilde{n}}$. According to (4.3), we have $\tilde{\Lambda} \le \Lambda$. Hence

$$\mathcal{N}_{K/\mathbb{Q}}(D_{K(\tilde{P})/K}) \le e^{[K(\tilde{P}):\mathbb{Q}]\cdot(\mathrm{h}(S)+(1+\log 1728)\Lambda)}. \tag{4.5}$$

Notice that the degree $[K(\tilde{P}) : K] = [K(\tilde{P}) : K(P)]$, which is not greater than the degree of ϕ. So we have $[K(\tilde{P}) : K] \le d_N$.

4.4 Proof of Theorems

Under the assumptions of Theorems 4.1 and 4.2, the curve $X_{\widetilde{\Gamma}}$ has at least three cusps. In this section, we fix an S-integral point P on X_Γ and a point $\tilde{P}$ on $X_{\widetilde{\Gamma}}$ such that $\phi(\tilde{P}) = P$.

Let $K_0 = K(\tilde{P})$ and $d_0 = [K_0 : \mathbb{Q}]$. Let S_0 be the set consisting of the extensions of the places from S to K_0, that is,

$$S_0 = \{v \in M_{K_0} : v|w \in S\},$$

where M_{K_0} is the set of all valuations (or places) of K_0 extending the standard infinite and p-adic valuations of $\mathbb{Q}$. Put $s_0 = |S_0|$. We define the following quantity

$$\Delta_0 = d_0^{-d_0}\sqrt{N^{d_0 N}|D_0|^{\varphi(N)}}\left(\log(N^{d_0 N}|D_0|^{\varphi(N)})\right)^{d_0\varphi(N)}\left(\prod_{\substack{v\in S_0\\ v\nmid\infty}}\log\mathcal{N}_{K_0/\mathbb{Q}}(v)\right)^{\varphi(N)}, \tag{4.6}$$

where D_0 is the absolute discriminant of K_0.

Notice that $d_0 \le dd_N$ and $s_0 \le sd_N$. Let $D_{K_0/K}$ be the relative discriminant of K_0/K. By Formula (4.5), we have

$$\begin{aligned} D_0 &= \mathcal{N}_{K/\mathbb{Q}}(D_{K_0/K})D_K^{[K_0:K]} \\ &\le D^*. \end{aligned}$$

Now let w be a non-archimedean place of K, and $v_1, \cdots, v_m$ all its extensions to K_0, their residue degrees over K being $f_1, \cdots, f_m$ respectively. Then $f_1 + \cdots + f_m \le [K_0 : K] \le d_N$, which implies that $f_1 \cdots f_m \le 2^{d_N}$. Since $\mathcal{N}_{K_0/\mathbb{Q}}(v_k) = \mathcal{N}_{K/\mathbb{Q}}(w)^{f_k}$ for $1 \le k \le m$, we have

$$\prod_{v|w}\log\mathcal{N}_{K_0/\mathbb{Q}}(v) \le 2^{d_N}(\log\mathcal{N}_{K/\mathbb{Q}}(w))^{d_N}.$$

Hence

$$\prod_{\substack{v\in S_0\\ v\nmid\infty}}\log\mathcal{N}_{K_0/\mathbb{Q}}(v) \le 2^{sd_N}\left(\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v)\right)^{d_N}. \tag{4.7}$$

Combining with $d_0 \geq d$, we have

$$\Delta_0 \leq 2^{s\varphi(N)d_N}\Delta_1.$$

First we assume that N is not a power of any prime. By Theorem 2.2, we have

$$\mathrm{h}(\tilde{P}) \leq \left(Cd_0 s_0 N^2\right)^{2s_0 N} (\log(d_0 N))^{3s_0 N} p^{d_0 N} \Delta_0,$$

where C is an absolute effective constant. Note that $j(P) = j(\tilde{P})$, we have $\mathrm{h}(P) = \mathrm{h}(\tilde{P})$. Then we have

$$\mathrm{h}(P) \leq \left(Cdsd_N^2 N^2\right)^{2sNd_N} (\log(dNd_N))^{3sNd_N} p^{dNd_N} \Delta_1, \tag{4.8}$$

the constant C being modified. So we prove Theorem 4.1.

For the case that N is a prime power, applying Theorem 2.3, we can easily prove Theorem 4.2.

Part II

Heuristics of Pairing-friendly Elliptic Curves

Chapter 5

Introduction

5.1 Motivation

In 1985, Koblitz [59] and Miller [70] independently proposed elliptic curve cryptography. At the same time, Lenstra [64] succeeded in using elliptic curves for integer factorization. Afterwards, elliptic curves over finite fields and their cryptographic applications are intensively studied by both mathematicians and computer scientists. Currently, elliptic curve cryptography is one of the most popular practical public-key cryptographic schemes.

In recent years, mainly inspired by the following pioneering works: three-party one-round key agreement [56], identity-based encryption [33, 75], short signature scheme [34], easing the cryptographic applications of pairings [91] and efficient computation of pairings associated to elliptic curves [71], there has been a flurry of activity in the design and analysis of cryptographic protocols by using pairings on elliptic curves over finite fields. For example, the Tate pairing and the Weil pairing have been used to construct many novel cryptographic systems for which no other practical implementation is known. More in-depth studies of pairing-based cryptography can be found in the expository articles [52, 74].

The elliptic curves suitable for implementing pairing-based systems should have a small embedding degree with respect to a large prime-order subgroup, we call them *pairing-friendly elliptic curves*. More precisely, a pairing-friendly elliptic curve over a finite field $\mathbb{F}_q$ contains a subgroup of large prime order r such that for some k, $r|q^k - 1$ and $r \nmid q^i - 1$ for $0 < i < k$, and the parameters q, r and k should satisfy the following conditions:

- r should be large enough so that the Discrete Logarithm Problem (DLP) in an order-r subgroup of $E(\mathbb{F}_q)$ is infeasible.
- k should be sufficiently large so that DLP in $\mathbb{F}_{q^k}^*$ is intractable.
- k should be small enough so that arithmetic in $\mathbb{F}_{q^k}$ is feasible.

Here, k is called the *embedding degree* of E with respect to r, and the ratio $\frac{\log q}{\log r}$ called the *rho-value* of E with respect to r. There is a specific definition for pairing-friendly elliptic curve in [48, Definition 2.3], that is, it should meet $r \geq \sqrt{q}$ and $k \leq \log_2(r)/8$.

Balasubramanian and Koblitz [15] showed that in general the embedding degree k can be expected to be around r. Thus, the above conditions make pairing-friendly curves rare, and they can not be constructed by random generation. This naturally produces two important problems:

- Finding efficient constructions of pairing-friendly curves.
- Analyzing these constructions, including the frequency of curves constructed, efficiency, security level, etc.

The earliest constructions of pairing-friendly curves involved supersingular curves. However, on the one hand due to MOV attack [69], Frey-Rück reduction [49] and most recently [54], supersingular curves are widely believed to have some cryptographic weaknesses; on the other hand, for supersingular curves the embedding degree k has only 5 choices, i.e. $k \in \{1, 2, 3, 4, 6\}$. Thus, it seems quite important to construct ordinary curves with the above properties.

After consecutive efforts of many researchers, many methods for constructing ordinary curves are found. An exhaustive survey can be found in [48], furthermore the authors gave a coherent framework of all existing constructions. Unfortunately, none of these constructions has been rigorously analyzed. Even heuristic analysis is far from sufficiency except for the so-called *MNT curves* [72]. For the heuristic analysis of MNT curves, see [66, 90]. Most recently, a heuristic asymptotic formula for the number of isogeny classes of pairing-friendly curves over prime fields was presented in [35], some heuristic arguments about *Barreto-Naehrig family* [19] were also given therein.

It is widely accepted that the *Cocks-Pinch method* [38] is one of the most flexible algorithms for constructing pairing-friendly curves, such as with many curves possible, with arbitrary embedding degree, with prime-order subgroups of nearly arbitrary size, and so on. We will recall it in Chapter 7. The other general algorithm is the *Dupont-Enge-Morain method* [42].

In addition, *pairing-friendly fields* were introduced by Koblitz and Menezes [60] as an efficient way to implement cryptographic bilinear pairings. They define a field $\mathbb{F}_{p^k}$ as being pairing-friendly if the prime characteristic $p \equiv 1 \pmod{12}$ and the embedding degree $k = 2^i 3^j, i > 0$. If $j = 0$, it only needs $p \equiv 1 \pmod 4$. Definitely pairing-friendly curves over pairing-friendly fields are attractive.

5.2 Structure of Part II

Firstly, we continue the counting approach of [66, 67, 90] for pairing-friendly curves. We give a new heuristic upper bound for the number of isogeny classes of ordinary pairing-friendly curves, which seems to have slight improvement upon the previous bounds.

Secondly, we give two different kinds of heuristics to justify the same asymptotic formula about the Cocks-Pinch method, which confirms some of its general consensuses, such as many curves possible and with rho-value around 2. One is based on the prime ideal theorem, the other is based on the Bateman-Horn conjecture. Finally, we will see that the formula is compatible with numerical data.

Thirdly, we illustrate the first known heuristics about pairing-friendly curves over pairing-friendly fields. The heuristics suggest that any efficient construction of pairing-friendly curves is also an efficient construction of such curves over pairing-friendly fields, naturally including the Cocks-Pinch method. Especially, the heuristics will be confirmed by the numerical data from the Cocks-Pinch method.

This part is based on the manuscript [80].

5.3 Preliminary and Notation

Let Φ_k be the k-th cyclotomic polynomial. The existing constructions of ordinary curves with small embedding degree typically work in the following two steps.

1. Find an odd prime r, integers $k \geq 2$ and t, and a prime power q such that

$$|t| \leq 2\sqrt{q}, \quad \gcd(q,t) = 1, \quad r | q+1-t, \quad r | \Phi_k(q). \tag{5.1}$$

2. Construct an elliptic curve E over $\mathbb{F}_q$ with $|E(\mathbb{F}_q)| = q+1-t$.

Since $r|\Phi_k(q)$, k is the multiplicative order of q modulo r and then $k|r-1$. For satisfying the practical requirements, k should be reasonably small, while the rho-value should be as small as possible, preferably close to 1.

Unfortunately, the second step above is feasible only if $t^2 - 4q$ has a very small square-free part; that is, if the so-called *CM equation*

$$4q = t^2 + Du^2 \tag{5.2}$$

with some integers u and D, where D is a small square-free positive integer. In this case, for example $D \le 10^{13}$ (see [88]), E can be efficiently constructed via the *CM method* (see [4, Section 18.1]). Here, D is called the *CM discriminant* of E.

For the imaginary quadratic field $\mathbb{Q}(\sqrt{-D})$, let h_D be the class number of $\mathbb{Q}(\sqrt{-D})$ and w_D the number of roots of unity in $\mathbb{Q}(\sqrt{-D})$. We denote its discriminant by D^*. Then put $e(k,D) = 2$ if $D^*|k$ (namely $\mathbb{Q}(\sqrt{-D}) \subseteq \mathbb{Q}(\zeta_k)$), otherwise put $e(k,D) = 1$.

Recall that a well-known kind of constructions of pairing-friendly curves with k and D fixed is called the *complete polynomial family*, which is due to [18, 36, 72, 76]. Briefly speaking, the idea is to parameterize t, r, q, u as polynomials and then choose $t(x), r(x), q(x), u(x)$ satisfying Conditions (5.1) and (5.2) for any x. Here we define the ratio $\frac{\deg q(x)}{\deg r(x)}$ as the *rho-value* of the family. See [48, Section 2.1] for more details.

Throughout this part, we use the Landau symbols O and o and the Vinogradov symbol $\ll$. We recall that the assertions $U = O(V)$ and $U \ll V$ are both equivalent to the inequality $|U| \le cV$ with some constant c, while $U = o(V)$ means that $U/V \to 0$.

In this part, we also use the asymptotic notation $\sim$. Let f and g be two real functions with respect to x, both of them are strictly positive for sufficiently large x. We say that f is asymptotically equivalent to g if $f(x)/g(x) \to 1$ when $x \to \infty$, denoted by $f(x) \sim g(x)$.

Chapter 6

Upper bound for isogeny classes of ordinary pairing-friendly elliptic curves

In this chapter, we will obtain a new heuristic upper bound for isogeny classes of ordinary pairing-friendly elliptic curves, see Theorem 6.1.

For positive real numbers x, y and z, let $Q_k(x, y, z)$ be the number of prime powers $q \le x$ for which there exist a prime $r \ge y$ and an integer t satisfying Conditions (5.1) and (5.2) with some square-free positive integer $D \le z$. We also denote by $I_k(x, y, z)$ the number of pairs (q, t) of prime powers $q \le x$ and integers t such that Conditions (5.1) and (5.2) are satisfied with some prime $r \ge y$ and some square-free positive integer $D \le z$. That is, $I_k(x, y, z)$ is exactly the number of isogeny classes of the corresponding ordinary elliptic curves.

The function $Q_k(x, y, z)$ was first introduced in [66], The authors provided an upper bound for it therein and improved it in [67]. In [90], by introducing and bounding the function $I_k(x, y, z)$ the authors obtained a better bound for $Q_k(x, y, z)$, namely,

$$Q_k(x, y, z) \le I_k(x, y, z) \ll \varphi(k)(xy^{-1} + x^{1/2})z^{1/2} \frac{\log x}{\log \log x}, \tag{6.1}$$

where φ is the Euler's totient function.

We will see that the new upper bound presented here gives slight improvement upon the inequality (6.1) in the instance of main practical interest.

Estimate 6.1. *For any integer $k \geq 2$ and positive real numbers x, y and z, we heuristically have*

$$I_k(x, y, z) \ll \frac{\varphi(k)xy^{-1}z}{\log\log x}. \tag{6.2}$$

Proof. First fixing D, we want to bound the number of pairs (q, t) with $q \leq x$ and satisfying Condition (5.2). Here we borrow an idea from [35, Section 1]. For a given positive square-free integer D, we consider the element

$$\alpha = \frac{t + u\sqrt{-D}}{2}$$

of the imaginary quadratic field $\mathbb{Q}(\sqrt{-D})$. Since α is a root of $X^2 - tX + q$, α is an algebraic integer. If we denote by $\mathcal{N}(\cdot)$ the absolute norm of $\mathbb{Q}(\sqrt{-D})$, then $\mathcal{N}(\alpha) = q$. We also notice that $\gcd(t, q) = 1$ from Condition (5.1). Thus, the condition that q is a prime power is equivalent to the condition that α generates a prime ideal power of $\mathbb{Q}(\sqrt{-D})$. Denote by $\pi(x)$ the number of prime ideals of $\mathbb{Q}(\sqrt{-D})$ with norm bounded by x, the prime ideal theorem gives

$$\pi(x) \sim x/\log x.$$

Then the number of prime ideal powers of $\mathbb{Q}(\sqrt{-D})$ with norm bounded by x is bounded by

$$\sum_{n=1}^{\log x} x^{1/n}/\log(x^{1/n}) \ll x/\log x + x^{1/2}\log x \ll x/\log x.$$

Hence, for fixed D, the number of such pairs (q, t) is $O(\frac{x}{\log x})$.

For a given pair (q, t) with $q \leq x$, we need to estimate that probability that there exists a prime r satisfying Condition (5.1). Let $\omega(n)$ denote the number of prime divisors of an integer n, we know that

$$\omega(n) \ll \frac{\log n}{\log\log n},$$

see the proof of [66, Theorem 1]. So $\omega(q+1-t) \ll \frac{\log x}{\log\log x}$. For a prime r, the probability that $r|\Phi_k(q)$ is at most $\varphi(k)/r$.

It is well-known that there are $(6/\pi^2 + o(1))z$ positive square-free integers $D \leq z$ as $z \to \infty$, for example see [53, Theorem 334].

Therefore, we get

$$I_k(x, y, z) \ll \frac{x}{\log x} \cdot \frac{\log x}{\log\log x} \cdot \frac{\varphi(k)}{y} \cdot z = \frac{\varphi(k)xy^{-1}z}{\log\log x}.$$

□

Assume that $y \geq x^{1/2+o(1)}$ and $z = x^{o(1)}$, which is the most interesting case from the cryptographic point of view. Then (6.2) becomes

$$I_k(x, y, z) \ll x^{1/2+o(1)},$$

which can be compared with the number $x^{3/2+o(1)}$ of all possible isogeny classes (i.e. of pairs (q, t)) of elliptic curves over finite fields with $q \leq x$. Thus, one can not expect to generate suitable elliptic curves by random selection.

In particular, under the assumption $z = x^{o(1)}$, the bound in (6.2) is slightly better than that in (6.1). Recall that there is a heuristic lower bound of $I_k(x, y, z)$ under some assumptions in [90, Section 2.3], that is, for any fixed k and $\varepsilon > 0$, we have

$$I_k(x, y, z) \geq c(\varepsilon, k) x y^{-1+\varepsilon} z^{1/2},$$

where $c(\varepsilon, k)$ depends only on ε and k. Compared with (6.2), this lower bound is tight.

Noticing the trivial inequality $Q_k(x, y, z) \leq I_k(x, y, z)$, we get the following corollary.

Estimate 6.2. *For any integer $k \geq 2$ and positive real numbers x, y and z, we heuristically have*

$$Q_k(x, y, z) \ll \frac{\varphi(k) x y^{-1} z}{\log \log x}. \tag{6.3}$$

Chapter 7

Heuristics of the Cocks-Pinch method

7.1 Background on the Cocks-Pinch method

In an unpublished manuscript [38], Cocks and Pinch proposed an algorithm for constructing pairing-friendly curves with arbitrary embedding degree. More precisely, see [48, Theorem 4.1] or [52, Algorithm IX.4], fix an embedding degree k and a CM discriminant D, then execute the following steps:

Step 1. Choose a prime r such that $k|r-1$ and $-D$ is square modulo r.

Step 2. Choose an integer g which is a primitive k-th root of unity in $(\mathbb{Z}/r\mathbb{Z})^*$.

Step 3. Put $t' = g+1$ and choose an integer $u' \equiv (t'-2)/\sqrt{-D} \pmod r$.

Step 4. Let $t \in \mathbb{Z}$ be congruent to t' modulo r, and let $u \in \mathbb{Z}$ be congruent to u' modulo r. Put $q = (t^2 + Du^2)/4$.

Step 5. If q is an integer and prime, then there exists an elliptic curve E over $\mathbb{F}_q$ with an order-r subgroup and embedding degree k. If D is not to large, then E can be efficiently constructed via the CM method.

Notice that for any pairing triple (r,t,q), it satisfies the Cocks-Pinch method. In other words, when executing the method, it can generate (r,t,q). This can explain why the Cocks-Pinch method is highly important.

Given a real number $\rho > 0$, let $F_{k,D,\rho}(x)$ be the number of triples (r,t,q) constructed by the Cocks-Pinch method with fixed k and D such that q is an odd prime, $r \leq x$ and

$q \le r^\rho$. The previous paragraph implies that there is a natural one to one correspondence between the triples (r,t,q) here and the triples in [35, Estimate 1]. The reason we use the parameter q in the triples here is that we want to underline its importance.

In the sequel, first we will extend [35, Estimate 1] to all $\rho > 1$ for $F_{k,D,\rho}(x)$, for the sake of completeness. Then we will give another approach to this heuristic formula by applying the Bateman-Horn conjecture. In Chapter 9, we will see that this formula is compatible with numerical data.

7.2 Heuristics from algebraic number theory

As the above discussions, Boxall [35, Estimate 1] actually got a heuristic asymptotic formula for $F_{k,D,\rho}(x)$ when $1 < \rho < 2$. In this section, we will extend this formula to all $\rho > 1$ by applying the same techniques.

First, we need the following lemma, which can be gathered from [93, Chapter 2].

Lemma 7.1. *Let $k \ge 1$ be an integer and $r \nmid k$ a prime. Then the following statements are equivalent.*

1. *$\Phi_k(X)$ has a root modulo r.*
2. *$\Phi_k(X)$ can be factored into distinct linear factors modulo r.*
3. *r splits completely over the cyclotomic field $\mathbb{Q}(\zeta_k)$.*
4. *$k | r-1$.*

Estimate 7.2. *Given an integer $k \ge 3$, a positive square-free integer D and a real $\rho > 1$. Suppose that*

1. *$(k,D) \ne (3,3),(4,1)$ and $(6,3)$;*
2. *If there exists a complete family $(t(x), r(x), q(x))$ of pairing-friendly curves with rho-value 1, embedding degree k and CM discriminant D, then $\rho > 1 + \frac{1}{\deg r(x)}$.*

Then we have the following heuristic asymptotic formula

$$F_{k,D,\rho}(x) \sim \frac{e(k,D)w_D}{2\rho h_D} \int_5^x \frac{dz}{z^{2-\rho}(\log z)^2}. \tag{7.1}$$

Proof. We investigate the first four steps of the Cocks-Pinch method one by one.

Let $r \geq 2$ be any integer. The probability that r is prime is $1/\log r$, here we use the regular heuristic that the probability of a random integer n to be prime is $1/\log n$. Since k has finitely many prime factors, for an arbitrary prime r, the probability that $r \nmid k$ is 1. Notice that there are $\varphi(k)$ residue classes modulo k which consist of integers prime to k, the probability that r is prime and $k|r-1$ is $\frac{1}{\varphi(k)\log r}$.

Since $k|r-1$, r is completely splitting over $\mathbb{Q}(\zeta_k)$ by Lemma 7.1. Therefore, if $\mathbb{Q}(\sqrt{-D}) \subseteq \mathbb{Q}(\zeta_k)$, i.e. the discriminant of $\mathbb{Q}(\sqrt{-D})$ divides k, then r is completely splitting over $\mathbb{Q}(\sqrt{-D})$, thus $-D$ is square modulo r. Otherwise, if $\mathbb{Q}(\sqrt{-D}) \not\subseteq \mathbb{Q}(\zeta_k)$, the probability that $-D$ is square modulo r is $1/2$. So the probability that $-D$ is square modulo r is $e(k,D)/2$.

When r is fixed, the number of choices of g is $\varphi(k)$. After fixing g, t' is fixed and u' has two choices.

Thus, for an arbitrary integer $r \geq 2$, the probability that r satisfies Steps 1, 2 and 3 is $e(k,D)/\log r$. Moreover, since $k|r-1$ and $k \geq 3$, we have $r \geq k+1 \geq 4$. So $r \geq 5$. In the sequel, we investigate Step 4.

We consider the element

$$\alpha = \frac{t + u\sqrt{-D}}{2}$$

of $\mathbb{Q}(\sqrt{-D})$. We have known that α is an algebraic integer, $\mathcal{N}(\alpha) = q$ and $\mathcal{N}(\alpha - 1) = q + 1 - t$. So the condition that q is prime is equivalent to the condition that α generates a principal prime ideal of $\mathbb{Q}(\sqrt{-D})$ whose underlying prime number is not inert in $\mathbb{Q}(\sqrt{-D})$. By the prime ideal theorem for ideal classes, the number of principal prime ideals of $\mathbb{Q}(\sqrt{-D})$ with norm bounded by x is asymptotically equivalent to $\frac{x}{h_D \log x}$ as $x \to \infty$. Notice that the number of prime ideals of $\mathbb{Q}(\sqrt{-D})$ with norm bounded by x and underlying prime number inert is $O(\frac{\sqrt{x}}{\log\sqrt{x}})$ as $x \to \infty$. So the number of principal prime ideals of $\mathbb{Q}(\sqrt{-D})$ with norm bounded by x and underlying prime number not inert is asymptotically equivalent to $\frac{x}{h_D \log x}$ as $x \to \infty$. In the ring of integers of $\mathbb{Q}(\sqrt{-D})$, the units are exactly the roots of unity in $\mathbb{Q}(\sqrt{-D})$. For any such root of unity $\beta \neq 1$, $\alpha\beta$ and α generate the same ideal but $\alpha\beta \neq \alpha$. Note that $\pm u$ correspond to the same triple (r,t,q). Here we also notice that if t' and u' are fixed, then the residue classes modulo r which t and u belong to are fixed. Thus, the expected number of pairs (t,q) associated to a triple (r,t',u') with $q \leq r^\rho$ is asymptotically equivalent to $\frac{w_D r^\rho}{2\rho h_D r^2 \log r}$ as $r \to \infty$.

Therefore, we have

$$\begin{aligned} F_{k,D,\rho}(x) &\sim \sum_{k+1 \le r \le x} \frac{e(k,D)}{\log r} \cdot \frac{w_D r^{\rho}}{2\rho h_D r^2 \log r} \\ &\sim \frac{e(k,D) w_D}{2\rho h_D} \int_{k+1}^{x} \frac{dz}{z^{2-\rho} (\log z)^2}. \end{aligned} \tag{7.2}$$

Notice that the above integral tends to infinity as $x \to \infty$. For uniformity, we can take

$$F_{k,D,\rho}(x) \sim \frac{e(k,D) w_D}{2\rho h_D} \int_{5}^{x} \frac{dz}{z^{2-\rho} (\log z)^2}.$$

$\square$

As explained in [35], without the two assumptions in Estimate 7.2, the asymptotic formula may not hold any more. In particular, if there exists a complete polynomial family with rho-value 1, embedding degree k and CM discriminant D, then this family can generate more triples than predicted by (7.1). For example, the Barreto-Naehrig family is currently the only known complete polynomial family with rho-value 1, for this family $k = 12$, $D = 3$ and $\deg r(x) = 4$, see Table 9.7 for numerical data.

Now we want to say more about the parameters in (7.1). It is well-known that w_D is given by the following formula:

$$w_D = \begin{cases} 4 & \text{if } D = 1, \\ 6 & \text{if } D = 3, \\ 2 & \text{if } D = 2 \text{ or } D > 3. \end{cases}$$

Furthermore, by the well-known Dirichlet's class number formula of imaginary quadratic fields (for example see [45, Exercise 10.5.12]), we know

$$h_D = \begin{cases} \sqrt{D} w_D L_D / \pi & \text{if } D \equiv 1, 2 \pmod 4, \\ \sqrt{D} w_D L_D / (2\pi) & \text{if } D \equiv 3 \pmod 4, \end{cases} \tag{7.3}$$

where $L_D = \sum\limits_{n=1}^{\infty} \left(\frac{D^*}{n}\right)/n = \prod\limits_{\text{prime } p} \left(1 - \left(\frac{D^*}{p}\right)/p\right)^{-1}$, D^* is the discriminant of $\mathbb{Q}(\sqrt{-D})$ and $\left(\frac{\cdot}{\cdot}\right)$ is the Kronecker symbol.

Based on the following lemma, we can get another version of the above proposition, that is,

$$F_{k,D,\rho}(x) \sim \frac{e(k,D) w_D}{2\rho(\rho - 1) h_D} \frac{x^{\rho-1}}{(\log x)^2}, \tag{7.4}$$

see also [35, Formula (0.1)]. We are sure that the lemma is well-known. It is more convenient to give a simple proof rather than find some references. We will use it later.

Lemma 7.3. *For any real numbers a, m, s with $a > 1$ and $s < 1$, we have*

$$\int_a^x \frac{dz}{z^s(\log z)^m} \sim \frac{x^{1-s}}{(1-s)(\log x)^m}.$$

Proof. Integrating by parts, we obtain

$$\int_a^x \frac{dz}{z^s(\log z)^m} = \frac{z^{1-s}}{(1-s)(\log z)^m}\Big|_a^x + \frac{m}{1-s}\int_a^x \frac{dz}{z^s(\log z)^{m+1}},$$

and

$$\int_a^x \frac{dz}{z^s(\log z)^{m+1}} = \frac{z^{1-s}}{(1-s)(\log z)^{m+1}}\Big|_a^x + \frac{m+1}{1-s}\int_a^x \frac{dz}{z^s(\log z)^{m+2}}.$$

We choose a positive real number A such that $A > a$ and $\log A > \frac{m+1}{1-s}$. Notice that for $x > A$, we have

$$\int_a^x \frac{dz}{z^s(\log z)^{m+2}} \leq \int_a^A \frac{dz}{z^s(\log z)^{m+2}} + \frac{1}{\log A}\int_A^x \frac{dz}{z^s(\log z)^{m+1}}.$$

Then we get

$$\int_a^x \frac{dz}{z^s(\log z)^{m+1}} \ll \frac{x^{1-s}}{(1-s)(\log x)^{m+1}}.$$

Finally, we have

$$\int_a^x \frac{dz}{z^s(\log z)^m} \sim \frac{x^{1-s}}{(1-s)(\log x)^m}.$$

□

It is widely accepted that the rho-value of curves produced by the Cocks-Pinch method tends to be around 2. From (7.4) we can easily see that when ρ is close to 1, the curves with relevant rho-value are rare among the whole family constructed by the Cocks-Pinch method.

As explained in [35], when $1 < \rho < 2$, Estimate 7.2 also predicts a heuristic asymptotic estimate for the number of isogeny classes of elliptic curves with given $k \geq 3$ and D defined over primes fields $\mathbb{F}_q$ and possessing a subgroup of prime order $r \leq x$ such that $q \leq r^\rho$. In addition, for this number, by applying the same arguments as the first two paragraphs of the proof of Estimate 6.1, we can get the following upper bound

$$\frac{c\varphi(k)x^{\rho-1}}{\log\log x},$$

where c is an absolute constant. By (7.4), it is easy to see that these two results are compatible.

7.3 Heuristics from the Bateman-Horn conjecture

The Bateman-Horn conjecture has been used to analyze some constructions of pairing-friendly elliptic curves, see [35, 90]. In this section, applying the Bateman-Horn conjecture we will give another approach to justify the heuristic asymptotic formula of $F_{k,D,\rho}(x)$ in Estimate 7.2.

The Bateman-Horn conjecture provides a conjectured density for the positive integers at which a given system of polynomials all have prime values, see [20]. We recall it here for the conveniences of readers.

Given any finite set $\mathcal{F} = \{f_1, f_2, \cdots, f_m\}$ consisting of irreducible polynomials $f_1(T), \cdots, f_m(T) \in \mathbb{Z}[T]$ with positive leading coefficients and such that there is no prime p with $p|f_1(n)\cdots f_m(n)$ for every integer $n \geq 1$, the Bateman-Horn conjecture says

$$|\{1 \leq n \leq X : f_1(n), \cdots, f_m(n) \text{ are all prime}\}| \sim \frac{C(\mathcal{F})}{\deg f_1 \cdots \deg f_m} \int_2^X \frac{dz}{(\log z)^m}, \quad (7.5)$$

where $C(\mathcal{F})$ is given by the conditionally convergent infinite product

$$C(\mathcal{F}) = \prod_{p \text{ prime}} \frac{1 - \omega_p(\mathcal{F})/p}{(1-1/p)^m},$$

and

$$\omega_p(\mathcal{F}) = |\{1 \leq n \leq p : f_1(n)\cdots f_m(n) \equiv 0 \pmod p\}|.$$

Based on Lemma 7.3, we can get another version of the Bateman-Horn conjecture, that is,

$$|\{1 \leq n \leq X : f_1(n), \cdots, f_m(n) \text{ are all prime}\}| \sim \frac{C(\mathcal{F})}{\deg f_1 \cdots \deg f_m} \frac{X}{(\log X)^m}, \quad (7.6)$$

which we will use in the sequel.

Notice that the ring of integer of $\mathbb{Q}(\sqrt{-D})$ is $\mathbb{Z} \oplus \mathbb{Z}\frac{1+\sqrt{-D}}{2}$ if $D \equiv 3 \pmod 4$, and it is $\mathbb{Z} \oplus \mathbb{Z}\sqrt{-D}$ if $D \equiv 1$ or $2 \pmod 4$. Since $\alpha = \frac{t+u\sqrt{-D}}{2}$ should be an algebraic integer of $\mathbb{Q}(\sqrt{-D})$, t and u must have the same parity if $D \equiv 3 \pmod 4$, and otherwise both of them must be even.

Estimate 7.4. *For any integer $k \geq 3$, and positive square-free integer $D \equiv 1, 2 \pmod 4$, under the same assumptions as Estimate 7.2, we heuristically have*

$$F_{k,D,\rho}(x) \sim \frac{e(k,D)w_D}{2\rho h_D} \int_5^x \frac{dz}{z^{2-\rho}(\log z)^2}. \quad (7.7)$$

Proof. As the proof of Estimate 7.2, for an arbitrary integer $r \geq 2$, the probability that r satisfies Steps 1, 2 and 3 is $e(k, D)/\log r$. Moreover, it also needs that $r \geq 5$. In the sequel, we investigate Step 4.

Since $D \equiv 1, 2 \pmod 4$, t and u must be even. So it is equivalent to count the number of integer pairs (t, u) such that $q = t^2 + Du^2$ is prime with $q \leq r^\rho$. Then for the integers t and u, we have $|t| \leq \sqrt{r^\rho}$ and $|u| \leq \sqrt{r^\rho/D}$. Notice that the ratio between the area of the ellipse $t^2 + Du^2 = r^\rho$ and that of the rectangle $\{(t, u) : |t| \leq \sqrt{r^\rho}, |u| \leq \sqrt{r^\rho/D}\}$ is $\pi/4$. Now we first count the number of (t, q) with $q = t^2 + Du^2$ prime, $t \leq \sqrt{r^\rho}$ and $u \leq \sqrt{r^\rho/D}$, and then to get the final result we need to multiply this amount by $\pi/4$.

For every positive integer $u \leq \sqrt{r^\rho/D}$, let $f_u(T) = T^2 + Du^2 \in \mathbb{Z}[T]$. For $\mathcal{F} = \{f_u\}$, it satisfies the required conditions. By the Bateman-Horn conjecture, we have

$$|\{1 \leq t \leq \sqrt{r^\rho} : f_u(t) \text{ is prime}\}| \sim \frac{C(f_u)\sqrt{r^\rho}}{\rho \log r},$$

where

$$C(f_u) = \prod_{p \text{ prime}} \frac{1 - \omega_p(f_u)/p}{1 - 1/p},$$

and

$$\omega_p(f_u) = |\{1 \leq n \leq p : n^2 \equiv -Du^2 \pmod p\}|.$$

It is easy to see that

$$\omega_p(f_u) = \begin{cases} 1 & \text{if } p = 2 \text{ or } p|u, \\ \left(\frac{-D}{p}\right) + 1 & \text{if } p \geq 3 \text{ and } p \nmid u. \end{cases}$$

Put

$$g(u) = \prod_{p \geq 3,\, p|u} \frac{p-1}{p - 1 - \left(\frac{-D}{p}\right)}.$$

We also set $g(2^n) = 1$ for any integer $n \geq 0$, this makes $g(u)$ a multiplicative function. Notice that

$$C(f_1) = C(f_2) = \prod_{\text{prime } p \geq 3} \frac{p - 1 - \left(\frac{-D}{p}\right)}{p-1}.$$

Obviously, $C(f_u) = C(f_1) \cdot g(u)$. Then we have

$$\sum_{1 \leq u \leq \sqrt{r^\rho/D}} \frac{C(f_u)\sqrt{r^\rho}}{\rho \log r} = \frac{C(f_1)\sqrt{r^\rho}}{\rho \log r} \sum_{1 \leq u \leq \sqrt{r^\rho/D}} g(u).$$

Here we need an asymptotic formula for

$$S(X) = \sum_{1 \le u \le X} g(u).$$

Notice that $g(u)$ is a multiplicative function and $1 - 1/p \le g(p) \le 1 + \frac{3}{p}$ for any prime p. Recall the Mertens' second theorem

$$\sum_{\text{prime } p \le X} \frac{1}{p} = \log\log X + B_1 + o(1),$$

where B_1 is an absolute constant, see [53, Theorem 427]. Then we get

$$\sum_{\text{prime } p \le X} g(p) = \frac{X}{\log X} + O(\log\log X).$$

Then by [47, Proposition 4] which concerns the sum of multiplicative functions, we have

$$S(X) = (C_g + o(1))X,$$

where $C_g = \prod_{\text{prime } p} (1 + \frac{g(p)}{p} + \frac{g(p^2)}{p^2} + \cdots)(1 - \frac{1}{p})$.

Notice that $g(p^n) = g(p)$ for any prime p and any $n \ge 1$. Then we have

$$C_g = \prod_{\text{prime } p \ge 3} \frac{p-1}{p} \left(1 + \frac{1}{p - 1 - \left(\frac{-D}{p}\right)}\right),$$

and thus

$$C(f_1)C_g = \prod_{\text{prime } p \ge 3} \left(1 - \left(\frac{-D}{p}\right)/p\right) = L_D^{-1},$$

where L_D has been defined in (7.3). Hence

$$\begin{aligned}\sum_{1 \le u \le \sqrt{r^\rho/D}} \frac{C(f_u)\sqrt{r^\rho}}{\rho \log r} &= (L_D^{-1} + o(1)) \frac{r^\rho}{\sqrt{D}\rho \log r} \\ &\sim \frac{r^\rho}{\rho L_D \sqrt{D} \log r} = \frac{w_D r^\rho}{\pi \rho h_D \log r}.\end{aligned}$$

Note that t can be taken negative integer. We also note that if t' and u' are fixed, then the residue classes modulo r which t and u belong to are also fixed. So the expected number of pairs (t, q) associated to a triple (r, t', u') with $q \le r^\rho$ is asymptotically equivalent to

$$\frac{\pi}{4} \cdot \frac{w_D r^\rho}{\pi \rho h_D \log r} \cdot 2 \cdot \frac{1}{r^2} = \frac{w_D}{2\rho h_D r^{2-\rho} \log r},$$

as $r \to \infty$.

Therefore, we have

$$\begin{aligned} F_{k,D,\rho}(x) &\sim \sum_{5 \le r \le x} \frac{e(k,D)}{\log r} \cdot \frac{w_D}{2\rho h_D r^{2-\rho} \log r} \\ &\sim \frac{e(k,D) w_D}{2\rho h_D} \int_5^x \frac{dz}{z^{2-\rho} (\log z)^2}. \end{aligned}$$

□

For the Cocks-Pinch method, it is lucky that we can apply two different kinds of heuristics. But in general, the Bateman-Horn conjecture is indispensable when investigating the constructions of pairing-friendly curves. Estimate 7.4 tells us that such investigations based on the Bateman-Horn conjecture are likely to be reasonable.

Chapter 8

Involving Pairing-friendly Fields

In this chapter, we want to heuristically count the number of triples (r, t, q) such that $q \equiv 1 \pmod{4 \text{ or } 12}$ constructed by the Cocks-Pinch method.

Similar as before, let $G_{k,D,\rho}(x)$ be the number of triples (r, t, q) constructed by the Cocks-Pinch method with fixed k and D such that q is an odd prime, $q \equiv 1 \pmod 4$, $r \leq x$ and $q \leq r^\rho$. If furthermore requiring $q \equiv 1 \pmod{12}$, let $H_{k,D,\rho}(x)$ be the number of such triples (r, t, q).

From the CM equation: $q = \frac{t^2+Du^2}{4}$, it is easy to see that $q \equiv 1 \pmod{12}$ if and only if $q \equiv 1 \pmod 4$ and $t^2 + Du^2 \equiv 1 \pmod 3$.

First, we study the probability that $t^2 + Du^2 \equiv 1 \pmod 3$.

Proposition 8.1. *If* $3|D$, *then we always have* $t^2 + Du^2 \equiv 1 \pmod 3$.

Proof. Since $3|D$, $t^2 + Du^2 \equiv t^2 \equiv 1 \pmod 3$ holds only if $3 \nmid t$. Assume that $3|t$. Then we have $3|q$, thus $q = 3$. Then $t = 0, D = 3$ and $u = \pm 2$. Since $r|q + 1 \pm t$ and $r \geq 5$, there is no possible r. So we must have $3 \nmid t$, and thus we always have $t^2 + Du^2 \equiv 1 \pmod 3$. □

Corollary 8.2. *If* $3|D$, *then we always have* $G_{k,D,\rho}(x) = H_{k,D,\rho}(x)$.

Proposition 8.3. *No matter* $D \equiv 1$ *or* $2 \pmod 3$, *the formula* $t^2 + Du^2 \equiv 1 \pmod 3$ *is true with the probability of* $1/2$, *under some assumptions.*

Proof. Suppose that $D \equiv 1 \pmod 3$. Then $t^2 + Du^2 \equiv t^2 + u^2 \equiv 1 \pmod 3$ holds only if 3 exactly divides one of t and u. For pairs (t, u), they can be divided into nine classes according to the residue classes modulo 3 which t and u belong to. Notice that

3 does not divide t and u at the same time. So there are only eight classes which can appear. Assume that all the eight classes have the same probability. Thus the desired probability is 1/2.

Suppose that $D \equiv 2 \pmod 3$. Then $t^2 + Du^2 \equiv t^2 + 2u^2 \equiv 1 \pmod 3$ holds only if $3 \nmid t$ and $3|u$. Assume that $3 \nmid t$ and $3 \nmid u$. Then $3|t^2 + Du^2$. Since q is a prime, $q = 3$, which contradicts $q \equiv 1 \pmod 4$. Thus it is impossible. So 3 must exactly divide one of t and u, which is naturally divided into two cases. Suppose that these two cases have the same probability. Then the desired probability is 1/2. □

Proposition 8.4. *Assume that $k \geq 3$ and $D \equiv 1 \pmod 4$. Then the followings hold.*

(1) $G_{k,D,\rho}(x) = F_{k,D,\rho}(x)$.

(2) *If furthermore $D \equiv 0 \pmod 3$, we have $H_{k,D,\rho}(x) = F_{k,D,\rho}(x)$.*

Proof. (1) Since $D \equiv 1 \pmod 4$, for a constructed prime $q = \frac{t^2+Du^2}{4}$, t and u must be even. Notice that since D and q are odd, $\frac{t}{2}$ and $\frac{u}{2}$ must have different parities. Thus it is always true that $q \equiv 1 \pmod 4$. So we prove (1).

(2) Since $q \equiv 1 \pmod 4$, we know that $q \equiv 1 \pmod{12}$ if and only if $t^2 + Du^2 \equiv 1 \pmod 3$. Then (2) follows from Proposition 8.1. □

Estimate 8.5. *Assume that $k \geq 3$, $D \equiv 1 \pmod 4$ and $D \equiv 1, 2 \pmod 3$. we heuristically have $H_{k,D,\rho}(x) \sim \frac{1}{2}F_{k,D,\rho}(x)$.*

Proof. Since $D \equiv 1 \pmod 4$, we have $q \equiv 1 \pmod 4$. So, $q \equiv 1 \pmod{12}$ if and only if $t^2 + Du^2 \equiv 1 \pmod 3$. Then the desired result follows from Proposition 8.3. □

For the case $D \equiv 2, 3 \pmod 4$, the heuristics are also straightforward.

Estimate 8.6. *Assume that $k \geq 3$ and $D \equiv 2, 3$ (mod 4). Then the followings hold heuristically.*

(1) $G_{k,D,\rho}(x) \sim \frac{1}{2}F_{k,D,\rho}(x)$.

(2) *If furthermore $D \equiv 0$ (mod 3), we have $H_{k,D,\rho}(x) \sim \frac{1}{2}F_{k,D,\rho}(x)$.*

(3) *If furthermore $D \equiv 1, 2$ (mod 3), we have $H_{k,D,\rho}(x) \sim \frac{1}{4}F_{k,D,\rho}(x)$.*

Proof. We divide the proof into three parts according to three cases.

(I) Assume that $D \equiv 2 \pmod 4$.

(1) Since $D \equiv 2 \pmod 4$, for a constructed prime $q = \frac{t^2+Du^2}{4}$, t and u must be even. Notice that since D is even and q is odd, $\frac{t}{2}$ must be odd. Then $(\frac{t}{2})^2 + D(\frac{u}{2})^2 \equiv 1 \pmod 4$ holds only if $\frac{u}{2}$ is even. Suppose that the even parity and odd parity of $\frac{u}{2}$ have the same probability. Then the probability that $q \equiv 1 \pmod 4$ is 1/2, which proves (1).

(2) and (3) By Propositions 8.1 and 8.3, under the same assumptions, the probability that $t^2 + Du^2 \equiv 1 \pmod 3$ is $1, 1/2$ or $1/2$ corresponding to $D \equiv 0, 1$ or $2 \pmod 3$, respectively. Suppose that the two events $q \equiv 1 \pmod 4$ and $t^2 + Du^2 \equiv 1 \pmod 3$ are independent. Then we can get the desired results.

(II) Assume that $D \equiv 7, 15 \pmod{16}$.

(1) Since $D \equiv 3 \pmod 4$, for a constructed prime $q = \frac{t^2+Du^2}{4}$, t and u must have the same parity. Furthermore, since $D \equiv 7, 15 \pmod{16}$, we claim that t and u must be even.

Suppose that t and u are odd. Consider the CM equation $4q = t^2 + Du^2$. Since q is odd, $4q$ is equal to 4 or 12 modulo 16. But $t^2 + Du^2$ is equal to 0 or 8 modulo 16 under the condition $D \equiv 7, 15 \pmod{16}$. This leads to a contradiction.

Since D and q are odd, $\frac{t}{2}$ and $\frac{u}{2}$ must have different parities, which is naturally divided into two cases. Suppose that these two cases have the same probability. Then the probability that $(\frac{t}{2})^2 + D(\frac{u}{2})^2 \equiv 1 \pmod 4$ is 1/2, which proves (1).

(2) and (3) Apply the same arguments as (I).

(III) Assume that $D \equiv 3, 11 \pmod{16}$.

(1) Since $D \equiv 3 \pmod 4$, for a constructed prime $q = \frac{t^2+Du^2}{4}$, t and u must have the same parity. Furthermore, the two parities may occur due to $D \equiv 3, 11 \pmod{16}$.

First suppose that both of t and u are even. The deduction and the result of this case are the same as (II).

Now suppose that both of t and u are odd. Notice that when n is an odd integer, then $n^2 \equiv 1, 9 \pmod{16}$. In this case, pairs (t^2, u^2) can be divided into four classes according to the residue classes modulo 16 which t^2 and u^2 belong to. Suppose that all the four classes have the same probability. Then, when $D \equiv 3, 11 \pmod{16}$, the probability that $t^2 + Du^2 \equiv 4 \pmod{16}$ is 1/2.

Notice that we obtain the same result for the two parities, then the probability that $q \equiv 1 \pmod 4$ is 1/2. So we prove (1).

(2) and (3) Apply the same arguments as (I).

□

From the above results, the heuristics suggest that pairing-friendly curves over pairing-friendly fields can be efficiently constructed by the Cocks-Pinch method. Notice that there are 18 cases in the above proofs according to D modulo 4 or 16 and D modulo 3. In the next chapter, we will see that the heuristic results of this section are compatible with numerical data.

Remark 8.7. Notice that the above heuristics are independent of the Cocks-Pinch method, they can be applied to any other constructions. So we can say that any efficient construction of pairing-friendly curves is also an efficient construction of pairing-friendly curves over pairing-friendly fields.

Chapter 9

Numerical Evidence

For testing Estimate 7.2 and the heuristic results in Chapter 8, we write a programme in PARI/GP [89] to executive the Cocks-Pinch method for searching all the triples (r, t, q) with k, D and ρ being given, and r in some interval $[a, b]$.

For given k, D, ρ, a and b, we denote by $N_1(k, D, \rho, a, b)$ the number of triples (r, t, q) as in Estimate 7.2 with $a \leq r \leq b$. If furthermore requiring $q \equiv 1 \pmod 4$ (resp. $q \equiv 1 \pmod{12}$), we denote the number of such triples by $N_2(k, D, \rho, a, b)$ (resp. $N_3(k, D, \rho, a, b)$). The outputs of the programme are these three quantities.

For $N_1(k, D, \rho, a, b)$, under some assumptions, there exists a heuristic formula from Estimate 7.2, stated as follows

$$I(k, D, \rho, a, b) = \frac{e(k, D)w_D}{2\rho h_D} \int_a^b \frac{dz}{z^{2-\rho}(\log z)^2}. \tag{9.1}$$

Let $I_0 = e(k, D)^{-1} I(k, D, \rho, a, b)$. Then I_0 depends only on D and ρ but not on k.

In Chapter 8, we present some definite or heuristic results about the relations among $N_i(k, D, \rho, a, b)$, $i = 1, 2, 3$. We list them as follows,

$$\begin{cases} N_2(k, D, \rho, a, b) = N_1(k, D, \rho, a, b) & \text{if } D \equiv 1 \pmod 4, \\ N_2(k, D, \rho, a, b) \approx \frac{1}{2} N_1(k, D, \rho, a, b) & \text{if } D \equiv 2, 3 \pmod 4; \end{cases} \tag{9.2}$$

$$\begin{cases} N_3(k, D, \rho, a, b) = N_1(k, D, \rho, a, b) & \text{if } D \equiv 1 \pmod 4 \text{ and } D \equiv 0 \pmod 3, \\ N_3(k, D, \rho, a, b) \approx \frac{1}{2} N_1(k, D, \rho, a, b) & \text{if } D \equiv 1 \pmod 4 \text{ and } D \equiv 1, 2 \pmod 3, \\ N_3(k, D, \rho, a, b) \approx \frac{1}{2} N_1(k, D, \rho, a, b) & \text{if } D \equiv 2, 3 \pmod 4 \text{ and } D \equiv 0 \pmod 3, \\ N_3(k, D, \rho, a, b) \approx \frac{1}{4} N_1(k, D, \rho, a, b) & \text{if } D \equiv 2, 3 \pmod 4 \text{ and } D \equiv 1, 2 \pmod 3; \end{cases} \tag{9.3}$$

$$N_2(k, D, \rho, a, b) = N_3(k, D, \rho, a, b), \quad \text{if } D \equiv 0 \pmod 3. \tag{9.4}$$

In this chapter, we will test all these results by numerical data.

In fact, [35, Table 1 and Table 2] gave the values of $N_1(k, D, 1.7, 10^6, 85\ 698\ 768)$ and $N_1(k, D, 1.5, 10^6, 2 \times 10^8)$ respectively, for $3 \le k \le 30$ and all square-free integer D with $D \le 15$. These two tables are compatible with (9.1). In the sequel, we will choose more narrow interval $[a, b]$ and even choose $a = 5$ for testing.

Here, for each entry in the following tables, if its actual value is not an integer, then it is rounded to the nearest whole number.

Table 9.1 gives the values of $N_1(k, D, 1.8, 5, 5 \times 10^5)$ for all k with $3 \le k \le 18$ and various square-free D. Notice that in Chapter 8 there are 18 cases according to D modulo 4 (or 16) and D modulo 3. The choices of D here exactly cover all these cases. The second line gives the value of I_0. The main part of the table contains the values of $N_1(k, D, 1.8, 5, 5 \times 10^5)$, the entries corresponding to values of (k, D) with $e(k, D) = 2$ are highlighted in bold; (9.1) predicts that they should be close to $2I_0$ and thus roughly twice as large as the other entries in the same column. The entries corresponding to values of $(k, D) = (3, 3), (4, 1)$ and $(6, 3)$ are left blank. The last line gives the average value of each column as k varies from 3 to 18, the cases where $e(k, D) = 2$ being counted with weight $\frac{1}{2}$ and the excluded values $(k, D) = (3, 3), (4, 1)$ and $(6, 3)$ omitted. (9.1) predicts that each of these averages should be close to I_0.

Table 9.2 gives the values of $N_2(k, D, 1.8, 5, 5 \times 10^5)$ for the same values of (k, D) as Table 9.1. When $D \equiv 1 \pmod 4$, (9.2) tells us that $N_2(k, D, 1.8, 5, 5 \times 10^5) = N_1(k, D, 1.8, 5, 5 \times 10^5)$ for each value of (k, D). Otherwise, when $D \equiv 2, 3 \pmod 4$, (9.2) predicts that $N_2(k, D, 1.8, 5, 5 \times 10^5)$ should be close to half of $N_1(k, D, 1.8, 5, 5 \times 10^5)$.

Table 9.3 gives the values of $N_3(k, D, 1.8, 5, 5 \times 10^5)$ for the same values of (k, D) as Table 9.1. (9.3) presents some definite or heuristic results about the relation between $N_3(k, D, 1.8, 5, 5 \times 10^5)$ and $N_1(k, D, 1.8, 5, 5 \times 10^5)$. For example, when $D \equiv 1 \pmod 4$ and $D \equiv 0 \pmod 3$, we have $N_3(k, D, 1.8, 5, 5 \times 10^5) = N_1(k, D, 1.8, 5, 5 \times 10^5)$. If $3 | D$, (9.4) says that $N_2(k, D, 1.8, 5, 5 \times 10^5) = N_3(k, D, 1.8, 5, 5 \times 10^5)$.

The explanations of Tables 9.4, 9.5 and 9.6 are the same as Tables 9.1, 9.2 and 9.3, respectively. Here, we choose another choices of D to exactly cover the 18 cases in Chapter 8.

Although Tables 9.1–9.6 show that (9.2)–(9.4) are supported by numerical data, there is some discrepancy between the expected values and the calculated values. For Tables 9.1 and 9.4, this is expected. Because for the Bateman-Horn conjecture, there seems to be no good conjecture for the remainder, for example see [61] for a discussion of the case of prime pairs. Thus, it may be also a hard problem to find one in the

TABLE 9.1: Values of $N_1(k, D, 1.8, 5, 5 \times 10^5)$ for various k and D

D	1	2	3	5	6	7	10	11	15	19	21	23	31	35	39	43	47	123
I_0	377	189	566	94	94	189	94	189	94	189	47	63	63	94	47	189	38	94
$k = 3$	403	184		101	89	174	85	196	88	222	44	75	62	105	43	198	42	94
4		174	583	112	107	221	97	211	87	196	58	49	68	101	49	203	32	126
5	429	217	570	105	96	218	101	184	92	213	48	60	63	100	53	212	37	103
6	388	193		95	105	199	109	180	88	182	52	57	62	107	60	206	44	116
7	420	193	627	96	92	**374**	94	195	104	202	42	75	74	88	44	218	34	109
8	**802**	**365**	592	130	85	172	88	200	103	200	57	71	54	89	51	176	44	111
9	371	182	**1190**	93	117	188	105	215	92	194	53	74	64	100	40	183	38	99
10	409	189	592	107	95	206	92	197	109	199	46	65	55	83	33	231	32	94
11	371	179	589	95	91	178	105	**395**	86	186	53	60	59	98	43	182	41	94
12	**846**	182	**1230**	85	87	206	101	181	85	189	50	57	69	91	49	197	28	96
13	380	197	622	99	79	180	102	200	89	206	47	60	61	93	40	172	35	106
14	413	190	582	78	83	**423**	99	197	89	202	55	68	57	94	49	217	29	97
15	405	184	**1167**	93	109	187	89	185	**173**	208	44	54	74	100	50	178	51	106
16	**800**	**386**	609	101	95	175	84	201	84	201	48	55	74	81	43	201	52	96
17	358	202	579	98	103	193	103	202	100	227	49	72	69	88	40	208	52	114
18	397	201	**1203**	87	91	195	100	209	90	195	54	55	79	106	51	190	43	91
Avg	398	190	596	98	95	193	97	197	97	201	50	63	65	95	46	198	40	103

TABLE 9.2: Values of $N_2(k, D, 1.8, 5, 5 \times 10^5)$ for various k and D

D	1	2	3	5	6	7	10	11	15	19	21	23	31	35	39	43	47	123
$k = 3$	403	84		101	52	84	34	101	48	109	44	38	28	46	25	93	18	41
4		83	305	112	50	96	38	111	43	99	58	22	27	51	26	105	15	59
5	429	118	290	105	42	107	55	95	43	97	48	31	33	48	22	86	19	57
6	388	104		95	62	103	48	89	45	97	52	28	35	50	26	94	20	64
7	420	95	304	96	49	**203**	40	94	49	96	42	34	29	47	23	97	12	56
8	**802**	**186**	297	130	42	87	40	84	57	101	57	33	27	52	30	83	17	57
9	371	86	**603**	93	60	90	47	109	54	109	53	38	32	41	23	100	15	59
10	409	105	289	107	45	103	45	103	49	96	46	34	24	48	19	120	20	50
11	371	99	260	95	44	89	47	**184**	43	102	53	31	31	53	21	92	24	41
12	**846**	91	**623**	85	36	81	56	90	39	109	50	30	37	48	24	96	14	53
13	380	100	312	99	32	96	49	110	56	102	47	30	23	46	17	80	11	54
14	413	92	271	78	47	**215**	52	104	49	110	55	38	35	42	26	118	13	41
15	405	93	**574**	93	61	103	41	93	**86**	112	44	30	32	49	16	93	22	48
16	**800**	**195**	314	101	43	89	38	111	44	102	48	25	38	46	25	109	26	46
17	358	96	296	98	55	93	50	94	49	113	49	33	34	40	26	112	28	55
18	397	105	**653**	87	47	101	51	96	51	102	54	28	45	53	24	97	18	34
Avg	398	96	297	98	48	96	46	99	48	104	50	31	32	48	23	98	18	51

context of Estimate 7.2. The discrepancy in Tables 9.2, 9.3, 9.5 and 9.6 arises from the assumptions made in Chapter 8, it seems also hard to make them more precisely. But most of the calculated values and all the average values are close to the expected values, this make us have confidence in the heuristic results.

Table 9.7 gives the values of $N_i(12, 3, \rho, 10^4, 10^8)$ for various ρ and $i = 1, 2, 3$. It shows that there is a big gap between $I(12, 3, \rho, 10^4, 10^8)$ and $N_1(12, 3, \rho, 10^4, 10^8)$ when $\rho < 1.25$, because in this case the Barreto-Naehrig family makes the assumptions in Estimate 7.2 not satisfied. But in this exceptional case, (9.2)–(9.4) are also compatible with numerical data.

TABLE 9.3: Values of $N_3(k, D, 1.8, 5, 5 \times 10^5)$ for various k and D

D	1	2	3	5	6	7	10	11	15	19	21	23	31	35	39	43	47	123
$k = 3$	193	42		46	52	43	14	53	48	59	44	20	9	25	25	45	8	41
4		35	305	54	50	48	17	55	43	47	58	9	16	27	26	59	8	59
5	233	69	290	46	42	51	24	43	43	40	48	11	17	25	22	40	8	57
6	193	45		50	62	42	20	42	45	48	52	8	16	29	26	45	10	64
7	215	51	304	55	49	**111**	19	43	49	50	42	20	13	19	23	49	6	56
8	**402**	**84**	297	60	42	40	21	40	57	59	57	13	12	26	30	45	6	57
9	186	40	**603**	43	60	46	25	54	54	56	53	18	17	18	23	41	6	59
10	198	55	289	56	45	55	18	47	49	45	46	19	6	18	19	63	10	50
11	187	42	260	50	44	46	25	**90**	43	61	53	18	12	21	21	49	11	41
12	**414**	37	**623**	44	36	43	28	52	39	55	50	21	21	25	24	46	6	53
13	203	53	312	42	32	37	24	59	56	47	47	13	10	23	17	31	3	54
14	209	50	271	42	47	**104**	27	50	49	53	55	17	17	22	26	66	6	41
15	185	57	**574**	46	61	49	15	49	**86**	45	44	16	18	34	16	50	10	48
16	**401**	**106**	314	45	43	43	13	54	44	45	48	12	13	24	25	64	14	46
17	179	45	296	52	55	41	23	41	49	58	49	18	20	22	26	57	14	55
18	199	49	**653**	46	47	45	23	49	51	54	54	13	24	26	24	42	3	34
Avg	199	48	297	49	48	46	21	49	48	51	50	15	15	24	23	50	8	51

TABLE 9.4: Values of $N_1(k, D, 2, 5, 10^5)$ for various k and D

D	13	14	17	22	30	33	51	55	59	67	71	79	83	87	91	95	111	219
I_0	236	118	118	236	118	118	236	118	157	472	67	94	157	79	236	59	59	118
$k = 3$	248	115	132	240	109	131	256	135	156	513	89	91	149	81	229	56	58	117
4	251	118	119	250	138	116	227	128	194	498	77	106	167	86	242	75	67	144
5	249	117	126	272	100	109	227	119	170	488	66	92	149	78	250	57	63	144
6	261	118	104	273	133	106	229	118	171	514	72	85	203	77	249	62	64	107
7	244	131	130	229	122	132	250	120	152	498	79	104	180	81	240	64	65	133
8	277	111	128	238	111	116	269	124	127	480	79	93	150	72	238	65	54	112
9	264	139	136	248	118	109	236	125	164	522	62	104	156	75	256	56	74	109
10	233	126	125	246	131	103	230	102	168	486	58	103	161	78	254	66	54	121
11	240	117	126	223	131	135	239	124	156	441	65	101	174	96	253	59	58	99
12	243	125	110	245	116	128	211	125	151	503	75	87	152	79	244	63	52	126
13	256	124	121	237	118	116	285	114	167	493	62	96	152	88	249	57	49	137
14	246	127	131	225	136	128	253	114	164	475	69	87	163	74	235	66	67	119
15	257	117	109	265	108	108	249	119	137	453	51	111	177	88	240	68	62	130
16	250	121	106	250	112	106	242	108	178	454	66	91	165	81	223	60	68	122
17	240	110	147	240	130	119	227	107	155	454	74	107	147	93	248	70	67	138
18	235	125	105	227	125	128	266	141	171	496	72	104	147	85	237	81	63	136
Avg	250	121	122	244	121	118	244	120	161	486	70	98	162	82	243	64	62	125

Table 9.5: Values of $N_2(k, D, 2, 5, 10^5)$ for various k and D

D	13	14	17	22	30	33	51	55	59	67	71	79	83	87	91	95	111	219
$k=3$	248	56	132	137	60	131	130	68	79	268	38	42	82	40	112	33	30	54
4	251	64	119	129	67	116	110	69	103	238	42	53	86	43	127	30	33	78
5	249	70	126	131	55	109	115	63	82	244	31	42	75	45	113	28	27	78
6	261	56	104	134	64	106	102	56	86	245	26	52	102	30	122	37	30	59
7	244	65	130	108	60	132	130	55	76	239	45	51	91	44	119	30	24	62
8	277	60	128	117	61	116	117	62	62	237	36	40	77	38	120	35	32	50
9	264	72	136	113	62	109	126	65	73	266	31	52	91	36	124	26	38	55
10	233	64	125	117	67	103	121	47	85	248	26	57	79	30	123	30	22	54
11	240	56	126	113	60	135	116	59	77	239	33	52	95	44	119	30	30	48
12	243	73	110	131	58	128	108	65	83	250	36	42	87	36	125	32	30	54
13	256	62	121	129	53	116	133	61	87	240	31	45	80	32	113	29	24	58
14	246	62	131	105	62	128	129	59	79	254	31	40	96	40	129	34	37	65
15	257	60	109	132	59	108	124	53	52	233	19	69	86	51	122	33	32	68
16	250	63	106	126	56	106	127	55	93	228	29	53	82	53	120	28	28	64
17	240	61	147	122	64	119	125	62	80	214	33	50	80	53	128	27	31	68
18	235	63	105	112	51	128	141	78	89	249	28	46	73	45	128	37	32	74
Avg	250	63	122	122	60	118	122	61	80	243	32	49	85	41	122	31	30	62

Table 9.6: Values of $N_3(k, D, 2, 5, 10^5)$ for various k and D

D	13	14	17	22	30	33	51	55	59	67	71	79	83	87	91	95	111	219
$k=3$	141	32	65	69	60	131	130	35	36	127	19	23	46	40	56	16	30	54
4	139	32	70	63	67	116	110	37	50	114	22	23	43	43	56	13	33	78
5	127	31	63	64	55	109	115	38	40	119	13	23	40	45	50	20	27	78
6	129	32	54	70	64	106	102	29	43	110	11	23	51	30	68	16	30	59
7	128	33	62	51	60	132	130	24	33	125	22	26	50	44	68	13	24	62
8	130	28	60	67	61	116	117	31	34	115	20	16	33	38	66	18	32	50
9	116	32	71	58	62	109	126	28	33	135	15	27	51	36	56	11	38	55
10	130	41	61	58	67	103	121	31	43	129	10	27	42	30	61	14	22	54
11	110	21	66	56	60	135	116	28	37	120	13	25	43	44	54	14	30	48
12	123	38	46	59	58	128	108	35	45	110	16	18	47	36	63	13	30	54
13	115	30	58	72	53	116	133	36	49	113	17	20	38	32	52	14	24	58
14	115	30	64	60	62	128	129	28	36	139	16	18	45	40	64	21	37	65
15	114	41	54	64	59	108	124	30	24	124	8	32	48	51	47	15	32	68
16	129	37	58	61	56	106	127	28	52	111	15	31	44	53	64	9	28	64
17	107	38	88	59	64	119	125	26	40	111	20	25	47	53	61	11	31	68
18	123	25	53	50	51	128	141	36	48	126	17	17	37	45	69	18	32	74
Avg	124	33	62	61	60	118	122	31	40	121	16	23	44	41	60	15	30	62

Table 9.7: Values of $N_i(12, 3, \rho, 10^4, 10^8), i = 1, 2, 3$, for various ρ

ρ	1.1	1.15	1.2	1.25	1.3	1.35	1.4	1.45	1.5	1.55
$I(12, 3, \rho, 10^4, 10^8)$	1	2	4	8	16	32	67	142	304	658
$N_1(12, 3, \rho, 10^4, 10^8)$	8	12	15	22	33	47	83	177	355	706
$N_2(12, 3, \rho, 10^4, 10^8)$	2	5	7	11	16	23	43	88	178	388
$N_3(12, 3, \rho, 10^4, 10^8)$	2	5	7	11	16	23	43	88	178	388

Bibliography

[1] M. Abouzaid, A. Bérczes, Yu. Bilu, and S. Najib. Effective bounds for the polynomial norm equation. *In preparation*, 2012.

[2] L. Ahlfors. *Complex Analysis*. Mcgraw-Hill, 1979.

[3] T.M. Apostol. *Introduction to Analytic Number Theory*. Springer-Verlag, 1976.

[4] R. Avanzi, H. Cohen, C. Doche, G. Frey, T. Lange, K. Nguyen, and F. Vercauteren. *Handbook of elliptic and hyperelliptic curve cryptography*. CRC Press, 2005.

[5] A. Bajolet and Yu. Bilu. Computing integral points on $X_{\mathrm{ns}}^{+}(p)$. *Preprint*, 2012. arXiv:1212.0665.

[6] A. Bajolet and M. Sha. Bounding the j-invariant of integral points on $X_{\mathrm{ns}}^{+}(p)$. *Proceedings of the American Mathematical Society*, to appear. arXiv:1203.1187.

[7] A. Baker. Linear forms in the logarithms of algebraic numbers I. *Mathematika*, 13: 204–216, 1966.

[8] A. Baker. Linear forms in the logarithms of algebraic numbers II. *Mathematika*, 14: 102–107, 1967.

[9] A. Baker. Linear forms in the logarithms of algebraic numbers III. *Mathematika*, 14: 220–224, 1967.

[10] A. Baker. Contribution to the theory of Diophantine equations. *Philosophical Transactions of the Royal Society*, A263: 173–208, 1968.

[11] A. Baker. Linear forms in the logarithms of algebraic numbers IV. *Mathematika*, 15: 204–216, 1968.

[12] A. Baker. The Diophantine equation $y^2 = ax^3 + bx^2 + cx + d$. *Journal of the London Mathematical Society*, 43: 1–9, 1968.

[13] A. Baker. Bounds for solutions of hyperelliptic equations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 65: 439–444, 1969.

[14] A. Baker and J. Coates. Integer points on curves of genus 1. *Mathematical Proceedings of the Cambridge Philosophical Society*, 67: 595–602, 1970.

[15] R. Balasubramanian and N. Koblitz. The improbability that an elliptic curve has subexponential discrete log problem under the Menezes-Okamoto-Vanstone algorithm. *Journal of Cryptology*, 11: 141–145, 1998.

[16] B. Baran. A modular curve of level 9 and the class number one problem. *Journal of Number Theory*, 129: 715–728, 2009.

[17] B. Baran. Normalizers of non-split Cartan subgroups, modular curves, and the class number one problem. *Journal of Number Theory*, 130: 2753–2772, 2010.

[18] P.S.L.M. Barreto, B. Lynn, and M. Scott. Constructing elliptic curves with prescribed embedding degrees. *In Security in Communication Networks—SCN 2002. Lecture Notes in Computer Science*, 2576: 263–273, 2002.

[19] P.S.L.M. Barreto and M. Naehrig. Pairing-friendly elliptic curves of prime order. *In Selected Areas in Cryptography 2005. Lecture Notes in Computer Science*, 3897: 319–331, 2006.

[20] P.T. Bateman and R.A. Horn. A heuristic asymptotic formula concerning the distribution of prime numbers. *Mathematics of Computation*, 16: 363–367, 1962.

[21] G.V. Belyĭ. On Galois extensions of a maximal cyclotomic field. *Mathematics of the USSR-Izvestiya*, 14: 247–256, 1980.

[22] A. Bérczes, J. Evertse, and K. Győry. Effective results for linear equations in two unknowns from a multiplicative division group. *Acta Arithmetica*, 136: 331–349, 2009.

[23] Yu. Bilu. Effective analysis of integral points on algebraic curves. *Israel Joural of Mathematics*, 90: 235–252, 1995.

[24] Yu. Bilu. Baker's method and modular curves. In *A Panorama of Number Theory or The View from Baker's Garden*, pages 73–88. Cambridge University Press, 2002.

[25] Yu. Bilu and G. Hanrot. Solving Thue equations of high degree. *Journal of Number Theory*, 60: 373–392, 1996.

[26] Yu. Bilu and M. Illengo. Effective Siegel's theorem for modular curves. *Bulletin of the London Mathematical Society*, 43: 673–688, 2011.

[27] Yu. Bilu and P. Parent. Runge's method and modular curves. *International Mathematics Research Notices*, 2011: 1997–2027, 2011.

[28] Yu. Bilu and P. Parent. Serre's uniformity problem in the split Cartan case. *Annals of Mathematics*, 173: 569–584, 2011.

[29] Yu. Bilu, P. Parent, and M. Rebolledo. Rational points on $X_0^+(p^r)$. *Annales de l'Institut Fourier*, to appear. arXiv:1104.4641.

[30] Yu. Bilu and M. Strambi. Quantitative Riemann existence theorem over a number field. *Acta Arithmetica*, 145: 319–339, 2010.

[31] Yu. Bilu, M. Strambi, and A. Surroca. Quantitative Chevalley-Weil theorem for curves. *Monatshefte für Mathematik*, to appear. arXiv:0908.1233.

[32] Yu. F. Belotserkovskiĭ (Yu. Bilu). Effective analysis of a new class of Diophantine equations. (Russian. English summary). *Vestsi Akad. Navuk BSSR Ser. Fiz.-Mat. Navuk*, 125: 34–39, 1988.

[33] D. Boneh and M. Franklin. Identity-based encryption from the Weil pairing. *In Crypto 2001. Lecture Notes in Computer Science*, 2139: 213–229, 2001.

[34] D. Boneh, B. Lynn, and H. Shacham. Short signatures from the Weil pairing. *In Asiacrypt 2001. Lecture Notes in Computer Science*, 2248: 514–532, 2001.

[35] J. Boxall. Heuristics on pairing-friendly elliptic curves. *Journal of Mathematical Cryptology*, 6: 81–104, 2012.

[36] F. Brezing and A. Weng. Elliptic curves suitable for pairing based cryptography. *Designs, Codes and Cryptography*, 37: 133–141, 2005.

[37] Y. Bugeaud and K. Győry. Bounds for the solutions of unit equations. *Acta Arithmetica*, 74: 67–80, 1996.

[38] C. Cocks and R.G.E. Pinch. Identity-based cryptosystems based on the Weil pairing. *Unpublished manuscript*, 2001.

[39] A. Costa and E. Friedman. Ratios of regulators in totally real extensions of number fields. *Journal of Number Theory*, 37: 288–297, 1991.

[40] F. Diamond and J. Shurman. *A First Course in Modular Forms*. Springer-Verlag, 2005.

[41] E. Dobrowolski. On a question of Lehmer and the number of irreducible factors of a polynomial. *Acta Arithmetica*, 34: 391–401, 1979.

[42] R. Dupont, A. Enge, and F. Morain. Building curves with arbitrary small MOV degree over finite prime fields. *Journal of Cryptology*, 18: 79–89, 2005.

[43] R. Dvornicich and U. Zannier. Fields containing values of algebraic function. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze*, 21: 421–443, 1994.

[44] N.D. Elkies. *ABC* implies Mordell. *International Mathematics Research Notices*, 7: 99–109, 1991.

[45] J. Esmonde and M. Ram Murty. *Problems in algebraic number theory.* Springer-Verlag, 2004.

[46] G. Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Inventiones mathematicae*, 73: 349–366, 1983.

[47] S. Finch, G. Martin, and P. Sebah. Roots of unity and nullity modulo n. *Proceedings of the American Mathematical Society*, 138: 2729–2743, 2010.

[48] D. Freeman, M. Scott, and E. Teske. A taxonomy of pairing-friendly elliptic curves. *Journal of Cryptology*, 23: 224–280, 2010.

[49] G. Frey and H. Rück. A remark concerning m-divisibility and the discrete logarithm in the divisor class group of curves. *Mathematics of Computation*, 62: 865–874, 1994.

[50] E. Friedman. Analytic formulas for regulators of number fields. *Inventiones mathematicae*, 98: 599–622, 1989.

[51] A. Fröhlich and T. Martin. *Algebraic Number Theory.* Cambridge University Press, 1993.

[52] S. Galbraith. Pairings. In *Advances in Elliptic Curve Cryptography*, pages 183–213. Cambridge University Press, 2005.

[53] G.H. Hardy and E.M. Wright. *An introduction to the theory of numbers.* Oxford University Press, 1979.

[54] T. Hayashi, N. Shinohara T. Shimoyama, and T. Takagi. Breaking pairing-based cryptosystems using η_T pairing over $GF(3^{97})$. *In Asiacrypt 2012. Lecture Notes in Computer Science*, 7658: 43–60, 2012.

[55] K. Heegner. Diophantische Analysis und Modulfunktionen. *Mathematische Zeitschrift*, 56: 227–253, 1952.

[56] A. Joux. A one round protocol for tripartite Diffie-Hellman. *In Algorithmic Number Theory Symposium 2000. Lecture Notes in Computer Science*, 1838: 385–393, 2000.

[57] M.A. Kenku. A note on the integral points of a modular curve of level 7. *Mathematika*, 32: 45–48, 1985.

[58] N. Koblitz. *p-adic Numbers, p-adic Analysis, and Zeta-functions*. Springer-Verlag, 1984.

[59] N. Koblitz. Elliptic curve cryptosystems. *Mathematics of Computation*, 48: 203–209, 1987.

[60] N. Koblitz and A.J. Menezes. Pairing-based cryptography at high security levels. *In Cryptography and Coding 2005. Lecture Notes in Computer Science*, 3796: 13–36, 2005.

[61] J. Korevaar and H. Te Riele. Average prime-pair counting formula. *Mathematics of Computation*, 79: 1209–1229, 2010.

[62] D.S. Kubert and S. Lang. *Modular Units*. Springer-Verlag, 1981.

[63] S. Lang. *Introduction to Modular Forms*. Springer-Verlag, 1976.

[64] H.W. Lenstra. Factoring integers with elliptic curves. *Annals of Mathematics*, 126: 649–673, 1987.

[65] S. Louboutin. Upper bounds on $|L(1, \chi)|$ and applications. *Canadian Journal of Mathematics*, 50: 794–815, 1998.

[66] F. Luca and I.E. Shparlinski. Elliptic curves with low embedding degree. *Journal of Cryptology*, 19: 553–562, 2006.

[67] F. Luca and I.E. Shparlinski. On finite fields for pairing based cryptography. *Advances in Mathematics of Communications*, 1: 281–286, 2007.

[68] E.M. Matveev. An explicit lower bound for a homogeneous rational linear form in the logarithms of algebraic numbers II. *Izvestiya Mathematics*, 64: 1217–1269, 2000.

[69] A. Menezes, T. Okamoto, and S. Vanstone. Reducing elliptic curve logarithms to logarithms in a finite field. *IEEE Transactions on Information Theory*, 39: 1639–1646, 1993.

[70] V. Miller. Use of elliptic curves in cryptography. *In Crypto 1985. Lecture Notes in Computer Science*, 218: 417–426, 1986.

[71] V. Miller. The Weil pairing, and its efficient calculation. *Journal of Cryptology*, 17: 235–261, 2004.

[72] A. Miyaji, M. Nakabayashi, and S. Takano. New explicit conditions of elliptic curve traces for FR-reduction. *IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences*, E84-A: 1234–1243, 2001.

[73] K. Győry and K. Yu. Bounds for the solutions of S-unit equations and decomposable form equations. *Acta Arithmetica*, 123: 9–41, 2006.

[74] K. Paterson. Cryptography from pairings. In *Advances in Elliptic Curve Cryptography*, pages 215–251. Cambridge University Press, 2005.

[75] R. Sakai, K. Ohgishi, and M. Kasahara. Cryptosystems based on pairing. In *Symposium on Cryptography and Information Security 2000*, Okinawa, Japan, 2000.

[76] M. Scott and P.S.L.M. Barreto. Generating more MNT elliptic curves. *Designs, Codes and Cryptography*, 38: 209–217, 2006.

[77] J.P. Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Inventiones mathematicae*, 15: 259–331, 1972.

[78] J.P. Serre. *Lectures on the Mordell-Weil Theorem*. Vieweg, 1997.

[79] M. Sha. Bounding the j-invariant of integral points on certain modular curves. *Preprint*, 2012. arXiv:1210.3224.

[80] M. Sha. On the number of isogeny classes of ordinary pairing-friendly elliptic curves and heuristics of the Cocks-Pinch method. *Advances in Mathematics of Communications*, accepted. arXiv:1211.0971.

[81] M. Sha. Bounding the j-invariant of integral points on modular curves. *International Mathematics Research Notices*, doi: 10.1093/imrn/rnt085. arXiv:1208.1337.

[82] G. Shimura. *Introduction to the Arithmetic Theory of Automorphic Functions*. Iwanami Shoten; Princeton University Press, 1971.

[83] C.L. Siegel. Über einige Anwendungen Diophantischer Approximationen. *Abhandlungen der Preussischen Akademie der Wissenschaften*, Nr. 1, 1929.

[84] C.L. Siegel. Zum Beweise des Starkschen Satzes. *Inventiones mathematicae*, 5: 180–191, 1968.

[85] C.L. Siegel. Abschätzung von Einheiten. *Nachr. Akad. Wiss. Gottingen II. Math.-Phys. Kl.*, 9: 71–86, 1969.

[86] A. Surroca. Siegel's theorem and the *abc* conjecture. *Rivista di Matematica della Università di Parma*, 3*: 323–332, 2004.

[87] A. Surroca. Sur l'effectivité du théorème de Siegel et la conjecture *abc*. *Journal of Number Theory*, 124: 267–290, 2007.

[88] A.V. Sutherland. Computing Hilbert class polynomials with the Chinese Remainder Theorem. *Mathematics of Computation*, 80: 501–538, 2011.

[89] PARI/GP, version `2.5.3`. Bordeaux, 2012. http://pari.math.u-bordeaux.fr/.

[90] J.J. Urroz, F. Luca, and I.E. Shparlinski. On the number of isogeny classes and pairing-friendly elliptic curves and statistics for MNT curves. *Mathematics of Computation*, 81: 1093–1110, 2012.

[91] E. Verheul. Evidence that XTR is more secure than supersingular elliptic curve cryptosystems. *In Eurocrypt 2001. Lecture Notes in Computer Science*, 2045: 195–210, 2001.

[92] M. Waldschmidt. *Diophantine Approximation on Linear Algebraic Groups*. Springer-Verlag, 2000.

[93] L.C. Washington. *Introduction to Cyclotomic Fields*. Springer-Verlag, 1982.

[94] K. Yu. P-adic logarighmic forms and group varieties III. *Forum Mathematicum*, 19: 187–280, 2007.