Integral Points on Modular Curves, Singular Moduli and Conductor-Discriminant Inequality

Yulin Cai

THÈSE EN COTUTELLE

présentée par

Yulin CAI

pour obtenir le grade de

DOCTEUR

DE L'UNIVERSITÉ DE BORDEAUX

École Doctorale de Mathématiques et d'Informatique

Integral Points on Modular Curves, Singular Moduli and Conductor-Discrminant Inequality

dirigée par Yuri BILU et Qing LIU

Soutenue le 24 Juille 2020 devant le jury composé de :

Prof.	Pascal AUTISSIER	Université de Bordeaux	Président
Prof.	Nicolas BILLEREY	Université Clermont Auvergne	Rapporteur
Prof.	Yuri BILU	Université de Bordeaux	Directeur
Prof.	Sara CHECCOLI	Université Grenoble Alpes	Examinateur
Prof.	Jean GILLIBERT	Université Toulouse 2	Eximanteur
Prof.	Ariyan JAVANPEYKAR	Johannes Gutenberg-Universität Mainz	Rapporteur
Prof.	Qing LIU	Université de Bordeaux	Directeur

THÈSE EN COTUTELLE

présentée par

Yulin CAI

pour obtenir le grade de

DOCTEUR

DE L'UNIVERSITÉ DE BORDEAUX

École Doctorale de Mathématiques et d'Informatique

Points entiers sur les courbes modulaires, les modules singuliers et l'inégalité conducteur-discriminant

dirigée par Yuri BILU et Qing LIU

Soutenue le 24 Juille 2020 devant le jury composé de :

Prof.	Pascal AUTISSIER	Université de Bordeaux	Président
Prof.	Nicolas BILLEREY	Université Clermont Auvergne	Rapporteur
Prof.	Yuri BILU	Université de Bordeaux	Directeur
Prof.	Sara CHECCOLI	Université Grenoble Alpes	Examinateur
Prof.	Jean GILLIBERT	Université Toulouse 2	Eximanteur
Prof.	Ariyan JAVANPEYKAR	Johannes Gutenberg-Universität Mainz	Rapporteur
Prof.	Qing LIU	Université de Bordeaux	Directeur

Institut de Mathématiques de Bordeaux
(UMR 5251)
Université de Bordeaux
351, cours de la Libération
- F 33 405 TALENCE

Acknowledgements

First of all, I would like to express my deep gratitude and respect to my supervisors, Prof. Yuri Bilu and Prof. Qing Liu. They gave me a lot of help, not just in academic study, but also in daily life. They provided me these topics to work on, which opened a door for me to study number theory. When I read the book "Algebraic geometry and arithmetic curves" of Prof. Liu, Prof. Liu became an excellent instructor. Without him, learning algebraic geometry would be a hard nut for me. Moreover, when I arrived to Bordeaux, Prof. Liu spent his time to help me go through many formalities, which helped me get used to this new life abraod. From the couses of Prof. Bilu and the discussions with him, I learned techniques in Diphantien approximation, which are essential in my PhD thesis. Without him, it would be hard for me to work with these techniques. Two professor are very kind and patient, with their rich experience and perspicacity, they always provided me necessary help from our discussions and communications, including many valuable suggestions for my papers and PhD thesis. This thesis would not been seen without them.

I would like to thanks Prof. Nicolas Billerey and Prof Ariyan Javanpeykar for being agree to become the repportors of my thesis and the members of the jury. Their reports are very detailed. In particular, Prof. Billerey gave me valuable suggestions for last version of my thesis. Following his suggestions, my thesis has been improved.

I want to thank Prof. Pascal Autissier, Prof. Sara Checoli and Prof. Jean Gillibert for being agree to paticipate the jury of my defence. In Particular, Prof. Autissier was in my "comité de suivi" in the last 3 year, he encouraged me and gave me some useful advices.

I want to thank Prof. Jean-François Quint. He was also in my "comité de suivi" in the last 3 year. He encouraged me and taught me some of his experience during his researches.

I want to thank Prof. Deni Benois and Prof. Dajano Tossici for allowing me to participate in their master courses. The course, algebraic number theory, of Prof. Benois improved my skills in number theory. On the other hand, Prof. Tossici's course, group cohomology, enabled me to understand class field theory and other topics via group cohomology.

I wish to thank Prof. Andreas Hartmann. He has been helping me go through the formality of preparing my defence. Without him, I would have a hard time finding correct information to prepare me defence.

I want to thank the secretaries and librarians in Institut de Mathématiques de

Bordeaux. They are very kind and efficient. They keep the institute run smoothly and help me in many different ways.

I would like to thank Prof. Peter Bruin for coming to Bordeaux to discuss problems with me. It was a great pleasure to learn from him.

I wish to thank Prof. Min Sha and Prof. Jilong Tong. They have been helpful when I first arrived in Bordeaux. In particular, Min Sha helped me deal with the apartement and some documents when my French was poor.

I would like to express my gratitude to Prof. Yanan Lin, Prof. Lu Lin. They agree to become my guarantors, so that I could have the scholarship of CSC to become a PhD student in Frence. Moreover Prof. Yanan Lin was my supervisor when I was a master student in Xiamen University. Besides great help he gave me when I was master student, he still encouraged me and gave me some advices after I became a PhD student.

I want to thanks Abhinandan, Thoung Tuan Dang, Roberto Gualdi, Jean Kieffer, Emanuele Tron, Lumming Zhao for having some great seminars with me. It was a great experience to learn and discuss math with them.

I wish to thanks Chinese friends in Bordeaux: Wei Chen, Zhenhui Chen, Xiaoming Fu, Wencan He, Yiye Jiang, Jalun Li, Lin Li, Pei Su, Lina Wang, Han Xiao, Zaicheng Zhang, Lin Zhu, etc. Because of you, my study in Bordeaux was not only rich, but also colorful.

In this difficult time of confinement of Covid-19, I also need to thank my friends in China for their caring and support: Zhen Shen, Ce Xu, Yingyue Yang, Chunying Zhang, Jianfeng Zhou, etc.

Last but not least, I deeply thank my parents and my sisters for their unconditional trust and support. It was their love that encouraged me to get out from frustration.

Résumé

Cette thèse traite de trois sujets en trois parties.

Dans la première partie, nous étudions les points S-entiers de la courbe modulaire $X_0(p)$. Yuri Bilu a montré qu'en utilisant la méthode de Baker, on peut donner une borne effective de la hauteur de ces points en fonction de p, du corps de base et de l'ensemble de places S.

Min Sha a rendu ce résultat explicite. avec une borne doublement exponentielle en dans p. Nous améliorons considérablement dans cette thèse le résultat de Sha, en obtenant une borne simplement exponentielle. Cela se fait en utilisant une version très explicite du principe de Chevalley-Weil basée sur des travaux de Qing Liu et Dino Lorenzini. Notre borne est non seulement plus nette que celle de Sha, mais également explicite en tous les paramètres.

Dans la deuxième partie, nous considérons des modules singuliers de courbes elliptiques. Pour un module singulier fixe α, nous donnons une borne supérieure effective de la norme de $x - \alpha$ pour un autre module singulier x avec un grand discriminant.

Dans la troisième partie, nous donnons une relation entre les conducteurs d'Artin d'un modèle Werestrass Y et ceux de deux modèles de Weierstrass donnés Y_1, Y_2. Avec cette relation, nous déduisons que l'inégalité conducteur-discriminant est valable pour Y si elle est valable pour Y_1 et Y_2.

Mots-clefs

Points entiers; courbe modulaire; module singulier; courbe hyperelliptique; conducteur d'Artin.

Abstract

This thesis discusses three topics, so it includes three parts.

In the first part, we study S-integral points on the modular curve $X_0(p)$. Bilu showed that, using Baker's method, they can be effectively bounded in terms of p, the base field and the set of places S. Sha made this result explicit, but the bound he obtained is double exponential in p. We drastically improve upon the result of Sha, obtaining a simple exponential bound. This is done using a very explicit version of the Chevalley-Weil principle based on the work of Liu and Lorenzini. Our bound is not only sharper than that of Sha, but is also explicit in all parameters.

In the second part, we consider singular moduli. For a fixed singular modulus α, we give an effective upper bound of norm of $x - \alpha$ for another singular modulus x with large discriminant.

In the third part, we give a relation between Artin conductors of a Weierstrass model Y and the ones of two given Weierstrass models Y_1, Y_2. With this relation, we know that the conductor-discriminant inequality holds for Y if it holds for Y_1 and Y_2.

Keywords

integral point; modular curve; singular modulus; hyperelliptic curve; Artin conductor.

Contents

Dedication

To my parents

Introduction

Cette thèse se concentre sur trois sujets différents, elle est donc divisée en trois parties. Dans la première partie, nous considérons les points entiers sur les courbes modulaires. La deuxième partie est consacrée à donner une borne de différence de deux modules singuliers en termes de discriminants. La troisième partie est plus algébrique, nous étudions les conducteurs d'Artin et discriminants des courbes hyperelliptiques.

Points entiers sur les courbes modulaires

Soit X une courbe algébrique projective, lisse et connectée définie sur un corp de nombres K, et que $x \in K(X)$ soit une fonction rationnelle, non constante sur X. Si R est une sous-anneau de K, nous désignons par $X(R, x)$ l'ensemble des points K-rationnels de X qui sont R-entiers par rapport à la coordonnée x:

$$X(R, x) = \{P \in X(K) \mid x(P) \in R\}.$$

En particulier, si S est un ensemble fini de places de K (y compris toutes les places infinies), nous considérons l'ensemble de *points S-entiers* $X(\mathcal{O}_S, x)$, où $\mathcal{O}_S = \mathcal{O}_{S,K}$ est l'anneau de S-entiers en K.

Selon le théorème classique de Siegel [57] (voir aussi [33, Part D] pour une exposition moderne), l'ensemble $X(\mathcal{O}_S, x)$ est fini si au moins l'un des les conditions suivantes est remplie:

$$g(X) \geq 1; \tag{1}$$

$$x \text{ admet au moins 3 pôles dans } X(\bar{\mathbb{Q}}). \tag{2}$$

Le théorème de Faltings [24] (voir aussi [33, Part E]) affirme que $X(K)$ est fini si $g(X) \geq 2$. Malheureusement, toutes les preuves connues de théorème de Siegel et Faltings ne sont pas efficaces, ce qui signifie qu'elles n'impliquent aucune expression explicite bornant les hauteurs de points entiers ou rationnels.

À partir des travaux révolutionnaires d'A. Baker en 1960, des preuves efficaces du théorème de Siegel ont été découvertes, par Baker et d'autres, pour de nombreuses paires (X, x), voir [4, 5] et les références dedans.

Un cas intéressant est lorsque $X = X_\Gamma$ est la courbe modulaire correspondant à un sous-groupe de congruence Γ de $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$, et $x = j$ est la fonction rationnelle définie par le j-invariant. Ce problème a été examiné par Bilu [5], Bilu et Parent [9] [10], Sha [55] [54] et bien d'autres. En particulier, Bilu et Parent résolvent le cas dit "de Cartan deloyé" du problème d'uniformité de Serre dans [10] en considérant ce problème pour X_{split}.

Bilu [4, Section 5] (voir aussi [5, Section 4]) a fait l'observation suivante.

Proposition 1. *Soit Γ un sous-groupe de congruence de $\mathrm{SL}_2(\mathbb{Z})$ de niveau N ayant au moins 3 cuspides. Soit K un corp de nombres tel que X_Γ admet un modèle géométriquement*

irréductible sur K et tel que $j \in K(X_\Gamma)$. Soit S un ensemble fini de places de K contenant toutes les places infinies. Il existe alors une constante effective $c = c(N, K, S)$ telle que pour tout $P \in X_\Gamma(\mathcal{O}_S, j)$ nous avons $\mathrm{h}(j(P)) \leq c$.

(Ici $\mathrm{h}(\cdot)$ est la hauteur logarithmique définie sur l'ensemble $\bar{\mathbb{Q}}$ de nombres algébriques.)

En d'autres termes, si la condition (2) est satisfaite pour le paire (X_Γ, j), alors le théorème de Siegel est efficace pour ce paire.

Sha [55] a rendu le borne dans la proposition 1 totalement explicite. Pour énoncer son résultat, nous introduisons quelques notations. Pour un sous-groupe de congruence Γ comme ci-dessus, le nombre de cuspides sur X_Γ est indiqué par $v_\infty(\Gamma)$. Pour un corp de nombres K, soit M_K l'ensemble de toutes les places de K, et $S \subseteq M_K$ un sous-ensemble fini contenant toutes les places infinies. Nous mettons $d = [K : \mathbb{Q}]$ et $s = \#S$. Soit $\mathcal{O}_K$ l'anneau d'entiers de K. Nous définissons la quantité suivante

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

en fonction de $N \in \mathbb{N}^+$, où D est le discriminant absolu de K, $\varphi(N)$ est la fonction totiente d'Euler, et la norme $\mathcal{N}_{K/\mathbb{Q}}(v)$ d'une place v, par définition, est égal à $\#(\mathcal{O}_K/\mathfrak{p}_v|)$ lorsque v est fini et $\mathfrak{p}_v$ est son idéal premier correspondant, et est fixé à 1 si v est infini.

Sha [55] a prouvé le théorème suivant.

Théorèm 3.2.1 ([55] Theorem 1.2)**.** *Soit Γ de niveau N et X_Γ la courbe modulaire correspondante sur un corp de nombres K avec $d = [K : \mathbb{Q}]$ et $S \subseteq M_K$ un ensemble fini contenant tous les places infinis. Si $v_\infty(\Gamma) \geq 3$, alors pour tout $P \in X_\Gamma(\mathcal{O}_S, j)$,*

$$\mathrm{h}(j(P)) \leq (CdsM^2)^{2sM}(\log(dM))^{3sM}\ell^{dM}\Delta(M),$$

où C est une constante effective, ℓ est le nombre premier maximal tel qu'il existe $v \in S$ avec $v \mid \ell$, ou $\ell = 1$ si S ne contient que l'infini places, et M est défini comme suit:

$$M = \begin{cases} N & \text{si } N \text{ n'est pas une puissance première;} \\ 3N & \text{si } N \text{ est une puissance de } 2; \\ 2N & \text{si } N \text{ est une puissance d'un nombre impair.} \end{cases}$$

Ici, nous remarquons seulement que la borne de Sha est de la forme $c(K, S)^{N \log N}$, où $c(K, S)$ est une constante efficace ne dépendant que de K et S. En gros, nous avons ici une dépendance de type exponentielle dans N.

Pour certaines applications, il est utile d'avoir une valeur explicite de la constante C de Théorème 3.2.1. Dans cette partie, nous prouvons le résultat suivant.

Théorèm 3.2.2. *La constante C dans Théorème 3.2.1 peut être considérée comme 2^{14}.*

Dans la preuve, nous suivons les idées principales de Sha, avec quelques modifications mineures. Nous calculons explicitement les constantes implicites qui y sont présentes, y compris l'inégalité de Baker, Théorème 1.3.1.

Proposition 1 s'applique également dans de nombreux cas importants: voir [5, 8] pour plus de détails. En particulier, elle s'applique à la courbe modulaire $X_0(N)$ de niveau composite N. Cependant, elle ne s'applique pas directement à la courbe $X_0(p)$ du niveau premier p, car elle n'a que 2 cuspides.

Néanmoins, en utilisant un argument de recourvement, Bilu [5, Theorem 10] a prouvé que le théorème de Siegel est également efficace pour $X_0(p)$. Notez que la courbe $X_0(N)$ a un modèle géométriquement irréductible standard sur $\mathbb{Q}$.

Théorèm 1 (Bilu). *Soit p un nombre premier distinct de $2, 3, 5, 7, 13$. Soit K un champ numérique et S un ensemble fini de places de K contenant toutes les places infinies. Il existe alors une constante effective $c = c(p, K, S)$ telle que pour tout $P \in X_0(p)(\mathcal{O}_S, j)$ nous avons $\mathrm{h}(j(P)) \leq c$.*

L'outil principal est *Principe de Chevalley-Weil*, ou Théorème de Chevalley-Weil dans cette thèse, utilisé sous la forme suivante.

Proposition 2 (Principe de Chevalley-Weil). *Soit $\widetilde{X} \xrightarrow{\pi} X$ un morphisme non constant étale de courbes algébriques projectives définies sur un champ numérique K. Il existe alors un ensemble fini T de places de K tel que le suivant soit valable. Soit $P \in X(\bar{K})$ et $\widetilde{P} \in \widetilde{X}(\bar{K})$ tels que $\pi(\widetilde{P}) = P$. Soit v une place finie du corp $K(P)$ ramifié en $K(\widetilde{P})$. Alors v étend une place de T.*

Nous discuterons de ce théorème dans la section 1.5.

Bilu a trouvé un sous-groupe $\widetilde{\Gamma}$ de $\Gamma_0(p)$ tel que le morphisme naturel $X_{\widetilde{\Gamma}} \to X_0(p)$ is étale et $X_{\widetilde{\Gamma}}$ a au moins trois cuspides, voir Proposition 3.3.2 pour les détails. Principe de Chevalley-Weil permet maintenant de réduire le problème de $X_0(p)$ à $X_{\widetilde{\Gamma}}$, où Proposition 1 s'applique.

Dans [54] Sha a donné une version explicite du Théorème 1. Nous ne reproduisons pas ici la déclaration complète de Sha, qui est très compliquée, et nous nous concentrons uniquement sur la dépendance au niveau p. On peut s'attendre ici à une dépendance de type exponentielle dans p, mais Sha obtient une borne supérieure de la forme $c(K, S)^{\exp(p^6 \log p)}$, doublement exponentielle dans p.

La borne de Sha est si grande parce qu'il utilise une version quantitative d Théorème de Chevalley-Weil de [12], voir Proposition 1.5.4, qui fournit des bornes supérieures extrêmement grandes pour les quantités impliquées.

Dans cette thèse, nous allons prouver une autre version du théorème de Chevalley-Weil, proposition 1.5.5, combinée avec le théorème d'Igusa, voir [22, Section 8.6], nous parvenons à améliorer le résultat de Sha. Nous prouverons le théorème suivant.

Théorèm 3.3.1. *Gardons les notations de Théorème 1. Alors pour $P \in X_0(p)(\mathcal{O}_S, j)$, on a*

$$\mathrm{h}(j(P)) \leq e^{9s^2p^4 \log p} C(K, S)^{p^2},$$

où $C(K, S)$ peut être effectivement déterminé en termes de K et S. Plus explicitement, $C(K, S)$ peut être choisi comme

$$C(K, S) = 2^{29s} d^{9s} s^{2s} \ell^d |D| (\log(|D| + 1))^d \prod_{\substack{v \in S \\ v \nmid infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v),$$

où $d = [K : Q]$, D est le discriminant absolu de K, $s = \#S$ et ℓ est le premier maximal tel qu'il existe $v \in S$ avec $v \mid \ell$.

Pour un corps de nombres K, $v \in M_K$, nous définissons la valorisation $|\cdot|_v$ sur K comme suit: pour tout $\alpha \in K$:

$$|\alpha|_v := |\sigma(\alpha)|, \text{ si } v \text{ est infini avec plongement } \sigma;$$

$$|\alpha|_v := \mathcal{N}_{K/\mathbb{Q}}(v)^{-\mathrm{ord}_v(\alpha)/[K_v:\mathbb{Q}_v]}, \text{ si } v \text{ est fini.}$$

Modules singuliers

Soit $\mathbb{H}$ le demi-plan de Poincaré, un point $\tau \in \mathbb{H}$ est appelé un point CM si $\mathrm{End}(E_\tau)$ est un ordre dans un corps quadratique imaginaire , où E_τ est la courbe elliptique sur $\mathbb{C}$ correspondant à τ. Il est bien connu que $\tau \in \mathbb{H}$ est CM si et seulement si τ est un nombre algébrique de degré 2. Nous appelons $j(\tau)$ un module singulier si $\tau \in \mathbb{H}$ est CM. De la théorie CM classique, nous savons que chaque module singulier est un entier algébrique. On appelle $j(\tau)$ unité singulière s'il est un module singulier et une unité algébrique.

Dans [31], Habegger a prouvé qu'il y a au plus un nombre fini d'unités singulières. Cependant sa preuve est inefficace. Après cela, dans [7], Bilu, Habegger et Kühne prouvent qu'il n'y a pas d'unités singulières. En effet, leur méthode peut être généralisée pour donner une borne effective de norme de différence entre deux modules singuliers, c'est exactement ce que nous faisons dans cet thèse.

D'autre part, Gross et Zagier [27] ont énoncé une formule explicite pour la norme absolue de différence entre deux modules singuliers. Avec leurs travaux, Li [37] a également réussi à donner une borne de norme de différence entre deux modules singuliers, sa borne est un nombre strictement positif, ce qui lui permet de prouver une version généralisée du résultat principal de Bilu, Habegger et Kühne [7]. Cependant, il n'est pas clair comment son borne se comporte comme $\Delta \to -\infty$. Dans cet thèse, nous allons prouver le résultat suivant:

Théorèm 5.1.1. *Soit α, x deux modules singuliers de discriminants Δ_α, Δ respectivement, et $K = \mathbb{Q}(\alpha, x)$.*

(1) Si $\Delta_\alpha \neq -3, -4$ et $|\Delta| \geq \max\{e^{3.12}(\mathcal{C}(\Delta_\alpha)|\Delta_\alpha|^4 e^{\mathrm{h}(\alpha)})^3, 10^{15} \cdot \mathcal{C}(\Delta_\alpha)^6\}$, puis

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)| > \frac{|\Delta|^{1/2}}{2};$$

(2) Si $\Delta_\alpha = -4$, c'est-à-dire $\alpha = 1728$ et $|\Delta| \geq 10^{15}$, puis

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| > \frac{|\Delta|^{1/2}}{2};$$

(3) Si $\Delta_\alpha = -3$, c'est-à-dire $\alpha = 0$ et $|\Delta| \geq 10^{15}$, puis

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x)| > \frac{|\Delta|^{1/2}}{20}.$$

Les notations sont expliquées dans la section 5.1.

L'idée de prouver Théorème 5.1.1 vient de [7]. Premièrement, nous donnons une borne inférieure effective de $\mathcal{C}_\varepsilon(\tau, \Delta)$, voir Section 5.2 pour la définition et le résultat. Ensuite, en utilisant cette borne et la borne inférieure pour la différence de deux modules singuliers de [6], nous parvenons à donner une borne supérieure pour la hauteur de la différence, voir Corollary 5.3.2 dans Section 5.3. La limite inférieure de la hauteur de la différence provient de [7], voir la Section 5.4. Avec ces deux bornes, en estimant chaque terme des deux côtés, on en déduit Théorème 5.1.1, voir Section 5.5, 5.6, 5.7.

Voici une remarque, puisque Bilu, Habegger et Kühne [7] ont donné la plupart des résultats dont nous avons besoin pour le cas où $\tau = \zeta_6$, c'est-à-dire $\Delta_\alpha = -3$ dans Théorème 5.1.1 (3), nous utiliserons directement leur résultat et nous concentrerons principalement sur le cas où $\tau \neq \zeta_6$.

Les conducteurs d'Artin et les discriminants des courbes hyperelliptiques

Soit R un anneau de valuation discrète avec valuation v, corps residuel k parfait et corps de fraction K, X un schéma propre, plat et régulier sur R. Le conducteur d'Artin de X est défini comme

$$\mathrm{Art}(X) := \chi(X_K) - \chi(X_k) - \delta(X),$$

où $\chi(X_K)$ et $\chi(X_k)$ sont les caractéristiques d'Euler de respectivement X_K et X_k par rapport à la topologie étale, et $\delta(X)$ est le conducteur de Swan associé à représentation ℓ-adique $\mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}_{\mathbb{Q}_\ell}(\mathrm{H}^1_{\text{ét}}(X_K, \mathbb{Q}_\ell))$, $\ell \neq \mathrm{Car}(k)$, voir Section 7.1 pour plus de détails. Le conducteur Artin est une quantité pour mesurer la dégénérescence de X: c'est un entier non positif et $\mathrm{Art}(X) = 0$ si et seulement si X/R est lisse ou $g(X_K) = 1$ et $(X_k)_{\text{red}}$ est lisse. Il est également utilisé pour construire l'équation fonctionnelle de la fonction L associée à X, voir [52] ou [13] pour plus de détails.

Pour une courbe elliptique C, considérons son modèle régulier minimal X, nous avons la formule d'Ogg-Saito [48]:

$$-\mathrm{Art}(X) = v(\Delta(C)),$$

où $v(\Delta(C))$ est le valeur du discriminant minimal de C. Pour une courbe hyperelliptique C, nous avons également la définition du discriminant minimal $v(\Delta(C))$, voir Définition 6.3.2. Cependant, cette formule n'est pas vraie pour toutes les courbes hyperelliptiques. Dans [39], Liu a prouvé que si $\mathrm{Car}(k) \neq 2$ et le genre $g(C) = 2$, alors

$$-\mathrm{Art}(X) \leq v(\Delta(C)),$$

et l'égalité peut ne pas tenir dans certains cas. Dans [61] et [62], Srinivasan a montré que l'inégalité est vraie dans les cas suivants:

(1) les points de Weierstrass de C sont K-rationnels;

(2) $\mathrm{Car}(k) \geq 2g(C) + 1$.

Enfin, Obus et Srinivasan [47] ont montré que cette inégalité est valable pour toute les courbes hyperelliptiques lorsque $\mathrm{Car}(k) \neq 2$.

Dans cette partie, nous prouvons en fait le processus inductif dans le article d'Obus et Srinivasan [47].

Théorèm 7.2.1. *Soit R un anneau de valuation discrète avec corps de fraction K et corps residuel k parfait. Supposons que R est strictement hensélien et $\mathrm{Car}(k) \neq 2$. Soit Y, Y_1 et Y_2 des modèles de Weierstrass sur R définis par des équations de Weierstrass intégrales dans l'un des cas suivants:*

1. *$Y : y^2 = f_1(x)f_2(x)$, $Y_1 : y^2 = f_1(x)$ et $Y_2 : y^2 = f_2(x)$,*
2. *$Y : y^2 = \pi f_1(x)f_2(x)$, $Y_1 : y^2 = \pi f_1(x)$ et $Y_2 : y^2 = \pi f_2(x)$,*

où, dans les deux cas, $\deg(f_i) = \deg(\overline{f}_i) \geq 1$ pour $i = 1, 2$ et $\overline{f}_1, \overline{f}_2 \in k[x]$ sont coprimes. Si pour $i = 1, 2$,

$$-\mathrm{Art}(X_i) - \delta(X_i) \leq v(\Delta(Y_i)),$$

puis

$$-\mathrm{Art}(X) - \delta(X) \leq v(\ Delta(Y)),$$

où X, X_1 et X_2 sont les désingularisations minimales de Y, Y_1 et Y_2 respectivement. De plus, si l'égalité est valable pour Y_1 et Y_2, elle est également valable pour Y.

C'est un travail indépendant et la méthode est différente de celle du article d'Obus et Srinivasan [47]. Avec ce résultat, nous pouvons prouver le cas (1) des resultats de Srinivasan. Bien que ce résultat soit plus faible que celui d'Obus et de Srinivasan dans [47], nous avons encore quelque chose d'intéressant dans la preuve. Nous pouvons calculer certaines quantités importantes de Y dans Théorème 7.2.1 à partir de celles de Y_1 et Y_2, telles que le rang abélien, le rang torique, etc., voir Théorème 7.2.2 et Théorème 7.2.3.

Chapitre 6 est le premier chapitre de cette partie, il donne des résultats de base pour les courbes hyperelliptiques. Dans la première section de Chapitre 7, nous définissons les conducteurs d'Artin des variétés arithmétiques à partir de ses représentations ℓ-adic correspondantes, et collectons quelques résultats pour les conducteurs d'Artin. Le reste de Chapitre 7 est consacré à prouver Théorème 7.2.1 et Corollaire 7.2.4, nous construisons des recourvements étale pour associer ces trois courbes hyperelliptiques et donner des relations entre quantités correspondantes.

Introduction

This Thesis focuses on three different topics, so it is divided into three parts. In the first part, we consider the integral points on modular curves. The second part is devoted to giving a bound of difference of two singular moduli in terms of discriminants. The third part is more algebraic, we study the Artin conductors and discriminants of hyperelliptic curves.

Integral Points on Modular Curves

Let X be a smooth, connected projective algebraic curve defined over a number field K, and let $x \in K(X)$ be a non-constant rational function on X. If R is a subring of K, we denote by $X(R, x)$ the set of R-integral K-rational points of X with respect to the coordinate x:

$$X(R, x) = \{P \in X(K) \mid x(P) \in R\}.$$

In particular, if S is a finite set of places of K (including all the infinite places), we consider the set of *S-integral points* $X(\mathcal{O}_S, x)$, where $\mathcal{O}_S = \mathcal{O}_{S,K}$ is the ring of S-integers in K.

According to the classical theorem of Siegel [57] (see also [33, Part D] for a modern exposition), the set $X(\mathcal{O}_S, x)$ is finite if at least one of the following conditions is satisfied:

$$g(X) \geq 1; \tag{3}$$

$$x \text{ admits at least 3 poles in } X(\bar{\mathbb{Q}}). \tag{4}$$

The theorem of Faltings [24] (see also [33, Part E]) asserts that $X(K)$ is finite if $g(X) \geq 2$. Unfortunately, all known proofs of the theorem of Siegel and Faltings are non-effective, which means that they do not imply any explicit expression bounding the heights of integral or rational points.

Starting from the ground-breaking work of A. Baker in 1960th, effective proofs of Siegel's theorem were discovered, by Baker and others, for many pairs (X, x), see [4, 5] and the references therein.

One interesting case is when $X = X_\Gamma$ is the modular curve corresponding to a subgroup Γ of $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$, and $x = j$ is the rational function defined by the j-invariant. This problem has been considered by Bilu[5], Bilu and Parent[9] [10], Sha [55][54] and many others. In particular, Bilu and Parent solve the split Cartan case of Serre's uniformity problem in [10] by consider this problem for X_{split}.

Bilu [4, Section 5] (see also [5, Section 4]) made the following observation.

Proposition 1. *Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$ of level N having at least 3 cusps. Let K be a number field such that X_Γ admits a geometrically irreducible model over K and such that $j \in K(X_\Gamma)$. Let S be a finite set of places of K containing all the infinite places. Then there exists an effective constant $c = c(N, K, S)$ such that for any $P \in X_\Gamma(\mathcal{O}_S, j)$ we have $\mathrm{h}(j(P)) \leq c$.*

(Here $\mathrm{h}(\cdot)$ is the standard absolute logarithmic height defined on the set $\bar{\mathbb{Q}}$ of algebraic numbers.)

In other words, if condition (4) is satisfied for the couple (X_Γ, j), then Siegel's theorem is effective for this couple.

Sha [55] made the bound in Proposition 1 totally explicit. To state his result, we introduce some notations. For a congruence subgroup Γ as above, the number of cusps on X_Γ is denoted by $v_\infty(\Gamma)$. For a number field K, let M_K be the set of all places of K, and $S \subseteq M_K$ a finite subset containing all infinite places. We put $d = [K : \mathbb{Q}]$ and $s = |S|$. Let $\mathcal{O}_K$ be the ring of integers of K. We define the following quantity

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

as a function of $N \in \mathbb{N}^+$, where D is the absolute discriminant of K, $\varphi(N)$ is Euler's totient function, and the norm $\mathcal{N}_{K/\mathbb{Q}}(v)$ of a place v, by definition, is equal to $\#(\mathcal{O}_K/\mathfrak{p}_v)$ when v is finite and $\mathfrak{p}_v$ is its corresponding prime ideal, and is set to be 1 if v is infinite.

Sha [55] proved the following theorem.

Theorem 3.2.1 ([55] Theorem 1.2)**.** *Let Γ be of level N and X_Γ be the corresponding modular curve over a number field K with $d = [K : \mathbb{Q}]$, and $S \subseteq M_K$ be a finite set containing all infinite places. If $v_\infty(\Gamma) \geq 3$, then for any $P \in X_\Gamma(\mathcal{O}_S, j)$,*

$$\mathrm{h}(j(P)) \leq (CdsM^2)^{2sM}(\log(dM))^{3sM}\ell^{dM}\Delta(M),$$

where C is an absolute effective constant, ℓ is the maximal prime such that there exists $v \in S$ with $v|\ell$, or $\ell = 1$ if S only contains infinite places, and M is defined as following:

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

Here we only notice that Sha's bound is of the shape $c(K,S)^{N\log N}$, where $c(K,S)$ is an effective constant depending only on K and S. Roughly speaking, we have here exponential type dependence in N.

For certain applications it is useful to have an explicit value of the constant C from Theorem 3.2.1. In this part we prove the following result.

Theorem 3.2.2. *The constant C in Theorem 3.2.1 can be taken to be 2^{14}.*

In the proof, we follow the main lines of Sha's argument, with some minor modifications. We calculate explicitly the implicit constants occurring therein, including the Baker's inequality, Theorem 1.3.1.

Also Proposition 1 applies in many important cases: see [5, 8] for further details. In particular, it applies to the modular curve $X_0(N)$ of composite level N. However, it does not directly apply to the curve $X_0(p)$ of prime level p, because it has only 2 cusps.

Nevertheless, using a covering argument, Bilu [5, Theorem 10] proved that Siegel's theorem is effective for $X_0(p)$ as well. Note that the curve $X_0(N)$ has a standard geometrically irreducible model over $\mathbb{Q}$.

Theorem 1 (Bilu). *Let p be a prime number distinct from $2, 3, 5, 7, 13$. Let K be a number field and S be a finite set of places of K containing all the infinite places. Then there exists an effective constant $c = c(p, K, S)$ such that for any $P \in X_0(p)(\mathcal{O}_S, j)$ we have $\mathrm{h}(j(P)) \leq c$.*

The main tool is the classical *Chevalley-Weil Principle*, or so-called Chevalley-Weil Theorem in this thesis, used in the following form.

Proposition 2 (Chevalley-Weil Principle). *Let $\widetilde{X} \xrightarrow{\pi} X$ be a non-constant étale morphism of projective algebraic curves defined over a number field K. Then there exists a finite set T of places of K such that the following holds. Let $P \in X(\bar{K})$ and let $\widetilde{P} \in \widetilde{X}(\bar{K})$ be such that $\pi(\widetilde{P}) = P$. Let v be a finite place of the field $K(P)$ ramified in $K(\widetilde{P})$. Then v extends a place from T.*

We will discuss this theorem in Section 1.5.

Bilu found a subgroup $\widetilde{\Gamma}$ of $\Gamma_0(p)$ such that the natural morphism $X_{\widetilde{\Gamma}} \to X_0(p)$ is étale and $X_{\widetilde{\Gamma}}$ has at least three cusps, see Proposition 3.3.2 for the details. The Chevalley-Weil principle now allows one to reduce the problem from $X_0(p)$ to $X_{\widetilde{\Gamma}}$, where Proposition 1 applies.

In [54] Sha gave an explicit version of Theorem 1. We do not reproduce here Sha's full statement, which is very involved, and only focus on the dependence on the level p. One can expect here exponential type dependence in p, but Sha obtains an upper bound of the form $c(K, S)^{\exp(p^6 \log p)}$, doubly exponential in p.

Sha's bound is so big because he uses a quantitative version of the Chevalley-Weil Theorem from [12], see Proposition 1.5.4, which provides extremely high upper bounds for the quantities involved.

In this thesis, we will prove another version of the Chevalley-Weil Theorem, Proposition 1.5.5, combined with Igusa's theorem, see [22, Section 8.6], we manage to improve the result of Sha. We will prove the following theorem.

Theorem 3.3.1. *Keep the notations of Theorem 1. Then for $P \in X_0(p)(\mathcal{O}_S, j)$, we have*

$$\mathrm{h}(j(P)) \leq e^{9s^2 p^4 \log p} C(K, S)^{p^2},$$

where $C(K, S)$ can be effectively determined in terms of K and S. More explicitly, $C(K, S)$ can be chose as

$$C(K, S) = 2^{29s} d^{9s} s^{2s} \ell^d |D| (\log(|D| + 1))^d \prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v),$$

where $d = [K : Q]$, D is the absolute discriminant of K, $s = \#S$, and ℓ is the maximal prime such that there exists $v \in S$ with $v \mid \ell$.

For a number field K, $v \in M_K$, we define the valuation $|\cdot|_v$ on K as following: for any $\alpha \in K$:

$$|\alpha|_v := |\sigma(\alpha)|, \text{ if } v \text{ is infinite with embedding } \sigma;$$

$$|\alpha|_v := \mathcal{N}_{K/\mathbb{Q}}(v)^{-\mathrm{ord}_v(\alpha)/[K_v:\mathbb{Q}_v]}, \text{ if } v \text{ is finite.}$$

Singular Moduli

Let $\mathbb{H}$ be the upper half plane, a point $\tau \in \mathbb{H}$ is called a CM-point if $\mathrm{End}(E_\tau)$ is an order in an imaginary quadratic field, where E_τ is the ellptic curve over $\mathbb{C}$ corresponding to τ. It is well-known that $\tau \in \mathbb{H}$ is CM if and only if τ is algebraic number

of degree 2. We call $j(\tau)$ a singular modulus if $\tau \in \mathbb{H}$ is CM. From the classical CM-theory, we know that every singular modulus is an algebraic integer. We call $j(\tau)$ singular unit if it is a singular modulus and an algebraic unit.

In [31], Habegger proved that there is at most finitely many singular units. However his proof is ineffective. After this, in [7], Bilu, Habegger and Kühne prove that there is no singular units. Indeed, their method can be generalized to give a effective bound of norm of difference between two singular moduli, that is exactly what we do in this thesis.

On the other hand, Gross and Zagier [27] stated explicit formula for absolute norm of difference between two singular moduli. With their works, Li [37] also managed to give a bound of norm of difference between two singular moduli, his bound is a strictly positive number, which allows him to prove a generalized version of the main result of Bilu, Habegger and Kühne [7]. However, it is not clear how his bound behaves as $\Delta \to -\infty$. In this thesis, we are going to prove the following result:

Theorem 5.1.1. *Let α, x be two singular moduli of discriminants Δ_α, Δ respectively, and $K = \mathbb{Q}(\alpha, x)$.*

(1) If $\Delta_\alpha \neq -3, -4$ and $|\Delta| \geq \max\{e^{3.12}(\mathcal{C}(\Delta_\alpha)|\Delta_\alpha|^4 e^{h(\alpha)})^3, 10^{15} \cdot \mathcal{C}(\Delta_\alpha)^6\}$, then

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)| > \frac{|\Delta|^{1/2}}{2};$$

(2) If $\Delta_\alpha = -4$, i.e. $\alpha = 1728$, and $|\Delta| \geq 10^{15}$, then

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| > \frac{|\Delta|^{1/2}}{2};$$

(3) If $\Delta_\alpha = -3$, i.e. $\alpha = 0$, and $|\Delta| \geq 10^{15}$, then

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x)| > \frac{|\Delta|^{1/2}}{20}.$$

The notations are explained in Section 5.1.

The idea of proving Theorem 5.1.1 is from [7]. Firstly, we give an effective lower bound of $\mathcal{C}_\varepsilon(\tau, \Delta)$, see section 5.2 for the definition and result. Then by using this bound and the lower bound for the difference of two singular moduli from [6], we manage to give an upper bound for the height of difference, see Corollary 5.3.2 in section 5.3. The lower bound for height of difference comes from [7], see section 5.4. With these two bounds, by estimating each term in the both sides, we deduce Theorem 5.1.1, see section 5.5, 5.6, 5.7.

Here is a remark, since Bilu, Habegger and Kühne [7] have given most of results we need for the case where $\tau = \zeta_6$, i.e. $\Delta_\alpha = -3$ in Theorem 5.1.1 (3), we will use their result directly and focus mainly on the case where $\tau \neq \zeta_6$.

The Artin Conductors and Discriminants of Hyperelliptic Curves

Let R be a discrete valuation ring with valuation v, perfect residue field k and fraction field K, X a proper, flat, regular scheme over R. The Artin conductor of X is defined as

$$\mathrm{Art}(X) := \chi(X_K) - \chi(X_k) - \delta(X),$$

where $\chi(X_K)$ and $\chi(X_k)$ are Euler's characteristic of X_K and X_k with respect to étale topology respectively, and $\delta(X)$ is the Swan conductor associated to ℓ-adic representation $\mathrm{Gal}(\overline{K}/K) \to \mathrm{Aut}_{\mathbb{Q}_\ell}(\mathrm{H}^1_{\text{ét}}(X_K, \mathbb{Q}_\ell))$, $\ell \neq \mathrm{Char}(k)$, see Section 7.1 for full details. Artin conductor is quantity to measure degeneracy of X: it is a non-positive integer and $\mathrm{Art}(X) = 0$ if and only if X/R is smooth or $g(X_K) = 1$ and $(X_k)_{\mathrm{red}}$ is smooth. It is also used to construct the functional equation of L-function associated to X, see [52] or [13] for details.

For an elliptic curve C, consider its minimal regular model X, we have the Ogg-Saito formula [48]:

$$-\mathrm{Art}(X) = v(\Delta(C)),$$

where $v(\Delta(C))$ is the valuation of minimal discriminant of C. For a hyperelliptic curve C, we also the definition of minimal discriminant $v(\Delta(C))$, see Definition 6.3.2. However, this formula is not true for all hyperelliptic curves. In [39], Liu proved that if $\mathrm{Char}(k) \neq 2$ and the genus $g(C) = 2$, then

$$-\mathrm{Art}(X) \leq v(\Delta(C)),$$

and the equality may fail to hold for some cases. In [61] and [62], Srinivasan showed that the inequality hold for following cases:

(1) the Weierstrass points of C are K-rational;

(2) $\mathrm{Char}(k) \geq 2g(C) + 1$.

Finally, Obus and Srinivasan [47] showed that this inequality holds for any hyperelliptic curve when $\mathrm{Char}(k) \neq 2$.

In this part, we actually prove the inductive process in Obus and Srinivasan's paper [47].

Theorem 7.2.1. *Let R be a discrete valuation ring with fraction field K and perfect residue field k. Assume that R is strictly henselian and $\mathrm{Char}(k) \neq 2$. Let Y, Y_1 and Y_2 be the Weierstrass models over R defined by integral Weierstrass equations in one of the following cases:*

1. *$Y : y^2 = f_1(x)f_2(x)$, $Y_1 : y^2 = f_1(x)$ and $Y_2 : y^2 = f_2(x)$,*
2. *$Y : y^2 = \pi f_1(x)f_2(x)$, $Y_1 : y^2 = \pi f_1(x)$ and $Y_2 : y^2 = \pi f_2(x)$,*

where, in both cases, $\deg(f_i) = \deg(\overline{f}_i) \geq 1$ for $i = 1, 2$, and $\overline{f}_1, \overline{f}_2 \in k[x]$ are coprime. If for $i = 1, 2$,

$$-\mathrm{Art}(X_i) - \delta(X_i) \leq v(\Delta(Y_i)),$$

then

$$-\mathrm{Art}(X) - \delta(X) \leq v(\Delta(Y)),$$

where X, X_1 and X_2 are the minimal desingularizations of Y, Y_1 and Y_2 respectively. Moreover, if the equality holds for Y_1 and Y_2, it also holds for Y.

It is an independent work and the method is different from the one in Obus and Srinivasan's paper [47]. . With this result, we can prove the case (1) of Srinivasan's result. Although this result is weaker than Obus and Srinivasan's result in [47], we still have something interesting in the proof. We are able to calculate some important quantities of Y in Theorem 7.2.1 from the ones of Y_1 and Y_2, such as the abelian rank, the toric rank etc, see Theorem 7.2.2 and Theorem 7.2.3.

Chapter 6 is the first chapter of this part, it gives basic results for hyperelliptic curves. In first section of Chapter 7, We define the Artin conductors of arithmetic

varieties from its corresponding ℓ-adic representations, and collect some results for Artin conductors. The rest of Chapter 7 is devoted to prove Theorem 7.2.1 and Corollary 7.2.4, we construct étale converings to associate these three hyperelliptic curves and give relations between corresponding quantities.

Notations

We will use the $\mathbb{N}, \mathbb{N}^+, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ for the set of non-negative integers, the set of positive integers, the ring of integers, and the fields of rational, real and complex numbers, respectively.

For a field K, we denote K^* the set of nonzero elements of K, $\overline{K}$ a fixed algebraic closure of K, G_K the absolute Galois group of K. In particular, $\overline{\mathbb{Q}}$ is the field of all algebraic numbers. We also denote $\overline{\mathbb{Z}}$ the ring of all algebraic integers.

For a (commutative) ring R (with identity), and n a positive number, we use $\mathrm{M}_n(R), \mathrm{GL}_n(R), \mathrm{SL}_n(R)$ for the set of $n \times n$ matrices over R, the general linear group of degree n over R, the special linear group of degree n over R, respectively.

For $x \in \mathbb{R}$, we set

$\lfloor x \rfloor, [x]$ – the maximal integer which is smaller than or equal to x;
$\lceil x \rceil$ – the minimal integer which is bigger than or equal to x.

Since Part 3 has a rather different topic than Part 1 and Part 2, we would like to separate their notations to avoid misunderstanding.

In Part 1 and Part 2, for a number field K, denote:

d_K = $[K : \mathbb{Q}]$;
$\mathcal{O}_K$ = the ring of integers of K;
D_K – the absolute discriminant of K;
M_K – the set of all places of K;
M_K^∞ – the set of all Archimedean places of K;
M_K^0 – the set of all non-Archimedean places of K;
$\mathcal{C}_K$ – the class number of K (in order to distinguish the notation of height);
K_v – the completion of K with respect to a place v;
$\mathcal{N}_{K/\mathbb{Q}}(v)$ – the norm of v.

For a finite set S of place of K containing all archimedean places, denote:

$\mathcal{O}_S$ = the ring of S-integers of K;
$R(S)$ – the S-regulator, see Definition 1.2.1;

For a finite extension L/K, denote

$d_{L/K}$ = $[K : \mathbb{Q}]$;
$D_{L/K}$ – the relative discriminant of L/K (which is an ideal of $\mathcal{O}_K$);
$\mathcal{N}_{L/K}$ – the relative norm map of L/K;

For an order $\mathcal{O}$ in K, denote:

$I(\mathcal{O})$ – the group of invertible fractional $\mathcal{O}$-ideals;
$\mathrm{Cl}(\mathcal{O})$ = the class group (or Picard group) of $\mathcal{O}$;

For an element $\alpha \in \overline{\mathbb{Q}}$, denote $\mathrm{h}(\alpha)$ its absolute height function, see Definition 1.1.1.

We will use $\mathbb{H}$ for the upper half plane, and $\mathbb{H}^* = \mathbb{H} \bigcup \mathbb{Q} \bigcup \{\infty\}$. For positive integers $N \in \mathbb{N}^+$ and $k > 2$ denote:

ζ_N = $e^{2\pi i/N}$;
$\zeta(s)$ – the Riemann zeta function;
$\Gamma(N)$ = $\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod N\right\}$;
$\Gamma_1(N)$ = $\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \bmod N\right\}$;
$\Gamma_0(N)$ = $\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod N\right\}$;
$G_k(\tau)$ – the k-th Eisenstein series, see Definition 2.1.10;
$g_2(\tau)$ = $60G_4(\tau)$;
$g_4(\tau)$ = $140G_6(\tau)$;
$\Delta(\tau)$ – the discriminant function, i.e. $\Delta(\tau) = (g_2(\tau))^3 - 27(g_3(\tau))^2$;
$j(\tau)$ – the j-invariant, i.e. $j(\tau) = 1728\frac{(g_2(\tau))^3}{\Delta(\tau)}$.

For $\tau \in \mathbb{H}$, and a lattice Λ in $\mathbb{C}$,
Λ_τ = $\langle \tau, 1\rangle$, a lattice in $\mathbb{C}$ generated by τ and 1;
$\wp_\tau(z)$ – the Weierstrass function with respect to Λ_τ;
$\sigma(z;\Lambda)$ – the Weierstrass sigma function with respect to a lattice, see Definition 3.1.1;
$\zeta(z;\Lambda)$ – the Weierstrass zeta function, see also Definition 3.1.1;
$\eta(z;\Lambda)$ – the Weierstrass eta function, see also Definition 3.1.1;

In Part 3, we use the following notations:
$\mathbb{Z}_\ell$ = $\varprojlim_n \mathbb{Z}/\ell^n\mathbb{Z}$
$\hat{\mathbb{Z}}$ = $\varprojlim_n \mathbb{Z}/n\mathbb{Z}$
G_K – Abosolute Galois Group of a field K

When K is a field with discrete valuation:
v_K – the discrete valuation K;
$\mathcal{O}_K$ – the ring of integer of K;
$\mathfrak{m}$ – the maximal ideal of $\mathcal{O}_K$;
$k(v_K)$ – the residue field of $\mathcal{O}_K$;
G_K – the absolute Galois group.

For a projective curve C over a field K, we set:
$\mathrm{n}(C)$ – the number of irreducible components of C;
$\mathrm{p}_a(C)$ – the arithmetic genus of C;
$\mathrm{a}(C)$ – the abelian rank of C;
$\mathrm{t}(C)$ – the toric rank of C;
$\mathrm{u}(C)$ – the unipotent rank of C.

Part I

Integral Points on Modular Curves

Chapter 1

Integral Points on Algebraic Curves

This chapter provides some results we need in Diophantine approximation, the monograph [14] of Bombieri, Gubler and the doctoral thesis [3] of Bilu will be good references. The new results in this chapter will be an explicit version of Baker's inequality, Theorem 1.3.1 and a quantitative Chevalley-Weil Theorem for curves, Proposition 1.5.6.

1.1 Heights on $\overline{\mathbb{Q}}$

We recall some notations in algebraic number theory. For a number field K, $v \in M_K$, the norm of v is defined as

$$\mathcal{N}_{K/\mathbb{Q}}(v) := \begin{cases} \#(\mathcal{O}_K/\mathfrak{p}_v) & \text{if } v \text{ is finite;} \\ 1 & \text{if } v \text{ is infinite.} \end{cases}$$

The normalized valuation $||\cdot||_v$ on K as following: for any $\alpha \in K$:

$$||\alpha||_v := \begin{cases} |\sigma(\alpha)|^{[K_v:\mathbb{R}]} & \text{if } v \text{ is infinite with embedding } \sigma; \\ \mathcal{N}_{K/\mathbb{Q}}(v)^{-\mathrm{ord}_v(\alpha)} & \text{if } v \text{ is finite.} \end{cases}$$

We also define $|\cdot|_v := ||\cdot||_v^{1/[K_v:\mathbb{Q}_v]}$.

Definition 1.1.1. *Let K be a number field, we define the (logarithmic) K-height $\mathrm{h}_K : K \to \mathbb{R}_{\geq 0}$ as following: for $\alpha \in K$,*

$$\mathrm{h}_K(\alpha) = \sum_{v \in M_K} \log \max\{1, ||\alpha||_v\} = \sum_{v \in M_K} [K_v : \mathbb{Q}_v] \log \max\{1, |\alpha|_v\},$$

where M_K is the set of places of K.

We define the absolute (logarithmic) height $\mathrm{h} : \overline{\mathbb{Q}} \to \mathbb{R}_{\geq 0}$ as following: for $\alpha \in \overline{\mathbb{Q}}$,

$$\mathrm{h}(\alpha) = \frac{\mathrm{h}_K(\alpha)}{[K : \mathbb{Q}]},$$

where K is a number field containing α, and $\mathrm{h}(\alpha)$ is independent of the choice of K.

We collect the main properties about the height function, the proofs can be found in [14, section 1.5].

Proposition 1.1.2. *Let $\mathrm{h} : \overline{\mathbb{Q}} \to R_{\geq 0}$ be the height function.*

(1) (Galois action) If $\alpha, \beta \in \overline{\mathbb{Q}}$ are conjugate over $\mathbb{Q}$, then $\mathrm{h}(\alpha) = \mathrm{h}(\beta)$.

(2) (Height of a quotient) Let K be an number field, and $\alpha, \beta \in K$ with $\beta \neq 0$. Then

$$\mathrm{h}(\alpha/\beta) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in M_K} \log \max\{||\alpha||_v, ||\beta||_v\}.$$

If moreover, α, β are algebraic integers, then

$$\mathrm{h}(\alpha/\beta) \leq \frac{1}{[K:\mathbb{Q}]} \sum_{\sigma: K \hookrightarrow \mathbb{C}} \log \max\{|\sigma(\alpha)|, |\sigma(\beta)|\}.$$

(3) (Heights of sums and products) Let $f \in \mathbb{Z}[X_1, \cdots, X_n]$ be a non-zero polynomial, and $\alpha_1, \cdots, \alpha_n \in \overline{\mathbb{Q}}$. Then

$$\mathrm{h}(f(\alpha_1, \cdots, \alpha_n)) \leq \log L(f) + \sum_{i=1}^{n} \deg_{X_i}(f) \mathrm{h}(\alpha_i),$$

where $L(f)$ is the sum of the modulus of the coefficients of f. In particular,

$$\mathrm{h}(\alpha_1 \cdots \alpha_n) \leq \mathrm{h}(\alpha_1) + \cdots + \mathrm{h}(\alpha_n)$$

$$\mathrm{h}(\alpha_1 + \cdots + \alpha_n) \leq \mathrm{h}(\alpha_1) + \cdots + \mathrm{h}(\alpha_n) + \log n.$$

(4) (Height of power) For any $\alpha \in \overline{\mathbb{Q}}^*$ and $n \in \mathbb{Z}$, we have

$$\mathrm{h}(\alpha^n) = |n| \mathrm{h}(\alpha).$$

(5) (Height of a linear fraction) Let $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{GL}_2(\overline{\mathbb{Q}})$, then for any $x \in \overline{\mathbb{Q}}$ with $x \neq -\frac{d}{c}$, we have

$$\mathrm{h}(\frac{ax+b}{cx+d}) = \mathrm{h}(x) + C(a,b,c,d),$$

where $C(a,b,c,d)$ is an effective constant.

(6) (Northcott's finiteness theorem) For any $C > 0$, there exist only finitely many algebraic number α of degree and height bounded by C.

(7) (Kronecker's first theorem) For $\alpha \in \overline{\mathbb{Q}}$, then $\mathrm{h}(\alpha) = 0$ if and only if $\alpha = 0$ or α is a root of unity.

(8) (Kronecker's second theorem) For any positive integer d, there exists $\varepsilon(d) > 0$ with the following property: for any $\alpha \in \overline{\mathbb{Q}}$ of $\deg(\alpha) \leq d$, we have $\mathrm{h}(\alpha) = 0$ or $\mathrm{h}(\alpha) \geq \varepsilon(d)$.

(9) (Liouville's inequality) For a number field K, a subset $S \subset M_K$ and $\alpha \in K^*$, we have

$$\sum_{v \in S} \log ||\alpha||_v \geq -[K:\mathbb{Q}] \mathrm{h}(\alpha).$$

1.2 Siegel's Theory of Convenient Units

In this section, we recall some useful results on S-units when we use Baker's method to calculate integral points on algebraic curves, see [3, section 1.4] for more details.

For a number field K, and a finite subset $S \subseteq M_K$ containing all archimedean places, we put $d = [K : \mathbb{Q}]$ and $s = |S|$, $r = s - 1$. Let $\mathcal{O}_K$ (resp. $\mathcal{O}_S$) be the ring of integers (resp. S-integer) in K.

Definition 1.2.1. *If $s \geq 2$, we fix a $v_0 \in S$, set $S' = S \setminus \{v_0\} = \{v_1, \cdots, v_r\}$, the S-regulator $R(S)$ is defined as*

$$R(S) = |\det(d_{v_i} \log |\eta_j|_{v_i})_{1 \leq i,j \leq r}|,$$

where $d_{v_i} = [K_{v_i} : \mathbb{Q}_{v_i}]$ is the local degree of v_i for each i, and $\{\eta_1, \cdots, \eta_r\}$ is a fundamental system of the S-units.

If $s = 1$, then there is no fundamental system, and we define $R(S) = 1$.

If S consists of all archimedean places, then the S-regulator is the regulator R_K of K.

The value $R(S)$ is a positive number which is independent of the choice ot v_0 and the fundamental system of S-units. We also set ω_K the number of roots of unit in K.

By Kronecker's second theorem, see (8) of Proposition 1.1.2, we can take $\zeta > 0$ such that $\mathrm{h}(\alpha) \geq \frac{1}{d\zeta}$ for any $\alpha \in K \setminus \{0\}$ which is not a root of unity. By [65, Theorem and the Corollary 2], ζ can be taken to be

$$\zeta = \begin{cases} \log 2 & \text{if } d = 1 \\ \dfrac{(\log 6)^3}{2} & \text{if } d = 2 \\ 4\left(\dfrac{\log d}{\log \log d}\right)^3 & \text{if } d \geq 3. \end{cases}$$

LEMMA 1.2.2. *Let $\beta_1, \cdots, \beta_m \in K$ be multiplicatively independent elements such that $|\beta_i|_v = 1$ for any $i = 1, \cdots, m$ and $v \notin S$. Then*

(1) the group $\Gamma = \{\beta_1^{n_1} \cdots \beta_m^{n_m} \mid n_1, \cdots n_m \in \mathbb{Z}\}$ is a free abelian group of rank m, and $m \leq s$;

(2) for any $\alpha = \beta_1^{b_1} \cdots \beta_m^{b_m} \in \Gamma$, we have

$$\max\{|b_1|, \cdots, |b_m|\} \leq 2d\mathrm{Ch}(\alpha),$$

where $C = C(\beta_1, \cdots, \beta_m)$ is a constant. More precisely, if $\Omega \in \mathrm{M}_m(\mathbb{R})$ (which is in $\mathrm{GL}_m(\mathbb{R})$ in fact) is any submatrix of $(d_{v_i} \log |\beta_j|_{v_i})_{0 \leq i \leq r, 1 \leq j \leq m}$ with $\Omega^{-1} = (a_{ij})_{1 \leq i,j \leq m}$, then C can be taken to be $\max\limits_{1 \leq i,j \leq m} \{|a_{i,j}|\}$.

Proof. Since Γ is torsion free, so Γ is free. Consider

$$l : \Gamma \to \mathbb{R}^s,\ \alpha \mapsto (d_{v_i} \log |\alpha|_{v_i})_{i=0,\cdots,r},$$

which is a injective map of groups. Indeed, if $\log |\alpha|_{v_i} = 0$ for any $0 \leq i \leq r$, then $\mathrm{h}(\alpha) = \frac{1}{d} \sum\limits_{i=0}^{r} d_{v_i} \max\{0, \log |\alpha|_{v_i}\} = 0$, which means that α is root of unity by Kronecker's first theorem. Since Γ is free, so $\alpha = 1$.

Let $w_j = (d_{v_i} \log |\beta_j|_{v_i})_{i=1,\cdots,r}$, $j = 1, \cdots, m$. To show that the rank of Γ is m, it is sufficient to show that $w_1, \cdots, w_m$ are linearly independent over $\mathbb{Q}$. Indeed, if

$\sum_{j=1}^{m} k_j w_j = 0$, for some $k_j \in \mathbb{Q}$. After multiplying an integer, we can assume that $\sum_{j=1}^{m} k_j d_{v_i} \log|\beta_j|_{v_i} = 0$ for some $k_j \in \mathbb{Z}$. Let $\beta = \prod_{j=1}^{m} \beta_j^{k_j}$. As before, we know that $\beta = 1$, so $k_j = 0$ for any $1 \le j \le m$. Since $l(\Gamma)$ is a lattice in $\mathbb{R}^s$, so $m \le s$.

For (2), let $T = (d_{v_i} \log|\beta_j|_{v_i})_{0\le i\le r, 1\le j\le m} \in M_{s\times m}(\mathbb{R})$. Then

$$\begin{pmatrix} d_{v_0}\log|\alpha|_{v_0} \\ \vdots \\ d_{v_r}\log|\alpha|_{v_r} \end{pmatrix} = T \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Let $\Omega \in M_m(\mathbb{R})$ is a submatrix of $(d_{v_i} \log|\beta_j|_{v_i})_{0\le i\le r, 1\le j\le m}$. Since the rank of T is m, so $\Omega \in \mathrm{GL}_m(\mathbb{R})$. Let $\Omega^{-1} = (a_{ij})_{1\le i,j\le m}$, and $C = \max_{1\le i,j\le m}\{|a_{i,j}|\}$. Then

$$\Omega^{-1} \begin{pmatrix} d_{v_{t_1}}\log|\alpha|_{v_{t_1}} \\ \vdots \\ d_{v_{t_m}}\log|\alpha|_{v_{t_m}} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

$$|b_i| \le C\sum_{j=1}^{m} d_{v_{t_j}} |\log|\alpha|_{v_{t_j}}| \le C\sum_{i=0}^{r} d_{v_i}|\log|\alpha|_{v_i}| = 2d\mathrm{Ch}(\alpha),$$

for any $1 \le i \le m$, where $0 \le t_1 \le \cdots \le t_m \le r$, and we used the fact

$$\mathrm{h}(\alpha) = \frac{1}{2d}\sum_{v\in S} d_v |\log|\alpha|_v|.$$

□

Remark. *(1) We have*

$$0 \le \mathrm{h}(\alpha) \le \max\{|b_1|,\cdots,|b_m|\}\sum_{i=1}^{m}\mathrm{h}(\beta_i),$$

$$0 \le \max\{|b_1|,\cdots,|b_m|\} \le 2d\mathrm{Ch}(\alpha),$$

so bounding $\mathrm{h}(\alpha)$ *and bounding* $\max\{|b_1|,\cdots,|b_m|\}$ *are equivalent if* $\sum_{i=1}^{m}\mathrm{h}(\beta_i)$ *and* C *given above are bounded.*

Proposition 1.2.3 ([23] Proposition 4.3.9)**.** *Let $s \ge 2$. Then there exists a fundamental system of S-units $\eta_1,\cdot,\eta_r$ satisfying the following properties:*

(1) $\mathrm{h}(\eta_1)\cdots\mathrm{h}(\eta_r) \le \frac{(r!)^2}{2^{r-1}d^r}R(S)$;

(2) $(d\zeta)^{-1} \le \mathrm{h}(\eta_i) \le \min\{\frac{s}{2}, \frac{r!}{2^{r-1}}\}\cdot\frac{r!}{d}\zeta^{r-1}R(S)$ *for* $1 \le i \le r$;

(3) if $\eta = \xi\eta_1^{b_1}\cdots\eta_r^{b_r} \in \mathcal{O}_S^*$ *with* $b_1,\cdots,b_r \in \mathbb{Z}$ *and* ξ *a root of unity, then*

$$\mathrm{h}(\eta) \le \max\{|b_1|,\cdots,|b_r|\}\cdot\min\{\frac{s}{2},\frac{r!}{2^{r-1}}\}\cdot r\frac{r!}{d}\zeta^{r-1}R(S),$$

$$\max\{|b_1|,\cdots,|b_r|\} \le d\zeta\frac{(r!)^2}{2^{r-2}}\mathrm{h}(\eta).$$

Proof. Fix a $v_0 \in S$, set $S' = S \setminus \{v_0\} = \{v_1, \cdots, v_r\}$, and define

$$l : \mathcal{O}_S^* \to \mathbb{R}^r,\ \eta \mapsto (d_{v_i} \log |\eta|_{v_i})_{i=1,\cdots,r}.$$

Then the image of l is a lattice Λ of rank r with determinant $R(S)$. By [17, Chapter V 4, Lemma 8], [17, Chapter VIII 1.2 Lemma 1] and [17, Chapter VIII 4.3, Theorem V], or [23, Theorem 4.3.1] and [23, Theorem 4.3.3], for function $F(x) = |x_1| + \cdots + |x_r|$, $x \in \mathbb{R}^r$, there exists a basis $\{w_1, \cdots, w_r\} \subset \Lambda$ such that

$$F(w_j) \leq \max\{1, \frac{j}{2}\}\lambda_j,\ j = 1, \cdots, r,$$

$$\lambda_1 \cdots \lambda_r \leq 2^r \frac{R(S)}{\mathrm{Vol}(\mathrm{B}_{F,1}(0))} = r!R(S),$$

where $\lambda_1, \cdots, \lambda_r$ are the successive minima of F with respect to Λ, and $\mathrm{B}_{F,1}(0) = \{x \in \mathbb{R}^r \mid F(x) < 1\}$, here we use the fact $\mathrm{Vol}(\mathrm{B}_{F,1}(0)) = \frac{2^r}{r!}$.

Let $w_j = (d_{v_i} \log |\eta_j|_{v_i})_{i=1,\cdots,r}$ for some $\eta_j \in \mathcal{O}_S^*, j = 1, \cdots, r$. Then $\{\eta_1, \cdots, \eta_r\}$ is a fundamental system of the S-units, and

$$\prod_{j=1}^{r} F(w_j) \leq 1 \cdot 1 \cdot \frac{3}{2} \cdot \cdots \cdot \frac{r}{2} \cdot \lambda_1 \cdots \lambda_r \leq \frac{(r!)^2}{2^{r-1}} R(S),$$

i.e.

$$\prod_{j=1}^{r} (\sum_{i=1}^{r} d_{v_i} |\log |\eta_j|_{v_i}|) \leq \frac{(r!)^2}{2^{r-1}} R(S).$$

Notice that $\mathrm{h}(\eta_j) = \frac{1}{2d} \sum\limits_{v \in S} d_v |\log |\eta_j|_v|$ and

$$d_{v_0} |\log |\eta_j|_{v_0}| = |\sum_{i=1}^{r} d_{v_i} \log |\eta_j|_{v_i}| \leq \sum_{i=1}^{r} d_{v_i} |\log |\eta_j|_{v_i}|,$$

then

$$\mathrm{h}(\eta_j) \leq \frac{1}{d} \sum_{i=1}^{r} d_{v_i} |\log |\eta_j|_{v_i}|,$$

$$\prod_{j=1}^{r} \mathrm{h}(\eta_j) \leq \frac{1}{d^r} \prod_{j=1}^{r} (\sum_{i=1}^{r} d_{v_i} |\log |\eta_j|_{v_i}|) \leq \frac{(r!)^2}{2^{r-1} d^r} R(S).$$

For (2), by [17, Chapter VIII 1.2 Lemma 1], there exist r linearly independent points $z_j = (d_{v_i} \log |\varepsilon_j|_{v_i})_{i=1,\cdots,r} \in \mathbb{R}^r$, $\varepsilon_j \in \mathcal{O}_S^*$, $j = 1, \cdots, r$, such that

$$F(z_j) = \lambda_j.$$

Then as before, we have

$$\frac{1}{d\zeta} \leq \mathrm{h}(\varepsilon_j) \leq \frac{1}{d} F(z_j),$$

which implies that $\lambda_j \geq 1/\zeta$. Hence, for any $1 \leq i \leq r$,

$$\mathrm{h}(\eta_i) \leq \frac{1}{d} F(w_i) \leq \frac{s}{2d} \lambda_i \leq \frac{s}{2d\lambda_1 \cdots \lambda_{i-1}\lambda_{i+1} \cdots \lambda_r} r!R(S) \leq \frac{s!}{2d} \zeta^{r-1} R(S).$$

On the other hand, since $\mathrm{h}(\eta_j) \geq \frac{1}{d\zeta}$, then for any $1 \leq 1 \leq r$,

$$\mathrm{h}(\eta_i) \leq (d\zeta)^{r-1} \cdot \frac{(r!)^2}{2^{r-1} d^r} R(S) = \frac{(r!)^2}{2^{r-1} d} \zeta^{r-1} R(S).$$

Combining these two bounds, we have (2).

It remains to prove (3). By (3) of Proposition 1.1.2,

$$\begin{aligned}\mathrm{h}(\eta) &\le \sum_{i=1}^{r} |b_i| \mathrm{h}(\eta_i) \\ &\le \max\{|b_1|, \cdots, |b_r|\} \sum_{i=1}^{r} \mathrm{h}(\eta_i) \\ &\le \max\{|b_1|, \cdots, |b_r|\} \cdot \min\{\frac{s}{2}, \frac{r!}{2^{r-1}}\} \cdot r\frac{r!}{d}\zeta^{r-1}R(S).\end{aligned}$$

For the second inequality, let $\Omega = (\Omega_{ij})_{1\le i,j\le r} := (d_{v_i} \log |\eta_j|_{v_i})_{0\le i,j\le r}$ which is invertible, and let $\Omega^{-1} = (a_{ij})_{1\le i,j\le r}$, where $a_{ij} = \frac{\det(\Omega_{ij}^*)}{\det(\Omega)}$ and Ω_{ij}^* is the (i,j)-th entry of the adjoint of Ω. By Hadamard's inequality,

$$|\det(\Omega_{ij}^*)| \le \prod_{\substack{p=1\\p\ne i}}^{r} \sqrt{\sum_{\substack{q=1\\q\ne j}}^{r} \Omega_{qp}^2} \le \prod_{\substack{p=1\\p\ne i}}^{r} (\sum_{\substack{q=1\\q\ne j}}^{r} |\Omega_{qp}|).$$

Since

$$\sum_{\substack{q=1\\q\ne j}}^{r} |\Omega_{qp}| = \sum_{\substack{q=1\\q\ne j}}^{r} d_{v_q} |\log |\eta_p|_{v_q}| \le \sum_{q=1}^{r} d_{v_q} |\log |\eta_p|_{v_q}|,$$

so

$$\begin{aligned}|a_{ij}| &= \frac{|\det(\Omega_{ij}^*)|}{|\det(\Omega)|} \\ &\le \frac{\prod_{p=1}^{r} (\sum_{q=1}^{r} d_{v_q} |\log |\eta_p|_{v_q}|)}{(\sum_{q=1}^{r} d_{v_q} |\log |\eta_i|_{v_q}|) R(S)} \\ &\le \frac{(r!)^2/2^{r-1}R(S)}{\zeta^{-1}R(S)} \\ &= \frac{(r!)^2}{2^{r-1}}\zeta.\end{aligned}$$

By Lemma 1.2.2,

$$\max\{|b_1|, \cdots, |b_m|\} \le 2d \max_{1\le i,j\le r} \{|a_{ij}|\} \mathrm{h}(\eta) \le d\zeta \frac{(r!)^2}{2^{r-2}} \mathrm{h}(\eta).$$

□

Remark. *(1) In [23, Proposition 4.3.9], by [43, Theorem 3], we have*

$$\mathrm{h}(\eta_i) \le 29e\sqrt{r-1}\frac{(r!)^2}{2^{r-1}} \log^+(d) R(S)$$

for $1 \le i \le r$ when $s \ge 3$, where $\log^+(d) = \max\{1, \log(d)\}$.

Proposition 1.2.4. *We have*

$$0.1 \le R(S) \le C_K R_K \prod_{\substack{v\in S\\v\nmid\infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v),$$

$$R(S) \leq \frac{\omega_K}{2}\left(\frac{2}{\pi}\right)^{r_2}\left(\frac{e\log|D|}{4(d-1)}\right)^{d-1}\sqrt{|D|}\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v),$$

where $\mathcal{C}_K$ is the class number of K, r_2 is the number of complex embeddings of K, and D is the absolute discriminant of K.

Proof. For the first inequality see [16, Lemma 3]; one may remark that the lower bound $R(S) \geq 0.1$ follows from Friedman's famous lower bound [26, Theorem B] for the usual regulator R_K. The second one follows from Siegel's estimate [58], or see [44, Theorem 1]

$$\mathcal{C}_K R_K \leq \frac{\omega_K}{2}\left(\frac{2}{\pi}\right)^{r_2}\left(\frac{e\log|D|}{4(d-1)}\right)^{d-1}\sqrt{|D|}\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v),$$

here, we replace $(1/d-1)^{d-1}$ with 1 when $d=1$. □

We will use the following lemma.

LEMMA 1.2.5. *$\omega_K \leq 2d^2$. Moreover, $\omega_K \leq d^2$ if $\zeta_N = e^{2\pi i/N} \in K$ for some $N \geq 6$.*

Proof. It's sufficient to show that $\varphi(n) \geq \sqrt{n}$ for $n \neq 2, 6$.

For $k \geq 1$, set $f_k(x) := x^k - x^{k-1} - x^{k/2}$, $g_k(x) := x^k - x^{k-1} - \sqrt{2}x^{k/2}$. Then

$$f_k(x) = x^{(k-1)/2}(x^{(k-1)/2}(x-1) - x^{1/2}) \geq x - 1 - x^{1/2} > 0,$$

if $x \geq 3$. Similarly, $g_k(x) > 0$ if $x \geq 5$ or $k \geq 2, x \geq 3$.

Let $n = 2^m \prod_p p^{e_p}$, where p runs through all odd prime numbers. If $m = 0$, then

$$\varphi(n) = \prod_{e_p\geq 1}(p^{e_p} - p^{e_p-1}) \geq \prod_{e_p\geq 1} p^{e_p/2} = \sqrt{n}.$$

It is similar for the case where $m \geq 2$.

If $m = 1$, then there exists a prime q such that $q \geq 5, e_q \geq 1$ or $q = 3, e_q \geq 2$. Hence

$$\varphi(n) = \prod_{e_p\geq 1}(p^{e_p} - p^{e_p-1}) \geq \sqrt{2}q^{e_q/2}\prod_{\substack{p\neq q\\ e_p\geq 1}} p^{e_p/2} = \sqrt{n}.$$

□

Proposition 1.2.6 ([4] Proposition 1.4.6)**.** *For any $\alpha \in K$ there exists $\eta \in U_K$ such that $\beta = \alpha\eta^{-1}$ satisfies*

$$\frac{1}{d}\sum_{i=1}^{r} d_{v_i}|\log|\beta|_{v_i}| \leq \frac{s!r!}{2^{r-1}d}\zeta^{r-1}R(S).$$

Proof. Let $\eta_1, \cdot, \eta_r$ be a fundamental system of S-units in Proposition 1.2.3. Denote

$$T = (d_{v_i}\log|\eta_j|_{v_i})_{1\leq i,j\leq r},$$

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix} = T^{-1}\begin{pmatrix} d_{v_1}\log|\alpha|_{v_1} \\ \vdots \\ d_{v_r}\log|\alpha|_{v_r} \end{pmatrix}.$$

Let b_i be the nearest integer of a_i, $1 \le i \le r$, and $\eta = \eta_1^{b_1} \cdot \eta_r^{b_r}$. Then by Proposition 1.2.3 (2), $\beta = \alpha\eta^{-1}$ satisfies

$$\begin{aligned}\frac{1}{d}\sum_{i=1}^{r} d_{v_i}|\log|\beta|_{v_i}| &= \frac{1}{d}\sum_{i=1}^{r} d_{v_i}|\log|\alpha|_{v_i} - \log|\eta|_{v_i}| \\ &\le \frac{1}{d}\sum_{i=1}^{r}\sum_{j=1}^{r} d_{v_j}|a_i - b_i||\log|\eta_i|_{v_j}| \\ &\le \frac{1}{2d}\sum_{i=1}^{r}\sum_{j=1}^{r} d_{v_j}|\log|\eta_i|_{v_j}| \\ &\le \mathrm{h}(\eta_1) + \cdots + \mathrm{h}(\eta_r) \\ &\le \frac{s!r!}{2^{r-1}d}\zeta^{r-1}R(S).\end{aligned}$$

□

1.3 Baker's Inequality

In this section, we state Baker's inequality in the following explicit form.

Theorem 1.3.1 (Baker's inequality)**.** *Let n be a positive integer bigger than 2, K be a number field of degree d. Let $\alpha_1, \cdots, \alpha_n \in K^*$, and $b_1, \cdots, b_n \in \mathbb{Z}$ be such that $\alpha_1^{b_1} \cdots \alpha_n^{b_n} \neq 1$. We define $A_1, \cdots, A_n, B_0$ by*

$$\log A_j := \max\{\mathrm{h}(\alpha_j), 1/d\}, 1 \le j \le n;$$
$$B_0 := \max\{3, |b_1|, \cdots, |b_n|\}.$$

Then for any $v \in M_K$, we have

$$|\alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1|_v \ge \exp\{-\mathrm{Y}\log A_1 \cdots \log A_n \log B_0\}, \tag{1.1}$$

where

$$\mathrm{Y} = \begin{cases} 2^{8n+29}d^{n+2}\log(ed) & \text{if } v|\infty \\ 2^{10n+10}\cdot e^{2n+2}d^{3n+4}p_v^d & \text{if } v|p_v < \infty \end{cases} \tag{1.2}$$

The proof of this theorem is based on [45, Corollary 2.3] and [67, Main Theorem,page 190-191].

For the convenience of readers, we state their results here.

Theorem 1.3.2 ([45], Corollary 2.3)**.** *Let $n \in \mathbb{N}^+$, K be a number field of degree d, $\alpha_1, \cdots, \alpha_n \in K^*$, and $b_1, \cdots, b_n \in \mathbb{Z}$ such that $\Lambda := b_1\log\alpha_1 + \cdots + b_n\log\alpha_n \neq 0$. We define $A_1^*, \cdots, A_n^*, B$ by*

$$\log A_j^* = \max\{\mathrm{h}(\alpha_j), \frac{|\log\alpha_j|}{d}\}, 1 \le j \le n,$$
$$B = \max\{3, \frac{|b_j|\log A_j^*}{\log A_n^*} : 1 \le j \le n\}.$$

Then

$$\log|\Lambda| \ge -C(n,\varkappa)d^{n+2}\log(ed)\log A_1^* \cdots \log A_n^* \log B,$$

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

$$\varkappa = \begin{cases} 1 & \text{if } \alpha_1, \cdots, \alpha_n \in \mathbb{R} \\ 2 & \text{otherwise} \end{cases}$$

Theorem 1.3.3 ([67] consequence of Main Theorem). *Let $n \in \mathbb{N}^+$, K be a number field of degree d, $\alpha_1, \cdots, \alpha_n \in K^*$, and $b_1, \cdots, b_n \in \mathbb{Z}$ such that $\alpha_1^{b_1} \cdots \alpha_n^{b_n} \neq 1$. We define $A_1, \cdots, A_n, B$ by*

$$\log A_j = \max\{\mathrm{h}(\alpha_j), \frac{1}{16ed}\}, 1 \leq j \leq n,$$
$$B_0 = \max\{3, |b_j| : 1 \leq j \leq n\}.$$

Then for each prime number p, and a prime ideal $\mathfrak{p} \subset \mathbb{Q}(\alpha_1, \cdots, \alpha_n)$ over p, we have

$$\mathrm{ord}_{\mathfrak{p}}(\alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1) < C_0(n, d, \mathfrak{p}) \log A_1 \cdots \log A_n \log B_0,$$

where $C_0(n, d, \mathfrak{p}) = (16ed)^{2(n+1)} n^{5/2} \log(2nd) \log(2d) \cdot e_{\mathfrak{p}}^n \frac{p^{f_{\mathfrak{p}}}}{(f_{\mathfrak{p}} \log p)^2}$, and $e_{\mathfrak{p}}, f_{\mathfrak{p}}$ are the ramification index and the residue degree of $\mathfrak{p}$ respectively.

Now we prove Theorem 1.3.1, the idea comes from [66, subsection 9.4.4].

Proof. If $v|p_v$ for some prime p_v, then from Theorem 1.3.3, we have

$$|\alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1|_v > \exp\{-C_1(n, d, \mathfrak{p}) \log A_1 \cdots \log A_n \log B_0\},$$

where $C_1(n, d, \mathfrak{p}) = (f_{\mathfrak{p}} \log p_v) C_0(n, d, \mathfrak{p}) = (16ed)^{2(n+1)} n^{5/2} \log(2nd) \log(2d) \cdot e_{\mathfrak{p}}^n \frac{p_v^{f_{\mathfrak{p}}}}{f_{\mathfrak{p}} \log p_v}$.
We have

$$\begin{aligned} C_1(n, d, \mathfrak{p}) &\leq (16e)^{2(n+1)} d^{2n+2} n^{5/2} \cdot 2nd \cdot 2d \cdot d^n \cdot p_v^d \\ &\leq 2^{10n+10} \cdot e^{2n+2} d^{3n+4} p_v^d, \end{aligned}$$

since $n^{7/2} \leq 4^n$.

If $v|\infty$, set $\log z = \log|z| + i \arg z$, with $-\pi < \arg z \leq \pi$. For $|z| < 1$, we have

$$\log(1+z) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} z^n,$$

and if $|z| \leq 1/2$, we have

$$1 + |z| + |z|^2 + \cdots = \frac{1}{1-|z|} \leq 2,$$

$$|\log(1+z)| \leq |z|(1 + |z| + |z|^2 + \cdots) \leq 2|z|. \tag{1.3}$$

To prove the corollary, without loss of generality, we may assume that $b_i \neq 0$ for $1 \leq i \leq n$, and $A_1 \leq \cdots \leq A_n$, and set $\alpha = \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1$.

(a) If $B_0 \leq 2nd$, with Liouville's inequality, we have

$$\mathrm{h}(\alpha) \leq \log 2 + \sum_{i=1}^{n} |b_i| \mathrm{h}(\alpha_i),$$

$$\log|\alpha| \geq -d\mathrm{h}(\alpha) \geq -d(\log 2 + nB_0 \log A_n),$$

that is

$$|\alpha| \geq \exp\{-(d \log 2 + 2n^2 d^2 \log A_n)\}.$$

Since $1 \leq d \log A_i$ for $1 \leq i \leq n$, and

$$\log 2 + 2n^2 \leq 2^{8n+29} \log(ed)$$

$$d \log 2 + 2n^2 d^2 \log A_n \leq (\log 2 + n^2) d^2 \log A_n \leq \Upsilon \log A_1 \cdots \log A_n \log B_0.$$

Hence we have inequality 1.1.

(b) If $B_0 > 2nd$, and $|\alpha| > 1/2$, since $\log 2 \leq 2^{8n+29}\log(ed)$, it is easy to deduce inequality 1.1 from this as above.

(c) If $B_0 > 2nd$, and $|\alpha| \leq 1/2$, this is main part of the proof. By 1.3, we have

$$|\alpha| \geq \frac{1}{2}|\log(1+\alpha)| = \frac{1}{2}|\log(\alpha_1^{b_1}\cdots\alpha_n^{b_n})| = \frac{1}{2}|\Lambda|,$$

where $\Lambda = b_0\log(-1) + b_1\log\alpha_1 + \cdots + b_n\log\alpha_n$, $b_0 = 2k$ for some integer k. Hence, it is sufficient to bound $|\Lambda|$.

To use Theorem 1.3.2, for $1 \leq i \leq n$, we set

$$\log A_i^* = \sqrt{\pi^2+1}\cdot\log A_i,$$

$$\log A_0^* = \frac{\pi}{d},$$

$$B = B_0^2.$$

We will show that for $1 \leq i \leq n$, we have

$$\log A_i^* \geq \max\{\mathrm{h}(\alpha_i), \frac{|\log\alpha_i|}{d}\},$$

$$\log A_0^* \geq \max\{\mathrm{h}(-1), \frac{|\log(-1)|}{d}\} = \frac{\pi}{d},$$

$$B \geq \max\{3, \frac{|b_j|\log|A_j^*|}{\log A_n^*} : 0 \leq j \leq n\}.$$

Indeed, notice that for $1 \leq i \leq n$, we have

$$|\log\alpha_i|^2 \leq \pi^2 + (\log|\alpha_i|)^2,$$

$$\frac{\log|\alpha_i|}{d} \leq \mathrm{h}(\alpha_i) \leq \log A_i < \log A_i^*,$$

so

$$|\log\alpha_i| \leq (\pi^2 + d^2(\log A_i)^2)^{1/2} \leq \sqrt{\pi^2+1}\cdot d\log A_i.$$

For $\log A_0^*$, it's obvious.

For B, we bound b_0 first. Since $|\alpha| \leq 1/2$, so $|\Lambda| \leq 1$ and

$$\begin{aligned}\pi|b_0| &\leq |\Lambda| + |b_1\log\alpha_1 + \cdots + b_n\log\alpha_n| \\ &\leq 1 + nB_0\sqrt{\pi^2+1}d\log A_n \\ &\leq 2\pi ndB_0\log A_n,\end{aligned}$$

for the final one, we use the fact that $\sqrt{\pi^2+1} \leq \pi+1$, $1 \leq (\pi-1)ndB_0\log A_n$.

Obviously, $B \geq 3$, and since $B_0 > 2nd \geq 2n$, so $B = B_0^2 > 2nB_0$,

$$B = B_0^2 \geq ed,$$

$$\frac{|b_0|\log A_0^*}{\log A_n^*} = \frac{\pi|b_0|}{\sqrt{\pi^2+1}\cdot d\log A_n} \leq \frac{2\pi}{\sqrt{\pi^2+1}}nB_0 < 2nB_0 < B,$$

$$\frac{|b_i|\log A_i^*}{\log A_n^*} = \frac{|b_i|\log A_i}{\log A_n} \leq |b_i| \leq B_0 < B$$

for $1 \leq i \leq n$.

By applying Theorem 1.3.2, we have

$$\begin{aligned}\log|\Lambda| &\geq -C(n+1,\varkappa)d^{n+3}\log(ed)\log A_0^*\log A_1^*\cdots\log A_n^*\log B\\ &= -2\pi(\pi^2+1)^{n/2}C(n+1,\varkappa)d^{n+2}\log(ed)\cdot\log A_1\cdots\log A_n\log B_0,\end{aligned}$$

$$|\alpha| \geq \frac{1}{2}|\Lambda| \geq \exp\{-(2\pi(\pi^2+1)^{n/2}C(n+1,\varkappa)+\log 2)d^{n+2}\log(ed)\cdot\log A_1\cdots\log A_n\log B_0\}$$

Hence it's sufficient to show that

$$2\pi(\pi^2+1)^{n/2}C(n+1,\varkappa)+\log 2 \leq 2^{2n+3}C(n+1,\varkappa) \leq 2^{8n+29}.$$

Indeed,

$$\begin{aligned}2(\pi^2+1)^{n/2}(2\cdot(\frac{4}{\sqrt{\pi^2+1}})^n-\pi)C(n+1,\varkappa) &\geq 2(\pi^2+1)^{1/2}(\frac{8}{\sqrt{\pi^2+1}}-\pi)C(2,\varkappa)\\ &\geq 11.28\cdot C(2,\varkappa)\\ &\geq \log 2,\end{aligned}$$

since $C(2,\varkappa) \geq \min\{2^{2.5}e\cdot 30^5, 2^{32}\} \geq \log 2$, and we have

$$\begin{aligned}C(n+1,\varkappa) &= \min\{\frac{1}{\varkappa}(\frac{1}{2}en)^{\varkappa}30^{n+3}n^{3.5}, 2^{6n+20}\}\\ &\leq 2^{6n+20}.\end{aligned}$$

□

The following lemma will be used when we apply Theorem 1.3.1:

LEMMA 1.3.4 ([49] Lemma 2.2). *Let $b \geq 0, h \geq 1, a > (e^2/h)^h$, and let $x \in \mathbb{R}^+$ such that*

$$x - a(\log x)^h - b \leq 0,$$

then $x < 2^h(b^{1/h} + a^{1/h}\log(h^h a))^h$. In particular, if $h = 1$, then $x < 2(b + a\log a)$.

1.4 Baker's Method on Algebraic Curves

One of the method to use Baker's inequality to solve Diopantine equations is to use S-unit equations. We will not go to it in this thesis, for more details about it, see [2] and [23].

In this section, we state Yu.Bilu's results and idea to calculate integral points on algebraic curves, see [5] and [4].

Definition 1.4.1. *Let X be a geometrically integral projective curves over a field K, and $\Sigma \subset X(K)$ a finite subset. A function $z \in K(X)$ is a Σ-unit (over K) if $\mathrm{Supp}(z) \subset \Sigma$. We denote the group of Σ-units (over K) by $U_{\Sigma,K}$.*

Remark. *1. For a function $z \in K(X)$ on an integral algebraic variety X, the support of z is defined as*

$$\mathrm{Supp}(z) = \{\}$$

LEMMA 1.4.2. *Keep the notations in Definition 1.4.1, then*

$$U_{\Sigma,K} \simeq K^* \oplus \mathbb{Z}^{\rho},$$

where $\rho = \rho(\Sigma, K)$ satisfies $0 \leq \rho(\Sigma, K) \leq \#\Sigma$.

Proof. Let $\Sigma = \{p_1, \cdots, p_n\}$, $U = X \setminus \Sigma$, and

$$H := \{\sum_{i=1}^{n} a_i[p_i] \in \bigoplus_{i=1}^{n} \mathbb{Z}[p_i] \mid \sum_{i=1}^{n} a_i = 0\} \subset \mathrm{Div}^0(X).$$

Then H is a free abelian group of rank $n-1$, and we have

$$\mathrm{div} : U_{\Sigma,K} \to H,$$

with kernel $\overline{K}^*$ and free image of rank $\rho = \rho(\Sigma, K) \leq n-1$. Hence

$$U_{\Sigma,K} \simeq K^* \oplus \mathbb{Z}^\rho.$$

□

Remark. *(1) For any field extension L/K, we have that $U_{\Sigma,K} \subset U_{\Sigma,L}$ and $0 \leq \rho(\Sigma, K) \leq \rho(\Sigma, L) \leq \#\Sigma$. Hence $\rho(\Sigma, L) = \rho(\Sigma, \overline{K})$ for some finite extension L/K. In this case, if L^* and $u_1, \cdots, u_\rho \in K(X)$ generates $U_{\Sigma,L}$, where $\rho = \rho(\Sigma, L)$, then $\overline{K}^*$ and $u_1, \cdots, u_\rho \in K(X)$ generates $U_{\Sigma,\overline{K}}$. In particular, if $\rho(\Sigma, K) = \#\Sigma$, then $\rho(\Sigma, K) = \rho(\Sigma, \overline{K})$.*

For a non-constant function $x \in \overline{K}(X)$, we denote $\Sigma_x \subset X(\overline{K})$ the set of support of x. Using the Baker's inequality, Yu.Bilu [4] proved the following result:

Theorem 1.4.3 ([4], Theorem 1B)**.** *Let X be an algebraic curve defined over a number field K, $x \in K(C)$ non-constant such that $\rho(\Sigma_x, \overline{K}) \geq 2$, then for any finite set S of places of K, containing all infinity places, we have*

$$\mathrm{h}(x(P)) \leq c(X, x, K, S),$$

for any $P \in X(\mathcal{O}_S, x)$, where c is effective.

We give the idea of the proof, which is useful when we want to calculate c in practice, see [5, section 3]: we can assume that $\Sigma_x \subset X(K)$ and $\rho(\Sigma_x, K) \geq 2$. Set $d = [K : \mathbb{Q}]$, $s = |S|$ and $r = s - 1$.

(1) For any $P \in X(\mathcal{O}_S, x)$, we have

$$\mathrm{h}(x(P)) = \frac{1}{d} \sum_{v \in S} d_v \log^+ |x(P)|_v \leq s \log^+ |x(P)|_v,$$

for some $v \in S$, so it's sufficient to bound $|x(P)|_v$ or $\mathrm{h}(x(P))$ for some $v \in S$.

(2) For $v \in S$, and $Q \in \mathrm{Supp}(x)_\infty$, take a neighborhood $V_{Q,v}$ of Q, and a large $A_v > 0$, such that

$$\{P \in X(K) : |x(P)|_v > A_v\} \subset \bigcup_{Q \in \Sigma_x} V_{Q,v}.$$

We can take $V_{Q,v} = \{P \in X(K) : \mathrm{d}_{v,Q}(P) \leq B_v\}$, where $\mathrm{d}_{v,Q}$ is a v-adic "distance" around Q, e.g. $\mathrm{d}_Q(P) = (x(P))^{-1/e_Q}$ and $B_v = A_v^{1/e_Q}$.

It is sufficient to bound $\mathrm{h}(x(P))$ for $P \in V_{Q,v}$ for fixed Q.

(3) There exists a Σ_x-unit $z = z_Q$, such that $z(Q) = \gamma \neq 0$. Such z exists, since $\rho(\Sigma_x - Q, K) \geq \rho(\Sigma_x, K) - 1 \geq 1$. We should bound $\mathrm{h}(\gamma)$, and suppose that $\mathrm{h}(\gamma) \leq c_1$. Moreover, since $\mathrm{h}(z(\cdot))$ and $\mathrm{h}(x(\cdot))$ are "quasi-equivalence", we will have

$$\mathrm{h}(x(\cdot)) \leq a\mathrm{h}(z(\cdot)) + b$$

for some positive constant a, b.

(4) Consider $\mu = \gamma^{-1}z(P)$, there exists a suitable $\mu_0 \in K$ such that $\mu = \mu_0\eta$ with $\eta \in U_S$. We may apply Proposition 1.2.6 or calculate μ_0 explicitly to bound $\mathrm{h}(\mu_0)$, and suppose that $\mathrm{h}(\mu_0) \leq c_2$.

(5) Apply Proposition 1.2.3 to take a fundamental system $\eta_1, \cdots, \eta_r$ of S-units, and up to a root of unity, we set $\eta = \eta_1^{b_1} \cdots \eta_{s-1}^{b_{s-1}}$. We should get a upper bound of $|\mu - 1|$ in the following form:

$$|\mu_0\eta_1^{b_1} \cdots \eta_{s-1}^{b_{s-1}} - 1| \leq e^{-c_2 B},$$

where $B = \max\{|b_1|, \cdots, |b_{s-1}|\}$;

(6) Apply Baker's inequality to get a upper bound of B, and apply Proposition 1.2.3 to bound $\mathrm{h}(\eta)$, i.e.

$$\mathrm{h}(\eta) \leq Br\zeta^{r-1}\frac{(r!)^2}{2^{r-1}d}R(S).$$

(7) Bound $\mathrm{h}(z(P))$ with the inequality

$$\mathrm{h}(z(P)) \leq \mathrm{h}(\gamma) + \mathrm{h}(\mu_0) + \mathrm{h}(\eta).$$

and bound $\mathrm{h}(x(P))$ with

$$\mathrm{h}(x(P)) \leq a\mathrm{h}(z(P)) + b.$$

Remark. *(1) Before this calculating process, we should calculate $\rho(\Sigma_x, K)$ first. In some cases, we have $\rho(\Sigma_x, K) = |\Sigma_x| - 1$, for example, when X is a modular curve and x is the j-invariant. If so, we only need that x has at least 3 poles. Otherwise, we may use the Chevalley-Weil Theorem in the next section.*

(2) Which method to generate such z in (3) in the process depends on what kind of information we have. For example, in [4], Bilu used the theory of Puiseux series, in [55], Sha used Siegel functions.

1.5 The Chevalley-Weil Theorem

The main reference of this section is [14, section 10.3]. There is an analogous statements for S-integral points without the completeness hypothesis for varieties.

1.5.1 Local Chevalley-Weil Theorem

Recall some notations: for a finite field extension L/K with discrete valuations, we denote $D_{L/K}$ the discriminant of $\mathcal{O}_L/\mathcal{O}_K$, i.e.

$$D_{L/K} := \{\det(Tr_{L/K}(a_i b_j))|a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathcal{O}_L\},$$

where $\mathcal{O}_K$, $\mathcal{O}_L$ are the ring of integers of K and L respectively and $n = [L : K]$. We also denote $\mathcal{O}_K$, v_K and $\hat{K}$ for the ring of integer, the discrete valuation and the completion of K with respect to the discrete valuation, respectively.

Proposition 1.5.1 ([14], Proposition 10.3.3)**.** *Let K be a field with a non-archimedean absolute value $|\cdot|$ on the algebraic closure $\overline{K} \supset K$. Let $\widetilde{X} \xrightarrow{\pi} X$ be a finite unramified morphism of varieties defined over K. If X complete, then there is $\alpha \in \mathcal{O}_K \setminus \{0\}$ such that $\alpha \in D_{\widehat{K(\widetilde{P})}/\hat{K}}$ whenever $\widetilde{P} \in \widetilde{X}(\bar{K})$ and $P := \pi(\widetilde{P}) \in X(K)$, where $K(P)$ and $K(\widetilde{P})$ are the residue fields of P and $\widetilde{P}$ respectively. In other words, $v_{\hat{K}}(D_{\widehat{K(\widetilde{P})}/\hat{K}}) \leq v_K(\alpha)$.*

1.5.2 Global Chevalley-Weil Theorem

Recall some notations: for a finite field extension L/K of number fields, we denote $D_{L/K}$ the discriminant of L/K, and $\partial_{L/K}$ the normalized logarithmic relative discriminant of L/K, i.e.

$$\partial_{L/K} = \frac{\log \mathcal{N}_{K/\mathbb{Q}}(D_{L/K})}{[L:\mathbb{Q}]}.$$

To fully understand Global Chevalley-Weil principle, we demonstrate the following lemma.

LEMMA 1.5.2. *Let L/K be a finite extension of number fields and T be a finite set of prime numbers such that every ramified place is above a prime from T. Then*

$$\left|\mathcal{N}_{K/\mathbb{Q}}(D_{L/K})\right| \leq \left(\prod_{p\in T} p\right)^{[L:\mathbb{Q}]^2},$$

where $D_{L/K}$ is the discriminant of L over K.

Proof. The "Dedekind Discriminant Formula" [14, Theorem B.2.12] implies that

$$v_{\mathfrak{p}}(D_{L/K}) = \sum_{\mathfrak{P}|\mathfrak{p}} (e_{\mathfrak{P}/\mathfrak{p}}(1+\delta_{\mathfrak{P}})-1) f_{\mathfrak{P}/\mathfrak{p}} \leq [K:\mathbb{Q}] \sum_{\mathfrak{P}|\mathfrak{p}} e^2_{\mathfrak{P}/\mathfrak{p}} f_{\mathfrak{P}/\mathfrak{p}} \leq [L:K][L:\mathbb{Q}],$$

where $\mathfrak{p}$ is a prime of K ramified in L, the sum is over the primes of L above $\mathfrak{p}$, and $0 \leq \delta_{\mathfrak{P}} \leq v_{\mathfrak{p}}(e_{\mathfrak{P}/\mathfrak{p}}) < [K:\mathbb{Q}]e_{\mathfrak{P}/\mathfrak{p}}$. For every such $\mathfrak{p}$ we have $\left|\mathcal{N}_{K/\mathbb{Q}}\mathfrak{p}\right| = p^{f_{\mathfrak{p}/p}}$, where p is the prime number below $\mathfrak{p}$. Hence

$$\left|\mathcal{N}_{K/\mathbb{Q}}(D_{L/K})\right| \leq \left(\prod_{p\in T} p^{\sum_{\mathfrak{p}|p} f_{\mathfrak{p}/p}}\right)^{[L:K][L:\mathbb{Q}]} \leq \left(\prod_{p\in T} p^{[K:\mathbb{Q}]}\right)^{[L:K][L:\mathbb{Q}]} = \left(\prod_{p\in T} p\right)^{[L:\mathbb{Q}]^2}.$$

□

Theorem 1.5.3 ([14], Theorem 10.3.11)**.** *Let $\widetilde{X} \xrightarrow{\pi} X$ be a finite unramified morphism of varieties defined over a number field K. If X is complete, then there exists a finite extension L/K such that $\widetilde{P} \in \widetilde{X}(L)$ for any $\widetilde{P} \in \widetilde{X}(\bar{K})$ and $P := \pi(\widetilde{P}) \in X(K)$.*

Remark. *(1) Under the hypothesis of the Proposition, the following statements are equivalent:*

- *(i) there exists a finite field extension L/K such that $\widetilde{P} \in \widetilde{X}(L)$ for any $\widetilde{P} \in \widetilde{X}(\bar{K})$ and $P = \pi(\widetilde{P}) \in X(K)$;*
- *(ii) there exists an $\alpha \in \mathcal{O}_K \setminus \{0\}$ such that $\alpha \in D_{K(\widetilde{P})/K}$ for any $\widetilde{P} \in \widetilde{X}(\bar{K})$ and $P = \pi(\widetilde{P}) \in X(K)$;*
- *(iii) there exist a finite set T of places of K such that $K(\widetilde{P})/K$ is unramified outside T for any $\widetilde{P} \in \widetilde{X}(\bar{K})$ and $P = \pi(\widetilde{P}) \in X(K)$;*
- *(iv) there exists constant $C > 0$ such that $\partial_{K(\widetilde{P})/K} \leq C$ for any $\widetilde{P} \in \widetilde{X}(\bar{K})$ and $P = \pi(\widetilde{P}) \in X(K)$.*

Hence the global Chevalley-Weil Theorem may be demonstrated in these four forms, and the quantitative Chevalley-Weil Theorem in general means to find T in (iii) or C in (iv).

Proof. Obviously, (i) implies (ii).

Since a place of K is ramified over L if and only if the corresponding prime divides $D_{L/K}$, and $D_{L/K}|(\alpha)$, so (ii) implies (iii).

If (iii) holds, let S be the restriction of T on $\mathbb{Q}$. Then by Lemma 1.5.2,

$$\partial_{K(\widetilde{P})/K} \leq [K(\widetilde{P}) : \mathbb{Q}] \sum_{p \in S} \log p \leq n[K : \mathbb{Q}] \sum_{p \in S} \log p,$$

where $[K(\widetilde{P}) : K] \leq n = \max_{x \in X} \{\dim_{k(x)} \mathcal{O}_{\widetilde{X}_x}(\widetilde{X}_x)\}$ is bounded. Hence (iii) implies (iv).

By the transitivity rule of discriminant,

$$D_{K(\widetilde{P})/\mathbb{Q}} = \mathcal{N}_{K/\mathbb{Q}}(D_{K(\widetilde{P})/K}) \cdot D_{K/\mathbb{Q}}^{[K(\widetilde{P}):K]} = e^{[K(\widetilde{P}):\mathbb{Q}]\partial_{K(\widetilde{P})/\mathbb{Q}}} D_{K/\mathbb{Q}}^{[K(\widetilde{P}):K]} \leq (e^{[K:\mathbb{Q}]C} D_{K/\mathbb{Q}})^n,$$

where n is defined as above. Then by the Hermite's discriminant theorem, see [14, Theorem B.2.14], there are finite possibilities of $K(\widetilde{P})$. Hence (iv) implies (i). □

1.5.3 The first version of quantitative Chevalley-Weil Theorem for curves

One version of quatitative Chevalley-Weil Theorem for curves is given in [12].

Proposition 1.5.4 ([12], Theorem 1.3)**.** *Let $\widetilde{C} \xrightarrow{\pi} C$ be a non-constant, unramified morphism of geometrically integral projective curves defined over a number field K. Let $x \in K(C) \subset K(\widetilde{C})$ be a non-constant function on C, and $f(X,Y), \tilde{f}(X,Y) \in K[X,Y]$ such that $K(C) \simeq K(x)[Y]/(f(x,Y))$ and $K(\widetilde{C}) \simeq K(x)[Y]/(\tilde{f}(x,Y))$. We put*

$$m = \deg_X f, \;\; n = \deg_Y f,$$

$$\tilde{m} = \deg_X \tilde{f}, \;\; \tilde{n} = \deg_Y \tilde{f},$$

$$\Omega = mn^2(\mathrm{h}_p(f) + 2m + 2n), \;\; \widetilde{\Omega} = \tilde{m}\tilde{n}^2(\mathrm{h}_p(\tilde{f}) + 2\tilde{m} + 2\tilde{n}),$$

$$\mathrm{Y} = 2\tilde{n}(\tilde{m}\mathrm{h}_p(f) + m\mathrm{h}_p(\tilde{f})).$$

Then for any $\widetilde{P} \in Y(\bar{K})$ and $P := \pi(\widetilde{P}) \in \widetilde{X}(K)$, we have

$$\partial_{K(\widetilde{P})/K} \leq 400(\Omega + \widetilde{\Omega}) + 2\mathrm{Y} + 6mn.$$

With quantitative Riemann's existence theorem, see [11], we can calculate $f(X,Y)$, $\tilde{f}(X,Y)$ and get a bound in terms of invariants of C and $\widetilde{C}$.

1.5.4 The second version of quantitative Chevalley-Weil Theorem for curves

Another version of quantitative Chevalley-Weil Theorem for curves is given algebraically. To do this, we need a result from [42].

Proposition 1.5.5 ([42] Corollary 4.10)**.** *Let K be a discrete valuation field with ring of integers $\mathcal{O}_K$, and $f : X \to Y$ be a finite morphism of smooth, connected projective curves over K. Assume that $g(Y) \geq 1$, and that X admits a smooth projective model $\mathcal{X}$. Then Y admits a smooth projective model $\mathcal{Y}$, and f extends to a finite morphism $\mathcal{X} \to \mathcal{Y}$.*

The following proposition is an explicit form of Chevalley-Weil Principle under some conditions.

Proposition 1.5.6. *Let $\widetilde{C} \xrightarrow{\pi} C$ be a non-constant, unramified morphism of smooth, connected projective curves over a number field K with $g(C) \geq 1$, and let $\mathfrak{p} \subset \mathcal{O}_K$ be a non-zero prime with residue field $k(\mathfrak{p})$. Suppose that*

(1) $\widetilde{C}$ admits a smooth projective model at $\mathfrak{p}$;

(2) $[K(\widetilde{C}) : K(C)] < \mathrm{Char}(k(\mathfrak{p}))$ or $K(\widetilde{C})/K(C)$ is Galois of degree prime to $\mathrm{Char}(k(\mathfrak{p}))$.

Then for every point $P \in C(K)$ and $\widetilde{P} \in \pi^{-1}(P)$, we have that $\mathfrak{p}$ is unramified in the residue field $K(\widetilde{P})$ of $\widetilde{P}$.

Proof. We should notice that π is finite and étale.

Suppose that $\mathcal{X}$ is the smooth model of C over $\mathrm{Spec}(\mathcal{O}_{K,\mathfrak{p}})$. Since f is finite, and $g(C) \geq 1$, then by Proposition 1.5.5, C admits a smooth model $\mathcal{Y}$ and π is extended to a finite morphism $\mathcal{X} \to \mathcal{Y}$. We still denote the extended morphism by π.

We endow the closure $\overline{\{P\}}$ of $\{P\}$ in $\mathcal{Y}$ with structure of reduced closed subscheme. It is a section of $\mathcal{Y}$ over $\mathrm{Spec}(\mathcal{O}_{K,\mathfrak{p}})$, that is because $P \in C(K)$, and $\overline{\{P\}}$ is finite, birational over $\mathrm{Spec}(\mathcal{O}_{K,\mathfrak{p}})$. Consider $\mathcal{X} \times_{\mathcal{Y}} \overline{\{P\}}$. It is finite over $\overline{\{P\}} \simeq \mathrm{Spec}(\mathcal{O}_{K,\mathfrak{p}})$, hence affine, denoted by $\mathrm{Spec}(A)$. Its underlying space is $\pi^{-1}(\overline{\{P\}})$. If $\mathcal{X} \to \mathcal{Y}$ is étale, then after the base change $\overline{\{P\}} \to \mathcal{Y}$, $\mathrm{Spec}(A) \to \mathrm{Spec}(\mathcal{O}_{K,\mathfrak{p}})$ is also étale. Since $\mathcal{O}_{K,\mathfrak{p}}$ is regular, so A is regular too. Suppose that $A = \bigoplus_{i=1}^{m} A_i$ such that A_i is connected for each i. The fact that A is regular and finite over $\mathcal{O}_{K,\mathfrak{p}}$ implies that A_i is normal and finite over $\mathcal{O}_{K,\mathfrak{p}}$ for each i. In particular, the affine ring corresponding to $\overline{\{\widetilde{P}\}}$ is the integral closure of $\mathcal{O}_{K,\mathfrak{p}}$ in $K(\widetilde{P})$. Any closed point x on $\overline{\{\widetilde{P}\}}$ is also a closed point on $\mathrm{Spec}(A)$. We know that $\overline{\{\widetilde{P}\}}$ and $\mathrm{Spec}(A)$ have the same local rings at x, so $\overline{\{\widetilde{P}\}} \to \overline{\{P\}}$ is étale at x. Hence $\mathfrak{p}$ is unramified in $K(\widetilde{P})$.

It remains to show that $\mathcal{X} \to \mathcal{Y}$ is étale. Let Z be the set of points in $\mathcal{X}$ at which f is not étale, then Z is closed in $\mathcal{X}$. If $Z \neq \emptyset$, since $Z \neq \mathcal{X}$, by Zariski-Nagata purity theorem in [29, Théorèm de pureté 3.1], it is purely of codimension 1. Any irreducible component W of Z is vertical, because $\widetilde{C} \xrightarrow{\pi} C$ is étale. Let η be the generic point of W, then $\xi = \pi(\eta) \in \mathcal{Y}$ is also a generic point in $\mathcal{X}_s$ from the fact that π is dominant and finite, where $\mathcal{X}_s$ is the special fiber of $\mathcal{X}$. Consider $\pi_\eta^\# : \mathcal{O}_{\mathcal{Y},\xi} \to \mathcal{O}_{\mathcal{X},\eta}$. We claim that the maximal ideals of $\mathcal{O}_{\mathcal{Y},\xi}$ and $\mathcal{O}_{\mathcal{X},\eta}$ are $\mathfrak{p}\mathcal{O}_{\mathcal{Y},\xi}$ and $\mathfrak{p}\mathcal{O}_{\mathcal{X},\eta}$ respectively. Indeed, we have that $\mathcal{O}_{\mathcal{X}_s,\eta} = \mathcal{O}_{\mathcal{X},\eta}/\mathfrak{p}\mathcal{O}_{\mathcal{X},\eta}$, and the special fiber $\mathcal{X}_s$ is smooth, so $\mathcal{O}_{\mathcal{X}_s,\eta}$ is integral with only one prime ideal. Hence $\mathcal{O}_{\mathcal{X}_s,\eta}$ is a field, and $\mathfrak{p}\mathcal{O}_{\mathcal{X},\eta}$ is the maximal ideal of $\mathcal{O}_{\mathcal{X},\eta}$. It is similar for $\mathcal{O}_{\mathcal{Y},\xi}$. On the other hand, $[k(\eta) : k(\xi)] \leq [K(\widetilde{C}) : K(C)]$ and $[k(\eta) : k(\xi)] | [K(\widetilde{C}) : K(C)]$ if $\widetilde{C} \xrightarrow{\pi} C$ is Galois. By the assumption (2), the residue degree $[k(\eta) : k(\xi)] < \mathrm{Char}(k(\mathfrak{p}))$ or $[k(\eta) : k(\xi)]$ prime to $\mathrm{Char}(k(\mathfrak{p}))$, so $k(\eta)/k(\xi)$ is separable. Hence $\mathcal{O}_{\mathcal{Y},\xi} \to \mathcal{O}_{\mathcal{X},\eta}$ is unramified. It is also flat since it is injective and $\mathcal{O}_{\mathcal{Y},\xi}$ is a Dedekind domain, hence also étale. Contradiction. □

Corollary 1.5.7. *Let $\widetilde{C} \xrightarrow{\pi} C$ be a non-constant, unramified morphism of smooth, connected projective curves over a number field K with $g(C) \geq 1$. We set*

$$T = \{\mathfrak{p} \in \mathrm{Spec}(\mathcal{O}_K) \mid \mathrm{Char}(k(\mathfrak{p})) \leq [K(\widetilde{C}) : K(C)] \text{ or } \widetilde{C} \text{ has bad reduction at } \mathfrak{p}\}.$$

Then $K(\widetilde{P})/K$ is unramified outside T for any $P =\in C(K)$ and $\widetilde{P} \in \pi^{-1}(P)$.

Chapter 2

Automorphic Forms and Modular Curves

This chapter defines the modular curves analytically and algebraically. Another important object of this Chapter is Corollary 2.5.3 in Section 2.5. This corollary implies that the automorphic forms over a number field are exactly the fractional functions on the modular curves as algebraic curves, which provides a theoretic support for studying modular units to bound integral points on modular curves. The main references are [22] and [56].

2.1 Automorphic Forms

In this section, we state basic properties of modular forms, see [22, Chapter 1] for full details.

2.1.1 Congruence subgroups

Proposition 2.1.1. *The group $SL_2(\mathbb{Z})$ is generated by $\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ and $\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$.*

Definition 2.1.2. *Let $N \in \mathbb{N}^+$, the principal congruence subgroup of level N is defined as*

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

It is a normal subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

A subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup of level N if $\Gamma(N) \subset \Gamma$. In particular, we have the following congruence subgroups of level N:

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

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

LEMMA 2.1.3. *The canonical morphism $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}(\mathbb{Z}/N\mathbb{Z})$ is surjective. Consequently, it induces a bijection*

$$\{\textit{congurence subgroups of } \mathrm{SL}_2(\mathbb{Z}) \textit{ of level } N\} \leftrightarrow \{\textit{subgroups of } \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})\}$$

There is a left $\mathrm{SL}_2(\mathbb{Z})$-action on $\mathbb{H}$ defined as: for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, $\tau \in \mathbb{H}$,

$$\gamma(\tau) := \frac{a\tau + b}{c\tau + d}.$$

Indeed, this action can be extended on $\mathbb{H}^* = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$ and induces the action of $\mathrm{PSL}(\mathbb{Z}) = \dfrac{\mathrm{SL}_2(\mathbb{Z})}{\{\pm I\}}$ on $\mathbb{H}^*$.

2.1.2 Automorphic forms and modular forms

Definition 2.1.4. *Let $\mathcal{M}(\mathbb{H})$ be the field of meromorphic functions on $\mathbb{H}$, $k \in \mathbb{Z}$. Then the weight-k action $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{M}(\mathbb{H})$ is a right action defined as*

$$f[\gamma]_k(\tau) := j(\gamma, \tau)^{-k} f(\gamma(\tau)), \;\; f \in \mathcal{M}(\mathbb{H}), \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

where $j(\gamma, \tau) = c\tau + d$.

Definition 2.1.5. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, $k \in \mathbb{Z}$, a meromorphic function $f \in \mathcal{M}(\mathbb{H})$ is weakly modular of weight k with respect to γ if*

$$f[\gamma]_k = f,$$

for any $\gamma \in \Gamma$.

Remark. *(1) If f is weakly modular of weight k with respect to a congruence subgroup Γ, then $f(\tau + h) = f(\tau)$ for some $h \in \mathbb{N}^+$.*

(2) If f is weakly modular of weight k with respect to Γ, then for any $\alpha \in SL_2(\mathbb{Z})$, the function $f[\alpha]_k$ is weakly modular of weight k with respect to $\alpha^{-1}\Gamma\alpha$.

Definition 2.1.6. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, $k \in \mathbb{N}^+$, and f be a weakly modular function of weight k. If h is the minimal positive integer such that $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \in \Gamma$, then we can write*

$$f(\tau) = g(q_h) = \sum_{n=-N}^{\infty} a_n q_h^n, \;\; q_h = e^{2\pi i \tau / h},$$

and call this the Fourier q_h-expansion of f at infinity. The coefficients a_n are called the Fourier coefficients of f with respect to Γ. If $a_{-N} \neq 0$, we call $-N$ the order of f at infinity, and denote it by $v_\infty(f)$. For any $\tau \in \mathbb{H}$, we denote the order of f at τ by $v_\tau(f)$.

We shall say that f is meromorphic (resp. holomorphic) at infinity (or at $i\infty$) if g is meromorphic (resp. holomorophic) at 0.

Remark. *(1) In fact, to show that $f(\tau)$ is meromorphic or holomorphic at ∞, it is sufficient to take any positive integer h such that $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \in \Gamma$, and consider the q_h-expansion of $f(\tau)$.*

Definition 2.1.7. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, we have a Γ-action on $\mathbb{H}^*$ defined as before. A Γ-equivalence class of points in $\mathbb{Q} \bigcup \{\infty\}$ is called a cusp of Γ.*

Remark. *(1) If $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, there is only one cusp. In general, for a congruence subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$,*

$$\#\{\textit{cusps of } \Gamma\} \leq [\mathrm{SL}_2(\mathbb{Z}) : \Gamma].$$

Proof. We have $\{\text{cusps of } \Gamma\} = \{\Gamma\gamma(\infty) \mid \gamma \in \mathrm{SL}_2(\mathbb{Z})\}$, which implies a surjection

$$\Gamma\backslash^{\mathrm{SL}_2\mathbb{Z}} \twoheadrightarrow \{\text{cusps of } \Gamma\}.$$

Hence $\#\{\text{cusps of } \Gamma\} \leq [\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$. □

Definition 2.1.8. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, $k \in \mathbb{N}^+$. A function $f : \mathbb{H} \to \mathbb{C}$ is an automorphic (resp. modular) form of weight k with respect to Γ if*

(a) f is meromorphic (resp. holomorphic);

(b) f is weakly mordular of weight k with respect to Γ;

(c) $f[\alpha]_k$ is meromorphic (resp. holomorphic) at $i\infty$ for any $\alpha \in \mathrm{SL}_2(\mathbb{Z})$.

If f is a modular form and in addition

(d) $a_0 = 0$ in the Fourier expansion of $f[\alpha]_k$ for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$,

then f is called a cusp form of weight k with respect to Γ.

The set of automorphic (resp. modular, resp. cusp) forms of weight k with respect to Γ is a $\mathbb{C}$-vector space, denoted by $\mathcal{A}_k(\Gamma)$ (resp. $\mathcal{M}_k(\Gamma)$, resp. $\mathcal{S}_k(\Gamma)$).

Remark. *(1) If $[\alpha(\infty)]$ is a cusp of Γ, $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, we say a weakly modular function f is meromorphic (resp. holomorphic) at $[\alpha(\infty)]$ if $f[\alpha]_k$ is meromorphic (resp. holomorphic) at $i\infty$, it is independent of the choice of α. Hence the condition (c) becomes that f is meromorphic (resp. holomorphic) at all cusps of Γ.*

If $\mathrm{SL}_2(\mathbb{Z}) = \bigcup_j \Gamma\alpha_j$ is a left coset decomposition, then condition (c) holds if and only if $f[\alpha_j]_k$ is meromorphic (resp. holomorphic) at $i\infty$ for each j.

(2) If $\Gamma' \subset \Gamma$ are congruence subgroups, then

$$\mathcal{A}_k(\Gamma) = \{f \in \mathcal{A}_k(\Gamma') \mid f \text{ is } \Gamma\text{-invariant}\},$$
$$\mathcal{M}_k(\Gamma) = \{f \in \mathcal{M}_k(\Gamma') \mid f \text{ is } \Gamma\text{-invariant}\},$$
$$\mathcal{S}_k(\Gamma) = \{f \in \mathcal{S}_k(\Gamma') \mid f \text{ is } \Gamma\text{-invariant}\}.$$

(3) If k is odd and $-I \in \Gamma$, then $\mathcal{A}_k(\Gamma) = 0$.

Proposition 2.1.9. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup of level N, $q_N = e^{2\pi i\tau}$ for $\tau \in \mathbb{H}$, and f be a weakly modular function of weight k with respect to Γ. Suppose that f is holomorphic on $\mathbb{H}$ and at $i\infty$, and there exists some constants C and r such that for any $n > 0$,*

$$|a_n| \leq Cn^r,$$

where a_n is the n-th coefficients of the Fourier q_N-expansion of f. Then $f \in \mathcal{M}_k(\Gamma)$.

Proof. Let $\tau = x + iy$, then

$$f(\tau) \leq |a_0| + C\sum_{n=1}^{\infty} n^r e^{-2\pi ny/N}.$$

For $r \geq 1$, set $g(t) = t^r e^{-2\pi ty/N}, t \geq 0$, then $g'(t) = (r - 2\pi ty/N)t^{r-1}e^{-2\pi ty/N},$. Hence $g(t)$ is increasing when $t \in [0, \frac{rN}{2\pi y}]$, and is decreasing when $t \in [\frac{rN}{2\pi y}, \infty)$. For any $n \geq \frac{rN}{2\pi y} + 1$, we have

$$n^r e^{-2\pi ny/N} \leq \int_{n-1}^{n} t^r e^{-2\pi ty/N} dt,$$

$$\sum_{n=m}^{\infty} n^r e^{-2\pi ny/N} \leq \int_{m-1}^{\infty} g(t)dt,$$

where $m-1 = \left\lceil \frac{rN}{2\pi y} \right\rceil$, the minimal integer which is bigger than $\frac{rN}{2\pi y}$. There exists $C_0 > 0$ such that

$$f(\tau) \leq C_0 + C\int_0^{\infty} g(t)dt.$$

We have

$$\begin{aligned}
\int_0^{\infty} g(t)dt &= \int_0^{\infty} t^r e^{-2\pi ty/N}dt \\
&= \int_0^{\infty} (\frac{t}{y})^r e^{-2\pi t/N} d(\frac{t}{y}) \\
&= y^{-(r+1)} \int_0^{\infty} t^r e^{-2\pi t/N}dt \\
&= (\frac{N}{2\pi})^{r+1} y^{-(r+1)} \int_0^{\infty} t^r e^{-t}dt \\
&= C_1 y^{-(r+1)},
\end{aligned}$$

since $\int_0^{\infty} t^r e^{-t}dt$ is the value of Gamma function $\Gamma(z)$ for $z = r$. Hence

$$|f(\tau)| \leq C_0 + C_1 y^{-(r+1)},$$

when we replace CC_1 by C_1.

For any $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, to show that $f[\alpha]_k(\tau)$ is holomorphic at $i\infty$, it is sufficient to show that $\lim_{q_N \to 0} |q_N f[\alpha]_k(\tau)| = 0$. Indeed,

$$\mathrm{Im}(\alpha(\tau) = \frac{y}{c\tau + d},$$

$$|f(\alpha(\tau))| \leq C_0 + C_1 \frac{|c\tau + d|^{r+1}}{y^{r+1}} \leq C_2(x) y^{r+1},$$

$$|f[\alpha]_k(\tau)| = |(c\tau + d)^{-k} f(\alpha(\tau))| \leq C_3(x) y^{r-k},$$

where $C_2(x)$ and $C_3(x)$ are positive constants only depending on x. Hence

$$\lim_{q_N \to 0} |q_N f[\alpha]_k(\tau)| \leq \lim_{q_N \to 0} C_3(x) y^{r-k} e^{-2\pi ny/N} = C_3(x) \lim_{y \to +\infty} y^{r-k} e^{-2\pi ny/N} = 0.$$

□

2.1.3 Eisenstein series and j-invariant

We will define some important modular forms and cusp forms.

Definition 2.1.10. *Let $k > 2$ be an integer. We define the Eisenstein series of weight k to be*

$$G_k(\tau) := \sum_{(c,d) \in \mathbb{Z}^2 \setminus \{(0,0)\}} \frac{1}{(c\tau + d)^k},\ \tau \in \mathbb{H}.$$

We set $g_2(\tau) = 60G_4(\tau), g_3(\tau) = 140G_6(\tau)$, and define the discriminant function and j-invariant (or modular invariant) as

$$\Delta(\tau) := (g_2(\tau))^3 - 27(g_3(\tau))^2,$$

$$j(\tau) := 1728 \frac{g_2(\tau)^3}{\Delta(\tau)} = \frac{(12g_2(\tau))^3}{\Delta(\tau)},$$

respectively.

Remark. *(1) Obviously, if k is odd, then $G_k(\tau) = 0$.*

We have the following classical results.

Proposition 2.1.11. *Let $k > 2$ be an integer, $q = e^{2\pi i\tau}$. Then the following statements hold:*

(1) The function $G_k(\tau)$ absolutely converges and converges uniformly on compact subsets of $\mathbb{H}$. In particular, G_k is holomorphic on $\mathbb{H}$.

(2) We have

$$G_k \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})),$$
$$\Delta \in \mathcal{S}_{12}(\mathrm{SL}_2(\mathbb{Z})),$$
$$j \in \mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z})).$$

Moreover, $\mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z})) = \mathbb{C}(j)$.

(3) We have the Fourier expansions:

$$G_k(\tau) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty}\sigma_{k-1}q^n = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty}\frac{n^{k-1}q^n}{1-q^n},$$

$$\Delta(\tau) = (2\pi)^{12}\sum_{n=1}^{\infty}b(n)q^n,$$

$$j(\tau) = q^{-1} + \sum_{n=0}^{\infty}c(n)q^n == q^{-1} + 744 + 196884q + 21493760q^2 + \cdots,$$

where $b(n), c(n) \in \mathbb{Z}$ and $b(1) = 1, b(2) = -24, \cdots$.

(4) For any $\tau, \tau' \in \mathbb{H}$, $j(\tau) = j(\tau')$ if and only if $\tau' = \gamma(\tau)$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. The function j has a simple pole at $i\infty$ and $j(i) = 1728, j(\zeta_3) = 0$.

2.2 Modular curves as Moduli Spaces

We do not intend to discuss too much about the theory of arithmetic moduli of elliptic curves. For readers who are interested in this theory, [21], [34] and [19] will be good references to read. Here we follow [22, Section 1.5] to give a rough sense that modular curses are moduli spaces.

Definition 2.2.1. *Let $N \in \mathbb{N}^+$ be a positive integer. An enhanced elliptic curve for $\Gamma_0(N)$ is an ordered pair (E, C) with E a complex elliptic curve and C a cyclic subgroup of $E(\mathbb{C})$ of order N. Two such pairs (E, C) and (E', C') are equivalent, written $(E, C) \sim (E', C')$, if there is an isomorphism $E \to E'$ of group varieties taking C to C'. The set of equivalent classes is denoted by $S_0(N)$, an element of $S_0(N)$ is denoted by $[E, C]$.*

An enhanced elliptic curve for $\Gamma_0(N)$ is an ordered pair (E, Q) with E a complex elliptic curve and Q an element in $E(\mathbb{C})$ of order N. Two such pairs (E, Q) and (E', Q') are equivalent, written $(E, Q) \sim (E', Q')$, if there is an isomorphism $E \to E'$ of group varieties taking Q to Q'. The set of equivalent classes is denoted by $S_1(N)$, an element of $S_1(N)$ is denoted by $[E, Q]$.

An enhanced elliptic curve for $\Gamma_0(N)$ is an ordered pair $(E, (P, Q))$ with E a complex elliptic curve and (P, Q) a pair of element in $E(\mathbb{C})$ that generates $E[N](\mathbb{C})$ with Weil pairing $e_N(P, Q) = e^{2\pi i/N}$. Two such pairs $(E, (P, Q))$ and $(E', (P', Q'))$ are equivalent, written $(E, (P, Q)) \sim (E', (P', Q'))$, if there is an isomorphism $E \to E'$ of group varieties taking C to C'. The set of equivalent classes is denoted by $S(N)$, an element of $S(N)$ is denoted by $[E, (P, Q)]$.

Definition 2.2.2. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup acting on $\mathbb{H}^*$, we define*

$$Y_\Gamma := \mathbb{H}/\Gamma,$$

$$X_\Gamma := \mathbb{H}^*/\Gamma,$$

the quotient spaces of orbits under Γ, and call X_Γ the modular curves for Γ. A point in X_Γ/Y_Γ is called a cusp on X_Γ.

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

For $\tau \in \mathbb{H}$, we denote E_τ the elliptic curve corresponding to lattice $\Lambda_\tau = \langle \tau, 1\rangle$, and via $\mathbb{C}/\Lambda_\tau \leftrightarrow E_\tau$, we don't distinguish the corresponding points.

Theorem 2.2.3 ([22], Theorem 1.5.1)**.** *Let $N \in \mathbb{N}^+$. Then the following statements hold:*

(1) $S_0(N) = \left\{ \left[E_\tau, \left\langle \frac{1}{N} + \Lambda_\tau \right\rangle \right] \mid \tau \in \mathbb{H} \right\}$, *and for any $\tau, \tau' \in \mathbb{H}$,* $\left[E_\tau, \left\langle \frac{1}{N} + \Lambda_\tau \right\rangle \right] = \left[E_{\tau'}, \left\langle \frac{1}{N} + \Lambda_{\tau'} \right\rangle \right]$ *if and only if $\Gamma_0(N)\tau = \Gamma_0(N)\tau'$. Thus there is a bijection*

$$S_0(N) \leftrightarrow Y_0(N),$$
$$\left[E_\tau, \left\langle \frac{1}{N} + \Lambda_\tau \right\rangle \right] \mapsto \Gamma_0(N)\tau.$$

(2) $S_1(N) = \left\{ \left[E_\tau, \frac{1}{N} + \Lambda_\tau \right] \mid \tau \in \mathbb{H} \right\}$, *and for any $\tau, \tau' \in \mathbb{H}$,* $\left[E_\tau, \frac{1}{N} + \Lambda_\tau \right] = \left[E_{\tau'}, \frac{1}{N} + \Lambda_{\tau'} \right]$ *if and only if $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$. Thus there is a bijection*

$$S_1(N) \leftrightarrow Y_1(N),$$
$$\left[E_\tau, \frac{1}{N} + \Lambda_\tau \right] \mapsto \Gamma_1(N)\tau.$$

(3) $S_0(N) = \left\{ \left[E_\tau, \left(\frac{\tau}{N} + \Lambda_\tau, \frac{1}{N} + \Lambda_\tau \right) \right] \mid \tau \in \mathbb{H} \right\}$, *and for any $\tau, \tau' \in \mathbb{H}$,* $\left[E_\tau, \left(\frac{\tau}{N} + \Lambda_\tau, \frac{1}{N} + \Lambda_\tau \right) \right] = \left[E_{\tau'}, \left(\frac{\tau'}{N} + \Lambda_{\tau'}, \frac{1}{N} + \Lambda_{\tau'} \right) \right]$ *if and only if $\Gamma(N)\tau = \Gamma(N)\tau'$. Thus there is a bijection*

$$S_0(N) \leftrightarrow Y_0(N),$$
$$\left[E_\tau, \left(\frac{\tau}{N} + \Lambda_\tau, \frac{1}{N} + \Lambda_\tau \right) \right] \mapsto \Gamma(N)\tau.$$

2.3 Modular curves as Riemann surfaces

With suitable charts, for any congruence subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$, X_Γ is a Riemann surfaces, and so is Y_Γ.

For further discussion, we fix some notations. For $\tau \in \mathbb{H}^*$, we set

$$\Gamma_\tau = \{\gamma \in \Gamma | \gamma(\tau) = \tau\}.$$

Definition 2.3.1. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. Each $\tau \in \mathbb{H}$ has an associated positive integer,*

$$h_\tau := \#\frac{\pm\Gamma_\tau}{\{\pm I\}} = \begin{cases} \frac{\#\Gamma_\tau}{2} & \text{if } -I \in \Gamma; \\ \#\Gamma_\tau & \text{if } -I \notin \Gamma. \end{cases}$$

It is called the period of τ.

For $s \in \mathbb{Q} \bigcup \{\infty\}$, we define the width of s to be

$$h_s := [\mathrm{SL}_2(\mathbb{Z})_s : \pm\Gamma_s].$$

Proposition 2.3.2 ([22], Proposition 2.4.2)**.** *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, then the modular curve X_Γ is a Hausdorff, connected and compact Riemann surface.*

Proof. We give a sketch of the proof. Set $\pi : \mathbb{H}^* \to X_\Gamma$. For any $\tau \in \mathbb{H}$, we can find a neighborhood U of τ in $\mathbb{H}$ such that

(i) for any $\gamma \in \Gamma$, if $\gamma(U) \bigcap U \neq \emptyset$, then $\gamma \in \Gamma_\tau$,

(ii) U has no elliptic point except possibly τ, i.e. if $z \in U$ such that $\pm\Gamma_z \neq \{\pm I\}$, then $z = \tau$.

We set $\delta_\tau(z) = \dfrac{z-\tau}{z-\overline{\tau}}$, and $\rho : \mathbb{C} \to \mathbb{C}, z \mapsto z^{h_\tau}$. Then $\rho \circ \delta_\tau : U \to \rho(\delta_\tau(U)) \subset \mathbb{C}$ induces a homeomorphism $\pi(U) \to \rho(\delta_\tau(U))$, which is a local coordinates around $\pi(\tau)$.

For $s \in \mathbb{Q} \bigcup \{\infty\}$, there exists $\delta_s \in SL_2(\mathbb{Z})$ such that $\delta_s(s) = \infty$. Let $U = \delta_s^{-1}(\{z \in \mathbb{H} \mid \mathrm{Im}(z) > 2\} \bigcup \{\infty\})$ which is an open neighborhood of s in $\mathbb{H}^*$. Set $\phi : U \to \mathbb{C}, \tau \mapsto e^{2\pi i \delta_s(\tau)/h_s}$. This will induce a homeomorphism $\pi(U) \to \phi(U) \subset \mathbb{C}$, which is a local coordinates around $\pi(s)$. □

Remark. *(1) For $\tau \in \mathbb{H}^*$, $\pi : \mathbb{H}^* \to X_\Gamma$, the ramification index of the morphism $X_\Gamma \to X(1)$ at $\pi(\tau)$ is h_τ. When there is no confustion, we will write $h_{\pi(\tau)}$ instead of h_τ.*

Definition 2.3.3. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, $k \in \mathbb{Z}$, $\pi : \mathbb{H}^* \to X_\Gamma$ and $f \in \mathcal{A}_k(\Gamma)$. For any $\tau \in \mathbb{H}$, we define*

$$v_\tau(f) := \textit{the order of } f \textit{ at } \tau \in \mathbb{H},$$

$$v_{\pi(\tau)}(f) := \frac{v_\tau(f)}{h_\tau},$$

where $h_\tau = \#\left(\dfrac{\pm\Gamma_\tau}{\pm I}\right)$ is the width of τ.

For $s \in \mathbb{Q} \bigcup \{\infty\}$ and $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ such that $\alpha(\infty) = s$, let $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \in \alpha^{-1}\Gamma\alpha$ be such that $h > 0$ and minimal. In fact h is the width of ∞ with respect to $\alpha^{-1}\Gamma\alpha$. The function $f[\alpha]_k \in \mathcal{A}_k(\alpha^{-1}\Gamma\alpha)$ has Fourier expansion

$$f[\alpha]_k(\tau) = \sum_{n=m}^{\infty} a_n q_h^n$$

with $q_h = e^{2\pi i\tau/h}$ and $a_m \neq 0$. We define

$$v_s(f) := m,$$

$$v_{\pi(s)}(f) := \begin{cases} \dfrac{v_s(f)}{2} & \textit{if } \alpha^{-1}\Gamma\alpha = \langle -\begin{pmatrix} 1 & h_s \\ 0 & 1 \end{pmatrix} \rangle \textit{ and } k \textit{ is odd}, \\ v_s(f) & \textit{otherwise}, \end{cases}$$

where $h_s = [\pm\mathrm{SL}_2(\mathbb{Z})_\infty : \alpha^{-1}\Gamma_s\alpha]$ is the width of s.

Remark. *(1) The values $v_{\pi(\tau)}$ and $v_{\pi(s)}$ may not be integers, but they have close connection with the orders for corresponding differential on X_Γ.*

The following proposition builds a connection between automorphic forms on $\mathbb{H}$ and differential forms on modular curves. With this, we can calculate the dimension of $\mathcal{M}_k(\Gamma)$ and $\mathcal{S}_k(\Gamma)$ with Riemann-Roch Theorem, see details in [22, Chapter 3].

Proposition 2.3.4 ([22], Theorem 3.3.1)**.** *Let $k \in \mathbb{N}$ an even non-negative integer, $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup and Ω_{X_Γ} be the sheaf of meromorphic forms on X_Γ. Then we have an isomorphism of $\mathbb{C}$-verctor spaces:*

$$\mathcal{A}_k(\Gamma) \xrightarrow{\sim} \Omega_{X_\Gamma}^{\otimes k/2}(X_\Gamma),$$

$$f \mapsto \omega,$$

where ω is a rational section of $\Omega^{\otimes k/2}$ such that $\pi^*\omega = f(\tau)(d\tau)^{k/2}$, and $\pi : \mathbb{H}^* \to X_\Gamma$ is the natural map. Under this isomorphism,

$$\mathcal{M}_k(\Gamma) = \{\omega \in \Omega_{X_\Gamma}^{\otimes k/2}(X_\Gamma) | v_{\pi(\tau)}(\omega) \geq -\frac{1}{2}(1-\frac{1}{h_\tau}), v_{\pi(s)}(\omega) \geq -\frac{k}{2} \text{ for } \tau \in \mathbb{H}, s \in \mathbb{Q} \cup \{\infty\}\},$$

$$\mathcal{S}_k(\Gamma) = \{\omega \in \Omega_{X_\Gamma}^{\otimes k/2}(X_\Gamma) | v_{\pi(\tau)}(\omega) \geq -\frac{1}{2}(1-\frac{1}{h_\tau}), v_{\pi(s)}(\omega) \geq 1-\frac{k}{2} \text{ for } \tau \in \mathbb{H}, s \in \mathbb{Q} \cup \{\infty\}\}.$$

Remark. *(1) In particular, we have*

$$\mathcal{A}_0(\Gamma) \simeq \mathbb{C}(X_\Gamma),$$

$$\mathcal{S}_2(\Gamma) \simeq \Omega^1_{\mathrm{hol}}(X_\Gamma),$$

where $\Omega^1_{\mathrm{hol}}(X_\Gamma)$ the $\mathbb{C}$-vector space of holomorphic forms on X_Γ.

(2) If $\Gamma = \Gamma(N)$, $f \in \mathcal{A}_k(\Gamma(N))$ and $s \in \mathbb{Q} \bigcup \{\infty\}$ with $\alpha(\infty) = s$ for some $\alpha \in \mathrm{SL}_2(\mathbb{Z})$, then

$$v_s(f) = v_\infty(f[\alpha]_k).$$

Corollary 2.3.5. *Keep the notations in Proposition 2.3.4. If k is even , $f \in \mathcal{A}_k(\Gamma)$ corresponds to $\omega \in \Omega_{X_\Gamma}^{\otimes k/2}(X_\Gamma)$. Then*

(1) for any $\tau \in \mathbb{H}$, the order $\mathrm{Ord}_{\pi(\tau)}(\omega)$ of ω at $\pi(\tau) \in X_\Gamma$ is $v_{\pi(\tau)}(f) - \frac{k}{2}(1-\frac{1}{h_\tau})$;

(1) for any $s \in \mathbb{Q} \bigcup \{\infty\}$, the order $\mathrm{Ord}_{\pi(s)}(\omega)$ of ω at $\pi(s) \in X_\Gamma$ is $v_s(f) - \frac{k}{2}$.

2.4 Modular curves as Algebraic Curves

By Riemann's existence Theorem, we know that every compact Riemann is indeed a smooth algebraic curve. Moreover, a modular curve is defined over a number field, Corollary 2.4.11, see also [22, Theorem 7.6.3].

2.4.1 Function fields over $\mathbb{C}$

This subsection describes the function fields for the curve $X(N)$, $X_1(N)$ and $X_0(N)$, where $N \in \mathbb{N}^+$.

Recall the definition of $g_2(\tau)$ and $g_3(\tau)$ in Definition 2.1.10 and

$$\wp_\tau(z) = \frac{1}{z^2} + \sum_{(c,d)\in\mathbb{Z}^2\setminus\{(0,0)\}} \left(\frac{1}{(z-(c\tau+d))^2} - \frac{1}{(c\tau+d)^2} \right).$$

Definition 2.4.1. *Let $N \in \mathbb{N}^+$, $\tau \in \mathbb{H}$.*

(1) For each non-zero element $\overline{v} \in \mathbb{Z}^2/N\mathbb{Z}^2$, we set

$$f_0^{\overline{v}}(\tau) := \frac{g_2(\tau)}{g_3(\tau)} \wp_\tau \left(\frac{c_v\tau + d_v}{N} \right),$$

$$f_{1,0}(\tau) := f_0^{\overline{(1,0)}}(\tau),$$

$$f_{0,1}(\tau) = f_1(\tau) := f_0^{\overline{(0,1)}}(\tau),$$

where $(c_v, d_v) \in \mathbb{Z}^2$ is a representative of $\overline{v}$.

(2) For each non-zero element $\overline{d} \in \mathbb{Z}/N\mathbb{Z}$, we set

$$f_0^{\overline{d}}(\tau) := \frac{g_2(\tau)}{g_3(\tau)} \wp_\tau \left(\frac{d}{N} \right) = f_0^{\overline{d}}(\tau),$$

where $d \in \mathbb{Z}$ is a representative of $\overline{d}$.

(3) We set

$$f_0(\tau) := \frac{g_2(\tau)}{g_3(\tau)} \sum_{d=1}^{N-1} \wp_\tau(\frac{d}{N}) = \sum_{d=1}^{N-1} f_0^{\overline{d}}(\tau).$$

LEMMA 2.4.2. *Keep the notations in Definition 2.4.1. Then*

$$f_0^{\overline{v}}(\tau) \in \mathbb{C}(X(N)) = \mathcal{A}_0(\Gamma(N)),$$

$$f_0^{\overline{d}}(\tau) \in \mathbb{C}(X_1(N)) = \mathcal{A}_0(\Gamma_1(N)),$$

$$f_0(\tau) \in \mathbb{C}(X_0(N)) = \mathcal{A}_0(\Gamma_0(N)).$$

Proof. It is not hard to show that they are weakly modular of weight 0 with respect to corresponding congruence subgroups, i.e. invariant under corresponding actions. It's sufficient to show that $f_0^{\overline{v}}$, $f_0^{\overline{d}}$, f_0 are meromorphic on $\mathbb{H}$ and at the cusps. Firstly, we consider $f_0^{\overline{v}}$.

For $v = (c_v, d_v) \in \mathbb{Z}^2$ with $v \not\equiv (0,0) \bmod N$, the function $f_0^{\overline{v}}(\tau)$ is meromorphic on $\mathbb{H}$ since $g_2(\tau)$, $g_3(\tau)$ are meromorphic on $\mathbb{H}$, and $\wp_\tau(z)$ is meromorphic on $\mathbb{H}\times\mathbb{C}$.

For any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathbb{H}$, let $m = (c\tau+d)^{-1}$, we have

$$\wp_{m\Lambda_\tau}(mz) = m^{-2}\wp_\tau(z),$$

$$m^{-1}(c_v\gamma(\tau) + d) = (ac_v + cd_v)\tau + (bc_v + dd_v).$$

Then

$$\begin{aligned} f_0^{\overline{v}}(\gamma(\tau)) &= \frac{g_2(\gamma(\tau))}{g_3(\gamma(\tau))}\wp_\gamma\left(\frac{c_v\gamma(\tau)+d_v}{N}\right) \\ &= \frac{m^{-4}g_2(\tau)}{m^{-6}g_3(\tau)}\wp_{m\Lambda_\tau}\left(\frac{c_v\gamma(\tau)+d_v}{N}\right) \\ &= \frac{g_2(\tau)}{m^{-2}g_3(\tau)}m^{-2}\wp_\tau\left(\frac{(ac_v+cd_v)\tau+(bc_v+dd_v)}{N}\right) \\ &= f_0^{\overline{v\gamma}}(\tau). \end{aligned}$$

Hence, it is sufficient to show that $f_0^{\overline{v}}$ is meromorphic at $i\infty$. We will show that $\lim_{\mathrm{Im}\,\tau\to\infty}\wp_\tau\left(\frac{c_v\tau+d_v}{N}\right)$ exists, which implies that $f_0^{\overline{v}}$ is meromorphic at $i\infty$.

$$\begin{aligned} \lim_{\mathrm{Im}\,\tau\to\infty}\wp_\tau\left(\frac{c_v\tau+d_v}{N}\right) &= \lim_{\mathrm{Im}\,\tau\to\infty}\frac{N^2}{(c_v\tau+d_v)^2}+ \\ &\sum_{(c,d)\in\mathbb{Z}\setminus\{(0,0)\}}\left(\lim_{\mathrm{Im}\,\tau\to\infty}\frac{N^2}{((c_v-Nc)\tau+(d_v-Nd))^2}-\lim_{\mathrm{Im}\,\tau\to\infty}\frac{1}{(c\tau+d)^2}\right) \\ &\longrightarrow \begin{cases} 2\sum_{d=1}^{\infty}\left(\frac{N^2}{(d_v-Nd)}-\frac{1}{d^2}\right) & \text{if } c_v\equiv 0 \bmod N; \\ -2\sum_{d=1}^{\infty}\frac{1}{d^2} & \text{if } c_v\not\equiv 0 \bmod N. \end{cases} \\ &= \begin{cases} -2\zeta(2)+2N^2\sum_{n\equiv d_v \bmod N}\frac{1}{n^2} & \text{if } c_v\equiv 0 \bmod N; \\ -2\zeta(2) & \text{if } c_v\not\equiv 0 \bmod N. \end{cases} \end{aligned}$$

Since $f_0^{\overline{d}}(\tau)=f_0^{\overline{(0,d)}}(\tau)$ for any non-zero element $\overline{d}\in\mathbb{Z}/N\mathbb{Z}$, so $f_0^{\overline{d}}(\tau)$ is meromorphic on $\mathbb{H}$, and $f_0^{\overline{d}}(\gamma(\tau))$ is meromorphic at $i\infty$ for any $\gamma\in\mathrm{SL}_2(\mathbb{Z})$. It is similar for $f_0(\tau)=\sum_{d=1}^{N-1}f_0^{\overline{d}}(\tau)$. □

Remark. *(1) For any $\gamma\in\mathrm{SL}_2(\mathbb{Z})$, $f_0^{\overline{v\gamma}}(\tau)=f_0^{\overline{v}}(\gamma(\tau))$. In particular, $f_0^{\overline{-v}}(\tau)=f_0^{\overline{v}}(\tau)\in\mathbb{C}(X(N))$. Similarly, $f_0^{\overline{-d}}(\tau)=f_0^{\overline{d}}(\tau)\in\mathbb{C}(X_1(N))$. More generally, for $\overline{v},\overline{w}\in\mathbb{Z}^2/N\mathbb{Z}^2$, $\overline{c},\overline{d}\in\mathbb{Z}/N\mathbb{Z}$,*

$$f_0^{\overline{v}}(\tau)=f_0^{\overline{w}}(\tau)\Longleftrightarrow v\equiv\pm w \bmod N,$$

$$f_0^{\overline{d}}(\tau)=f_0^{\overline{c}}(\tau)\Longleftrightarrow c\equiv\pm d \bmod N.$$

We also have

$$\#\{f_0^{\overline{v}}\mid 0\neq\overline{v}\in\mathbb{Z}^2/N\mathbb{Z}^2\}=\begin{cases}\frac{N^2-1}{2} & \text{if } N \text{ is odd,}\\ \frac{N^2}{2}+1 & \text{if } N \text{ is even,}\end{cases}$$

$$\#\{f_0^{\overline{d}}\mid 0\neq\overline{v}\in\mathbb{Z}/N\mathbb{Z}\}=\left[\frac{N-1}{2}\right]+1.$$

Proof. It is sufficient to show that $f_0^{\overline{v}}(\tau)=f_0^{\overline{w}}(\tau)\Longrightarrow v\equiv\pm w \bmod N$. Indeed, notice that $\wp_\tau(z)=\wp_\tau(z')$ if and only if $z\equiv z' \bmod \Lambda_\tau$, hence

$$\frac{c_v\tau+d_v}{N}\equiv\frac{c_w\tau+d_w}{N} \bmod \Lambda_\tau.$$

That is $N|(v-w)$ or $N|(v+w)$, i.e. $\overline{v}=\pm\overline{w}$. □

To make the following statement clearer, for a field extension L/K, we will let $\mathrm{Aut}_K(L)$ right act on L, and an element $x \in L$ acted by $\sigma \in \mathrm{Aut}_K(L)$ is denoted by x^σ. By the way, if $\mathrm{Aut}_K(L)$ left acts on L, we will use the notation $\sigma(x)$.

Proposition 2.4.3. *Keep the notations in Definition 2.4.1. Then the following statements hold:*

(1) $\mathbb{C}(X(N)) = \mathbb{C}\left(j, \left\{ f_0^{\pm\overline{v}} \mid \pm\overline{v} \in \dfrac{(\mathbb{Z}/N\mathbb{Z})^2 \setminus \{(0,0)\}}{\{\pm 1\}} \right\}\right) = \mathbb{C}(j, f_{1,0}, f_{0,1})$;

$\mathbb{C}(X_1(N)) = \mathbb{C}\left(j, \left\{ f_0^{\pm\overline{d}} \mid \pm\overline{d} \in \dfrac{(\mathbb{Z}/N\mathbb{Z}) \setminus \{0\}}{\{\pm 1\}} \right\}\right) = \mathbb{C}(j, f_1)$;

$\mathbb{C}(X_0(N)) = \mathbb{C}(j, f_0) = \mathbb{C}(j, j_N)$, *where* $j_N(\tau) = j(N\tau)$.

(2) The fields extensions $\mathbb{C}(X(N))/\mathbb{C}(X(1))$, $\mathbb{C}(X(N))/\mathbb{C}(X_1(N))$ *and* $\mathbb{C}(X(N))/\mathbb{C}(X_0(N))$ *are Galois extensions, and*

$$\theta^{-1} : \mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(X(1))) \simeq \frac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}} \simeq \frac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma(N)}$$

is an isomorphism of groups, where θ *is defined as following: for any* $[\gamma] \in \mathrm{SL}_2(Z)/\pm\Gamma(N)$ *with* $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, *and* $f \in \mathbb{C}(X(N))$,

$$f^{\theta([\gamma])} := f \circ \gamma.$$

Moreover, θ^{-1} *induces isomorphisms*

$$\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(X_1(N))) \simeq \left\{ \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in \frac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}} \right\} \simeq \frac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma_1(N)},$$

$$\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(X_0(N))) \simeq \left\{ \pm \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \frac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}} \right\} \simeq \frac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma_0(N)}.$$

Proof. By previous lemma, we have

$$\mathbb{C}\left(j, \left\{ f_0^{\pm\overline{v}} | \pm\overline{v} \in \frac{(\mathbb{Z}/N\mathbb{Z})^2 \setminus \{(0,0)\}}{\{\pm 1\}} \right\}\right) \subset \mathbb{C}(X(N)).$$

We define

$$\theta : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{Aut}_{\mathbb{C}(j)}(\mathbb{C}(X(N))),$$
$$\gamma \mapsto f \circ \gamma = f[\gamma]_0.$$

It is well-defined since $\Gamma(N) \subset \mathrm{SL}_2(\mathbb{Z})$ is normal. It defines a left group action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{C}(X(N))$.

We claim that $\mathrm{Ker}\,\theta = \pm\Gamma(N)$. Obviously, $\pm\Gamma(N) \subset \mathrm{Ker}\,\theta$. Conversely, if $\gamma \in \mathrm{Ker}\,\theta$, then for any non-zero element $\overline{v} \in (\mathbb{Z}/N\mathbb{Z})^2$, $f_0^{\overline{v}} = f_0^{\overline{v}} \circ \gamma = f_0^{\overline{v\gamma}}$. Hence $v\gamma = \pm\overline{v}$ for any $\overline{v} \in (\mathbb{Z}/N\mathbb{Z})^2$. Take $\overline{v} = (0,1), (1,0)$, then we know that $\overline{\gamma} = \pm I$, i.e. $\gamma \equiv I \bmod N$, $\gamma \in \pm\Gamma(N)$. Hence θ induces an injection $\dfrac{SL_2(\mathbb{Z})}{\pm\Gamma(N)} \hookrightarrow \mathrm{Aut}_{\mathbb{C}(j)}(\mathbb{C}(X(N)))$.

Before further discussion, we recall a fact: for a field extension L/K, if $G \subset \mathrm{Aut}_K(L)$ is a subgroup such that $L^G = K$, then L/K is Galois and $G = \mathrm{Aut}_K(L)$ is the Galois group of L/K. This is because that $K \subset L^{\mathrm{Aut}_K(L)} \subset L^G = K$.

To prove that $\mathbb{C}(X(N))/\mathbb{C}(X(1))$ is Galois with $\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(X(1))) \simeq \dfrac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma(N)}$, it is sufficient to show that the fixed subfield of $\theta(SL_2(\mathbb{Z}))$ is contained $\mathbb{C}(j)$. This is from the fact that $f \circ \gamma = f$ for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ if and only if f is weakly modular

with respect to $\mathrm{SL}_2(\mathbb{Z})$, and the fact that $f \in \mathbb{C}(X(N)) = \mathcal{A}_0(\Gamma(N))$ is automatically meromorphic on $\mathbb{H}$ and at all cusps.

We have $\mathbb{C}(j, f_{1,0}, f_{0,1}) \subset \mathbb{C}(X(N))$ and

$$\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(j, f_{1,0}, f_{0,1})) = \{\gamma \in SL_2(\mathbb{Z}) | f_{1,0} \circ \gamma = f_{0,1}, f_{0,1} \circ \gamma = f_{0,1}\} / \pm\Gamma(N).$$

Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ with $f_{1,0} \circ \gamma = f_0^{\overline{(1,0)}} \circ \gamma = f_0^{\overline{(a,b)}} = f_0^{\overline{(1,0)}}$ and $f_{0,1} \circ \gamma = f_0^{\overline{(0,1)}} \circ \gamma = f_0^{\overline{(c,d)}} = f_0^{\overline{(0,1)}}$. Then $\gamma \equiv \pm I \bmod N$. Hence $\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(j, f_{1,0}, f_{0,1}))$ is trivial and

$$\mathbb{C}(X(N)) = \mathbb{C}(j, f_{1,0}, f_{0,1}) = \mathbb{C}(j, \{f_0^{\pm\overline{v}} | \pm\overline{v} \in \frac{(\mathbb{Z}/N\mathbb{Z})^2 \setminus \{(0,0)\}}{\{\pm 1\}}\}).$$

For $X_1(N)$, the proof is similar. We define a left group action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{C}(X_1(N))$ by

$$\theta : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{Aut}_{\mathbb{C}(j)}(\mathbb{C}(X_1(N))),$$
$$\gamma \mapsto f \circ \gamma = f[\gamma]_0.$$

We can prove that $\mathrm{Ker}\,\theta = \pm\Gamma_1(N)$ by the fact that $\mathbb{C}(j, \{f_0^{\pm\overline{d}} | \pm\overline{d} \in \frac{(\mathbb{Z}/N\mathbb{Z})\setminus\{0\}}{\{\pm 1\}}\}) \subset \mathbb{C}(X_1(N))$ and $f_0^{\overline{d}} = f_0^{\overline{c}}$ if and only if $\overline{d} = \pm\overline{c}$. Then we have $\dfrac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma_1(N)} \simeq \theta(\mathrm{SL}_2(\mathbb{Z})) \subset \mathrm{Aut}_{\mathbb{C}(j)}(\mathbb{C}(X_1(N)))$. Moreover, as before, we can show that the fixed subfield of $\theta(SL_2(\mathbb{Z}))$ is $\mathbb{C}(j)$, which implies that $\mathbb{C}(X_1(N))/\mathbb{C}(j)$ is Galois and

$$\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(X_1(N))) \simeq \{\pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in \frac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}}\} \simeq \frac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma_1(N)}.$$

We have $\mathbb{C}(j, f_1) \subset \mathbb{C}(X_1(N))$ and

$$\mathrm{Gal}(\mathbb{C}(X_1(N))/\mathbb{C}(j, f_1)) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) | f_1 \circ \gamma = f_1\} / \pm\Gamma_1(N),$$

Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ with $f_1 \circ \gamma = f_0^{\overline{(0,1)}} \circ \gamma = f_0^{\overline{(c,d)}} = f_0^{\overline{(0,1)}}$. Then $c \equiv 0 \bmod N$, and $d \equiv \pm 1 \bmod N$. Hence $\mathrm{Gal}(\mathbb{C}(X_1(N))/\mathbb{C}(j, f_1))$ is trivial and

$$\mathbb{C}(X_1(N)) = \mathbb{C}(j, f_1) = \mathbb{C}(j, \{f_0^{\pm\overline{d}} | \pm\overline{d} \in \frac{(\mathbb{Z}/N\mathbb{Z}) \setminus \{0\}}{\{\pm 1\}}\}).$$

It is similar for $X_0(N)$. □

Remark. *(1) This proposition tells us that $X_1(N)$ (resp. $X_0(N)$) is birationally equivalent to a plane curve defined by the complex polynomial φ_1 (resp. φ_0) $\in \mathbb{C}[X, Y]$ such that $\varphi_1(j, f_1) = 0$ (resp. $\varphi_0(j, f_0)$) in $\mathbb{C}(X_1(N))$ (resp. $\mathbb{C}(X_0(N))$). We will see that the polynomials have rational coefficients.*

Corollary 2.4.4. *For any $N \in \mathbb{N}^+$, there is a bijection*

$$\{X_\Gamma \mid \Gamma \subset \mathrm{SL}_2(\mathbb{Z}) \text{ congruence subgroup of level } N\} \leftrightarrow \left\{\text{subgroups of } \frac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}}\right\},$$

$$X_\Gamma \mapsto \frac{\pm\Gamma}{\pm\Gamma(N)},$$
$$X_{\pi^{-1}(\overline{\Gamma})} \hookleftarrow \overline{\Gamma},$$

and $\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(X_\Gamma)) \simeq \dfrac{\pm\Gamma}{\pm\Gamma(N)}$, *where* $\pi : \mathrm{SL}_2(\mathbb{Z}) \to \dfrac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}}$ *is the quotient map.*

In general, if $\Gamma \subset \Gamma' \subset \mathrm{SL}_2(\mathbb{Z})$ *are congruence subgroups of level* N, *then* $\mathbb{C}(X_{\Gamma'})/\mathbb{C}(X_\Gamma)$ *is Galois if and only if* $\pm\Gamma' \subset \pm\Gamma$ *is normal. In this case, we have*

$$\mathrm{Gal}(\mathbb{C}(X_{\Gamma'})/\mathbb{C}(X_\Gamma)) \simeq \frac{\pm\Gamma}{\pm\Gamma'}.$$

Proof. The bijection is obviously from Lemma 2.1.3. Let $\theta : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{Aut}_{\mathbb{C}(j)}(\mathbb{C}(X(N)))$ define as in the Proposition 2.4.3. Then for any subgroup $\overline{\Gamma} \subset \dfrac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}}$, we have $\mathbb{C}(X(N))^{\theta(\overline{\Gamma})} = \mathbb{C}(X(\Gamma))$. Indeed, this comes from the fact that for any $f \in \mathbb{C}(X(N))$, $\theta(\overline{\gamma}) = f \circ \gamma^{-1} = f$ for any $\gamma \in \Gamma$ if and only if $f \in \mathbb{C}(X_\Gamma) = \mathcal{A}_0(\Gamma)$. Hence $\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(X_\Gamma)) \simeq \overline{\Gamma} = \dfrac{\pm\Gamma}{\pm\Gamma(N)}$.

In general, if $\Gamma \subset \Gamma' \subset \mathrm{SL}_2(\mathbb{Z})$ are congruence subgroups of level N, then by Galois theory, $\mathbb{C}(X_{\Gamma'})/\mathbb{C}(X_\Gamma)$ is Galois if and only if $\pm\Gamma' \subset \pm\overline{\Gamma}$ is normal i.e. $\pm\Gamma' \subset \pm\Gamma$ is normal. In this case, $\mathrm{Gal}(\mathbb{C}(X_{\Gamma'})/\mathbb{C}(X_\Gamma)) \simeq \dfrac{\pm\Gamma}{\pm\Gamma'}$. □

We have seen that $\mathrm{Gal}(\mathbb{C}(X(N))/\mathbb{C}(j)) \simeq \dfrac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}}$, and $\mathbb{C}(X(1)) = \mathbb{C}(j)$, we may wonder if we can find a Galois extension $K/\mathbb{C}(j)$ such that $\mathrm{Gal}(K/\mathbb{C}(j)) \simeq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. The answer is yes. With the construction of such K, we can see the field $\mathbb{C}(X(N))$ from a different point of view.

Definition 2.4.5. *For* $\tau \in \mathbb{H} \setminus j^{-1}(\{0, 1728\})$, *we define the universal elliptic curve by Weierstrass equation:*

$$E_j : y^2 = 4x^3 - (\frac{27j}{j-1728})x - (\frac{27j}{j-1728}).$$

Remark. *(1) We can view* E_j *as an elliptic curve over* $\mathbb{C}(j)$, *so we can talk about its* N*-torsion points.*

Proposition 2.4.6. *For any* $\tau \in \mathbb{H} \setminus j^{-1}\{0, 1728\}$, *we have isomorphisms of Riemann surfaces*

$$\varphi : \mathbb{C}/\Lambda_\tau \simeq E_{j(\tau)},$$

where $\Lambda_\tau = \langle \tau, 1 \rangle \subset \mathbb{C}$ *is a lattice. Moreover, the j-invariant of* $E_{j(\tau)}$ *is* $j(\tau)$, *and for any* $N \in \mathbb{N}^+$, φ *takes the canonical generators* $\dfrac{\tau}{N} + \Lambda_\tau$, $\dfrac{1}{N} + \Lambda_\tau$ *of* $\mathbb{C}/\Lambda_\tau$ *to*

$$\varphi\left(\frac{\tau}{N} + \Lambda_\tau\right) = P_\tau := \left(\frac{g_2(\tau)}{g_3(\tau)}\wp_\tau\left(\frac{\tau}{N}\right), \left(\frac{g_2(\tau)}{g_3(\tau)}\right)^{3/2}\wp'_\tau\left(\frac{\tau}{N}\right)\right),$$

$$\varphi\left(\frac{1}{N} + \Lambda_\tau\right) = Q_\tau := \left(\frac{g_2(\tau)}{g_3(\tau)}\wp_\tau\left(\frac{1}{N}\right), \left(\frac{g_2(\tau)}{g_3(\tau)}\right)^{3/2}\wp'_\tau\left(\frac{1}{N}\right)\right),$$

where $\left(\dfrac{g_2(\tau)}{g_3(\tau)}\right)^{3/2}$ *is fixed with respect to* φ.

Proof. Without confusion, for $\tau \in \mathbb{H}$, we denote $E_\tau : y^2 = 4x^3 - g_2(\tau)x - g_3(\tau)$, the elliptic curve over $\mathbb{C}$ defined by the affine equation. We fix $u = \left(\frac{g_2(\tau)}{g_3(\tau)}\right)^{1/2}$ which is in $\mathbb{C}^*$ since $\tau \notin j^{-1}(\{0, 1728\}) = \{\zeta_3, i\}$, and we take an admissible change of variables

$$(x, y) \mapsto (u^2, u^3 y).$$

Then we get the elliptic curve defined by $y^2 = 4x^3 - \frac{g_2(\tau)^3}{g_3(\tau)^2}x - \frac{g_2(\tau)^3}{g_3(\tau)^2}$, which is $E_{j(\tau)}$, since $j(\tau) = \frac{g_2^3}{g_2^3 - 27g_3^2}$, i.e. $\frac{g_2(\tau)^3}{g_3(\tau)^2} = \frac{27j(\tau)}{j(\tau) - 1728}$. The morphism is given as following:

$$\begin{aligned} \varphi : \mathbb{C}/\Lambda_\tau &\to E_{j(\tau)} \\ z &\mapsto (u^2 \wp(z), u^3 \wp'(z)), \end{aligned}$$

so φ maps $\frac{\tau}{N} + \Lambda_\tau$, $\frac{1}{N} + \Lambda_\tau$ to P_τ, Q_τ respectively. □

Recall that for an elliptic curve $E : y^2 = x^3 + ax + b$ over a field K, the x-coordinates of N-torsion points are characterized by a polynomial $\psi_N(a, b, x) = 0$ with $\psi_N \in \mathbb{Z}[a, b, x]$.

Corollary 2.4.7. *Let $N \in \mathbb{N}^+$, and E_j be the universal elliptic curve. Then the non-zero x-coordinates of points in $E_j[N](\overline{\mathbb{C}(j)})$ are $\{f_0^{\pm\overline{v}} | \pm\overline{v} \in \frac{(\mathbb{Z}/N\mathbb{Z})^2 \setminus \{(0,0)\}}{\{\pm 1\}}\}$. Moreover, if $P, Q \in E_j[N](\overline{\mathbb{C}(j)})$ such that $x(P) = f_0^{\overline{v}}$ and $x(Q) = f_0^{\overline{u}}$, then $x(P + Q) = f_0^{\overline{v+u}}$.*

Proof. Let $g = \frac{27j}{j - 1728}$, and $\psi_N(g, g, x) \in \mathbb{Z}[g, x]$ be such that the x-coordinates points in $E_j[N]$ are characterized by $\psi_N(g, g, x) = 0$. For $\overline{v} = (\overline{v}_1, \overline{v}_2) \in \mathbb{Z}/N\mathbb{Z} \setminus \{0\}$, consider $\psi_N(g, g, f_0^{\pm\overline{v}}) \in \mathbb{C}(X(N))$. We claim that $\psi_N(g, g, f_0^{\pm\overline{v}}) = 0$. Indeed, for any $\tau \notin j^{-1}(\{0, 1728\})$, we have $\psi_N(g(\tau), g(\tau), f_0^{\pm\overline{v}}(\tau)) = 0$, since $f_0^{\pm\overline{v}}(\tau)$ is the x-coordinate of $\overline{v}_1 P_\tau + \overline{v}_2 Q_\tau \in E_{j(\tau)}[N](\mathbb{C})$, where P_τ, Q_τ are defined in Proposition 2.4.6. Hence $\psi_N(g, g, f_0^{\pm\overline{v}}) = 0$ since it has infinitely many zeros on $X(N)$, and $X(N)$ is compact.

To prove $\{f_0^{\pm\overline{v}} | \pm\overline{v} \in \frac{(\mathbb{Z}/N\mathbb{Z})^2 \setminus \{(0,0)\}}{\{\pm 1\}}\} = \{x\text{-coordinates of points in } E_j[N](\overline{\mathbb{C}(j)})\}$, it remains to show that they have the same cardinality. This comes from the remark of Lemma 2.4.2 and the fact that $\#E_j[N] \simeq \mathbb{Z}^2/N\mathbb{Z}^2$ and the zero of E_j is the infinity which has no x-coordinate. □

Proposition 2.4.8. *Let $N \in \mathbb{N}^+$, and E_j be the universal elliptic curve. Then*

$$\mathbb{C}(X(N)) = \mathbb{C}(j, x(E_j[N])).$$

In particular,

$$\mathrm{Gal}(\mathbb{C}(j, x(E_j[N]))/\mathbb{C}(j)) \simeq \frac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}}.$$

Proof. It comes from Proposition 2.4.3 and Corollary 2.4.7 □

Proposition 2.4.9. *Let $N \in \mathbb{N}^+$, and E_j be the universal elliptic curve. Then the field extension $\mathbb{C}(j, E_j[N])/\mathbb{C}(j)$ is Galois and there is an isomorphism*

$$\theta^{-1} : \mathrm{Gal}(\mathbb{C}(j, E_j[N])/\mathbb{C}(j)) \simeq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}).$$

Moreover, this isomorphism is compatible with θ^{-1} in Proposition 2.4.3.

Proof. Let $\sigma : \mathbb{C}(j, E_j[N]) \to \overline{\mathbb{C}(j)}$ be a embedding fixing $\mathbb{C}(j)$. Then σ permutes points in $E_j[N](\overline{\mathbb{C}(j)})$, hence $\mathbb{C}(j, E_j[N])/\mathbb{C}(j)$ is Galois.

Denote $\mathrm{Gal}(\mathbb{C}(j, E_j[N])/\mathbb{C}(j))$ by H. Fix an ordered basis (P_τ, Q_τ) of $E_j[N]$, we define a homomorphism of groups:

$$\rho : H \to \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$$

such that

$$\begin{pmatrix} \sigma(P_\tau) \\ \sigma(Q_\tau) \end{pmatrix} = \rho(\sigma) \begin{pmatrix} P_\tau \\ Q_\tau \end{pmatrix}.$$

If $\sigma \in H$ such that $\rho(\sigma) = I$, then $\sigma(P_\tau) = P_\tau$, and $\sigma(Q_\tau) = Q_\tau$. Notice that σ induces a endomorphism of $E_j[N]$, so $\sigma(P) = P$ for any $P \in E_j[N]$ and $\sigma = \mathrm{id} \in H$. This proves the injectivity.

We claim that $\rho(H) = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. For any $\sigma \in H$, consider the Weil pairing: since $e_N(P_\tau, Q_\tau) \in \mathbb{C}$ is a primitive N-th root of unity, so

$$e_N(P_\tau, Q_\tau) = \sigma(e_N(P_\tau), Q_\tau) = e_N(\sigma(P_\tau), \sigma(Q_\tau)) = e_N(P_\tau, Q_\tau)^{\det \rho(\sigma)},$$

which implies $\det \rho(\sigma) \in \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, $\rho(H) \subset \mathrm{SL}_2(\mathbb{Z}/NZ)$.

We have that $[\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho(H)] \leq 2$. To prove this, we set

$$W := \mathrm{Gal}(\mathbb{C}(j, E_j[N])/\mathbb{C}(j, x(E_j[N]))).$$

For any $\sigma \in W$, we have $\sigma(P) = \pm P$ for any $P \in E_j[N]$, then $\rho(\sigma) \in \{\pm I\} \subset \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, i.e. $W \subset \rho^{-1}(\{\pm I\})$. On the other hand, if $\sigma \in H$ such that $\rho(\sigma) = \pm I$, we also know that σ fixes $x(E_j[N])$. Hence $W = \rho^{-1}(\{\pm I\})$,

$$H/W \simeq \mathrm{Gal}(\mathbb{C}(j, x(E_j[N]))/C(j)) \simeq \frac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\pm I},$$

$$\frac{2}{\#W} = \frac{\#(\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}))}{\#H} = [\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho(H)] \leq 2.$$

If $[\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) : \rho(H)] = 2$, then $\#W = \#\rho^{-1}(\{\pm I\})$ and $-I \notin \rho(H)$. Hence $(-\rho(H)) \bigcup \rho(H) = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. One of $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, -\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ belongs to $\rho(H)$, so $-I = \left(\pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right)^2 \in \rho(H)$. Contradiction. This proves that $\rho(H) = \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. □

2.4.2 Function Fields over $\mathbb{Q}$

Proposition 2.4.10. *Let $N \in \mathbb{N}^+$, and E_j be the universal elliptic curve. Then*

(1) $\zeta_N \in \mathbb{Q}(j, E_j[N])$ and $\overline{\mathbb{Q}} \bigcap \mathbb{Q}(j, E_j[N]) = \mathbb{Q}(\zeta_N)$.

(2) $\mathbb{Q}(j, E_j[N])/\mathbb{Q}(j)$ is Galois with

$$\mathrm{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(j)) \overset{\rho}{\simeq} \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}),$$

where $\begin{pmatrix} P_\tau^\sigma \\ Q_\tau^\sigma \end{pmatrix} = \rho(\sigma) \begin{pmatrix} P_\tau \\ Q_\tau \end{pmatrix}$, (P_τ, Q_τ) is an ordered basis of $E_j[N]$ over $\mathbb{Z}/N\mathbb{Z}$, and $\sigma(\zeta_N) = \zeta_N^{\det \rho(\sigma)}$ for any $\sigma \in \mathrm{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(j))$. Hence $(f_0^{\overline{v}})^\sigma = f_0^{\overline{v}\rho(\sigma)}$ for any $0 \neq \overline{v} \in \mathbb{Z}/N\mathbb{Z}$. Moreover, ρ induces an isomorphism of exact sequences:

$$\begin{array}{ccccccccc}
1 & \longrightarrow & \mathrm{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(\zeta_N, j)) & \longrightarrow & \mathrm{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(j)) & \longrightarrow & \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) & \longrightarrow & 1 \\
 & & \downarrow & & \downarrow{\scriptstyle \rho} & & \downarrow & & \\
1 & \longrightarrow & \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) & \longrightarrow & \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) & \xrightarrow[\det]{} & (\mathbb{Z}/N\mathbb{Z})^* & \longrightarrow & 1
\end{array}$$

(3) for any subgroup $G \subset \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, we have $K \bigcap \overline{\mathbb{Q}} = \mathbb{Q}(\zeta_N)^{\det G}$, where K is the intermediate field with $\mathrm{Gal}(\mathbb{Q}(j, E_j[N])/K) \simeq G$ via ρ. Thus K is a fraction field of a geometrically integral smooth projective curve over $\mathbb{Q}(\zeta_N)^{\det G} = K \bigcap \overline{\mathbb{Q}}$.

Proof. (1) will come from the proof of (2) and (3).

For (2), since $P^\sigma \in E_j[N]$ for any embedding $\sigma : \mathbb{Q}(j, E_j[N]) \hookrightarrow \overline{\mathbb{Q}(j)}$ fixing $\mathbb{Q}(j)$ and $P \in E_j[N]$, so $x(P^\sigma), y(P^\sigma) \in \mathbb{Q}(j, E_j[N])$. Hence $\mathbb{Q}(j, E_j[N])$ is normal and separable over $\mathbb{Q}(j)$, i.e. Galois.

We set $H_{\mathbb{Q}} = \mathrm{Gal}(\mathbb{Q}(\zeta_N, j, E_j[N])/\mathbb{Q}(j))$ and a representation

$$\rho : H_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$$

such that

$$\begin{pmatrix} P_\tau^\sigma \\ Q_\tau^\sigma \end{pmatrix} = \rho(\sigma) \begin{pmatrix} P_\tau \\ Q_\tau \end{pmatrix},$$

where(P_τ, Q_τ) is a fixed basis of $E_j[N]$ over $\mathbb{Z}/N\mathbb{Z}$. Then we claim $\zeta_N^\sigma = \zeta_N^{\det \rho(\sigma)}$ for any $\sigma \in H_{\mathbb{Q}}$. Indeed, since Weil paring is bilinear and commutes with Galois actions, so $e_N(P_\tau, Q_\tau)^\sigma = e_N(P_\tau^\sigma, Q_\tau^\sigma) = e_N(P_\tau, Q_\tau)^{\det \rho(\sigma)}$. Combining this with the fact that there exists $k \in \mathbb{N}$ such that $\zeta_N = e_N(P_\tau, Q_\tau)^k$, we have our claim. Hence

$$\mathrm{Gal}(\mathbb{Q}(\zeta_N, j, E_j[N])/\mathbb{Q}(j, E_j[N])) = 1,$$

i.e. for any $\sigma \in H_{\mathbb{Q}}$ fixing $E_j[N]$, $\zeta_N^\sigma = \zeta_N^{\det \rho(\sigma)} = \zeta_N$. We conclude that $\zeta_N \in \mathbb{Q}(j, E_j[N])$ and $H_{\mathbb{Q}} = \mathrm{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(j))$.

Firstly, we prove that ρ is injective. If $\sigma \in \mathrm{Ker}\, \rho$, and $P = aP_\tau + bQ_\tau \in E_j[N]$, then $P^\sigma = aP_\tau^\sigma + bQ_\tau^\sigma = aP_\tau + bQ_\tau = P$, since the map $E \to E$ induced by σ is a homomorphism of elliptic curves. Hence $\sigma = \mathrm{id}$ and ρ is injective.

Next we will show that ρ is surjective. We have the following diagram of field extensions:

$$\begin{array}{ccccc} & & \mathbb{C}(j, E_j[N]) & & \\ & \swarrow & & \searrow & \\ \mathbb{C}(j) & & & & \mathbb{Q}(j, E_j[N]). \\ & \searrow & & \swarrow & \\ & & \mathbb{Q}(\zeta_N, j) & & \end{array}$$

By the restriction lemma in Galois theory, we have an injective homomorphism induced by restriction

$$\mathrm{Gal}(\mathbb{C}(j, E_j[N])/\mathbb{C}(j)) \hookrightarrow H_{\mathbb{Q}(\zeta)} := \mathrm{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(\zeta_N, j)).$$

Notice that $\mathrm{Gal}(\mathbb{Q}(\zeta_N, j)/\mathbb{Q}(j)) = H_{\mathbb{Q}}/H_{\mathbb{Q}(\zeta_N)}$, i.e. $(\mathbb{Z}/N\mathbb{Z})^* \simeq \mathrm{Im}\, \rho / H_{\mathbb{Q}(\zeta_N)}$, and by Proposition 2.4.9, then

$$\#(\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})) \leq \#(H_{\mathbb{Q}(\zeta_N)}) \leq \frac{\#(\mathrm{Im}\, \rho)}{\#((\mathbb{Z}/N\mathbb{Z})^*)} \leq \frac{\#(\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}))}{\#((\mathbb{Z}/N\mathbb{Z})^*)} \leq \#(\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})).$$

Thus they all equal and $\mathrm{Im}\, \rho = \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$, $H_{\mathbb{Q}(\zeta_N)} \simeq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. Again, by the restriction lemma, we deduce that $\mathbb{C}(j) \bigcap \mathbb{Q}(j, E_j[N]) = \mathbb{Q}(\zeta_N, j)$.

To prove the isomorphism of exact sequences, we have the following commutative diagram:

$$\begin{array}{ccc} \mathrm{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(j)) & \longrightarrow & \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q}) \\ \downarrow \rho & & \downarrow \tau \\ \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) & \xrightarrow{\det} & (\mathbb{Z}/N\mathbb{Z})^* \end{array},$$

where $\zeta_N^\sigma = \zeta_N^{\tau(\sigma)}$ for any $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_N)/\mathbb{Q})$. Since the vertical maps are isomorphisms, then the morphism of exact sequences is an isomorphism.

To show that $(f_0^{\overline{v}})^\sigma = f_0^{\overline{v}\rho(\sigma)}$ for any $0 \neq \overline{v} = (\overline{v}_1, \overline{v}_2) \in \mathbb{Z}/N\mathbb{Z}$, we take P_τ, $Q_\tau \in E_j[N]$ such that $x(P_\tau) = f_0^{\overline{(1,0)}}$, $x(Q_\tau) = f_0^{\overline{(0,1)}}$. We set $T_\tau = \overline{v}_1 P_\tau + \overline{v}_2 Q_\tau$, then $x(T_\tau) = f_0^{\overline{v}}$ by Corollary 2.4.7. Hence

$$T_\tau^\sigma = \overline{v} \begin{pmatrix} P_\tau^\sigma \\ Q_\tau^\sigma \end{pmatrix} = \overline{v}\rho(\sigma) \begin{pmatrix} P_\tau \\ Q_\tau \end{pmatrix}.$$

In particular, $(f_0^{\overline{v}})^\sigma = x(T_\tau^\sigma) = f_0^{\overline{v}\rho(\sigma)}$.

For (3), we have that $\mathbb{Q}(\zeta_N)^{\det G} \subset \mathbb{Q}(j, E_j[N])^G = K$. □

Remark. *(1) From the proof, we have $\mathbb{C}(j) \bigcap \mathbb{Q}(j, E_j[N]) = \mathbb{Q}(\zeta_N, j)$ and the following commutative diagram:*

$$\begin{array}{ccc} \mathrm{Gal}(\mathbb{C}(j, E_j[N])/\mathbb{C}(j)) & \xrightarrow{\quad\sim\quad} & \mathrm{Gal}(\mathbb{Q}(j, E_j[N])/\mathbb{Q}(\zeta_N, j)) \\ & {\scriptstyle\theta^{-1}}\searrow \quad \swarrow{\scriptstyle\rho} & \\ & \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) & \end{array}$$

where θ^{-1} is the map in Proposition 2.4.9. Basically, θ^{-1} and ρ are the same.

(2) For sub-extension $\mathbb{Q}(j, x(E_j[N]))/\mathbb{Q}(j)$, the inverse of the isomorphism

$$\mathrm{Gal}(\mathbb{Q}(j, x(E_j[N]))/\mathbb{Q}(j)) \overset{\rho}{\simeq} \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm I\},$$

is similar with θ in Proposition 2.4.10. That is, for any $\gamma \in \mathrm{GL}_2(\mathbb{Z})$ and $f \in \mathbb{Q}(j, x(E_j[N])) \subset \mathbb{C}(X(N))$,

$$f^{\rho^{-1}(\overline{\gamma})} = f \circ [\gamma]$$

Corollary 2.4.11. *For any congruence subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ of level N, let X be the corresponding projective curves of $\overline{\Gamma} \subset \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}) \subset \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ in (3) of Proposition 2.4.10. Then X is a model of modular curve X_Γ over $\mathbb{Q}(\zeta_N)$, i.e. $X(\mathbb{C}) \simeq X_\Gamma$ as Riemann surfaces. Moreover, we have the following commutative diagram*

$$\begin{array}{ccc} \mathrm{Gal}(\mathbb{C}(X_\Gamma)/\mathbb{C}(j)) & \xrightarrow{\quad\sim\quad} & \mathrm{Gal}(K/\mathbb{Q}(\zeta_N, j)) \\ & {\scriptstyle\theta^{-1}}\searrow \quad \swarrow{\scriptstyle\rho} & \\ & \dfrac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma} & \end{array}$$

We will still denote X by $X_{\Gamma,\mathrm{alg}}$, or simply by X_Γ if there is no confusion.

Proof. By (2) of Proposition 2.4.10,

$$\mathrm{Gal}(K/\mathbb{Q}(\zeta_N, j)) \simeq \frac{\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})}{\pm\Gamma},$$

and we have the following diagram of field extensions:

$$\begin{array}{ccccc} & & \mathbb{C}K & & \\ & \swarrow & & \searrow & \\ \mathbb{C}(j) & & & & K. \\ & \searrow & & \swarrow & \\ & & \mathbb{Q}(\zeta_N, j) & & \end{array}$$

Notice that $\mathbb{C}(j) \bigcap K = \mathbb{Q}(\zeta_N, j)$, so $\mathrm{Gal}(\mathbb{C}K/\mathbb{C}(j)) \simeq \mathrm{Gal}(K/\mathbb{Q}(\zeta_N, j))$. Hence we have the following commutative diagram

$$\begin{array}{ccc} \mathrm{Gal}(\mathbb{C}K/\mathbb{C}(j)) & \xrightarrow{\quad\sim\quad} & \mathrm{Gal}(K/\mathbb{Q}(\zeta_N, j)) \\ & {\scriptstyle\theta^{-1}}\searrow \quad \swarrow{\scriptstyle\rho} & \\ & \dfrac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma} & \end{array}$$

On the other hand, by Corollary 2.4.4, $\theta^{-1} : \mathrm{Gal}(\mathbb{C}(X_\Gamma)/\mathbb{C}(j)) \simeq \dfrac{\mathrm{SL}_2(\mathbb{Z})}{\pm\Gamma}$. Hence $\mathbb{C}K = \mathbb{C}(X_\Gamma)$. The fact that X is geometrically integral over $\mathbb{Q}(\zeta_N)$ implies that the function field of $X_{\mathbb{C}} := X \times_{\mathrm{Spec}(\mathbb{Q}(\zeta_N))} \mathrm{Spec}(\mathbb{C})$ is exactly $K \otimes_{\mathbb{Q}(\zeta_N)} \mathbb{C} = \mathbb{C}K$. Hence $X(\mathbb{C})$ and X_Γ have the same functional fields as projective Riemann surfaces, which implies that $X(\mathbb{C}) \simeq X_\Gamma$. □

Remark. *(1) In general, a model of X_Γ over $\mathbb{Q}(\zeta_N)$ is not unique.*

(2) The model $X_{\Gamma,\mathrm{alg}}$ may be defined over a subfield of $\mathbb{Q}_N$. For example, $X_0(N)$ is defined over $\mathbb{Q}$. In general, if there exists a subgroup $G \subset \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ such that $\overline{\Gamma} = G \bigcap \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, then $X_{\Gamma,\mathrm{alg}}$ is defined over $\mathbb{Q}(\zeta_N)^{\det G}$.

(3) Notice that the models $X(N)$, $X_1(N)$ and $X_0(N)$ correspond to the fields $\mathbb{Q}(j, x(E_j[N])) = \mathbb{Q}(j, f_{1,0}, f_{0,1})$, $\mathbb{Q}(j, f_1)$ and $\mathbb{Q}(j, f_0)$ respectively.

We know that $X(N)$ is defined over $\mathbb{Q}(\zeta_N)$, and

$$\mathrm{Gal}(\mathbb{Q}(\zeta_N)(X(N))/\mathbb{Q}(j)) \simeq \frac{\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})}{\{\pm I\}}.$$

Example 2.4.12 (T.Weston)**.** $X_0(11) : y^2 + y = x^3 - x^2 - 10x - 20$.

2.5 The Field of Modular Functions over a Number Field

This is main part we want to discuss in this Chapter, the main reference is [56, Section 6.2].

Definition 2.5.1. *Let $N \in \mathbb{N}^+, k \in \mathbb{Z}$, and $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup of level N. We say an automorphic form $f \in \mathcal{A}_k(\Gamma)$ is defined over a field $K \subset \mathbb{C}$ if the coefficients of its q_N-expansion lie in K, where $q_N = e^{2\pi i\tau/N}$.*

We denote the set of automorphic forms (resp. modular forms, resp. cusp forms) of weight k (with respect to Γ) defined over K by $\mathcal{A}_k(K, \Gamma)$ (resp. $\mathcal{M}_k(K, \Gamma)$, resp. $\mathcal{S}_k(K, \Gamma)$), and set

$$\mathcal{A}_k(K) = \bigcup_\Gamma \mathcal{A}_k(K, \Gamma),$$

$$\mathcal{M}_k(K) = \bigcup_\Gamma \mathcal{M}_k(K, \Gamma),$$

$$\mathcal{S}_k(K) = \bigcup_\Gamma \mathcal{S}_k(K, \Gamma).$$

Remark. *(1) Since $\mathcal{A}_k(\Gamma) \subset \mathcal{A}_k(\Gamma(N))$ for some N, we mainly consider automorphic forms with respect to $\Gamma(N)$. For a automorphic form $f \in \mathcal{A}_k = \bigcap_N \mathcal{A}_k(\Gamma(N)) = \bigcap_\Gamma \mathcal{A}_k(\Gamma)$, whether f is defined over K or not is independent of the choice of N.*

Proof. If $f \in \mathcal{A}_k(\Gamma(N))$, we take minimal $h \in \mathbb{N}^+$ such that $f \in \mathcal{A}_k(\Gamma(h))$, then $h|N$ and the assertion comes from the fact $q_h = q_N^{N/h}$. □

(2) If $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup of level N, we easily have

$$\mathcal{A}_k(K, \Gamma) = \{f \in \mathcal{A}_k(K, \Gamma(N)) | f \text{ is } \Gamma\text{-invariant}\}.$$

In algebraic geometry, for a geometrically integral variety X defined over a number field K, we say that a rational function $f \in \mathbb{C}(X)$ is defined over K if $f \in K(X) \subset \mathbb{C}(X)$. We have know that, for a modular curve X_Γ, $\mathbb{C}(X) = \mathcal{A}_0(\Gamma)$. Hence we may wonder, when X_Γ is defined over K, if the field of rational functions $K(X\Gamma)$ is exactly $\mathcal{A}_0(K, \Gamma)$ or not. If the answer is yes, then there is no confusion with the definition above and the one in algebraic geometry. Although I cannot find the answer for general cases we still have the following Proposition, see [56, Proposition 6.9] for the proof.

Proposition 2.5.2. *Let $N \in \mathbb{N}^+$. Then we have*

(1) $\mathcal{A}_0(\mathbb{Q}(\zeta_N), \Gamma(N)) = \mathbb{Q}(\zeta_N)(X(N))$;

(2) $\mathcal{A}_0(\mathbb{Q}, \Gamma(N)) = \mathbb{Q}(j, j_N, f_{1,0}) \subset \mathbb{Q}(\zeta_N)(X(N))$ *with*

$$\mathrm{Gal}(\mathbb{Q}(\zeta_N)(X(N))/\mathbb{Q}(j, j_N, f_{1,0})) \simeq \{\begin{pmatrix} \pm 1 & 1 \\ 0 & x \end{pmatrix} | x \in (\mathbb{Z}/N\mathbb{Z})^*\}/\{\pm I\},$$

where $j_N(\tau) = j(N\tau)$.

Corollary 2.5.3. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup of level N, K be a field containing $\mathbb{Q}(\zeta_N)$. Then $K(X_\Gamma) = \mathcal{A}_0(K, \Gamma)$.*

Proof. By the proposition above, the corollary holds for $\Gamma(N)$, i.e. $\mathcal{A}_0(K, \Gamma(N)) = K(X(N))$. We have

$$\mathcal{A}_0(K, \Gamma) = \{f \in \mathcal{A}_0(K, \Gamma(N)) \mid f \text{ is } \Gamma\text{-invariant}\}.$$

On the other hand, since $\mathbb{Q}(\zeta_N) \subset K$, so $\mathrm{Gal}(K(X(N))/K(X_\Gamma)) \simeq \overline{\pm\Gamma}$, where $\overline{\pm\Gamma}$ is the image of Γ in $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. Notice that the right $\overline{\pm\Gamma}$-action on $K(X(N))$ is defined as following: for an element $[\gamma] \in \overline{\pm\Gamma}$ with $\gamma \in \pm\Gamma$,

$$(f)^{[\gamma]} = f[\gamma^{-1}]_0.$$

Hence

$$K(X_\Gamma) = K(X(N))^{\overline{\pm\Gamma}} = \{f \in K(X(N)) \mid f \text{ is } \Gamma\text{-invariant}\} = \mathcal{A}_0(K, \Gamma(N)).$$

□

Chapter 3

Integral Points on Modular Curves

In this chapter, we are going to prove our main results of the first part, Theorem 3.2.2 and Theorem 3.3.1 by using Baker's method and Chevalley-Weil principle.

3.1 Modular Units

In order to use Baker's inequality to bound integral points on an algebraic curve, one important step is to know its group of Σ_x-units. For a modular curve, that is the group of modular units.

In this section, we recall some ingredients to define modular units and some facts about them. The main references are [35] and [36], see also [1].

3.1.1 The Weierstrass sigma and zeta functions

Definition 3.1.1. *Let $\Lambda \subset \mathbb{C}$ be a lattice, we write down the Weierstrass sigma function, which has zeros of order 1 at all lattice points, by the Weierstrass product*

$$\sigma(z;\Lambda) := z \prod_{0\neq w\in\Lambda} (1 - z/w)e^{z/w+1/2(z/w)^2}.$$

Taking the logarithmic derivative of $\sigma(z;\Lambda)$ formally yields the Weierstrass zeta function

$$\zeta(z;\Lambda) := \frac{\sigma'(z;\Lambda)}{\sigma(z;\Lambda)} = \frac{1}{z} + \sum_{0\neq w\in\Lambda} (\frac{1}{z-w} + \frac{1}{w} + \frac{z}{w^2}).$$

We see that, for any $w \in \Lambda$, $\zeta(z+w;\Lambda) - \zeta(z;\Lambda)$ is independent of the choice of $z \in \mathbb{C}$. It is denoted by $\eta(w;\Lambda)$, called the Weierstrass eta function.

For $\tau \in \mathbb{H}$, we will denote $\sigma(z;\Lambda_\tau)$, $\zeta(z;\Lambda_\tau)$ and $\eta(w;\Lambda_\tau)$ by $\sigma(z;\tau)$, $\zeta(z;\tau)$ and $\eta(w;\tau)$ respectively.

Remark. *(1) By the Weierstrass factorization Theorem, see [64, Theorem 2.2.2], it is easy to see that $\sigma(z;\Lambda)$ uniformly converges on any compact subset of $\mathbb{C}$, and it is analytic on $\mathbb{C}$. Its zeros (of order 1) are the points on Λ.*

(2) From above, $\zeta(z;\Lambda)$ converges absolutely and uniformly on any compact subset of $\mathbb{C}$ not containing any lattice point. We also have

$$\zeta'(z;\Lambda) = -\wp(z;\Lambda) = -\frac{1}{z^2} - \sum_{0\neq w\in\Lambda} (\frac{1}{(z-w)^2} - \frac{1}{w^2}).$$

(3) *For any $\lambda \in \mathbb{C}^*$ and $z \in C$, we have*

$$\sigma(\lambda z; \lambda\Lambda) = \lambda\sigma(z; \Lambda),$$

$$\zeta(\lambda z; \lambda\Lambda) = \frac{1}{\lambda}\zeta(z; \Lambda).$$

In particular, $\sigma(\zeta; \Lambda)$ and $\zeta(z; \Lambda)$ are odd functions.

(4) *For any $w \in \Lambda$, the function $\zeta(z + w; \Lambda) - \zeta(z; \Lambda)$ is independent of the choice of $z \in \mathbb{C}$, i.e. $\eta(w; \Lambda)$ is well-defined.*

Proof. We fix $w_0 \in \Lambda$, then

$$\begin{aligned}\frac{d}{dz}(\zeta(z + w_0, \Lambda) - \zeta(z, \Lambda)) &= \zeta'(z + w_0; \Lambda) - \zeta'(z; \Lambda) \\ &= -\wp(z + w_0; \Lambda) + \wp(z; \Lambda) \\ &= 0,\end{aligned}$$

since the Weierstrass elliptic function $\wp(z; \Lambda)$ is Λ-periodic. Hence $\zeta(z + w_0, \Lambda) - \zeta(z, \Lambda)$ is a constant for a fixed w_0. □

(5) *For any $\lambda \in \mathbb{C}^*$ and $w \in \Lambda$, we have*

$$\eta(\lambda w; \lambda\Lambda) = \frac{1}{\lambda}\eta(w; \Lambda).$$

The map $\eta(\cdot; \Lambda) : \Lambda \to \mathbb{C}$ is a homomorphism of groups. Hence, it can be extended to an $\mathbb{R}$-linear map $\eta(\cdot; \Lambda) : \mathbb{C} \to \mathbb{C}$.

Proof. By definition, for $\lambda \in \mathbb{C}^*$, $w, w' \in \Lambda$,

$$\begin{aligned}\eta(\lambda w, \lambda\Lambda) &= \zeta(\lambda z + \lambda w; \lambda\Lambda) - \zeta(\lambda z; \lambda\Lambda) \\ &= \frac{1}{\lambda}\zeta(z + w; \Lambda) - \frac{1}{\lambda}\zeta(z; \Lambda) \\ &= \frac{1}{\lambda}\eta(w; \Lambda),\end{aligned}$$

$$\begin{aligned}\eta(w + w'; \Lambda) &= (\zeta(z + w + w'; \Lambda) - \zeta(z + w; \Lambda)) + (\zeta(z + w; \Lambda) - \zeta(z; \Lambda)) \\ &= \eta(w; \Lambda) + \eta(w'; \Lambda).\end{aligned}$$

□

For simplicity, when there is no confusion, we will omit Λ in $\sigma(z; \Lambda)$, $\zeta(z; \Lambda)$ and $\eta(w; \Lambda)$.

Theorem 3.1.2. *Let $\Lambda \subset \mathbb{C}$ be a lattice. Then for any $z \in \mathbb{C}$ and $w \in \Lambda$*

$$\frac{\sigma(z + w)}{\sigma(z)} = \psi(w)e^{\eta(w)(z + w/2)},$$

where

$$\psi(w) = \begin{cases} 1 & \text{if } w/2 \in \Lambda, \\ -1 & \text{if } w/2 \notin \Lambda. \end{cases}$$

Proof. For $z \in \mathbb{C}$ and $w \in \Lambda$, we have

$$\frac{d}{dz}\log(\frac{\sigma(z+w)}{\sigma(z)}) = \zeta(z+w) - \zeta(z) = \eta(w),$$

hence $\log(\frac{\sigma(z+w)}{\sigma(z)}) = \eta(w)z + c(w)$ for some function $c(w)$, i.e. $\sigma(z+w) = \sigma(z)e^{\eta(w)z+c(w)}$. By

$$\begin{aligned}\frac{\sigma(z+2w)}{\sigma(z)} &= \frac{\sigma(z+2w)}{\sigma(z+w)}\frac{\sigma(z+w)}{\sigma(z)} \\ &= e^{\eta(w)(z+w)+c(w)}e^{\eta(w)z+c(w)} \\ &= e^{\eta(2w)z+c(2w)},\end{aligned}$$

if we put $z = -w$, then $e^{-\eta(w)w+2c(w)} = e^{-\eta(2w)w+c(2w)}$. Set $\psi(w) = e^{-\eta(w)wfrm-e+c(w)}$, then $\psi(w)^2 = \psi(2w)$ and

$$\frac{\sigma(z+w)}{\sigma(z)} = \psi(w)e^{\eta(w)(z+w/2)}.$$

It remains to calculate $\psi(w)$. Since $\sigma(z)$ is odd, if $w/2 \notin \Lambda$, then $\sigma(-w/2) \neq \infty$, and

$$-1 = \frac{\sigma(w/2)}{\sigma(-w/2)} = \psi(w).$$

If $w/2 \in \Lambda$, there exists $n \geq 1$ such that $w/2^n \in \Lambda, w/2^{n+1} \notin \Lambda$. Hence

$$\psi(w) = \psi(w/2)^2 = \cdots = \psi(w/2^n)^{2n} = (-1)^{2n} = 1.$$

□

Proposition 3.1.3. *For $\tau \in \mathbb{H}, z \in \mathbb{C}$, let $q_\tau = e^{2\pi i\tau}, q_z = e^{2\pi iz}$. Then*

$$\sigma(z;\tau) = (2\pi i)^{-1}e^{\frac{1}{2}\eta z^2}(q_z^{1/2} - q_z^{-1/2})\prod_{n=1}^{\infty}\frac{(1-q_\tau^n q_z)(1-q_\tau^n/q_z)}{(1-q_\tau^n)^2},$$

where $\eta = \eta(1;\tau)$.

Proof. We give a sketch of the proof, see [36, Page 247, Theorem 4] for full details.

For fixed τ, we set

$$\varphi(z) = e^{-\frac{1}{2}\eta z^2}q_z^{1/2}\sigma(z;\tau),$$

$$g(z) = (2\pi i)^{-1}(q_z - 1)\prod_{n=1}^{\infty}\frac{(1-q_\tau^n q_z)(1-q_\tau^n/q_z)}{(1-q_\tau^n)^2}.$$

It is sufficient to show that $\varphi(z) = g(z)$. We have the following claims without proofs:

(i) $g(z)$ uniformly converges on any compact subset of $\mathbb{C}$, and it is analytic on $\mathbb{C}$. Its zeros (of order 1) are the points on Λ_τ;

(ii) $\varphi(z+1) = \varphi(z), \varphi(z+\tau) = -\dfrac{1}{q_z}\varphi(z)$, and it is similar for $g(z)$;

(iii) $\lim\limits_{z\to 0}\varphi(z)/g(z) = 1$.

If these claims hold, notice that the zeros (of order 1) of $\varphi(z)$ are also the points on Λ_τ, then $\varphi(z)/g(z)$ has a period lattice Λ_τ and is holomorphic on $\mathbb{C}$. THis imply that $\varphi(z)/g(z)$ is a constant, which is 1 by (iii).

□

For a lattice $\Lambda \subset \mathbb{C}$, we write $\Lambda = \langle \omega_1, \omega_2 \rangle$ if Λ is generated by $\omega_1, \omega_2 \in \mathbb{C}$, and $\omega_1/\omega_2 \in \mathbb{H}$.

Definition 3.1.4. *Let $\Lambda = \langle \omega_1, \omega_2 \rangle \subset \mathbb{C}$ be a lattice. We call $\eta_1 = \eta(\omega_1)$ and $\eta_2 = \eta(\omega_2)$ a pair of basic quasi periods of ζ.*

Remark. *(1) For any $a_1, a_2 \in \mathbb{R}$, we have $\eta(a_1\omega_1 + a_2\omega_2) = a_1\eta_1 + a_2\eta_2$.*

Theorem 3.1.5 (Legendre Relation)**.** *Keep the notations in Definition 3.1.4, we have*

$$\eta_2\omega_1 - \eta_1\omega_2 = 2\pi i.$$

Proof. Let P be a fundamental parallelogram with vertexes $\alpha, \alpha + w_2, \alpha + w_1 + w_2, \alpha + w_1 \in \mathbb{C}$ and arrows s_1, s_2, s_3, s_4. We have $0 \in P$. By residue theorem,

$$\oint_{\partial P} \zeta(z)dz = 2\pi i \sum_{p \in P} \mathrm{Res}_p(\zeta(z)) = 2\pi i.$$

On the other hand,

$$\begin{aligned}\oint_{\partial P} \zeta(z)dz &= \int_{s_1+s_3} \zeta(z)dz + \int_{s_2+s_4} \zeta(z)dz \\ &= \int_{s_1} (\zeta(z) - \zeta(z+w_1))dz + \int_{s_2} (\zeta(z) - \zeta(z-w_2))dz \\ &= -\eta_1\omega_2 + \eta_2\omega_1.\end{aligned}$$

□

3.1.2 The Klein forms and Siegel functions

Definition 3.1.6. *Let $\Lambda \subset \mathbb{C}$ be a lattice, we define the Klein forms as*

$$\mathfrak{k}(z;\Lambda) := e^{-\eta(z;\Lambda)z/2}\sigma(z;\Lambda) : \mathbb{C} \to \mathbb{C}.$$

Let $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$, and $W = \begin{pmatrix}\omega_1 \\ \omega_2\end{pmatrix}$ such that $\omega_1/\omega_2 \in \mathbb{H}$, we set

$$\mathfrak{k}_{\mathbf{a}}(W) := \mathfrak{k}(\mathbf{a}W; \langle \omega_1, \omega_2 \rangle).$$

For $\tau \in \mathbb{H}$, we set

$$\mathfrak{k}_{\mathbf{a}}(\tau) := \mathfrak{k}(\mathbf{a}W_\tau; \langle \tau, 1 \rangle),$$

where $W_\tau = \begin{pmatrix}\tau \\ 1\end{pmatrix}$.

Remark. *(1) The zeros of $\mathfrak{k}(z;\Lambda)$ are the points on Λ, and they are of order 1. In particular, $\mathfrak{k}_{\mathbf{a}}(\tau) \equiv 0$ if $\mathbf{a} \in \mathbb{Z}^2$, and $\mathfrak{k}_{\mathbf{a}}(\tau) \neq 0$ for any $\tau \in \mathbb{H}$ if $\mathbf{a} \in \mathbb{R} \setminus \mathbb{Z}^2$.*

Proof. Notice that $\mathfrak{k}(z;\Lambda) = 0$ if and only if $\sigma(z;\Lambda) = 0$, whose zeros are the points on Λ. □

We call a positive integer N the denominator of $\mathbf{a} = (a_1, a_2) \in \mathbb{Q}^2$, if $Na_1, Na_2 \in \mathbb{Z}$ and $\gcd(Na_1, Na_2) = 1$. It is exactly the order of $\mathbf{a}$ in $\mathbb{Q}^2/\mathbb{Z}^2$.

Proposition 3.1.7. *Let $\mathbf{a} = (a_1, a_2) \in \mathbb{Q}^2 \setminus \mathbb{Z}^2$. Then the following statements hold:*

(1) The Klein form $\mathfrak{k}_{\mathbf{a}}(\tau)$ does not vanish on $\mathbb{H}$.

(2) For any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathbb{H}$, we have

$$\mathfrak{k}_{\mathbf{a}}(\gamma(\tau)) = (c\tau + d)^{-1}\mathfrak{k}_{\mathbf{a}\gamma}(\tau).$$

In particular, $\mathfrak{k}_{-\mathbf{a}}(\tau) = -\mathfrak{k}_{\mathbf{a}}(\tau)$.

(3) Let $N \geq 2$ be the denominator of $\mathbf{a} = (a_1, a_2)$. Let $\mathbf{b} = (b_1, b_2) \in \mathbb{Z}^2$. Then

$$\mathfrak{k}_{\mathbf{a}+\mathbf{b}}(\tau) = \varepsilon(\mathbf{a}, \mathbf{b})\mathfrak{k}_{\mathbf{a}}(\tau),$$

where $\varepsilon(\mathbf{a}, \mathbf{b}) = (-1)^{b_1 b_2 + b_1 + b_2} \cdot e^{\pi i(a_1 b_2 - a_2 b_1)}$ is a $2N$-th root of unity.

(4) Let $N \geq 2$ be the denominator of $\mathbf{a} = (a_1, a_2) = (r/N, s/N)$. Then for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(N)$, we have

$$\mathfrak{k}_{\mathbf{a}}(\gamma(\tau)) = \varepsilon'(\mathbf{a}, \gamma)(c\tau + d)^{-1}\mathfrak{k}_{\mathbf{a}}(\tau),$$

where $\varepsilon'(\mathbf{a}, \gamma) = -(-1)^{((a-1)r/N + cs/N + 1)(br/N + (d-1)s/N + 1)} \cdot e^{\pi i(br^2 + (d-a)rs - cs^2)/N^2}$ is a $2N$-th root of unity.

Proof. (1) comes from the remark of Definition 3.1.6.

For (2), we claim: for $W = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$ such that $\omega_1/\omega_2 \in \mathbb{H}$, we have

(i) $\mathfrak{k}_{\mathbf{a}}(\lambda W) = \lambda \mathfrak{k}_{\mathbf{a}}(W)$ for any $\lambda \in \mathbb{C}^*$;

(ii) $\mathfrak{k}_{\mathbf{a}}(\gamma W) = \mathfrak{k}_{\mathbf{a}\gamma}(W)$ for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Indeed, let $\Lambda = \langle \omega_1, \omega_2 \rangle$, $z = \mathbf{a}W$. Then

$$\mathfrak{k}_{\mathbf{a}}(\lambda W) = e^{-\eta(\lambda z; \lambda\Lambda)\lambda z/2}\sigma(\lambda z; \lambda\Lambda) = \lambda e^{-\eta(z;\Lambda)z/2}\sigma(z; \Lambda) = \lambda \mathfrak{k}_{\mathbf{a}}(W).$$

Let $W' = \begin{pmatrix} \omega_1' \\ \omega_2' \end{pmatrix} := \gamma W$, $z' := \mathbf{a}W'$. Then $\Lambda =< \omega_1', \omega_2' >$ and

$$\begin{aligned} \mathfrak{k}_{\mathbf{a}}(\gamma W) &= e^{-\eta(z';\Lambda)z'/2}\sigma(z'; \Lambda) \\ &= e^{-\eta(\mathbf{a}\gamma W; \Lambda)\mathbf{a}\lambda W/2}\sigma(\mathbf{a}\gamma W; \Lambda) \\ &= \mathfrak{k}_{\mathbf{a}\gamma}(W). \end{aligned}$$

These claims implies that

$$\mathfrak{k}_{\mathbf{a}}(\gamma(\tau)) = \mathfrak{k}_{\mathbf{a}}((c\tau + d)^{-1}\gamma W_\tau) = (c\tau + d)^{-1}\mathfrak{k}_{\mathbf{a}}(\gamma W_\tau) = (c\tau + d)^{-1}\mathfrak{k}_{\mathbf{a}\gamma}(\tau).$$

For (3), under the notations in (3), it is sufficient to prove the following claim: for $W = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}$ such that $\omega_1/\omega_2 \in \mathbb{H}$, we have

$$\mathfrak{k}_{\mathbf{a}+\mathbf{b}}(W) = \varepsilon(\mathbf{a}, \mathbf{b})\mathfrak{k}_{\mathbf{a}}(W).$$

Indeed, again, let $\Lambda = \langle \omega_1, \omega_2 \rangle$. Then by Theorem 3.1.2,

$$\begin{aligned} \mathfrak{k}_{\mathbf{a}+\mathbf{b}}(W) &= e^{-\eta((\mathbf{a}+\mathbf{b})W;\Lambda)(\mathbf{a}+\mathbf{b})W/2} \cdot \sigma((\mathbf{a} + \mathbf{b})W; \Lambda) \\ &= e^{-\eta((\mathbf{a}+\mathbf{b})W;\Lambda)(\mathbf{a}+\mathbf{b})W/2} \cdot \sigma(\mathbf{a}W; \Lambda)\psi(\mathbf{b}W)e^{\eta(\mathbf{b}W;\Lambda)(\mathbf{a}+\mathbf{b})W/2} \\ &= e^{-\eta(\mathbf{a}W;\Lambda)(\mathbf{a}+\mathbf{b})W/2} \cdot e^{\eta(\mathbf{b}W;\Lambda)\mathbf{a}W/2}\sigma(\mathbf{a}W; \Lambda)\psi(\mathbf{b}W) \\ &= \psi(\mathbf{b}W)\mathfrak{k}_{(a)}(W)e^{-\eta(\mathbf{a}W;\Lambda)\mathbf{b}W/2 + \eta(\mathbf{b}W;\Lambda)\mathbf{a}W/2}. \end{aligned}$$

It is sufficient to show that $\varepsilon(\mathbf{a},\mathbf{b}) = \psi(\mathbf{b}W)e^{-\eta(\mathbf{a}W;\Lambda)\mathbf{b}W/2+\eta(\mathbf{b}W;\Lambda)\mathbf{a}W/2}$. Let $\eta_1 = \eta(\omega_1;\Lambda), \eta_2 = \eta(\omega_2;\Lambda)$. Then by Theorem 3.1.5, i.e. Legendre Relation,

$$\begin{aligned}
&\eta(\mathbf{b}W;\Lambda)\mathbf{a}W/2 - \eta(\mathbf{a}W;\Lambda)\mathbf{b}W/2\\
=&(b_1\eta_1 + b_2\eta_2)(a_1\omega_1 + a_2\omega_2)/2 - (a_1\eta_1 + a_2\eta_2)(b_1\omega_1 + b_2\omega_2)/2\\
=&(a_1b_2\eta_2\omega_1 + a_2b_1\eta_1\omega_2 - a_1b_2\eta_1\omega_2 - a_2b_1\eta_2\omega_1)/2\\
=&\pi i(a_1b_2 - a_2b_1).
\end{aligned}$$

Also notice that

$$\psi(\mathbf{b}W) = \begin{cases} 1 & \text{if } 2 \nmid b_1 \text{ or } 2 \nmid b_2,\\ -1 & \text{if } 2 \mid b_1, 2 \mid b_2, \end{cases}$$

i.e. $\psi(\mathbf{b}W) = (-1)^{(b_1+1)(b_2+1)+1} = (-1)^{b_1b_2+b_1+b_2}$. Hence we have our claim.

For (4), by (1), it is sufficient to prove that, for any $\gamma \in \Gamma(N)$,

$$\mathfrak{k}_{\mathbf{a}\gamma}(\tau) = \varepsilon'(\mathbf{a},\gamma)\mathfrak{k}_{\mathbf{a}}(\tau).$$

Indeed, $\mathbf{a}\gamma = (a_1a + a_2c, a_1b + a_2d)$ and

$$a_1a + a_2c = a_1 + \left(\frac{(a-1)r}{N} + \frac{cs}{N}\right) \in a_1 + \mathbb{Z},$$

$$a_1b + a_2d = a_2 + \left(\frac{br}{N} + \frac{(d-1)s}{N}\right) \in a_2 + \mathbb{Z}.$$

Hence by (3),

$$\mathfrak{k}_{\mathbf{a}\gamma}(\tau) = \varepsilon(\mathbf{a}, (\frac{(a-1)r}{N} + \frac{cs}{N}, \frac{br}{N} + \frac{(d-1)s}{N}))\mathfrak{k}_{\mathbf{a}}(\tau),$$

and $\varepsilon(\mathbf{a}, (\frac{(a-1)r}{N} + \frac{cs}{N}, \frac{br}{N} + \frac{(d-1)s}{N})) = \varepsilon'(\mathbf{a},\gamma)$ is a $2N$-th root of unity. □

Corollary 3.1.8. *Let N be a positive integer and $\mathbf{a} \in (\frac{1}{N}\mathbb{Z})^2 \setminus \mathbb{Z}^2$. Then $\mathfrak{k}_{\mathbf{a}}^{2N}(\tau)$ depends only on $\overline{\mathbf{a}} \in (\frac{1}{N}\mathbb{Z})^2/\mathbb{Z}^2$, and $\mathfrak{k}_{\mathbf{a}}^{2N}(\tau) \in \mathcal{A}_{-2N}(\Gamma(N))$.*

Proof. This statement directly comes from (3) and (4) of Proposition 3.1.7. □

Next we will study another type of functions, the Siegel functions.

Definition 3.1.9. *For $\mathbf{a} \in \mathbb{Q}^2$, we define the Siegel function (associated to $\mathbf{a}$) as*

$$g_{\mathbf{a}}(\tau) := \mathfrak{k}_{\mathbf{a}}(\tau)\Delta^{1/12}(\tau) : \mathbb{H} \to \mathbb{C},$$

where $\Delta^{1/12}(\tau) = 2\pi i \cdot \eta^2(\tau)$, and $\eta(\tau) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n)$ with $q = e^{2\pi i\tau}$ is the Dedekind eta function.

Remark. *(1) Recall that, for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, we have*

$$\Delta^{1/12}(\gamma(\tau)) = \varepsilon(\gamma)(c\tau + d)\Delta^{1/12}(\tau),$$

where $\varepsilon(\gamma)$ is a 12-th root of unity.

Proposition 3.1.10. *Let $\mathbf{a} = (a_1, a_2) \in \mathbb{Q}^2 \setminus \mathbb{Z}^2$. Then the following statements hold:*

(1) The Siegel function $g_{\mathbf{a}}(\tau)$ does not vanish on $\mathbb{H}$.

(2) *For any* $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ *and* $\tau \in \mathbb{H}$, *we have*

$$g_{\mathbf{a}}(\gamma(\tau)) = \varepsilon(\gamma) g_{\mathbf{a}\gamma}(\tau),$$

where $\varepsilon(\gamma)$ *is a* 12*-th root of unity..*

(3) *Let* $N \geq 2$ *be a denominator of* $\mathbf{a} = (a_1, a_2)$, *and* $\mathbf{b} = (b_1, b_2) \in \mathbb{Z}^2$. *Then*

$$g_{\mathbf{a}+\mathbf{b}} = \varepsilon(\mathbf{a}, \mathbf{b}) g_{\mathbf{a}}(\tau),$$

where $\varepsilon(\mathbf{a}, \mathbf{b}) = (-1)^{b_1 b_2 + b_1 + b_2} \cdot e^{-2\pi i (b_1 a_2 - b_2 a_1)/2}$ *is a* $2N$*-th root of unity.*

Proof. For (1), notice that $\mathfrak{k}_{\mathbf{a}}(\tau) and \Delta^{12}(\tau)$ don't vanish on H, then so does $g_{\mathbf{a}}(\tau)$.

For (2), by (2) of Proposition 3.1.7 and the remark of Definition 3.1.9, we have

$$\begin{aligned} g_{\mathbf{a}}(\gamma(\tau)) &= \mathfrak{k}_{\mathbf{a}}(\gamma(\tau)) \Delta^{1/12}(\gamma(\tau)) \\ &= (c\tau + d)^{-1} \mathfrak{k}_{\mathbf{a}\gamma}(\tau) \cdot \varepsilon(\gamma)(c\tau + d) \Delta^{1/12}(\tau) \\ &= \varepsilon(\gamma) g_{\mathbf{a}\gamma}(\tau), \end{aligned}$$

where $\varepsilon(\gamma)$ is a 12-th root of unity.

Similarly, (3) comes from (3) of Proposition 3.1.7. □

Proposition 3.1.11. *For* $\mathbf{a} \in \mathbb{Q}^2 \setminus \mathbb{Z}^2$, *we have the q-product for the Siegel function:*

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

where $q = e^{2\pi i \tau}$ *and* $B_2(T) = T^2 - T + 1/6$ *is the* 2*-nd Bernoulli polynomial. In particular,* $g_{\mathbf{a}}$ *has zeros at* $i\infty$ *of order* $\ell_{\mathbf{a}} := B_2(a_1 - \lfloor a_1 \rfloor)/2$, *i.e.* $\lim_{\tau \to i\infty} q^{-\ell_{\mathbf{a}}} g_{\mathbf{a}}(\tau)$ *exists and is nonzero.*

Proof. Set $q_z = e^{2\pi(a_1 \tau + a_2)} = q^{a_1} e^{2a_2 \pi i}$. By q-product of $\sigma(z; \tau)$ in Proposition 3.1.3 and of $\Delta^{1/12}(\tau)$,

$$\begin{aligned} g_{\mathbf{a}}(\tau) &= \mathfrak{k}_{\mathbf{a}}(\tau) \Delta^{1/12}(\tau) \\ &= e^{-\frac{1}{2}\eta(a_1\tau + a_2; \tau)(a_1 \tau + a_2)} \sigma(a_1 \tau + a_2; \tau) \cdot \Delta^{1/12}(\tau) \\ &= e^{-\frac{1}{2}(a_1 \eta(\tau;\tau) + a_2 \eta)(a_1 \tau + a_2)} \cdot e^{\frac{1}{2}\eta(a_1\tau + a_2)^2} (q_z^{1/2} - q_z^{-1/2}) \prod_{n=1}^{\infty} \frac{(1 - q^n q_z)(1 - q^n / q_z)}{(1 - q^n)^2} \\ &\quad \cdot q_\tau^{1/12} \prod_{n=1}^{\infty} (1 - q^n)^2 \\ &= e^{-\frac{1}{2}\eta(\tau;\tau) a_1 (a_1 \tau + a_2) + \frac{1}{2}\eta(a_1^2 \tau^2 + a_1 a_2 \tau)} q^{1/12} (q^{a - 1/2} e^{a_2 \pi i} - q^{-a_1/2} e^{-a_2 \pi i}) \\ &\quad \cdot \prod_{n=1}^{\infty} (1 - q^{n + a_1} e^{2a_2 \pi i})(1 - q^{n - a_1} e^{-2a_2 \pi i}) \end{aligned}$$

By Legendre relation, we have $\eta(\tau; \tau) = \tau \eta - 2\pi i$, so

$$\begin{aligned} &-\frac{1}{2}\eta(\tau;\tau) a_1 (a_1 \tau + a_2) + \frac{1}{2}\eta(a_1^2 \tau^2 + a_1 a_2 \tau) \\ =&\frac{1}{2} a_1 (2\pi i - \tau \eta)(a_1 \tau + a_2) + \frac{1}{2}\eta(a_1^2 \tau^2 + a_1 a_2 \tau) \\ =&2\pi i \cdot \frac{1}{2} a_1^2 + \pi i a_1 a_2. \end{aligned}$$

Hence

$$g_{\mathbf{a}}(\tau) = -q^{\frac{1}{2}(a_1^2-a_1+1/6)}e^{\pi i a_2 a_1 - \pi i a_2}(1-q^{a_1}e^{2\pi i a_2})\prod_{n=1}^{\infty}(1-q^{n+a_1}e^{2a_2\pi i})(1-q^{n-a_1}e^{-2a_2\pi i})$$

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

If $0 \le a_1, a_2 < 1$ and $a_1^2 + a_2^2 > 0$, since for any $n \ge 0$

$$\lim_{q\to 0}(1-q^{n+a_1}e^{2\pi i a_2}) \neq 0,$$

$$\lim_{q\to 0}(1-q^{n+1-a_1}e^{-2\pi i a_2}) \neq 0,$$

so

$$\lim_{\tau\to i\infty} q^{-\ell_{\mathbf{a}}} g_{\mathbf{a}}(\tau) = -e^{\pi i a_2(a_1-1)} \neq 0.$$

□

3.1.3 Modular units

Definition 3.1.12. *Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup of level N, we know that $\mathbb{Q}(j) \subset \mathbb{Q}(\zeta_N)(X_\Gamma) = \mathcal{A}_0(\mathbb{Q}(\zeta_N), \Gamma)$. Let $R_{\Gamma,N}$ be the integral closure of $\mathbb{Z}[j]$ in $\mathbb{Q}(\zeta_N)(X_\Gamma)$, and $\mathbb{Q}R_{\Gamma,N}$ be the integral closure of $\mathbb{Q}[j]$ in $\mathbb{Q}(\zeta_N)(X_\Gamma)$. Elements in $(\mathbb{Q}R_{\Gamma,N})^*$ will be called modular units, and elements in $(R_{\Gamma,N})^*$ will be called modular units over $\mathbb{Z}$.*

Remark. *(1) The ring $\mathbb{Q}R_{\Gamma,N}$ is exactly $\mathbb{Q}\otimes_{\mathbb{Z}} R_{\Gamma,N}$.*

(2) Notice that $\mathbb{Q}R_{\Gamma,N}$ is also the closure of $\mathbb{Q}(\zeta_N)[j]$ in $\mathbb{Q}(\zeta_N)(X_\Gamma)$, and $\mathrm{Spec}(\mathbb{Q}(\zeta_N)[j]) \subset X(1)$ is the affine open subset not containing the cusp, where we view $X(1)$ as a projective line over $\mathbb{Q}(\zeta_N)$. Hence $\mathrm{Spec}(\mathbb{Q}R_{\Gamma,N}) \subset X_\Gamma$ is the affine open subset not containing the cusps. We conclude that $(\mathbb{Q}R_{\Gamma,N})^*$ is the group of Σ_j-units of X_Γ over $\mathbb{Q}(\zeta_N)$, i.e. $U_{\Sigma_j,\mathbb{Q}(\zeta_N)} = (\mathbb{Q}R_{\Gamma,N})^*$. Here we view $X(1)$ and X_Γ as smooth projective curves over $\mathbb{Q}(\zeta_N)$.

Proof. Since $\mathrm{Spec}(\mathbb{Q}R_{\Gamma,N}) \subset X_\Gamma$ is the affine open subset not containing the cusps, then

$$\mathbb{Q}R_{\Gamma,N} = \{f \in \mathbb{Q}(\zeta_N)(X_\Gamma) \mid \text{the only possible poles of } f \text{ are cusps}\},$$

which implies that $(\mathbb{Q}R_{\Gamma,N})^*$ is the group of Σ_j-units of X_Γ. □

LEMMA 3.1.13. *If $f \in \mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z}))$ is holomorphic on $\mathbb{H}$ with q-expansion $f(\tau) = \sum_{n=-N}^{\infty} c_n q^n$, $q = e^{2\pi i\tau}$, then $f \in \mathbb{Z}[c_{-N}, c_{-N+1}, \cdots][j] \subset \mathbb{C}[j]$.*

Proof. Induction on N. If $N = 0$, then f is holomorphic on $X(1)$. That means that $f = c_0$ is a constant, and $f \in \mathbb{Z}[c_0][j]$.

For $N \ge 1$, assume that the lemma holds for $N-1$. Let $g(\tau) = f(\tau) - c_{-N}j(\tau)^N = \sum_{n=-N+1}^{\infty} b_n q^n \in \mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z}))$. Then by the inductive hypothesis, the function $g \in \mathbb{Z}[b_{-N+1}, b_{-N+2}, \cdots][j] \subset \mathbb{Z}[c_N, c_{-N+1}, \cdots][j]$. Hence $f \in \mathbb{Z}[c_N, c_{-N+1}, \cdots][j]$. □

Remark. *(1) In particular, from this lemma, we can also know that, for $f \in \mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z}))$, f is holomorphic on $\mathbb{H}$ if and only if $f \in \mathbb{C}[j]$.*

LEMMA 3.1.14. *Let $f \in \mathcal{A}_0(\Gamma(N))$ which is holomorphic on $\mathbb{H}$. If for each $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ the coefficients of q_N-expansion of $f \circ [\gamma]$ are algebraic integers. Then f is integral over $\mathbb{Z}[j]$.*

Proof. The coefficients of $F(X) := \prod_{\gamma \in \Gamma(N)\backslash \mathrm{SL}_2(\mathbb{Z})} (X - f \circ [\gamma])$ are in $\mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z}))$ and holomorphic on $\mathbb{H}$. To verify the coefficients of their q-expansions are algebraic integers, it is sufficient to notice that $q = q_N^N$ and consider their q_N expansions. Indeed the coefficients of $F(X)$ are in $\overline{\mathbb{Z}}[j]$ by Lemma 3.1.13. If follows that f is integral over $\overline{\mathbb{Z}}[j]$, hence over $\mathbb{Z}[j]$. □

From Section 1.4, we know that in order to bound integral points on modular curves X_Γ, we should calculate the group $U_{\Sigma_j,K}$ of Σ_j-units of X_Γ. We will construct its generator out of Siegel functions.

We fix some notations. Let Γ be a congruence subgroup of level N, $\overline{\Gamma}$ be the image of Γ in $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ and X_Γ be the corresponding modular curve. We set

$$\mathcal{A}_N := \{\mathbf{a} \in (\mathbb{Z}/N\mathbb{Z})^2 \mid \mathrm{ord}(\mathbf{a}) = N\},$$

then $\overline{\Gamma}$ acts on $\mathcal{A}_N$ naturally, and we have $|\mathcal{A}_N/\Gamma| = v_\infty(\Gamma)$, where $v_\infty(\Gamma)$ is the cardinality of cusps on X_Γ. We can identify $\mathcal{A}_N$ and the set $\{\mathbf{a} \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2 \mid \mathrm{ord}(\mathbf{a}) = N\}$. Moreover, for a representative element of $\mathbf{a} = (a_1, a_2) \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$ satisfying $0 \le a_1, a_2 < 1$, and let $g_\mathbf{a}$ be the corresponding Siegel function.

Definition 3.1.15. *Keep the notations as Definition 3.1.12. Let $\mathbf{a} \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$, we denote $g_\mathbf{a}^{12N}$ by $u_\mathbf{a}$.*

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

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

Let $\mathcal{O} \in \mathcal{A}_N/\overline{\Gamma}$ be an orbit, we have

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

LEMMA 3.1.16. *We have $u_\mathbf{a} \in \mathcal{A}_0(\mathbb{Q}(\zeta_N), \Gamma(N)) = \mathbb{Q}(\zeta_N)(X(N))$ for any $\mathbf{a} \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$. Moreover, via $\rho^{-1} : \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm I\} \xrightarrow{\sim} \mathrm{Gal}(\mathbb{Q}(\zeta_N)(X(N))/\mathbb{Q}(j))$ in Proposition 2.4.3,*

$$u_\mathbf{a}^{\overline{\gamma}} = u_\mathbf{a} \circ [\gamma] = u_{\mathbf{a}\gamma}$$

for any $\gamma \in \mathrm{GL}_2(\mathbb{Z})$.

Proof. By Corollary 3.1.8 and Proposition 3.1.11, $u_\mathbf{a} \in \mathcal{A}_0(\mathbb{Q}(\zeta_N), \Gamma(N))$. By Proposition 2.4.3, Corollary 2.4.11 and Proposition 3.1.10 (2), we have

$$u_\mathbf{a}^{\overline{\gamma}} = u_\mathbf{a} \circ [\gamma] = g_\mathbf{a}^{12N} \circ [\gamma] = g_{\mathbf{a}\gamma}^{12N} = u_{\mathbf{a}\gamma}.$$

□

Remark. *(1) With this lemma, for any $T \subset \mathcal{A}_N$, $u_T \in \mathbb{Q}(\zeta_N)(X_\Gamma)$ if and only T is invariant under $\overline{\Gamma}$-action.*

Proposition 3.1.17. *Keep the notations as Definition 3.1.12. We have the following properties:*

(1)

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

where Φ_N is the N-th cyclotomic polynomial.

(2) Put $\lambda = (1-\zeta_N)^{12N^2\varphi(N)}$, then the functions $u_{\mathcal{O}}$ and $\lambda u_{\mathcal{O}}^{-1}$ are integral over $\mathbb{Z}[j]$, where φ is Euler's totient function.

(3) For the cusp $c_\infty \in X_\Gamma$ at infinity, we have

$$Ord_{c_\infty}(u_{\mathcal{O}}) = 12Nh_{c_\infty}\sum_{\mathbf{a}\in\mathcal{O}}\ell_{\mathbf{a}}.$$

where h_{c_∞} is the width of c_∞, see definition 2.3.1. For any cusp c, we have $|Ord_c(u_{\mathcal{O}})| < N^4$.

(4) $u_{\mathcal{O}}$ is a modular unit on X_Γ, moreover, the group generated by the principal divisor $(u_{\mathcal{O}})$, where $\mathcal{O}$ runs over the orbits of $\mathcal{A}_N/\overline{\Gamma}$, is of rank $v_\infty(\Gamma)-1$. In particular, $U_{\Sigma_j,\mathbb{Q}(\zeta_N)} = (\mathbb{Q}R_{\Gamma,N})^*$ is generated by $\{u_{\mathcal{O}} \mid \mathcal{O}\in\mathcal{A}_N/\overline{\Gamma}\}$ and $\rho(\Sigma_j,\mathbb{Q}(\zeta_N)) = v_\infty(\Gamma)-1$.

Proof. By Lemma 3.1.16, we have $u_{\mathcal{O}} \in \mathcal{A}_0(\mathbb{Q}(\zeta_N),\Gamma) = \mathbb{Q}(\zeta_N)(X_\Gamma)$ for any $\mathcal{O}\in\mathcal{A}_N/\overline{\Gamma}$. For (1), let $u := \prod_{\mathbf{a}\in\mathcal{A}_N} u_{\mathbf{a}}$. Then we have $u \in \mathbb{Q}(\zeta_N)(X(1))$by Lemma above. Since u doesn't vanish outside the cusp of $X(1)$ by Proposition 3.1.10 (1), then it must be a constant. By Proposition 3.1.11, we have

$$\begin{aligned} u(\tau) &= \prod_{(a_1,a_2)\in\mathcal{A}_N} (q^{6NB_2(a_1)}e^{12N\pi i a_2(a_1-1)}\prod_{n=0}^{\infty}(1-q^{n+a_1}e^{2\pi i a_2})^{12N}(1-q^{n+1-a_1}e^{-2\pi i a_2})^{12N}) \\ &\overset{q=0}{=\!=} \pm\prod_{\substack{(a_1,a_2)\in\mathcal{A}_N \\ a_0=0}}(1-e^{2\pi i a_2})^{12} \\ &= \pm\prod_{\substack{1\le k<N \\ \gcd(k,N)=1}}(1-e^{2\pi i a_2})^{12} \\ &= \pm\Phi_N(1)^{12N}. \end{aligned}$$

(2) is [10, Proposition 2.2], we prove it here. We have

$$u_{\mathbf{a}} = q^{6NB_2(a_1)}e^{12N\pi i a_2(a_1-1)}\prod_{n=0}^{\infty}(1-q^{n+a_1}e^{2\pi i a_2})^{12N}(1-q^{n+1-a_1}e^{-2\pi i a_2})^{12N},$$

so the coefficients of the q_N-expansion is algebraic integers. By Lemma 3.1.16 and Lemma 3.1.14, $u_{\mathbf{a}}$ is integral over $\mathbb{Z}[j]$.

For $\lambda u_{\mathbf{a}}^{-1}$, we consider the product expansion of $u_{\mathbf{a}}^{-1}$. The only problem is the term $(1-q^{n+a_1}e^{2\pi i a_2})^{-12N}$ when $n=0$ and $a_1=0$. If it is not this case, we can take expansion of each term by the fact $1/(1-z) = \sum_{k=0}^{\infty} z^k$, $|z|<1$. If it is this case, i.e. $n=0, a_1=0$, then $r := Na_2$ is coprime with N. Hence we know that $(1-\zeta_N^r)^{12N}u_{\mathbf{a}}^{-1}$ is integral over $\mathbb{Z}[j]$. Since $\gcd(r,N)=1$, so $(1-\zeta_N^r)/(1-\zeta_N)$ is a unit in $\mathbb{Z}[\zeta_N]$, and $(1-\zeta_N)^{12N}u_{\mathbf{a}}^{-1}$ is integral over $\mathbb{Z}[j]$. Hence $(1-\zeta_N)^{12N^2\varphi(N)}u_{\mathcal{O}}$ is integral over $\mathbb{Z}[j]$.

For (3), notice that the ramification index of $\pi : X_\Gamma \to X(1)$ at c is h_c and the order of $u_{\mathcal{O}}$ at c on X_Γ equals to the order is it comes from Corollary 2.3.5 and Proposition 3.1.11.

(4) comes from [35, Theorem 3.1]. □

3.1.4 Bounding modular units

In this subsection, we set

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

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

$$\gamma_T = \prod_{\mathbf{a}\in T} \gamma_{\mathbf{a}}.$$

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

$$\gamma_{\mathcal{O}} = \prod_{\mathbf{a}\in \mathcal{O}} \gamma_{\mathbf{a}}.$$

Proposition 3.1.18 ([55], Proposition 3.1)**.** *Let $\mathbf{a} \in \mathcal{A}_N$, $v \in M_K$. If $q \in \overline{K_v}$ 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

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

for each $k \geq 0$.

Corollary 3.1.19. *Let $\mathbf{a}_1, \cdots, \mathbf{a}_n \in \mathcal{A}_N$. If $q \in \overline{K_v}$ satisfies $|q|_v < 1$, then we have*

$$(-1)^n \prod_{i=1}^{n} q^{-\ell_{\mathbf{a}_i}}\gamma_{\mathbf{a}_i}^{-1} g_{\mathbf{a}_i}(q) = 1 + \sum_{k=1}^{\infty} \phi(k) q^{k/N},$$

where

$$|\phi(k)|_v \leq 2^{k+n} e^k.$$

Proof.

$$(-1)^n \prod_{i=1}^{n} q^{-\ell_{\mathbf{a}_i}}\gamma_{\mathbf{a}_i}^{-1} g_{\mathbf{a}_i}(q) = 1 + \sum_{k=1}^{\infty} \Big(\sum_{i_1+\cdots+i_n=k} \phi_{\mathbf{a}_1}(i_1)\cdots\phi_{\mathbf{a}_n}(i_n) \Big) q^{k/N},$$

so

$$|\phi(k)|_v = \Big| \sum_{i_1+\cdots+i_n=k} \phi_{\mathbf{a}_1}(i_1)\cdots\phi_{\mathbf{a}_n}(i_n) \Big|_v \leq 2^{n+k} e^k.$$

□

For each cusp c of X_Γ, let t_c be its local parameter defined in [9][Section 3], and $q_c = t_c^{h_c}$, where h_c is the width of c, that is ramification index of the covering $X_\Gamma \to X(1)$, see Definition 2.3.1 and Remark of Proposition 2.3.2. Moreover, for $v \in M_K$, $\Omega_{c,v}$ is a neighborhood of c on $X_\Gamma(\overline{K_v})$ defined in [9, Section 3].

Proposition 3.1.20 ([9] Proposition 3.1, or [55], Proposition 3.3)**.** *Put*

$$X_\Gamma(\overline{K_v})^+ = \begin{cases} \{P \in X_\Gamma(\overline{K_v}) \mid |j(P)|_v > 3500\} & \text{if } v \in M_K^\infty; \\ \{P \in X_\Gamma(\overline{K_v}) \mid |j(P)|_v > 1\} & \text{if } v \in M_K^0. \end{cases}$$

Then

$$X_\Gamma(\overline{K_v})^+ \subset \bigcup_{c\in X_\Gamma(\overline{K_v})\ cusp} \Omega_{c,v}$$

with equality for the non-Archimedean v. Moreover, for $P \in \Omega_{c,v}$, we have

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

if $v \in M_K^\infty$, and $|j(P)|_v = |q_c(P)^{-1}|_v$ if $v \in M_K^0$.

For any cusp c, recall that the the vanishing order of $u_\mathcal{O}$ at c is denoted by $\mathrm{Ord}_c(u_\mathcal{O})$. For a number field K and $v \in M_K$, define

$$\rho_v = \begin{cases} 12N^3 \log N & \text{if } v \mid \infty, \\ 0 & \text{if } v \nmid \infty \text{ and } |N|_v = 1, \\ \dfrac{12N^3 \log p_v}{p_v - 1} & \text{if } v \mid p_v < \infty \text{ and } p_v \mid N. \end{cases} \tag{3.1}$$

Proposition 3.1.21 ([55], Proposition 3.6)**.** *Let K be a number field where the modular curve X_Γ is defined. We have the following properties:*

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

$$|q_c(P)^{-Ord_c(u_\mathcal{O})/h_c}\gamma_{\mathcal{O},c}^{-1}u_\mathcal{O}(P) - 1|_v \leq 4^{12N^3}|q_c(P)|_v^{1/N},$$

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

(2) Let c be a cusp of X_Γ and $v \in M_K$. For $P \in \Omega_{c,v}$, we have

$$|\log|u_\mathcal{O}(P)|_v - \frac{Ord_c(u_\mathcal{O})}{h_c}\log|q_c(P)|_v| \leq \rho_v.$$

(3) For $v \in M_K^\infty$ and $P \in X_\Gamma(K_v)$, we have

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

3.2 Integral Points on Modular Curves

For a number field K, and $S \subseteq M_K$ a finite subset containing all infinite places. We put $d = [K : \mathbb{Q}]$ and $s = |S|$. We define the following quantity

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

as a function of $N \in \mathbb{N}^+$, where D is the absolute discriminant of K, $\varphi(N)$ is Euler's totient function, and the norm $\mathcal{N}_{K/\mathbb{Q}}(v)$ of a place v, by definition, is equal to $\#(\mathcal{O}_K/\mathfrak{p}_v)$ when v is finite and $\mathfrak{p}_v$ is its corresponding prime ideal, and is set to be 1 if v is infinite.

Sha [55] proved the following theorem:

Theorem 3.2.1 ([55] Theorem 1.2)**.** *Let Γ be of level N. If $v_\infty(\Gamma) \geq 3$, then for any $P \in X_\Gamma(\mathcal{O}_S, j)$,*

$$\mathrm{h}(j(P)) \leq (CdsM^2)^{2sM}(\log(dM))^{3sM}\ell^{dM}\Delta(M),$$

where C is an absolute effective constant, ℓ is the maximal prime such that there exists $v \in S$ with $v|\ell$, or $\ell = 1$ if S only contains infinite places, and M is defined as following:

$$M = \begin{cases} N & \text{if } N \text{ is not a power of any prime;} \\ 3N & \text{if } N \text{ is a power of } 2; \\ 2N & \text{if } N \text{ is a power of an odd prime.} \end{cases}$$

(Here $\mathrm{h}(\cdot)$ is the standard absolute logarithmic height defined on the set $\bar{\mathbb{Q}}$ of algebraic numbers.)

For certain applications it is useful to have an explicit value of the constant C from Theorem 3.2.1. In this note we prove the following result.

Theorem 3.2.2. *The constant C in Theorem 3.2.1 can be taken to be 2^{14}.*

In the proof, we follow the main lines of Sha's argument, with some minor modifications. We calculate explicitly the implicit constants occurring therein.

3.2.1 Proof of Theorem 3.2.2

We only consider the case of mixed level, i.e Theorem 3.2.1, since if N is a power of some prime p, we can replace N by $3N$ if $p = 2$, and by $2N$ if $p \neq 2$. From the assumption, we have that $N \geq 6$.

We consider the case where $\mathbb{Q}(\zeta_N) \subset K$ at first, then consider the general case.

For $P \in X_\Gamma(\mathcal{O}_S, j)$, since $j(P) \in \mathcal{O}_S$, we have

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

for some $w \in S$. Hence, it suffices to bound $\log |j(P)|_w$.

If $|j(P)|_w \leq 3500$, then $\mathrm{h}(j(P)) \leq 16s$, which is a better bound than that given in Theorem 3.2.1 (1) when $C = 2^{14}$.

If $|j(P)|_w > 3500$, then by [55, Proposition 3.3] or [9, Proposition 3.1], we have $P \in \Omega_{c,w}$ for some cusp c, and $|j(P)|_w \leq 2|q_c(P)^{-1}|_w$, where $\Omega_{c,w}$ and q_c are defined in [9, Section 3]. Hence, we only need to bound $\log |q_c(P)^{-1}|_w$.

Notice that if moreover $|q_c(P)|_w > 10^{-N}$, then $\log |j(P)|_w \leq 2N \log 10$ and $\mathrm{h}(j(P)) < 6sN$, which is better than that given in Theorem 3.2.1 when $C = 2^{14}$.

In the sequel, we consider the case where $P \in \Omega_{c,w}$ and $|q_c(P)|_w \leq 10^{-N}$.

We have the following lemma:

LEMMA 3.2.3. *There exists a modular unit W on X_Γ which is integral over $\mathbb{Z}[j]$, and a constant $\gamma_w \in \mathbb{Q}(\zeta_N)$ such that*

$$|\gamma_w^{-1}W(P) - 1|_w \leq 4^{24N^7}|q_c(P)|_w^{1/N},$$

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

If moreover $P \in X_\Gamma(\mathcal{O}_S, j)$ for some $S \subset M_K$ containing all infinity places, then $W(P)$ can be a unit of $\mathcal{O}_S$.

Proof. After taking a transformation, we can suppose that c is the infinity cusp.

We fix an orbit $\mathcal{O}$ of the group action of G on $\mathcal{A}_N$. Put $U = u_{\mathcal{O}}$, where $u_{\mathcal{O}}$ is defined in Definition 3.1.15.

If $\mathrm{Ord}_c U \neq 0$, by Proposition 3.1.17 (1)(4) and the assumption that $v_\infty(G) \geq 3$, we can choose $\mathcal{O}$ with $\mathrm{Ord}_c U < 0$ and another orbit $\mathcal{O}'$ with $\mathrm{Ord}_c V > 0$, where $V = u_{\mathcal{O}'}$, moreover U and V are multiplicatively independent modulo constants.

Define the following function:

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

So we always have $\mathrm{Ord}_c W = 0$ and W is integral over $\mathbb{Z}[j]$ since U, V is integral over $\mathbb{Z}[j]$ by Proposition 3.1.17 (2). If $P \in X_\Gamma(\mathcal{O}_S, j)$ for some $S \subset M_K$ containing all infinity places, then $W(P) \in \mathcal{O}_S$, and by Proposition 3.1.17 (1) and (2), $W(P)$ is a unit of O_S. Moreover, W is not a constant by Proposition 3.1.17(4).

If $W = U$, the bounds follow from Proposition 3.1.21(1).

If $W = U^{\mathrm{Ord}_c V} V^{-\mathrm{Ord}_c U}$, then by Corollary 3.1.19, Proposition 3.1.17(3), Proposition 3.1.21(1), and

$$|\mathcal{A}_N / G| \leq N^2 \prod_{p|N} (1 - p^{-2}) \leq N^2 - 1,$$

we have

$$\gamma_{\mathcal{O},c}^{-\mathrm{Ord}_c V} \gamma_{\mathcal{O}',c}^{\mathrm{Ord}_c U} W = 1 + \sum_{k=1}^{\infty} \phi(k) q_c^{k/N},$$

where

$$|\phi(k)|_w \leq 2^{k + 12N^3(\mathrm{Ord}_c V - \mathrm{Ord}_c U)} e^k \cdot 2^{-24}.$$

Hence

$$\begin{aligned} |\gamma_{\mathcal{O},c}^{-\mathrm{Ord}_c V} \gamma_{\mathcal{O}',c}^{\mathrm{Ord}_c U} W(P) - 1|_w &\leq 2^{12N^3(\mathrm{Ord}_c V - \mathrm{Ord}_c U)} |q_c(P)|_w^{1/N} 2e \sum_{k=0}^{\infty} e^k 5^{-k} \cdot 2^{-24} \\ &\leq 4^{24N^7} |q_c(P)|_w^{1/N}, \\ \mathrm{h}(\gamma_{\mathcal{O},c}^{\mathrm{Ord}_c V} \gamma_{\mathcal{O}',c}^{-\mathrm{Ord}_c U}) &\leq \mathrm{Ord}_c V \mathrm{h}(\gamma_{\mathcal{O},c}) + \mathrm{Ord}_c U \mathrm{h}(\gamma_{\mathcal{O}',c}) \\ &\leq 24N^7 \log 2. \end{aligned}$$

Combine these two cases, we set

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

then the lemma is proved. □

Hence $W(P) = \omega \eta_1^{b_1} \cdots \eta_r^{b_r}$ for some $b_1, \cdots, b_r \in \mathbb{Z}$, where ω is a root of unity and $\{\eta_1, \cdots, \eta_r\}$ is a fundamental system of S-units from [55, Proposition 4.1]. We set

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

where $\eta_0 = \omega \gamma_w^{-1}$. Then we have

$$|\Lambda - 1|_w \leq 4^{24N^7} |q_c(P)|_w^{1/N}. \tag{3.2}$$

If $\Lambda \neq 1$, we will use this upper bound and the lower bound from Theorem 1.3.1 to get a bound of $|q_c(P)|_w$ which gives an upper bound of $\mathrm{h}(j(P))$. For the case where $\Lambda = 1$, see [55, Section 8].

To state the following lemma, we set $r^r = 1$ when $r = 0$, i.e $s = 1$.

LEMMA 3.2.4. *If $\mathbb{Q}(\zeta_N) \subset K$ and $\Lambda \neq 1$, then we have*

$$\mathrm{h}(j(P)) \leq 40dsr^{2r}\zeta^r N^8 \tilde{\Upsilon} R(S) \log(d^2 s r^{4r} \zeta^s N^{16} \tilde{\Upsilon} R(S)),$$

where $\tilde{\Upsilon} = 2^{13s+22} d^{2s+3} \ell^d$, and ζ has been defined in Section 1.2.

Proof. We define $A_0, \cdots, A_r, B_0$ by

$$\log A_i := \max\{\mathrm{h}(\eta_i), 1/d\}, 0 \leq i \leq r;$$

$$B_0 := \max\{3, |b_1|, \cdots, |b_r|\}.$$

Since $\Lambda = \eta_0 \eta_1^{b_1} \cdots \eta_r^{b_r} \neq 1$, by Theorem 1.3.1, we have

$$|\Lambda - 1|_w \geq \exp\{-\Upsilon \log A_0 \cdots \log A_r \log B_0\},$$

where

$$\Upsilon = \begin{cases} 2^{8s+29} d^{s+2} \log(ed), & \text{if } w|\infty, \\ 2^{10s+10} \cdot e^{2s+2} d^{3s+3} p_w^d, & \text{if } w|p_w < \infty. \end{cases} \tag{3.3}$$

Obviously $2^{10s+19} \cdot 2^{3s+3} d^{3s+3} \ell^d = 2^{13s+22} d^{3s+3} \ell^d$ is larger than Υ in each case since $d \geq 2, s \geq 1$, so we can take $\Upsilon = 2^{13s+22} d^{3s+3} \ell^d$.

By (3.2), we have

$$\exp\{-\Upsilon \log A_0 \cdots \log A_r \log B_0\} \leq 4^{24N^7} |q_c(P)|_w^{1/N},$$

that is

$$\log |q_c(P)^{-1}|_w \leq N\Upsilon \log A_0 \cdots \log A_r \log B_0 + 48N^8 \log 2. \tag{3.4}$$

By [55, Proposition 4.1], we have $\zeta \mathrm{h}(\eta_k) \geq 1/d$ and $\zeta \geq 1$, so

$$\log A_k \leq \zeta \mathrm{h}(\eta_k), \;\; k = 1, \cdots, r,$$

$$\log A_1 \cdots \log A_r \leq d^{-r} r^{2r} \zeta^r R(S).$$

Notice that the both sides are 1 when $r = 0$. On the other hand, since

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

we have

$$\log A_0 \leq 24N^7 \log 2.$$

For B_0, we set $B^* = \max\{|b_1|, \cdots, |b_r|\}$ if $r \geq 1$, and $B^* = 0$ if $r = 0$. By [55, Corollary 4.2 and Proposition 6.1] we have

$$\begin{aligned} B^* &\leq 2dr^{2r} \zeta \mathrm{h}(W(P)) \\ &\leq 2dr^{2r} \zeta (2sN^8 \log |q_c^{-1}(P)|_w + 94sN^8 \log N), \end{aligned} \tag{3.5}$$

so

$$B_0 \leq 2dr^{2r} \zeta (2sN^8 \log |q_c^{-1}(P)|_w + 94sN^8 \log N).$$

We write

$$\alpha = 4dsr^{2r} \zeta N^8,$$

$$\beta = 188dsr^{2r} \zeta N^8 \log N = 47\alpha \log N,$$

$$C_1 = \alpha N \Upsilon \log A_0 \cdots \log A_r,$$

$$C_2 = 48\alpha N^8 \log 2 + \beta.$$

Hence, inequalities (3.4) and (3.5) yield

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

By Lemma 1.3.4, we obtain

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

Hence,

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

so we have

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

Next we bound each term on the right-hand side:

$$\begin{aligned} 2s\alpha^{-1} C_1 \log C_1 &= 2sN\mathrm{Y} \log A_0 \cdots \log A_r \log(4dsr^{2r}\zeta N^9 \mathrm{Y} \log A_0 \cdots \log A_r) \\ &\leq 48 \log 2 \cdot d^{-r} s r^{2r} \zeta^r N^8 \mathrm{Y} R(S) \log(96 \log 2 \cdot d^{-r+1} s r^{4r} \zeta^{r+1} N^{16} \mathrm{Y} R(S)) \\ &\leq 39 d^{-r} s r^{2r} \zeta^r N^8 \mathrm{Y} R(S) \log(d^{-r+1} s r^{4r} \zeta^{r+1} N^{16} \mathrm{Y} R(S)), \end{aligned}$$

here we use the fact that $48 \log 2 \times \log(96 \log 2) \leq 140 < 5 \log(d^{-r+1}\mathrm{Y})$; we also have

$$\begin{aligned} s\alpha^{-1}(2C_2 - \beta) + s \log 2 &= 96 \log 2 \cdot sN^8 + 47s \log N + s \log 2 \\ &\leq 98 \log 2 \cdot sN^8. \end{aligned}$$

After replacing $d^{-s}\mathrm{Y} = 2^{13s+22} d^{2s+3} \ell^d$ by $\tilde{\mathrm{Y}}$, we have

$$\mathrm{h}(j(P)) \leq 40 dsr^{2r} \zeta^r N^8 \tilde{\mathrm{Y}} R(S) \log(d^2 s r^{4r} \zeta^s N^{16} \tilde{\mathrm{Y}} R(S)).$$

□

We will use the bound $\zeta \leq 2^{13}(\log d)^3$ subsequently. If $d = 2$,

$$\zeta = \frac{(\log 6)^3}{2} = \frac{(\log_2 6)^3}{2} (\log d)^3 \leq 2^4 (\log d)^3;$$

if $d \geq 3$, then

$$\begin{aligned} \zeta &= 4 \left(\frac{\log d}{\log \log d} \right)^3 \\ &\leq 4 \left(\frac{\log d}{\log \log 3} \right)^3 \\ &\leq 4809 (\log d)^3 \\ &\leq 2^{13} (\log d)^3. \end{aligned}$$

By Lemma 1.2.4 and Lemma 1.2.5, we have

$$R(S) \leq \frac{\omega_K}{2} \frac{1}{(d-1)^{d-1}} (\log |D|)^{d-1} \sqrt{|D|} \prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v),$$

$$\omega_K \leq 2d^2,$$
$$\begin{aligned} \log R(S) &\leq \log(\frac{\omega_K}{2}) + d \log |D| + s \log(d\ell) \\ &\leq 2 \log d + d \log |D| + s \log(d\ell). \end{aligned}$$

We have $d \leq 2s$ and $\log s \leq s/2$. Then we have

$$\begin{aligned}\log \tilde{\Upsilon} &= (13s+22)\log 2 + (2s+3)\log d + d\log \ell \\ &\leq (15s+25)\log 2 + (2s+3)\log s + d\log \ell \\ &\leq 28s + (s+2)s + s\ell \\ &\leq 32s^2\ell\end{aligned}$$

and

$$\begin{aligned}\log(d^2 s r^{4r}\zeta^{r+1}N^{16}\tilde{\Upsilon}R(S)) &\leq 2\log d + 4s\log s + 13s\log 2 + 3s\log\log d + 16\log N + \log\tilde{\Upsilon} \\ &\quad + 2\log d + d\log|D| + s\log(d\ell) \\ &\leq 2s + 2s^2 + 10s + 2s^2 + 16\log N + 32s^2\ell + 2s + 2s\log|D| + s^2\ell \\ &\leq 8N + 2s\log|D| + 51s^2\ell \\ &\leq 61s^2\ell N\log|D| \\ &\leq 2^6 s^2 N\ell\log|D|.\end{aligned}$$

Hence combining with Lemma 3.2.4, we have

$$\begin{aligned}\mathrm{h}(j(P)) &\leq 2^6\cdot d s^{2s-1}\zeta^r N^8\tilde{\Upsilon}R(S)\log(d^2 s r^{4r}\zeta^{r+1}N^{16}\tilde{\Upsilon}R(S)) \\ &\leq 2^{26s+15}\cdot d^{2s+4}(\log d)^{3r}s^{2s-1}N^8\ell^d\frac{\omega_K}{2}\frac{1}{(d-1)^{d-1}}(\log|D|)^{d-1}\sqrt{|D|}\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v) \\ &\quad\cdot(2^6 s^2 N\ell\log|D|) \\ &= 2^{26s+20}d^{2s+4}(\log d)^{3r}s^{2s+1}N^9\ell^{d+1}\omega_K\frac{1}{(d-1)^{d-1}}(\log|D|)^d\sqrt{|D|}\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v)\end{aligned} \tag{3.6}$$

Next we deal with the general case. Set $\widetilde{K} = K\cdot\mathbb{Q}(\zeta_N) = K(\zeta_N)$. Let $\widetilde{S}$ be the set consisting of the extensions of the places from S to $\widetilde{K}$, that is,

$$\widetilde{S} = \{\tilde{v}\in M_{\widetilde{K}} : \tilde{v}|v, v\in S\}.$$

Then $P\in X_\Gamma(\mathcal{O}_{\widetilde{S}}, j)$. Put $\tilde{d} = [\widetilde{K}:\mathbb{Q}]$, $\tilde{s} = |\widetilde{S}|$, $\tilde{r} = \tilde{s}-1$, and let $\widetilde{D}$ be the absolute discriminant of $\widetilde{K}$.

LEMMA 3.2.5.

$$\begin{gathered}N - \varphi(N) \geq 4 \\ \tilde{s} \leq s\varphi(N), \\ \tilde{d} \leq d\varphi(N), \\ \omega_{\widetilde{K}} \leq 2d^2\varphi(N)^2, \\ |\widetilde{D}| \leq N^{dN}|D|^{\varphi(N)}, \\ \prod_{\substack{v\in\widetilde{S}\\ v\nmid\infty}}\log\mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v) \leq 4^{s\varphi(N)}\left(\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v)\right)^{\varphi(N)},\end{gathered}$$

Proof. The first three inequalities come directly from the definition of $\widetilde{K}$ and $\widetilde{S}$ and $N\geq 6$ has at least two prime factors. The fourth inequality comes from $\omega_{\widetilde{K}} \leq 2\tilde{d}^2 \leq 2d^2\varphi(N)^2$.

Let $D_{\widetilde{K}/K}$ be the relative discriminant of $\widetilde{K}/K$. We have

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

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

$$\mathcal{O}_K \subset \mathcal{O}_K(\zeta_N) \subset \mathcal{O}_{\widetilde{K}}.$$

Note that the absolute value of the discriminant of the polynomial $x^N - 1$ is N^N, we obtain

$$D_{\widetilde{K}/K} | N^N,$$

so

$$|\mathcal{N}_{K/\mathbb{Q}}(D_{\widetilde{K}/K})| \leq N^{dN}.$$

Hence,

$$|\widetilde{D}| \leq N^{dN}|D|^{\varphi(N)}.$$

Notice that $\widetilde{K}/K$ is Galois. Let v be a non-Archimedean place of K, and let $v_1, \ldots, v_g$ be all its extensions to $\widetilde{K}$ with residue degree f over K. Then $gf \leq [\widetilde{K} : K] \leq \varphi(N)$, which implies $g \log_2 f \leq gf \leq \varphi(N)$, i.e. $f^g \leq 2^{\varphi(N)}$. Note that $2 \log \mathcal{N}_{K/\mathbb{Q}}(v) > 1$ and $\mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v_k) = \mathcal{N}_{K/\mathbb{Q}}(v)^f$ for $1 \leq k \leq g$, $g \leq \varphi(N)$, we have

$$\begin{aligned}\prod_{k=1}^{g} \log \mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v_k) &\leq 2^{\varphi(N)} (\log \mathcal{N}_{K/\mathbb{Q}}(v))^g \\ &\leq 2^{\varphi(N)} (2 \log \mathcal{N}_{K/\mathbb{Q}}(v))^g \\ &\leq 4^{\varphi(N)} (\log \mathcal{N}_{K/\mathbb{Q}}(v))^{\varphi(N)}.\end{aligned}$$

Hence

$$\prod_{\substack{v \in \widetilde{S} \\ v \nmid \infty}} \log \mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v) \leq 4^{s\varphi(N)} \left(\prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v) \right)^{\varphi(N)}.$$

□

Combine the lemma above with the bound (3.6), we have

$$\begin{aligned}\mathrm{h}(j(P)) &\leq 2^{26\tilde{s}+20} \tilde{d}^{2\tilde{s}+4} (\log \tilde{d})^{3\tilde{r}} \tilde{s}^{2\tilde{s}+1} N^9 \ell^{\tilde{d}+1} \omega_{\widetilde{K}} \frac{1}{(\tilde{d}-1)^{\tilde{d}-1}} (\log |\widetilde{D}|)^{\tilde{d}} \sqrt{|\widetilde{D}|} \prod_{\substack{v \in \widetilde{S} \\ v \nmid \infty}} \log \mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v) \\ &\leq 2^{28s\varphi(N)+21} d^{2s\varphi(N)+6} (\log d\varphi(N))^{3s\varphi(N)} s^{2s\varphi(N)+1} \varphi(N)^{4s\varphi(N)+7} N^9 \ell^{d\varphi(N)+1} \Delta(N) \\ &\leq 2^{28sN} d^{2sN} (\log dN)^{3sN} s^{2sN} N^{4sN} \ell^{dN} \Delta(N) \\ &\leq (2^{14} dsN^2)^{2sN} (\log dN)^{3sN} \ell^{dN} \Delta(N).\end{aligned}$$

This completets the proof of Theorem 3.2.2 when $\Lambda \neq 1$.

$\Lambda = 1$

In this subsection, we keep the assumptions as last subsection and follow the idea of [55, Section 8]. For the convenience of readers, I illustrate them here: N is not a prime power, Γ is a congruence subgroup of level N with $v_\infty(\Gamma) \geq 3$, and as before, firstly, we will assume that $\mathbb{Q}(\zeta_N) \subset K$ and $S \subset M_K$ containing all infinity places, $w \in S$. Let $P \in \Omega_{c,w} \cap X_\Gamma(\mathcal{O}_S, j)$ with $|q_c(P)|_w \leq 10^{-N}$ and c a cusp, and W, γ_w are

those in Lemma 3.2.3, that is, W is a modular unit on X_Γ which is integral over $\mathbb{Z}[j]$, and $\gamma_w \in \mathbb{Q}(\zeta_N)$ such that $W(P)$ is a unit of $\mathcal{O}_S$

$$|\gamma_w^{-1}W(P) - 1|_w \leq 4^{24N^7}|q_c(P)|_w^{1/N},$$

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

We further assume that $\Lambda = \gamma_w^{-1}W(P) = 1$.

In the following, we view W as a function of q_c.

LEMMA 3.2.6. *There exists an integer-valued function f 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:*

$$\log \frac{W(q_c)}{\gamma_w} = 2\pi f(q_c)i + \sum_{k=1}^{\infty} \lambda_k^c q_c^{k/N}$$

and

$$|\lambda_k^c|_v = \begin{cases} |k|_v^{-1} & \text{if } v \text{ is finite,} \\ 48N^6(k+N) & \text{if } v \text{ is infinite,} \end{cases}$$

and $\lambda_k^c \neq 0$ for some $k \leq N^{10}$. In particular, for every $k \geq 1$, we have

$$\mathrm{h}(\lambda_k^c) \leq \log(48N^7 + 48kN^6) + \log k.$$

Proof. Firstly, we show that for any $U = u_{\mathcal{O}}$, $\mathcal{O}$ is an orbit of the action of Γ on $\mathcal{A}_N$, there exists f and an integer-valued function f with respect to q_c and $\lambda_1^c, \lambda_2^c, \lambda_3^c \cdots \in \mathbb{Q}(\zeta_N)$ such that

$$\log \frac{U(q_c)}{\gamma_{\mathcal{O},c} q_c^{\frac{\mathrm{Ord}_c U}{h_c}}} = 2\pi f_U(q_c)i + \sum_{k=1}^{\infty} \lambda_{k,U}^c q_c^{k/N}$$

and

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

If so, when $W = U^{\mathrm{Ord}_c V} V^{-\mathrm{Ord}_c U}$,

$$\begin{aligned} \log \frac{W(q_c)}{\gamma_w} &= 2\pi(f_U(q_c)\mathrm{Ord}_c V - f_V(q_c)\mathrm{Ord}_c U)i + \sum_{k=1}^{\infty} (\lambda_{k,U}^c \mathrm{Ord}_c V - \lambda_{k,V}^c \mathrm{Ord}_c U) q_v^{k/N} \\ &= 2\pi f(q_c)i + \sum_{k=1}^{\infty} \lambda_k^c q_c^{k/N}, \end{aligned}$$

where f_U and f_V, $\lambda_{k,U}$ and $\lambda_{k,V}$ are functions and constants for U and V respectively and

$$\begin{aligned} f(q_c) &= f_U(q_c)\mathrm{Ord}_c V - f_V(q_c)\mathrm{Ord}_c U, \\ \lambda_k^c &= \lambda_{k,U}^c \mathrm{Ord}_c V - \lambda_{k,V}^c \mathrm{Ord}_c U. \end{aligned}$$

Hence, if v is finite,

$$\begin{aligned} |\lambda_k^c|_v &\leq \max\{|\lambda_{k,U}^c \mathrm{Ord}_c V|_v, |\lambda_{k,V}^c \mathrm{Ord}_c U|_v\} \\ &= |k|_v^{-1}, \end{aligned}$$

since $|\mathrm{Ord}_c V|_v, |\mathrm{Ord}_c U|_v \leq 1$; and if v is infinite,

$$\begin{aligned} |\lambda_k^c|_v &\leq |\lambda_{k,U}^c \mathrm{ord}_c V|_v + |\lambda_{k,V}^c \mathrm{ord}_c U|_v \\ &= 48N^6(k+N), \end{aligned}$$

since $|\mathrm{Ord}_c V|_v, |\mathrm{Ord}_c U|_v \leq N^4$ by Proposition 3.1.17 (3). In this case,

$$\begin{aligned}
\mathrm{h}(\lambda_k^c) &= \frac{1}{\varphi(N)} \sum_{v \in M_{\mathbb{Q}(\zeta_N)}} [\mathbb{Q}(\zeta_N)_v : \mathbb{Q}_v] \log^+ |\lambda_k^c|_v \\
&\leq \log(48N^7 + 48N^6 k) + \frac{1}{\varphi(N)} \sum_{v \in M^0_{\mathbb{Q}(\zeta_N)}} [\mathbb{Q}(\zeta_N)_v : \mathbb{Q}_v] \log^+ |k^{-1}|_v \\
&\leq \log(48N^7 + 48N^6 k) + \mathrm{h}(k^{-1}) \\
&= \log(48N^7 + 48N^6 k) + \log k.
\end{aligned}$$

We will prove our assertion. By definition, we have

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

Hence

$$\begin{aligned}
\log \frac{U(q_c)}{\gamma_{\mathcal{O},c} q_c^{\frac{\mathrm{Ord}_c U}{e_c}}} &= 2\pi f(q_c) i \\
&\quad + \sum_{\mathbf{a} \in \mathcal{O}} \sum_{\substack{n=0 \\ n+a_1 \neq 0}}^{\infty} 12N \log(1 - q_c^{n+a_1} e^{2\pi i a_2}) + \sum_{n=0}^{\infty} 12N \log(1 - q_c^{n+1-a_1} e^{-2\pi i a_2}),
\end{aligned}$$

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 obtain the desire formula for $\log \dfrac{U(q_c)}{\gamma_{\mathcal{O},c} q_c^{\frac{\mathrm{Ord}_c U}{h_c}}}$.

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

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

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 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 true for the q-series for $\log(1 - q_c^{n+1-a_1} e^{-2\pi i a_2})$. The bound for $|\lambda_{k,U}^c|_v$ now follows by summation.

Finally, we will show that $\lambda_k^c \neq 0$ for some $k \leq N^1 0$. Since W is not a constant, there must exist some $\lambda_k^c \neq 0$. Since $\mathrm{Ord}_c W = 0$, then by Corollary 3.1.19, we have $W(c) = \gamma_w$, then $f(q_c(c)) = 0$. We extend the additive valuation Ord_c from the field $K(X_\Gamma)$ to the field of formal power series $K((q_c^{1/h_c}))$. Then $\mathrm{Ord}_c q_c^{1/h_c} = 1$, $\mathrm{Ord}_c q_c^{k/N} = k\dfrac{h_c}{N} \leq k$, and $\min\{k \mid \lambda_k^c \neq 0\}\dfrac{h_c}{N} = \mathrm{Ord}_c(-2\pi f(q_c) i + \log W/\gamma_w) \leq \mathrm{Ord}_c(\log W/\gamma_w) = \mathrm{Ord}_c(W/\gamma_w - 1)$. The latter quantity is bounded by the degree of $W/\gamma_w - 1$, which is equal to the degree of W.

The degree of W is equal to $1/2 \sum\limits_{c_0} |\mathrm{Ord}_{c_0} W|$, here the sum runs through all the cusps of X_Γ. From Proposition 3.1.17 (3), for $W = U$, $|\mathrm{Ord}_{c_0} W| < N^4$; for $W = U^{\mathrm{Ord}_c V} V^{-\mathrm{Ord}_c U}$, $|\mathrm{Ord}_{c_0} W| < 2N^8$. Notice that the number of cusps is $|\mathcal{M}_N/G|$, see section 3, which is bounded by N^2. Hence we get the result. □

Proposition 3.2.7. *For $P \in \Omega_{c,w}$ such that $W(P) = \gamma_w$, we have*

$$\log |q_c(P)^{-1}|_w \leq N\varphi(N)\log(48N^{26}+48N^{17}) + N\log(96N^6(N^{10}+N+1)).$$

Proof. Let n be the smallest k such that $\lambda_k^c \neq 0$. Then $n \leq N^{10}$. We assume that $|q_c(P)|_w \leq 10^{-N}$, otherwise, $\log |q_c(P)|_w^{-1} \leq N\log 10$, there is nothing to prove. Since $W(P) = \gamma_w$, it follows from last lemma that $2\pi f(q_c(P))i + \sum_{k=1}^{\infty} \lambda_k^c q_c(P)^{k/N} = 0$.

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

$$\begin{aligned}
|\sum_{k=n+1}^{\infty} \lambda_k^c q_c(P)^{k/N}|_w &\leq |\sum_{k=n+1}^{\infty} |\lambda_k^c|_v |q_c(P)|_w^{k/N} \\
&\leq \sum_{k=n+1}^{\infty} 48N^6(N+k)|q_c(P)|_w^{k/N} \\
&= 48N^6(n+N+1+\frac{|q_c(P)|_c^{1/N}}{1-|q_c(P)|_w^{1/N}})\frac{|q_c(P)|_w^{(n+1)/N}}{1-|q_c(P)|_w^{1/N}} \\
&\leq 48N^6(n+N+11/10)\cdot 10/9|q_c(P)|_w^{(n+1)/N} \\
&\leq 96N^6(n+N+1)|q_c(P)|_w^{(n+1)/N}.
\end{aligned} \tag{3.7}$$

On the other hand, using Liouville's inequality, we obtain

$$|\lambda_n^c|_w \geq e^{-[\mathbb{Q}(\zeta_N):\mathbb{Q}]h(\lambda_n^c)} \geq (48nN^7+48nN^6)^{-\varphi(N)}.$$

Then we have

$$96N^6(n+N+1)|q_c(P)|_w^{(n+1)/N} \leq |q_c(P)|_w^{n/N}(48nN^7+48n^2N^6)^{-\varphi(N)},$$

so

$$\log |q_c(P)^{-1}|_w \leq N\varphi(N)\log(48N^{26}+48N^{17}) + N\log(96N^6(N^{10}+N+1)).$$

Suppose that $f(q_c(P)) \neq 0$. Then $2\pi \leq |\sum_{k=n}^{\infty} \lambda_k^c q_c(P)^{k/N}|_w \leq 96N^6(n+N)|q_c(P)|_w^{n/N}$ by inequality 3.7. Then we obtain

$$\begin{aligned}
\log |q_c(P)^{-1}|_w &\leq -\frac{N}{n}\log(2\pi) + \frac{N}{n}\log(96N^6(N^{10}+N)) \\
&\leq N\log(96N^6(N^{10}+N))
\end{aligned}$$

□

For general number field K, we set $\widetilde{K} = K(\zeta_N)$, and $\tilde{s}$ as before. With the proposition above and $N \geq 6$, we have

$$\begin{aligned}
\mathrm{h}(j(P)) &\leq \tilde{s}(\log |q_c(P)^{-1}|_w + \log 2) \\
&\leq sN(N\varphi(N)\log(48N^{26}+48N^{17}) + N\log(96N^6(N^{10}+N+1)) + \log 2) \\
&\leq sN^3 \cdot 2\log(192N^{26}) \\
&\leq 58sN^3\log(N),
\end{aligned}$$

which is obviously better than the result in Theorem 3.2.1.

3.3 Integral Points on $X_0(p)$

Theorem 3.3.1. *Let p be a prime number other than $2,3,5,7,13$, K be a number field, and $S \subseteq M_K$ be a finite set containing all archimedean places. Then for $P \in X_0(p)(\mathcal{O}_S, j)$, we have*

$$\mathrm{h}(j(P)) \leq e^{9s^2p^4 \log p} C(K,S)^{p^2},$$

where $C(K,S)$ can be effectively determined in terms of K and S. More explicitly, $C(K,S)$ can be chose as

$$C(K,S) = 2^{29s} d^{9s} s^{2s} \ell^d |D| (\log(|D|+1))^d \prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v),$$

where $d = [K:Q]$, D is the absolute discriminant of K, $s = |S|$, ℓ is the maximal prime q such that there exists $v \in S$ with $v|q$.

Let $\widetilde{\Gamma}$ be the subgroup of $\Gamma_0(p)$ defined as follows: set $A = \{a \in \mathbb{F}_p^* \mid a^{12} = 1\}$, and

$$\widetilde{\Gamma} = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_0(p) \mid a \bmod p \in A \right\}. \tag{3.8}$$

It is not hard to see that the curve $X_{\widetilde{\Gamma}}$ and the natural morphisms $X_1(p) \to X_{\widetilde{\Gamma}} \xrightarrow{\pi} X_0(p)$ are defined over $\mathbb{Q}$.

Proposition 3.3.2. *1. We have* $\deg \pi \leq \dfrac{p-1}{2}$.

2. *When $p \notin \{2,3,5,7,13\}$, the curve $X_{\widetilde{\Gamma}}$ has at least 3 cusps.*

3. *The morphism π is étale.*

Proof. Set $\overline{\widetilde{\Gamma}}$ the image of $\widetilde{\Gamma}$ in $\mathrm{SL}_2(\mathbb{F}_p)$, then we have

$$\deg\pi = [\Gamma_0(p) : \widetilde{\Gamma}] = [\mathrm{ST}_2(\mathbb{F}_p) : \overline{\widetilde{\Gamma}}] = p(p-1)/(p|A|) \leq \frac{p-1}{2}.$$

The second assertion is proved in [5, page 84].

About the third assertion, it is only proved in [5] that π is étale outside the cusps. In fact, π is étale at the cusps as well. Indeed, the j-map $X(p) \xrightarrow{j} \mathbb{P}^1$ has ramification index p at every cusp. Hence 1 and p are the only possible ramification indices for π. Since $\deg \pi \leq (p-1)/2 < p$, the ramification indices at the cusps are all 1. □

Corollary 3.3.3. *Let K be a number field, $P \in X_0(p)(K)$ and $\widetilde{P} \in \pi^{-1}(P)$. Then*

$$[\widetilde{K} : K] \leq \frac{p-1}{2}, \tag{3.9}$$

$$|\mathcal{N}_{K/\mathbb{Q}}(D_{\widetilde{K}/K})| \leq p^{d^2(p-1)^3/8}, \tag{3.10}$$

where $\widetilde{K} = K(\widetilde{P})$, the residue field of $\widetilde{P}$, and $d = [K:\mathbb{Q}]$.

Proof. It follows from Proposition 3.3.2 and the formula

$$\deg\pi = \sum_{Q \in \pi^{-1}(P)} [K(Q) : K]$$

that

$$[\widetilde{K} : K] \leq \deg \pi \leq (p-1)/2.$$

We know that the modular curve $X_1(p)$ has good reductions outside p by Igusa's Theorem, see [22, Section 8.6]. Now by Proposition 1.5.5, $X_{\widetilde{\Gamma}}$ also admits good reduction outside p. Combining this with Proposition 3.3.2, Lemma 1.5.6 and the fact that $[K(X_{\widetilde{\Gamma}}) : K(X_0(p))] = \deg \pi \leq \frac{p-1}{2}$, we apply Lemma 1.5.2 with $T = \{q : q \leq (p-1)/2, q \text{ is prime}\} \cup \{p\}$, we obtain (3.10). □

3.3.1 Calculations

For a number field K, and a finite subset $S \subseteq M_K$ containing all infinite places, we put $d = [K : \mathbb{Q}]$ and $s = |S|$. Let $\mathcal{O}_K$ be the ring of integers of K. We define the following quantity

$$\Delta_0(N) := \sqrt{N^{dN}|D|^{\varphi(N)}}(\log(N^{dN}|D|^{\varphi(N)}))^{d\varphi(N)} \times \left(\prod_{\substack{v\in S\\ v\nmid\infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v)\right)^{\varphi(N)}$$

as a function of $N \in \mathbb{N}^+$, where D is the absolute discriminant of K, $\varphi(N)$ is the Euler's totient function, and the norm $\mathcal{N}_{K/\mathbb{Q}}(v)$ of a place v, by definition, is equal to $|\mathcal{O}_K/\mathfrak{p}_v|$ when v is finite and $\mathfrak{p}_v$ is its corresponding prime ideal, and is set to be 1 if v is infinite.

With these notations above, the main tool to prove Theorem 3.3.1 are Chevalley-Weil Principle, Theorem 3.2.1 and Theorem 3.2.2.

Proof. Recall the congruence subgroup $\widetilde{\Gamma} \subset \Gamma_0(p)$ defined in subsection 7.4.3, i.e.

$$\widetilde{\Gamma} = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_0(p) \mid a \bmod p \in A\},$$

where $A = \{a \in \mathbb{F}_p^* : a^{12} = 1\}$, and the natural map $\pi : X_{\widetilde{\Gamma}} \to X_0(p)$. For any $P \in X_0(p)(\mathcal{O}_S, j)$, there exist a finite extension $\widetilde{K}$ of K and $\widetilde{P} \in X_{\widetilde{\Gamma}}(\widetilde{K})$ such that $\pi(\widetilde{P}) = P$. For a non-constant morphism between projective curves, it's always dominant and finite, so $\pi(X_{\widetilde{\Gamma}}) = X_0(p)$. Obviously, $\mathrm{h}(j(\widetilde{P})) = \mathrm{h}(j(P))$, so it's sufficient to bound $\mathrm{h}(j(\widetilde{P}))$. Hence we consider the points in $X_{\widetilde{\Gamma}}(\mathcal{O}_{\widetilde{S}}, j)$, where

$$\widetilde{S} = \{v \in M_{\widetilde{K}} : v|w \text{ for some } w \in S\}.$$

By Proposition 3.3.2, we know that $X_{\widetilde{\Gamma}}$ has at least three cusps.

To apply Theorem 3.2.1 for $\mathrm{h}(j(\widetilde{P}))$, we should bound some invariants of $\widetilde{K}$ and $\widetilde{S}$. We fix some notations before proceeding with the proof, we set

$$\tilde{\Delta}_0 := \sqrt{(2p)^{2\tilde{d}p}|\widetilde{D}|^{p-1}}(\log((2p)^{2\tilde{d}p}|\widetilde{D}|^{p-1}))^{\tilde{d}(p-1)} \times \left(\prod_{\substack{v\in\widetilde{S}\\ v\nmid\infty}} \log \mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v)\right)^{p-1},$$

$$D^* := p^{d^2\frac{(p-1)^3}{8}}|D|^{\frac{p-1}{2}},$$

$$\Delta(p) := \sqrt{(2p)^{dp(p-1)}|D^*|^{p-1}}(\log((2p)^{dp(p-1)}|D^*|^{p-1}))^{d\frac{(p-1)^2}{2}} \times \left(\prod_{\substack{v\in S\\ v\nmid\infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v)\right)^{\frac{(p-1)^2}{2}},$$

where $\tilde{d} := [\widetilde{K} : \mathbb{Q}]$, and $\widetilde{D}$ is the absolute discriminant of $\widetilde{K}$.

Follow the idea of [55]. Let $\tilde{s} = |\widetilde{S}|$, then $\tilde{s} \le [\widetilde{K} : K]s \le \dfrac{p-1}{2}s$ and $\tilde{d} \le d\dfrac{p-1}{2}$. For the absolute discriminant $\widetilde{D}$ of $\widetilde{K}$, we have

$$\begin{aligned} |\widetilde{D}| &= |\mathcal{N}_{K/\mathbb{Q}}(D_{\widetilde{K}/K})||D|^{[\widetilde{K}:K]} \\ &\le p^{d^2\frac{(p-1)^3}{8}}|D|^{\frac{p-1}{2}} \\ &= D^*. \end{aligned}$$

Now let w be a non-Archimedean place of K, and $v_1, \dots, v_m$ be all its extensions to $\widetilde{K}$ with residue degrees $f_1, \dots, f_m$ respectively over K. Then $f_1 + \cdots + f_m \leq [\widetilde{K} : K] \leq \dfrac{p-1}{2}$, which implies $\log_2 f_1 + \cdots + \log_2 f_m \leq f_1 + \cdots + f_m \leq \dfrac{p-1}{2}$, i.e. $f_1 \dots f_m \leq 2^{\frac{p-1}{2}}$. Since $\mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v_k) = \mathcal{N}_{K/\mathbb{Q}}(w)^{f_k}$ for $1 \leq k \leq m$, we have

$$\prod_{v|w} \log \mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v) \leq 2^{\frac{p-1}{2}} (\log \mathcal{N}_{K/\mathbb{Q}}(w))^{\frac{p-1}{2}}.$$

Hence

$$\prod_{\substack{v \in \widetilde{S} \\ v \nmid \infty}} \log \mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v) \leq 2^{s\frac{p-1}{2}} \left(\prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v) \right)^{\frac{p-1}{2}},$$

and

$$\begin{aligned}
\tilde{\Delta}_0 &= \sqrt{(2p)^{2\tilde{d}p} |\widetilde{D}|^{p-1}} (\log((2p)^{2\tilde{d}p} |\widetilde{D}|^{p-1}))^{\tilde{d}(p-1)} \times \left(\prod_{\substack{v \in \widetilde{S} \\ v \nmid \infty}} \log \mathcal{N}_{\widetilde{K}/\mathbb{Q}}(v) \right)^{p-1} \\
&\leq \sqrt{(2p)^{dp(p-1)} |D^*|^{p-1}} (\log((2p)^{dp(p-1)} |D^*|^{p-1}))^{d\frac{(p-1)^2}{2}} \times 2^{s\frac{(p-1)^2}{2}} \\
&\quad \times \left(\prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v) \right)^{\frac{(p-1)^2}{2}} \\
&= 2^{s\frac{(p-1)^2}{2}} \Delta(p).
\end{aligned}$$

By Theorem 3.2.1, we have

$$\begin{aligned}
\mathrm{h}(j(P)) &= \mathrm{h}(j(\widetilde{P})) \\
&\leq (C\tilde{d}\tilde{s}(2p)^2)^{4\tilde{s}p} (\log(2\tilde{d}p))^{6\tilde{s}p} \ell^{2\tilde{d}p} \tilde{\Delta}_0 \\
&\leq 2^{s\frac{(p-1)^2}{2}} (Cds(p-1)^2 p^2)^{2sp(p-1)} (\log(dp(p-1)))^{3sp(p-1)} \ell^{dp(p-1)} \Delta(p)
\end{aligned}$$

where ℓ is the maximal prime such that there exists $v \in S$ with $v|\ell$.

This bound can be made clearer. Indeed, we have the inequalities

$$D^* \leq e^{d^2 p^3 / 8 \log p} |D|^{p/2},$$

$$\begin{aligned}
\Delta(p) &\leq e^{d^2 p^4 \log p} (2^d |D|)^{p^2} \cdot (d^2 p^4 \log p + p^2 \log |D|)^{dp^2/2} \times \left(\prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v) \right)^{p^2} \\
&\leq e^{4s^2 p^4 \log p} (2^d |D|)^{p^2} \cdot (d^2 p^5 / 2 \log(|D| + 1))^{dp^2} \times \left(\prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v) \right)^{p^2} \\
&\leq e^{7s^2 p^4 \log p} \left((d^2 \log(|D| + 1))^d |D| \prod_{\substack{v \in S \\ v \nmid \infty}} \log \mathcal{N}_{K/\mathbb{Q}}(v) \right)^{p^2} \\
&= e^{7s^2 p^4 \log p} C_1(K, S)^{p^2},
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{h}(j(P)) &\leq 2^{sp^2}(Cdsp^4)^{2sp^2}(\log d + 2\log p)^{3sp^2}\ell^{dp^2}e^{7s^2p^4\log p}C_1(K,S)^{p^2} \\
&\leq e^{9s^2p^4\log p}\cdot 2^{sp^2}(Cds)^{2sp^2}(2d)^{3sp^2}\ell^{dp^2}C_1(K,S)^{p^2} \\
&\leq e^{9s^2p^4\log p}\left(2^s\cdot C^{2s}d^{9s}s^{2s}\ell^d|D|(\log(|D|+1))^d\prod_{\substack{v\in S\\ v\nmid\infty}}\log\mathcal{N}_{K/\mathbb{Q}}(v)\right)^{p^2} \\
&= e^{9s^2p^4\log p}C(K,S)^{p^2}.
\end{aligned}$$

Hence we get Theorem 3.3.1 if we take $C = 2^{14}$ by Theorem 3.2.2. □

Part II

Singular Moduli

Chapter 4

Complex Multiplication

There is nothing new in this Chapter, but it provides the background for the study of singular moduli in the next chapter. The main reference is [59, Chapter II]. For elliptic curves over $\mathbb{C}$, we refer to [60, Chapter VI] or [22, Section 1.3, Section 1.4]. Because of the aim of this chapter, we will just give sufficient results for the next chapter.

4.1 CM Elliptic Curves over $\mathbb{C}$

For a lattice $\Lambda \subset \mathbb{C}$, we denote by E_Λ the corresponding elliptic curve. For $\tau \in \mathbb{H}$, we denote $\Lambda_\tau = \langle \tau, 1 \rangle$, and $E_\tau = E_{\Lambda_\tau}$. For a homomorphism of elliptic curves (over $\mathbb{C}$), we mean a morphism of algebraic varieties which is also a homomorphism of groups. For elliptic curves E_Λ and $E_{\Lambda'}$, we have

$$\mathrm{Hom}(E_\Lambda, E_{\Lambda'}) \simeq \{\lambda \in \mathbb{C} \mid \lambda\Lambda \in \Lambda'\}.$$

Theorem 4.1.1. *Let E be an elliptic curve over $\mathbb{C}$. Then exactly one of the following statements is true:*

(1) $\mathrm{End}(E) \simeq \mathbb{Z}$.

(2) $\mathrm{End}(E) \simeq \mathcal{O}$, *where $\mathcal{O}$ is isomorphic to an order in some imaginary quadratic field* $\mathbb{Q}(\sqrt{-D})/\mathbb{Q}$, $D \geq 1$.

Proof. Let $E \simeq \mathbb{C}/\Lambda$ with $\Lambda = \langle \omega_1, \omega_2 \rangle$, $\omega_1/\omega_2 \in \mathbb{H}$. We have $\mathrm{End}(E) = \{\lambda \in \mathbb{C} \mid \lambda\omega_1, \lambda\omega_2 \in \Lambda\}$. If $\mathrm{End}(E) \neq \mathbb{Z}$, there exists $\alpha \in \mathrm{End}(E) \setminus \mathbb{Z}$ such that

$$\alpha\omega_1 = m\omega_1 + n\omega_2,$$

$$\alpha\omega_2 = r\omega_1 + s\omega_2,$$

for some $m, n, r, s \in \mathbb{Z}$. Let $\tau = \omega_1/\omega_2$, so

$$\alpha\tau = m\tau + n,$$

$$\alpha = r\tau + s.$$

Since $\tau \in \mathbb{R}$ and $r \neq 0$, so $\alpha \notin \mathbb{R}$, and

$$r\tau^2 + (s - m)\tau - n = 0.$$

Hence $r\tau$ is integral over $\mathbb{Z}$ of degree 2, $K = \mathbb{Q}(\tau)$ is a imaginary quadratic field, and $\mathbb{Q}(\alpha) = \mathbb{Q}(\tau)$, $\alpha \in \mathcal{O}_K$. We have that $\mathrm{End}(E) \subset \mathcal{O}_K$ and $\mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q} = K$, so $\mathrm{End}(E)$ is an order in K. □

Definition 4.1.2. *An elliptic curve $E/\mathbb{C}$ is said to have complex multiplication if* $\mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$ *is an imaginary quadratic field when embedded into* $\mathbb{C}$*. We will call E a CM elliptic curve for abbreviation.*

Remark. *(1) From the theorem above, we have bijections*

$$\begin{aligned}\{\textit{CM elliptic curves over } \mathbb{C}\} &\leftrightarrow_{SL_2(\mathbb{Z})} \backslash\{\tau \in \mathbb{H} \mid \tau \textit{ is algebraic of degree } 2\} \\ &\leftrightarrow \{\tau \in \mathcal{F} \mid \tau \textit{ is algebraic of degree } 2\}\end{aligned}$$

where $\mathcal{F}$ is the standard fundamental domain.

Proof. It is sufficient to show the first bijection. If $E_\tau = \mathbb{C}/\Lambda_\tau$ with $\Lambda_\tau = \langle \tau, 1\rangle$ is of CM, then from proof of the theorem above, τ is algebraic of degree 2. Hence the map is well-defined and injective. Conversely, for every $\tau \in \mathbb{H}$ algebraic of degree 2, suppose that $a\tau^2 + b\tau + c = 0$ with $a, b, c \in \mathbb{Z}, a \neq 0$. Let $\alpha = a\tau$. Then

$$\alpha\tau = a\tau^2 = -b\tau - c \in \Lambda_\tau,$$

$$\alpha = a\tau \in \Lambda_\tau.$$

Hence $\alpha \in \mathrm{End}(E_\tau) \setminus \mathbb{Z}$, E_τ is a CM elliptic curve. □

(2) For an order $\mathcal{O}$ in some imaginary quadratic field K, we set

$$\mathcal{ELL}(\mathcal{O}) := \{\textit{elliptic curve } E/\mathbb{C} \textit{ with } \mathrm{End}(E) \simeq \mathcal{O}\}/\simeq_{\mathbb{C}},$$

here quotient by $\simeq_{\mathbb{C}}$ means taking isomorphic class over $\mathbb{C}$. Notice that, for a lattice Λ, $E_\Lambda \in \mathcal{ELL}(\mathcal{O}_K)$ if and only if $\mathcal{O}_K\Lambda = \Lambda$.

For a CM elliptic curve E, since $\mathrm{End}(E)$ is an order of an imaginary quadratic field, then there are two ways to embed the order $\mathrm{End}(E)$ into $\mathbb{C}$. We can pin down one of these embeddings in a canonical way.

Proposition 4.1.3. *Let $E/\mathbb{C}$ be an elliptic curve with complex multiplication by the ring $\mathcal{O}$. Then there exists a unique isomorphism*

$$[\cdot] : \mathcal{O} \to \mathrm{End}(E)$$

such that for any differential form $\omega \in H^0(E, \Omega_E)$ on E, and $\alpha \in \mathcal{O}$

$$[\alpha]^*\omega = \alpha\omega.$$

We say in this case that the pair $(E, [\cdot])$ is normalized.

Proof. Let $\Lambda \subset \mathbb{C}$ be a lattice such that $E \simeq E_\Lambda$. Then

$$\mathcal{O} = \{\alpha \in \mathbb{C} \mid \alpha\Lambda \subset \Lambda\}.$$

Set $[\alpha] : E_\Lambda \to E_\Lambda$ such that

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\phi_\alpha} & \mathbb{C} \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ E_\Lambda & \xrightarrow[[\alpha]]{} & E_\Lambda \end{array}$$

commutes, where $\pi : \mathbb{C} \to E_\Lambda$ is the quotient map, and $\phi_\alpha(z) = \alpha z$ for any $z \in \mathbb{C}$. We claim that $[\alpha]^*\omega = \alpha \cdot \omega$ for any $\omega \in \mathrm{H}^0(E_\Lambda, \Omega_{E_\Lambda})$. Indeed,

$$\pi^*[\alpha]^*\omega = \phi_\alpha^*\pi^*\omega = \alpha\pi^*\omega = \pi^*(\alpha\omega),$$

and $\pi^* : \mathrm{H}^0(E_\Lambda, \Omega_{E_\Lambda}) \hookrightarrow \mathrm{H}^0(\mathbb{C}, \Omega_{\mathbb{C}})$ is injective. Hence we have our claim. □

Corollary 4.1.4. *Let $(E_1, [\cdot]_1)$ and $(E_2, [\cdot]_2)$ be normalized elliptic curves with complex multiplication by $\mathcal{O}$, and let $\phi : E_1 \to E_2$ be an isogeny. Then*

$$\phi \circ [\alpha]_1 = [\alpha]_2 \circ \phi$$

for any $\alpha \in \mathcal{O}$.

Proof. Let $\Lambda_1, \Lambda_2 \subset \mathbb{C}$ be lattices such that $E_1 \simeq E_{\Lambda_1}$ and $E_2 \simeq E_{\Lambda_2}$. We can lift $\phi : E_{\Lambda_1} \to E_{\Lambda_2}$ to a map $\phi_\beta : \mathbb{C} \to \mathbb{C}, z \mapsto \beta z$, i.e. the diagram

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\phi_\beta} & \mathbb{C} \\ {\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ E_{\Lambda_1} & \xrightarrow[\phi]{} & E_{\Lambda_2} \end{array}$$

commutes. Similarly, let $\phi_{\alpha,1}, \phi_{\alpha,2}$ be the lifts of $[\alpha]_1, [\alpha]_2$. Then $\phi_\beta \circ \phi_{\alpha,1} = \phi_{\alpha,2} \circ \phi_\beta$, and $\phi \circ [\alpha]_1 = [\alpha]_2 \circ \phi$. □

4.2 Integrality of j

For a positive integer n, we define $S_n, D_n \subset \mathrm{M}_2(\mathbb{Z})$ as

$$D_n := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z}) \mid ad - bc = n \right\},$$

$$S_n := \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z}) \mid ad = n, d > 0, 0 \le b < d \right\}.$$

We have left $\mathrm{SL}_2(\mathbb{Z})$-action on D_n by multiplication on left, and obviously

$$\#S_n = \sum_{d|n} d = \sigma_1(n).$$

Proposition 4.2.1. *For $n \in \mathbb{N}^+$, S_n is a complete set of orbits of the left $\mathrm{SL}_2(\mathbb{Z})$-action on D_n, i.e. the natural map $S_n \to (_{\mathrm{SL}_2(\mathbb{Z})}\backslash D_n)$ is bijective.*

Proof. At first, we prove that the map is injective. If $\alpha_1 = \begin{pmatrix} a_1 & b_1 \\ 0 & d_1 \end{pmatrix}, \alpha_2 = \begin{pmatrix} a_2 & b_2 \\ 0 & d_2 \end{pmatrix} \in S_n$ such that there exists $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ with $\alpha_1 = \gamma\alpha_2$, then

$$\alpha_1\alpha_2^{-1} = \frac{1}{n} \begin{pmatrix} a_1d_2 & a_2b_1 - a_1b_2 \\ 0 & a_2d_1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Hence $n \mid (a_2b_1 - a_1b_2)$ and $a_1d_2 = a_2d_1 = \pm n$. Notice that $a_1, a_2, d_1, d_2 > 0$ and $a_1d_1 = n$, then $a_1 = a_2, d_1 = d_2$. Since $|b_1 - b_2| < d_1 = d_2$, and $n \mid a_1(b_1 - b_2) < a_1d_1 = n$, so $b_1 = b_2$. That is $\alpha_1 = \alpha_2$.

For $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in D_n$ such that $c \neq 0$, we will prove that there exists $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma\alpha \in S_n$. We can reduct to the case where $c = 0$ and $a, d > 0$. Indeed, let $\frac{a}{c} = \frac{q}{p}$ with $\gcd(p, q) = 1$, and let $s, r \in \mathbb{Z}$ be such that $ps + qr = 1$. We set $\gamma' = \begin{pmatrix} r & s \\ -p & q \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, then $\gamma'\alpha = \begin{pmatrix} ar + cs & br + ds \\ 0 & -bp + dq \end{pmatrix}$. Then we replace α by $\gamma\alpha$ and multiply by $-I$ if necessary. Let $\gamma_m = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ for $m \in \mathbb{Z}$. Then

$\gamma_m\alpha = \begin{pmatrix} a & b+dm \\ 0 & d \end{pmatrix}$. Hence, there exists an $m \in \mathbb{Z}$ such that $0 \leq b+dm < d$, i.e. $\gamma_m\alpha \in S_n$. That proves the surjectivity. □

Definition 4.2.2. *For $n \in \mathbb{N}^+$, the modular polynomial of order n is defined to be*

$$\Phi_n(X) = \prod_{\alpha \in S_n} (X - j \circ \alpha) \in \mathbb{C}[j, X].$$

We will write $s_m(\tau)$, $m = 1, \cdots, \#S_n$ to denote the m-th elementary symmetric function of the $(j \circ \alpha)(\tau)$, i.e.

$$\Phi_n(X) = X^{\#S_n} + \sum_{m=1}^{\#S_n} (-1)^m s_m(\tau) X^{\#S_n - m}.$$

Remark. *(1) For $1 \leq m \leq \#S_n$, $s_m \in \mathbb{C}[j]$, i.e. $s_m \in \mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z}))$ and is holomorphic on $\mathbb{H}$.*

Proof. The set $\{j \circ \alpha \mid \alpha \in S_n\}$ has a $\mathrm{SL}_2(\mathbb{Z})$-action on right:

$$(j \circ \alpha)\gamma := j \circ (\alpha\gamma)$$

for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $\alpha \in S_n$. It is well-defined, since there uniquely exists $\gamma' \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma'\alpha\gamma \in S_n$ by Proposition 4.2.1, which implies that $(j \circ \alpha)\gamma = j \circ (\gamma'\alpha\gamma)$. For any $\gamma \in SL_2(\mathbb{Z})$, γ induces a permutation on $\{j \circ \alpha \mid \alpha \in S_n\}$. Since s_m is the m-th polynomial in terms of $j \circ \alpha, \alpha \in S_n$, so $s_m(\tau) = s_m(\gamma\tau)$ for any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, and $s_m \in \mathcal{A}_0(\mathrm{SL}_2(\mathbb{Z}))$.

For any $\alpha = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in S_n$, since $\alpha(\tau) = (a\tau + b)/d$, so $j \circ \alpha$ is holomorphic on $\mathbb{H}$. Hence s_m is holomorphic on $\mathbb{H}$. □

(2) We have $\deg \Phi_n(X) = \#S_n = \sigma_1(n) = \sum_{d|n} d$.

LEMMA 4.2.3. *Let $n \in \mathbb{N}^+$ and $1 \leq m \leq \#S_n$. Then $s_m \in \mathbb{Z}[j]$. In particular, the Fourier coefficients a_k in the q-expansion*

$$s_m(\tau) = \sum_{k=-N}^{\infty} a_k q^k$$

are integers, and $\Phi_n \in \mathbb{Z}[j, X]$.

Proof. By Lemma 3.1.13, it is sufficient to show that the $a_k \in \mathbb{Z}$. Let $\alpha = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in S_n$, $j(\tau) = q^{-1} \sum_{k=0}^{\infty} c_k q^k$. Then

$$q \circ \alpha(\tau) = \zeta_n^{ab} \cdot q^{a^2/n},$$

$$j \circ \alpha(\tau) = \zeta_n^{-ab} q^{-a^2/n} + \sum_{k=0}^{\infty} c_k \zeta_n^{kab/n} q^{ka^2/n}.$$

The set $\{j \circ \alpha \mid \alpha \in S_n\}$ has a $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$-action on the left:

$$\sigma(j \circ \alpha)(\tau) := \sigma(\zeta_n^{-ab}) q^{-a^2/n} + \sum_{k=0}^{\infty} c_k \sigma(\zeta_n^{kab}) q^{ka^2/n},$$

for $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$. It is well-defined. Indeed, for any $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$, there exists $G(\sigma) \in (\mathbb{Z}/n\mathbb{Z})^*$ such that $\sigma(\zeta_n) = \zeta_n^{G(\sigma)}$. We view $G(\sigma) \in \mathbb{Z}$ such that $0 \leq G(\sigma)b < d$ and let $\beta_\sigma = \begin{pmatrix} a & G(\sigma)b \\ 0 & d \end{pmatrix} \in S_n$. Then obviously $\sigma(j \circ \alpha) = j \circ \beta_\sigma \in \{j \circ \alpha \mid \alpha \in S_n\}$. It is obviously a group action on the left. Since any $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ is a permutation on $\{j \circ \alpha \mid \alpha \in S_n\}$, so $\sigma(s_m) = s_m$. Hence the Fourier coefficients $a_k \in \mathbb{Q}$. Notice that they are also $\mathbb{Z}[\zeta_n]$, so $a_k \in \mathbb{Q} \bigcap \mathbb{Z}[\zeta_n] = \mathbb{Z}$. □

Corollary 4.2.4. *Let $\beta \in \mathrm{M}_2(\mathbb{Z})$ with $\det\beta \in \mathbb{N}^+$. Then $j \circ \beta$ is integral over $\mathbb{Z}[j]$.*

Proof. Let $n = \det\beta$. By Proposition 4.2.1 , there exists $\gamma \in SL_2(\mathbb{Z})$ such that $\gamma\beta = \alpha \in S_n$, so $j \circ \beta = j \circ \alpha$. By Lemma 4.2.3, $\Phi_n(X) \in \mathbb{Z}[j][X]$ and it is monic with $\Phi_n(j \circ \alpha) = 0$. □

We are ready to prove the integrality of $j(E)$ for an elliptic curve E with CM.

Theorem 4.2.5. *Let E be an elliptic over $\mathbb{C}$ with CM. Then $j(E)$ is an algebraic integer.*

Proof. For $n \in \mathbb{N}^+$ which is not a square, by Lemma 4.2.3, there exists $F_n(X, Y) \in \mathbb{Z}[X, Y]$ such that $F_n(j, X) = \Phi_n(X)$. We claim that $F_n(X, X)$ is a non-zero polynomial with leading coefficient -1 or 1. To show our claim, let $\alpha = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in S_n$. Then $n = ad$, $a \neq d$, and

$$j(\tau) - j \circ \alpha(\tau) = \frac{1}{q} + \sum_{k=0}^{\infty} c_k q^k - \frac{1}{\zeta_n^{ab} q^{a^2/n}} - \sum_{k=0}^{\infty} c_k \zeta_n^{kab} q^{ka^2/n}.$$

The leading coefficient is a root of unity, i.e either 1 if $a^2 < n$ or $-\zeta_n^{-ab}$. Let $F_n(j, j) = \prod_{\alpha \in S_n} (j - j \circ \alpha) = b_M j^M + \cdots + b_0 \in \mathbb{Z}[j]$. Notice that b_M is the product of leading coefficients of each $j - j \circ \alpha$, then $b_M \in \mathbb{Z}$ is a root of unity, i.e. $b_M = \pm 1$.

Let $\tau \in \mathbb{H}$ be such that $E \simeq \mathbb{C}/\langle \tau, 1 \rangle$, and $d \in \mathbb{N}^+$ be a square-free integer such that $\mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q} \simeq \mathbb{Q}(\sqrt{-d}) = \mathbb{Q}(\tau)$. Set $K = \mathbb{Q}(\tau)$.

If $\mathrm{End}(E) = \mathcal{O}_K$, there exists $\alpha \in \mathcal{O}_K$ such that $n := \mathcal{N}_{K/\mathbb{Q}}(\alpha) > 0$ is not a square (e.g. $\alpha = \sqrt{-d}$ if $d \neq 1$ and $\alpha = i + 1$ if $d = 1$). Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z})$ be such that

$$\alpha \begin{pmatrix} \tau \\ 1 \end{pmatrix} = \gamma \begin{pmatrix} \tau \\ 1 \end{pmatrix}.$$

Since $\mathcal{N}_{K/Q}(\alpha) = \det\gamma = n$, so $\gamma \in D_n$. By Proposition 4.2.1 and Lemma 4.2.3, $j \circ \gamma$ is integral over $\mathbb{Z}[j]$, i.e. $\Phi_n(j \circ \gamma) = F_n(j, j \circ \gamma) = 0$. Notice that $\gamma(\tau) = \tau$, so $F_n(j(\tau), j(\gamma(\tau))) = F_n(j(\tau), j(\tau)) = 0$. Hence $j(\tau) = j(E)$ is an algebraic integer.

In general case, $\mathrm{End}(E) \simeq \mathcal{O}$ with $\mathcal{O}$ an order in K. Let $\omega_1, \omega_2 \in \mathcal{O}_K$ such that $\omega_1/\omega_2 \in \mathbb{H}$ and $E \simeq \mathbb{C}/\langle \omega_1, \omega_2 \rangle$. Such ω_1 and ω_2 exist, since $E \simeq \mathbb{C}/\langle \tau, 1 \rangle$ with $r\tau^2 + s\tau + t = 0$ and $r, s, t \in \mathbb{Z}$, $r \neq 0$, so $r\tau, r \in \mathcal{O} \subset \mathcal{O}_K$ and we can take $\omega_1 = r\tau, \omega_2 = r$. Let $\tau' \in \mathbb{H}$ such that $\mathcal{O}_K = \langle \tau', 1 \rangle$. Then

$$\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = \beta \begin{pmatrix} \tau' \\ 1 \end{pmatrix},$$

where $\beta = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z})$. Let $n = \det\beta$ which is a positive integer. Then $\beta \in D_n$, $j(E) = j(\omega_1/\omega_2) = j(\beta(\tau'))$ and $F_n(j(\tau'), j(\beta(\tau'))) = 0$. Since $j(\tau')$ integral over $\mathbb{Z}$ from the discussion above, then so is $j(E)$. □

4.3 Group Actions on $\mathcal{ELL}(\mathcal{O})$

Recall that for an order $\mathcal{O}$ in a number field K, an fractional $\mathcal{O}$-ideal, i.e finitely generated sub-$\mathcal{O}$-module of K, $\mathfrak{a}$ is said to be proper if $\mathcal{O} = \{x \in K \mid x\mathfrak{a} \subset \mathfrak{a}\}$, see [20, Page 122]. We say that $\mathfrak{a}$ invertible if it is locally free $\mathcal{O}$-module, see [20, Page 122] or [46, Page 74]. This is equivalent to say that there exist another fractional $\mathcal{O}$-ideal $\mathfrak{b}$ such that $\mathfrak{ab} = \mathcal{O}$, and such $\mathfrak{b}$ is denoted by $\mathfrak{a}^{-1}$.

We will denote by $\mathcal{I}(\mathcal{O})$ the group of invertible fractional $\mathcal{O}$-ideals, and by $\mathrm{Cl}(\mathcal{O})$ is the class group of $\mathcal{O}$, or the Picard group of $\mathcal{O}$, see [46, Definition I.12.5].

LEMMA 4.3.1 ([20], Lemma 7.5)**.** *Let $K = \mathbb{Q}(\tau)$ be a quadratic field, and $aX^2 + bX + c \in \mathbb{Z}[X]$ be the minimal polynomial of τ with $\gcd(a, b, c) = 1$. Then $\mathbb{Z} + \mathbb{Z}\tau$ is a proper fraction ideal for the order $\mathbb{Z} + \mathbb{Z}\tau$ in K.*

Proof. Let $\mathcal{O} = \mathbb{Z} + \mathbb{Z} \cdot a\tau$. Obviously, $\mathcal{O}$ is an order of K and $\mathcal{O}\mathbb{Q} = K$. Set $\mathfrak{a} = \mathbb{Z} + \mathbb{Z}\tau$. For any $\beta = m + n\tau \in K$, $\beta\mathfrak{a} \subset \mathfrak{a}$ if and only if $m, n \in \mathbb{Z}$ and

$$\beta\tau = m\tau + n\tau^2 = -\frac{cn}{a} + (-\frac{bn}{a} + m)\tau \in \mathfrak{a},$$

i.e. a | cn and a | bn. Since $\gcd(a, b, c) = 1$, this is also equivalent to that $m, n \in \mathbb{Z}$ and $a|n$, i.e. $\beta \in \mathcal{O}$. Hence, $\mathcal{O} = \{x \in K \mid x\mathfrak{a} \subset \mathfrak{a}\}$. □

Proposition 4.3.2 ([20], Proposition 7.4)**.** *Let $\mathcal{O}$ be an order in a quadratic field K, and let $\mathfrak{a}$ be a fractional $\mathcal{O}$-ideal. Then $\mathfrak{a}$ is proper if and only if $\mathfrak{a}$ is invertible $\mathcal{O}$-module. In this case, $\mathfrak{a} \subset \mathbb{C}$ is a lattice.*

Proof. Since $\mathfrak{a}$ is an $\mathcal{O}$-module, so $\mathcal{O} \subset \{x \in K \mid x\mathfrak{a} \subset \mathfrak{a}\}$. If $\mathfrak{a}$ is invertible, then for any $\beta \in K$ such that $\beta\mathfrak{a} \subset \mathfrak{a}$, we have $\beta\mathcal{O} = \beta\mathfrak{a}\mathfrak{a}^{-1} \subset \mathfrak{a}\mathfrak{a}^{-1} = \mathcal{O}$, so $\beta \in \mathcal{O}$.

Conversely, if $\mathfrak{a} \subset K$ is a proper fractional $\mathcal{O}$-ideal, then $\mathfrak{a} \subset$ is a lattice, i.e. a free $\mathbb{Z}$-module of rank 2 and $\mathbb{R}\mathfrak{a} = \mathbb{C}$. Indeed, notice that $a\mathfrak{a} \subset \mathcal{O}$ for some $a \in \mathbb{Z}$, and $\mathcal{O}$ is a free $\mathbb{Z}$-module of rank 2, so $\mathfrak{a}$ is a free $\mathbb{Z}$-module whose rank is less than 2. Also, since $\alpha\mathcal{O} \subset \mathfrak{a}$ for any non-zero $\alpha \in \mathfrak{a}$, we conclude that the rank of $\mathfrak{a}$ ia more than 2, hence exactly 2 and $\mathbb{C} = \alpha\mathbb{R}\mathcal{O} \subset \mathbb{R}\mathfrak{a} \subset \mathbb{C}$. Let $\mathfrak{a} = \mathbb{Z}\alpha + \mathbb{Z}\beta = \alpha(\mathbb{Z} + \mathbb{Z}\tau)$, where $\tau = \beta/\alpha$, and $aX^2 + bX + c \in \mathbb{Z}[X]$ be the minimal polynomial of τ with $\gcd(a, b, c) = 1$. Then by Lemma 4.3.1, we have

$$\begin{aligned}\mathcal{O} &= \{x \in K \mid x\mathfrak{a} \subset \mathfrak{a}\} \\ &= \{x \in K \mid x(\mathbb{Z} + \mathbb{Z}\tau) \subset \mathbb{Z} + \mathbb{Z}\tau\} \\ &= \mathbb{Z} + \mathbb{Z} \cdot a\tau.\end{aligned}$$

We consider $\overline{\mathfrak{a}} = \overline{\alpha}(\mathbb{Z} + \mathbb{Z}\overline{\tau})$, it is easy to see that it is also proper fractional $\mathcal{O}$-ideal. We have

$$\begin{aligned}a\mathfrak{a}\overline{\mathfrak{a}} &= a\alpha\overline{\alpha}(\mathbb{Z} + \mathbb{Z}\tau + \mathbb{Z}\overline{\tau} + \mathbb{Z}\tau\overline{\tau}) \\ &= \mathcal{N}_{K/\mathbb{Q}}(\alpha)(a\mathbb{Z} + \mathbb{Z} \cdot a\tau + \mathbb{Z} \cdot (-b) + \mathbb{Z}c) \\ &= \mathcal{N}_{K/\mathbb{Q}}(\mathbb{Z} + \mathbb{Z} \cdot a\tau) \\ &= \mathcal{N}_{K/\mathbb{Q}}(\alpha)\mathcal{O},\end{aligned}$$

which means that $\mathfrak{a}$ is invertible. □

Remark. *(1) In the proof, we have $\mathfrak{a} \subset \mathbb{C}$ is a lattice for an invertible fractional $\mathcal{O}$-module.*

Proposition 4.3.3. *Keep the notations as Proposition 4.3.2. We have a well-defined and bijective map*

$$\begin{aligned} \mathrm{Cl}(\mathcal{O}) &\to \mathcal{ELL}(\mathcal{O}), \\ [\mathfrak{a}] &\mapsto E_{\mathfrak{a}}, \end{aligned}$$

where the invertible ideal $\mathfrak{a} \subset \mathbb{C}$ is viewed as a lattice. In particular, the class number $C_K = \#\mathcal{ELL}(\mathcal{O}_K)$.

Proof. For any invertible ideal $\mathfrak{a}$, we have

$$\begin{aligned} \mathrm{End}(E_{\mathfrak{a}}) &= \{\alpha \in \mathbb{C} \mid \alpha\mathfrak{a} \subset \mathfrak{a}\} \\ &= \{\alpha \in K \mid \alpha\mathfrak{a} \subset \mathfrak{a}\} \\ &= \mathcal{O}, \end{aligned}$$

i.e. $E_{\mathfrak{a}} \in \mathcal{ELL}(\mathcal{O})$. Obviously, $E_{c\mathfrak{a}} \simeq E_{\mathfrak{a}}$ for any $c \in K^*$. Hence the map is well defined.

The map is surjective: every $E \in \mathcal{ELL}(\mathcal{O})$ is isomorphic to E_τ for some $\tau \in K^*$, and $\{\alpha \in K \mid \alpha\Lambda_\tau \subset \Lambda_\tau\} = \mathcal{O}$. Hence $\Lambda_\tau \subset K$ is a proper fractional $\mathcal{O}$-ideal. That proves the surjectivity of the map.

The map is injective: for any invertible $\mathcal{O}$-ideals $\mathfrak{a}, \mathfrak{b}$, $E_{\mathfrak{a}} \simeq E_{\mathfrak{b}}$ if and only if there exist $c \in \mathbb{C}^*$ such that $\mathfrak{a} = c\mathfrak{b}$, that is exactly means that $[\mathfrak{a}] = [\mathfrak{b}] \in \mathrm{Cl}(\mathcal{O})$. □

Via this bijection, the multiplication on $\mathrm{Cl}(\mathcal{O})$ defines an action on $\mathcal{ELL}(\mathcal{O})$.

Proposition&Definition 4.3.4. *Keep the notations as Proposition 4.3.2. We have a* $\mathrm{Cl}(\mathcal{O})$*-action on $\mathcal{ELL}(\mathcal{O})$ as following:*

$$\begin{aligned} \mathrm{Cl}(\mathcal{O}) \times \mathcal{ELL}(\mathcal{O}) &\to \mathcal{ELL}(\mathcal{O}) \\ ([\mathfrak{a}], E_\Lambda) &\mapsto E_{\mathfrak{a}^{-1}\Lambda}. \end{aligned}$$

This action is free and transitive.

Proof. The bijection in Proposition 4.3.3 will induce the following commutative diagram:

$$\begin{array}{ccc} \mathrm{Cl}(\mathcal{O}) \times \mathcal{ELL}(\mathcal{O}) & \longrightarrow & \mathcal{ELL}(\mathcal{O}) \\ \downarrow & & \downarrow \\ \mathrm{Cl}(\mathcal{O}) \times \mathrm{Cl}(\mathcal{O}) & \longrightarrow & \mathrm{Cl}(\mathcal{O}), \end{array}$$

$$(\mathfrak{a}, \mathfrak{b}) \mapsto \mathfrak{a}^{-1}\mathfrak{b}.$$

Obviously, the $\mathrm{Cl}(\mathcal{O})$-action on itself is free and transitively, then is the one on $\mathcal{ELL}(\mathcal{O})$. □

Before going further, we set

$$\mathcal{ELL}_{\overline{\mathbb{Q}}}(\mathcal{O}) := \{\text{elliptic curve } E/\overline{\mathbb{Q}} \text{ with } \mathrm{End}(E) \simeq \mathcal{O}\}/\simeq_{\mathbb{Q}},$$

here $\simeq_{\mathbb{C}}$ means taking isomorphic classes over $\overline{\mathbb{Q}}$.

Proposition 4.3.5. *Keep the notations as Proposition 4.3.2. Then the natural map*

$$\mathcal{ELL}_{\overline{\mathbb{Q}}}(\mathcal{O}) \to \mathcal{ELL}(\mathcal{O})$$

is a bijection.

Proof. For an elliptic curve E with CM, by Proposition 4.2.5, $j(E) \in \overline{\mathbb{Q}}$. Then there exists an elliptic curve $E'/\overline{\mathbb{Q}}$ such that $j(E') = j(E)$ and $E' \simeq E$ over $\mathbb{C}$, see [60, Proposition III.1.4 (b),(c)]. This proves the surjectivity of the map. The injectivity also comes from [60, Proposition III.1.4 (b)]. □

Without confusion, we will always identify these two set. Hence $\mathcal{ELL}(\mathcal{O})$ has a $G_{\mathbb{Q}}$-action on left via this bijection.

LEMMA 4.3.6. *Let $\mathcal{O}$ be an order in a number field K, $\mathfrak{a}$ a invertible fractional ideal of $\mathcal{O}$ and M a torsion free $\mathcal{O}$-module. Then the natural map*

$$\begin{aligned} \phi : \mathfrak{a}^{-1}M &\to \mathrm{Hom}_{\mathcal{O}}(\mathfrak{a}, M) \\ x &\mapsto (\phi_x : \alpha \mapsto \alpha x), \end{aligned}$$

is an isomorphism of R-modules.

Proof. Recall that $\mathfrak{a}^{-1} = \{a \in \mathrm{Frac}(R) \mid a\mathfrak{a} \subset R\}$ and $\mathfrak{a}^{-1}M = \{a \in \mathrm{Frac}(R) \otimes_R M \mid \mathfrak{a}x \subset M\}$. Then ϕ is a well-defined morphism of R-modules.

If $x, y \in \mathfrak{a}^{-1}M$ such that $\alpha x = \alpha y$ for any $\alpha \in \mathfrak{a}$, we take one $\alpha \neq 0$. Then there exist $a \in R$ such that $a\alpha \in R$. Hence $a\alpha x = a\alpha y$ will imply that $x = y$, since $\mathfrak{a}^{-1}M$ is also a torsion free R-module. This prove injectivity.

For surjectivity, let $\varphi \in \mathrm{Hom}_R(\mathfrak{a}, M)$, and $x = \varphi(\alpha)/\alpha \in \mathrm{Frac}(R) \otimes_R M$ for some $0 \neq \alpha \in \mathfrak{a}$. Then $\phi(\beta) = \beta x$ for any $\beta \in \mathfrak{a}$. Indeed, there exists $a \in R$ such that $a\beta/\alpha \in R$, so

$$\varphi(\beta) = \varphi(\frac{\beta}{\alpha} \cdot \alpha) = \frac{\varphi(a\frac{\beta}{\alpha} \cdot \alpha)}{a} = \frac{\beta}{\alpha} \cdot \varphi(\alpha) = \beta x.$$

Hence we have $x \in \mathfrak{a}^{-1}M$ and $\varphi(\beta) = \beta x$ for any $\beta \in \mathfrak{a}$. □

Proposition 4.3.7. *Keep the notations as Proposition 4.3.2. Let $E \in \mathcal{ELL}_{\overline{\mathbb{Q}}}(\mathcal{O}) \to \mathcal{ELL}(\mathcal{O})$, $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ and $\sigma \in G_{\mathbb{Q}}$. Then*

$$\sigma([\mathfrak{a}] \cdot E) = [\sigma\mathfrak{a}] \cdot \sigma(E) \in \mathcal{ELL}(\mathcal{O}).$$

Proof. See proof of [59, Proposition 2.5]. Here we give a sketch of the proof. We have $E \simeq E_\Lambda$ for some lattice Λ. For an invertible $\mathcal{O}$-ideal $\mathfrak{a}$, $[\mathfrak{a}] \cdot E = E_{\mathfrak{a}^{-1}\Lambda}$ and we have an exact sequence:

$$\mathcal{O}^m \xrightarrow{A} \mathcal{O}^m \longrightarrow \mathfrak{a} \longrightarrow 0\ ,$$

where A is an $m \times n$ matrix with coefficients in $\mathcal{O}$. Furthermore, we have the following diagram:

$$\begin{array}{ccccccccc} & & 0 & & 0 & & 0 & & \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathfrak{a}, \Lambda) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathfrak{a}, \mathbb{C}) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathfrak{a}, E) & & \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathcal{O}^n, \Lambda) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathcal{O}^n, \mathbb{C}) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathcal{O}^n, E) & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathcal{O}^m, \Lambda) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathcal{O}^m, \mathbb{C}) & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathcal{O}^m, E) & \longrightarrow & 0 \end{array}$$

For any $\mathcal{O}$-module M, we have $\mathrm{Hom}(\mathcal{O}, M) \simeq M^n$. By Lemma 4.3.6 and $K \otimes_{\mathcal{O}} \mathbb{C} = \mathbb{C}$, we get

$$\mathrm{Hom}_{\mathcal{O}}(\mathfrak{a}, \Lambda) = \mathfrak{a}^{-1}\Lambda,$$

$$\mathrm{Hom}_{\mathcal{O}}(\mathfrak{a}, \mathbb{C}) = \mathfrak{a}^{-1}\mathbb{C} = \mathbb{C}.$$

Then the diagram becomes

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \mathfrak{a}^{-1}\Lambda & \longrightarrow & \mathbb{C} & \longrightarrow & \mathrm{Hom}_{\mathcal{O}}(\mathfrak{a}, E) & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \Lambda^n & \longrightarrow & \mathbb{C}^n & \longrightarrow & E^n & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle {}^tA} & & \downarrow{\scriptstyle {}^tA} & & \downarrow{\scriptstyle {}^tA} & & \\
0 & \longrightarrow & \Lambda^m & \longrightarrow & \mathbb{C}^m & \longrightarrow & E^m & \longrightarrow & 0
\end{array}$$

Here ${}^tA$ is the transpose of the matrix of A. Hence we see that $\mathrm{Hom}_{\mathcal{O}}(\mathfrak{a}, E) \simeq \mathrm{Ker}(E^n \overset{{}^tA}{\to} E^m)$, which is an algebraic group defined $\overline{\mathbb{Q}}$.

By the snake lemma, we have

$$0 \to \mathfrak{a}^{-1}\Lambda \to \mathbb{C} \to \mathrm{Ker}(E^n \overset{{}^tA}{\to} E^m) \to \Lambda^n / {}^tA\Lambda^m.$$

We view E as an elliptic curve over $\overline{\mathbb{Q}}$ and $E^n \to E^m$ a morphism over $\overline{\mathbb{Q}}$.

On the other hand, $\Lambda^n / {}^tA\Lambda^m$ is discrete and $\mathbb{C}/\mathfrak{a}^{-1}\Lambda = [\mathfrak{a}] \cdot E$ is connected. Hence

$$[\mathfrak{a}] \cdot E = \text{identity component of } \mathrm{Ker}(E^n \overset{{}^tA}{\to} E^m).$$

Acted by $\sigma \in G_{\mathbb{Q}}$, we have

$$\begin{aligned}
\sigma([\mathfrak{a}] \cdot E) &= \sigma(\text{identity component of } \mathrm{Ker}(E^n \overset{{}^tA}{\to} E^m)) \\
&= \text{identity component of } \mathrm{Ker}(\sigma(E^n) \overset{\sigma({}^tA)}{\to} \sigma(E^m)) \\
&= \sigma\mathfrak{a} \cdot \sigma(E).
\end{aligned}$$

□

Proposition 4.3.8. *Keep the notations as Proposition 4.3.2. Then there exists a homomorphism*

$$F : G_K \to \mathrm{Cl}(\mathcal{O})$$

uniquely characterized by the condition: for any $\sigma \in G_K$ and $E \in \mathcal{ELL}(\mathcal{O})$,

$$\sigma(E) = F(\sigma) \cdot E.$$

Proof. Since the action is free and transitive, so for a fixed element $E \in \mathcal{ELL}(\mathcal{O})$ and $\sigma \in G_K$, there uniquely exists $[\mathfrak{a}] \in \mathrm{Cl}(\mathcal{O})$ such that $\sigma(E) = [\mathfrak{a}] \cdot E$. Then we define $F(\sigma) := [\mathfrak{a}]$. This map is independent of the choice of E. Consider another

element $E' \in \mathcal{ELL}(\mathcal{O})$, there exists $[\mathfrak{b}] \in \mathrm{Cl}(\mathcal{O})$ such that $E' = [\mathfrak{b}] \cdot E$. Then by Proposition 4.3.7 for any $\sigma \in G_K$ and $F(\sigma)$ defined as above, we have

$$\begin{aligned}\sigma(E') &= \sigma([\mathfrak{b}] \cdot E) \\ &= [\sigma\mathfrak{b}] \cdot \sigma(E) \\ &= [\mathfrak{b}] \cdot F(\sigma) \cdot E \\ &= F(\sigma) \cdot E'.\end{aligned}$$

It remains to show that F is a homomorphism. For any $\sigma, \tau \in G_K$, and E as above,

$$F(\sigma\tau) \cdot E = \sigma(\tau(E)) = F(\sigma)(F(\tau) \cdot E) = (F(\sigma)F(\tau)) \cdot E,$$

so $F(\sigma\tau) = F(\sigma)F(\tau)$ by the uniqueness. □

Remark. *(1) The map $F : G_K \to \mathrm{Cl}(\mathcal{O})$ is also characterized by the condition: for any $\sigma \in G_K$, and lattice $\Lambda \subset \mathbb{C}$ such that E_Λ is with CM,*

$$\sigma(j(\Lambda)) = j(F(\sigma)^{-1}\Lambda).$$

Proof. Notice that for any two lattices Λ and Λ', $E_\Lambda \simeq E_{\Lambda'}$ if and only if $j(\Lambda) = j(\Lambda')$, also we have $j(\sigma(E_\Lambda)) = \sigma(j(E_\Lambda))$, then $\sigma(E_\Lambda) = F(\sigma) \cdot E_\Lambda = E_{F(\sigma)^{-1}\Lambda}$ if and only if $\sigma(j(\Lambda)) = j(F(\sigma)^{-1}\Lambda)$. □

(2) We know that the map in the proposition will induce a homomorphism

$$\mathrm{Gal}(K^{ab}/K) \to \mathrm{Cl}(\mathcal{O}),$$

it is natural to guess that it induces the inverse of Artin maps in some cases.

4.4 The Ring Class Fields for Imaginary Quadratic Fields

Recall that, for a number field K and an order $\mathcal{O} \subset K$ with $\mathfrak{f}$ as its conductor, we set $\mathcal{I}_\mathfrak{f}(K)$ as the group of ideals which are coprime to $\mathfrak{f}$. The ring class group of $\mathcal{O}$ is the class field for $\mathcal{I}_\mathfrak{f}(K)$, see [18, Page 53]. We have

$$\mathcal{I}_\mathfrak{f}(K) \simeq \mathcal{I}(\mathcal{O}),$$

$$\mathrm{Cl}_\mathfrak{f}(K) \simeq \mathrm{Cl}(\mathcal{O}),$$

where $\mathcal{I}(\mathcal{O})$ is the group of invertible fractional $\mathcal{O}$-ideals, and $\mathrm{Cl}_\mathfrak{f}(K)$ is the ray class group with respect to the modulus $\mathfrak{f}$. By global class field theory, see [46, Chapter VI], there exists a unique Abelian extension L/K such that the Artin map

$$\left(\frac{\cdot}{L/K}\right) : \mathrm{Cl}_\mathfrak{f}(K) \to \mathrm{Gal}(L/K)$$

is an isomorphism. Such extension L/K is call the ring class field of the order $\mathcal{O}$.

Theorem 4.4.1 (First Main Theorem, [60], Theorem 4.3)**.** *Let K be imaginary quadratic field, and $\mathcal{O} \subset K$ an order with conduct $\mathfrak{f}$. Let $E/\mathbb{C}$ be an elliptic curve representing an isomorphism class in $\mathcal{ELL}(\mathcal{O})$. Then the following statements hold:*

(1) The field $L = K(j(E))$ is the ring class group of $\mathcal{O}$.

(2) *We have*

$$[\mathbb{Q}(j(E)) : \mathbb{Q}] = [K(j(E)) : K] = h,$$

where $h = \#\mathrm{Cl}(\mathcal{O})$.

(3) *Let* $E_1, \cdots, E_h$ *be a complete set of representatives for* $\mathcal{ELL}(\mathcal{O})$. *Then* $j(E_1), \cdots, j(E_h)$ *is a complete set of* G_K*-conjugates for* $j(E)$.

(4) *For every prime ideal of* $\mathcal{O}_K$ *with* $\mathfrak{p} \nmid \mathfrak{f}$, *we have*

$$\left(\frac{\mathfrak{p}}{L/K}\right)(j(E)) = j([\mathfrak{p}] \cdot E).$$

More generally, for any $[\mathfrak{a}] \in \mathrm{Cl}_{\mathfrak{f}}(K)$, *we have*

$$\left(\frac{\mathfrak{a}}{L/K}\right)(j(E)) = j([\mathfrak{a}] \cdot E).$$

Chapter 5

The Difference of Singular Moduli

In this chapter, we prove the main result of Part 2, Theorem 5.1.1. It gives an explicit lower bound of norm difference of two singular moduli. In particular, this implies that the difference of two singular moduli is not a unit when their discriminants are large.

5.1 Main theorem and general setting

Recall that a point $\tau \in \mathbb{H}$ a CM-point if its corresponding elliptic curve E_τ over $\mathbb{C}$ is a CM elliptic curve. We have known that $\tau \in \mathbb{H}$ is CM if and only if τ is algebraic number of degree 2, see the remark of Definition 4.1.2. We call $j(\tau)$ a singular modulus if $\tau \in \mathbb{H}$ is CM. We have known that every singular modulus is an algebraic integer, see Theorem 4.2.5. We call $j(\tau)$ singular unit if it is a singular modulus and an algebraic unit.

As we said in the introduction, the main result of this part is the following theorem, and we will explain notations afterwards:

Theorem 5.1.1. *Let α, x be two singular moduli of discriminants Δ_α, Δ respectively, and $K = \mathbb{Q}(\alpha, x)$.*

(1) If $\Delta_\alpha \neq -3, -4$ and $|\Delta| \geq \max\{e^{3.12}(\mathcal{C}(\Delta_\alpha)|\Delta_\alpha|^4 e^{h(\alpha)})^3, 10^{15} \cdot \mathcal{C}(\Delta_\alpha)^6\}$, then

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)| > \frac{|\Delta|^{1/2}}{2};$$

(2) If $\Delta_\alpha = -4$, i.e. $\alpha = 1728$, and $|\Delta| \geq 10^{15}$, then

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| > \frac{|\Delta|^{1/2}}{2};$$

(3) If $\Delta_\alpha = -3$, i.e. $\alpha = 0$, and $|\Delta| \geq 10^{15}$, then

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x)| > \frac{|\Delta|^{1/2}}{20}.$$

In this theorem, the bound is effective. Next, we will explain notations.
For a number field K, $x \in K$, we denote $\mathcal{N}_{K/\mathbb{Q}}(x)$ the absolute norm of x.
Let Δ be a negative integer satisfying $\Delta \equiv 0, 1 \bmod 4$ and

$$\mathcal{O}_\Delta = \mathbb{Z}[(\Delta + \sqrt{\Delta})/2],$$

the imaginary quadratic order of discriminant Δ. We suppose that D is the discriminant of $\mathbb{Q}(\sqrt{\Delta})$, and $f = [\mathcal{O}_D : \mathcal{O}_\Delta]$ is the conductor of $\mathcal{O}_\Delta$, so we have $\Delta = f^2 D$. We also denote the class number of the order $\mathcal{O}_\Delta$ by $\mathcal{C}(\Delta)$, since h is used for height of an algebraic number. For further uses, we define the modified conductor $\tilde{f}$ of $\mathcal{O}_\Delta$ by

$$\tilde{f} = \begin{cases} f, & D \equiv 1 \bmod 4, \\ 2f, & D \equiv 0 \bmod 4. \end{cases}$$

On the other hand, let $\mathcal{F}$ be the standard fundamental domain in the Poincaré plane, that is, the open hyperbolic triangle with vertices ζ_3, ζ_6 , and $i\infty$, together with the geodesics $[i, \zeta_6]$ and $[\zeta_6, i\infty)$; here $\zeta_3 = e^{2\pi i/3}$ and $\zeta_6 = e^{\pi i/3}$. Then the Klein j-invariant $j : \mathbb{H} \to \mathbb{C}$ induces a bijection

$$j : \mathcal{F} \to \mathbb{C}.$$

For each CM-point τ in the standard fundamental domain $\mathcal{F}$, i.e. quadratic imaginary number in $\mathcal{F}$, the discriminant Δ_τ of τ is defined to be the discriminant of the primitive polynomial of τ over $\mathbb{Z}$, it is also the discriminant of the order $\mathrm{End}(\mathbb{C}/\Lambda_\tau)$, i.e. $\mathrm{End}(\mathbb{C}/\Lambda_\tau) = \mathcal{O}_{\Delta_\tau}$, where Λ_τ is the lattice generated by 1 and τ. Since the j-invariant $j : \mathcal{F} \to \mathbb{C}$ is a bijection, we call Δ_τ the discriminant of $\alpha = j(\tau)$, also denoted by Δ_α.

By classical CM-theory, we know that $\mathbb{Q}(\sqrt{\Delta_\tau}, j(\tau))$ is the ring class field of $\mathcal{O}_{\Delta_\tau}$, hence $\mathbb{Q}(\sqrt{\Delta_\tau}, j(\tau))/\mathbb{Q}(\sqrt{\Delta_\tau})$ is Galois and $\mathcal{C}(\Delta_\tau) = [\mathbb{Q}(\sqrt{\Delta_\tau}, j(\tau)) : \mathbb{Q}(\sqrt{\Delta_\tau})] = [\mathbb{Q}(j(\tau)) : \mathbb{Q}]$.

For $n \in \mathbb{N}^+$, we denote

$$\omega(n) = \sum_{p|n} 1, \quad \sigma_0(n) = \sum_{d|n} 1, \quad \sigma_1(n) = \sum_{d|n} d.$$

5.2 An Estimate for $C_\varepsilon(\tau, \Delta)$

For each $\tau \in \mathcal{F}$ and $\varepsilon \in (0, 1/2)$, we define

$$S_\varepsilon(\tau, \Delta) = \{z \in \mathbb{H} \mid z \text{ is a imaginary quadratic number of discriminant } \Delta \text{ and } |z - \tau| < \varepsilon\},$$

$$\mathcal{C}_\varepsilon(\tau, \Delta) = \#S_\varepsilon(\tau, \Delta),$$

here # means the cardinality of a set.

Let S_Δ be the set of primitive positive definite forms of discriminant Δ, that is, a quadratic form $ax^2 + bxy + cy^2 \in S_\Delta$ if $a, b, c \in \mathbb{Z}$ and

$$a > 0, \ \gcd(a, b, c) = 1, \ \Delta = b^2 - 4ac < 0$$

For $ax^2 + bxy + cy^2 \in S_\Delta$, we set

$$\tau(a, b, c) = \frac{b + \sqrt{\Delta}}{2a}.$$

Notice that $\sqrt{\Delta} = i\sqrt{|\Delta|}$, then the map $ax^2 + bxy + cy^2 \mapsto \tau(a, b, c)$ defines a bijection from S_Δ to the set of imaginary number on $\mathbb{H}$ of discriminant Δ.

We will prove the following theorem and corollary:

Theorem 5.2.1. *Let $\tau \in \mathcal{F}$ and $\varepsilon \in (0, 1/4)$, then*

$$\mathcal{C}_\varepsilon(\tau, \Delta) \le F\left(\frac{48 + 16\sqrt{3}}{3}\frac{\sigma_1(\tilde{f})}{\tilde{f}}|\Delta|^{1/2}\varepsilon^2 + \frac{12 + 4\sqrt{3}}{3}|\Delta|^{1/2}\varepsilon + \frac{8|\Delta|^{1/4}}{(\sqrt{3} - 1)^{1/2}}\sigma_0(\tilde{f})\varepsilon + 2\right),$$

where

$$F = F(\Delta) = \max\{2^{\omega(a)} \mid a \le |\Delta|^{1/2}\}. \tag{5.1}$$

Corollary 5.2.2. *In the set-up of Theorem 5.2.1, assume that* $|\Delta| \geq 10^{14}$. *Then*

$$\mathcal{C}_\varepsilon(\tau, \Delta) \leq F\left(46.488|\Delta|^{1/2}\varepsilon^2 \log\log|\Delta|^{1/2} + 7.752|\Delta|^{1/2}\varepsilon + 2\right)$$

5.2.1 Some lemmas

We say that $d \in \mathbb{Z}$ is a quadratic divisor of $n \in \mathbb{Z}$ if $d^2 \mid n$. We denote by $\gcd_2(m, n)$ the greatest common quadratic divisor of integers m and n.

We will use some lemmas from [7], for the reader's convenience, we restate them here:

LEMMA 5.2.3 ([7], Lemma 2.4)**.** *Let a be a positive integer and Δ a non-zero integer. Then the set of $b \in \mathbb{Z}$ satisfying $b^2 \equiv \Delta \bmod a$ consists of at most $2^{\omega(a/\gcd(a,\Delta))+1}$ residue classes modulo $a/\gcd_2(a, \Delta)$, where.*

LEMMA 5.2.4 ([7], Lemma 2.5)**.** *Let $\alpha, \beta \in \mathbb{R}$ be such that $\alpha < \beta$, and m a positive integer. Then every residue class modulo m has at most $(\beta - \alpha)/m + 1$ elements in the interval $[\alpha, \beta]$.*

LEMMA 5.2.5. *Let $\tau \in \mathcal{F}$, and $\varepsilon \in (0, 1/4)$, and let $ax^2 + bxy + cy^2 \in S_\Delta$ be such that $|\tau(a, b, c) - \tau| < \varepsilon$. Then*

$$\frac{|\Delta|^{1/2}}{2(\operatorname{Im}\tau + \varepsilon)} < a < \frac{|\Delta|^{1/2}}{2(\operatorname{Im}\tau - \varepsilon)}, \tag{5.2}$$

$$2a(\operatorname{Re}\tau - \varepsilon) < b < 2a(\operatorname{Re}\tau + \varepsilon). \tag{5.3}$$

Proof. Set $z = \tau(a, b, c)$, then from $|z - \tau| < \varepsilon$, we have

$$|\operatorname{Im} z - \operatorname{Im}\tau| < \varepsilon, \ \ |\operatorname{Re} z - \operatorname{Re}\tau| < \varepsilon,$$

that is,

$$\left|\frac{|\Delta|^{1/2}}{2a} - \operatorname{Im}\tau\right| < \varepsilon, \ \ \left|\frac{b}{2a} - \operatorname{Re}\tau\right| < \varepsilon,$$

so we have 5.2 and 5.3. □

5.2.2 Proof of Theorem 5.2.1

Set

$$I = \left(\frac{|\Delta|^{1/2}}{2(\operatorname{Im}\tau + \varepsilon)}, \frac{|\Delta|^{1/2}}{2(\operatorname{Im}\tau - \varepsilon)}\right),$$

$$\tau(a, b, c) = \frac{b + \sqrt{\Delta}}{2a}.$$

By Lemma 5.2.5, if $\tau(a, b, c) \in S_\varepsilon(\tau, \Delta)$, then $a \in I$ and $b \in (2a(\operatorname{Re}\tau - \varepsilon), 2a(\operatorname{Re}\tau + \varepsilon))$.

For a fixed a, by Lemma 5.2.3 and Lemma 5.2.4 and $\omega(a/\gcd(a, \Delta)) \leq \omega(a)$, there are at most $(4\varepsilon \gcd_2(a, \Delta) + 1) \cdot 2^{\omega(a)+1}$ possible b's. Since $\varepsilon < 1/4, \operatorname{Im}\tau \geq \sqrt{3}/2$, then $\frac{|\Delta|^{1/2}}{2(\operatorname{Im}\tau - \varepsilon)} \leq |\Delta|^{1/2}$. Hence

$$\begin{aligned}\mathcal{C}_\varepsilon(\tau, \Delta) &\leq 8\varepsilon \sum_{a \in I \cap \mathbb{Z}} \gcd_2(a, \Delta) \cdot 2^{\omega(a)} + 2 \sum_{a \in I \cap \mathbb{Z}} 2^{\omega(a)} \\ &\leq 8\varepsilon F \sum_{a \in I \cap \mathbb{Z}} \gcd_2(a, \Delta) + 2F\#(I \cap \mathbb{Z}).\end{aligned}$$

Note that

$$\sum_{a\in I\cap\mathbb{Z}} \gcd{}_2(a,\Delta) \le \sum_{d^2|\Delta} d\cdot\#(I\cap d^2\mathbb{Z}),$$

and the length of I is

$$\begin{aligned}\frac{|\Delta|^{1/2}}{2(\operatorname{Im}\tau-\varepsilon)}-\frac{|\Delta|^{1/2}}{2(\operatorname{Im}\tau+\varepsilon)} &= |\Delta|^{1/2}\frac{\varepsilon}{(\operatorname{Im}\tau+\varepsilon)(\operatorname{Im}\tau-\varepsilon)}\\ &\le |\Delta|^{1/2}\frac{\varepsilon}{\sqrt{3}/2(\sqrt{3}/2-1/2)}\\ &= \frac{6+2\sqrt{3}}{3}|\Delta|^{1/2}\varepsilon.\end{aligned}$$

When $d > \frac{|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}}$, we have $\frac{|\Delta|^{1/2}}{2(\operatorname{Im}\tau-\varepsilon)} < d^2$. Combine this with Lemma 5.2.4, we have

$$\#(I\cap d^2\mathbb{Z}) \le \begin{cases}\frac{6+2\sqrt{3}}{3}\frac{|\Delta|^{1/2}}{d^2}\varepsilon+1 & d\le\frac{|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}},\\ 0 & d>\frac{|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}}.\end{cases}$$

Since $\Delta/\tilde{f}^2$ is square-free, so for a positive integer d, $d^2\mid\Delta$ if and only if $d\mid\tilde{f}$, hence

$$\begin{aligned}\sum_{d^2|\Delta} d\cdot\#(I\cap d^2\mathbb{Z}) &\le \sum_{\substack{d|\tilde{f}\\ d\le\frac{|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}}}} d\left(\frac{6+2\sqrt{3}}{3}\frac{|\Delta|^{1/2}}{d^2}\varepsilon+1\right)\\ &\le \frac{6+2\sqrt{3}}{3}|\Delta|^{1/2}\varepsilon\sum_{d|\tilde{f}}1/d+\sum_{\substack{d|\tilde{f}\\ d\le\frac{|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}}}} d\\ &\le \frac{6+2\sqrt{3}}{3}\frac{\sigma_1(\tilde{f})}{\tilde{f}}|\Delta|^{1/2}\varepsilon+\frac{|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}}\sigma_0(\tilde{f}).\end{aligned}$$

Again, by Lemma 5.2.4, we have

$$\#(I\cap\mathbb{Z}) \le \frac{6+2\sqrt{3}}{3}|\Delta|^{1/2}\varepsilon+1.$$

Hence,

$$\begin{aligned}\mathcal{C}_\varepsilon(\tau,\Delta) &\le 8\varepsilon F\left(\frac{6+2\sqrt{3}}{3}\frac{\sigma_1(\tilde{f})}{\tilde{f}}|\Delta|^{1/2}\varepsilon+\frac{|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}}\sigma_0(\tilde{f})\right)+2F\left(\frac{6+2\sqrt{3}}{3}|\Delta|^{1/2}\varepsilon+1\right)\\ &\le F\left(\frac{48+16\sqrt{3}}{3}\frac{\sigma_1(\tilde{f})}{\tilde{f}}|\Delta|^{1/2}\varepsilon^2+\frac{12+4\sqrt{3}}{3}|\Delta|^{1/2}\varepsilon+\frac{8|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}}\sigma_0(\tilde{f})\varepsilon+2\right).\end{aligned}$$

5.2.3 Proof of Corollary 5.2.2

The following lemma estimate $\sigma_0(\tilde{f})$ and $\sigma_1(\tilde{f})$ in terms of $|\Delta|$:

LEMMA 5.2.6 ([7], Lemma 2.8). *For $|\Delta|\ge 10^{14}$, we have*

$$\sigma_0(\tilde{f}) \le |\Delta|^{0.192},$$

$$\sigma_1(\tilde{f})/\tilde{f} \le 1.842\log\log|\Delta|^{1/2}.$$

With this lemma, we have

$$\frac{48+16\sqrt{3}}{3}\frac{\sigma_1(\tilde{f})}{\tilde{f}} \leq \frac{48+16\sqrt{3}}{3}\cdot 1.842\log\log|\Delta| \leq 46.488\log\log|\Delta|,$$

$$\frac{8|\Delta|^{1/4}}{(\sqrt{3}-1)^{1/2}}\sigma_0(\tilde{f}) \leq \frac{8}{(\sqrt{3}-1)^{1/2}}|\Delta|^{0.442} \leq \frac{8}{(\sqrt{3}-1)^{1/2}\cdot 10^{0.812}}|\Delta|^{1/2} \leq 1.442|\Delta|^{1/2}.$$

$$\frac{12+4\sqrt{3}}{3}+1.442 \leq 7.752$$

With these bounds and Theorem 5.2.1, we have Corollary 5.2.2.

5.3 An Upper Bound for the Height of the difference of Singular Moduli

Let $\alpha = j(\tau), x = j(z)$ be two different singular moduli with $\tau, z \in \mathcal{F}$, and Δ_α, $\Delta = \Delta_x$ be their discriminants respectively. Let $K = \mathbb{Q}(x-\alpha)$, $d = [K:\mathbb{Q}]$, then we have $K = \mathbb{Q}(\alpha, x)$, see [25, Theorem 4.1]. Hence we can assume that $d = s\mathcal{C}(\Delta_\tau)$, where Δ_α is the discriminant of τ and $s = [K:\mathbb{Q}(\alpha)]$. Notice that $\mathbb{Q}(\alpha)/\mathbb{Q}$ and $\mathbb{Q}(x)/\mathbb{Q}$ are Galois, so is $K/\mathbb{Q}$. We suppose that $\mathrm{Gal}(K/\mathbb{Q}) = \{\sigma_1, \cdots, \sigma_d\}$. For each k, set $\alpha_k = \sigma_k(\alpha) = j(\tau_k)$ with $\tau_k \in \mathcal{F}$, and set $x_k = \sigma_k(x) = j(z_k)$ such that $z_k \in \mathbb{H}$ is the nearest point to τ_k among $\mathrm{SL}_2(\mathbb{Z})z_k$ with respect to the absolute norm. Then $\alpha_k \neq x_k$ for each k, and we have

$$\mathrm{h}(x-\alpha) = \mathrm{h}((x-\alpha)^{-1}) = \frac{1}{d}\sum_{k=1}^{d}\log^+|x_k-\alpha_k|^{-1} + \frac{1}{d}\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|, \qquad (5.4)$$

where $\log^+(\cdot) = \max\{1, \cdot\}$.

In this section, we are going to prove that following theorem and corollary:

Theorem 5.3.1. *Let $\alpha = j(\tau), x = j(z)$ be two different singular moduli with $\tau, z \in \mathcal{F}$, and Δ_α, $\Delta = \Delta_x$ be their discriminants respectively. Let $K = \mathbb{Q}(x-\alpha)$, $d = [K:\mathbb{Q}]$,*

(1) if $\tau \neq i, \zeta_6$ and $0 < \varepsilon < \min\{\frac{1}{3|\Delta_\alpha|^2}, 10^{-8}\}$, then

$$\mathrm{h}(x-\alpha) \leq \sum_{1\leq k\leq \mathcal{C}(\Delta_\alpha)} 4\frac{\mathcal{C}_\varepsilon(\tau_k,\Delta)}{d}\log(\max\{|\Delta|,|\Delta_\alpha|\}) + \log(\varepsilon^{-1}) + 2\log|\Delta_\alpha| - 7.783 + \frac{1}{d}\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|;$$

(2) if $\tau = i$ and $0 < \varepsilon \leq 7\cdot 10^{-3}$, then

$$\mathrm{h}(x-1728) \leq 2\frac{\mathcal{C}_\varepsilon(i,\Delta)}{\mathcal{C}(\Delta)}\log|\Delta| + 2\log\varepsilon^{-1} - 9.9 + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)|.$$

We don't discuss the case where $\tau = \zeta_6$, since the bound for this case in the following corollary can be get directly from [7].

Corollary 5.3.2. *In the setup of Theorem 5.3.1, assume that $|\Delta| \geq 10^{14}$,*

(1) if $\tau \neq i, \zeta_6$, then

$$\mathrm{h}(x-\alpha) \leq \frac{8A\mathcal{C}(\Delta_\alpha)}{d} + \log(\frac{A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}}{d}) + 4\log|\Delta_\alpha| + 0.33 + \frac{1}{d}\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|;$$

(2) if $\tau = i$, then

$$\mathrm{h}(x-1728) \leq \frac{4A}{\mathcal{C}(\Delta)} + 2\log\frac{A|\Delta|^{1/2}}{\mathcal{C}(\Delta)} - 2.68 + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)|;$$

(3) if $\tau = \zeta_6$, then

$$\mathrm{h}(x) \leq \frac{12A}{\mathcal{C}(\Delta)} + 3\log\frac{A|\Delta|^{1/2}}{\mathcal{C}(\Delta)} - 3.77 + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x)|,$$

where $A = F\log\max\{|\Delta|, |\Delta_\tau|\}$ and F is defined in 5.1.

5.3.1 Proof of Theorem 5.3.1

The following lemmas and theorems are needed.

LEMMA 5.3.3. *In the set-up of Theorem 5.3.1,*

1) if $\mathrm{Im}\,\tau \geq 1.3$, then there exist $z' \in \mathbb{H}$ with $x = j(z')$ such that

$$|x-\alpha| \geq e^{2.6\pi}\min\{0.4|z'-\tau|, 0.04\};$$

2) if $\mathrm{Im}\,\tau \leq 1.3$ and $\tau \neq i, \zeta_6$, then there exist $z' \in \mathbb{H}$ with $x = j(z')$ such that

$$|x-\alpha| \geq \min\{5\cdot 10^{-7}, 800|\Delta_\alpha|^{-4}, 2400|\Delta_\alpha|^{-2}|z'-\tau|\}.$$

Proof. Combine Proposition 4.1 and Proposition 4.2 in [6]. □

Theorem 5.3.4 ([6] Theorem 1.1)**.** *In the set-up of Theorem 5.3.1, we have*

$$|x-\alpha| \geq 800\max\{|\Delta|, |\Delta_\alpha|\}^{-4}.$$

LEMMA 5.3.5. *For $i \neq z \in \mathcal{F}$ with discriminant Δ, we have*

$$|j(z)-1728| \geq 20000\min\{|z-i|, 0.01\}^2,$$

$$|j(z)-1728| \geq 2000|\Delta|^{-2}.$$

Proof. Combine Proposition 3.7 and Corollary 5.3 in [6]. □

We start to prove Theorem 5.3.1 (1). Let $\tau_k, z_k, \alpha_k, x_k$ be as the begining of this section. Then we have

$$\sum_{k=1}^{d}\log^+|x_k-\alpha_k|^{-1} = \sum_{\substack{1\leq k\leq d\\ z_k\in S_\varepsilon(\tau_k,\Delta)}}\log^+|x_k-\alpha_k|^{-1} + \sum_{\substack{1\leq k\leq d\\ z_k\notin S_\varepsilon(\tau_k,\Delta)}}\log^+|x_k-\alpha_k|^{-1}$$

For the first sum, by Theorem 5.3.4, each term in the sum has

$$\log^+|x_k-\alpha_k|^{-1} \leq \max\{0, 4\log(\max\{|\Delta|, |\Delta_\alpha|\}) - \log(800)\} \leq 4\log(\max\{|\Delta|, |\Delta_\alpha|\}),$$

so we have

$$\sum_{\substack{1\le k\le d\\ z_k\in S_\varepsilon(\tau_k,\Delta)}} \log^+|x_k-\alpha_k|^{-1} \le \sum_{1\le k\le \mathcal{C}(\Delta_\alpha)} 4\mathcal{C}_\varepsilon(\tau_k,\Delta)\log(\max\{|\Delta|,|\Delta_\alpha|\}). \tag{5.5}$$

For the second sum, we claim that if $|z_k-\tau_k|\ge\varepsilon$, then

$$|x_k-\alpha_k|\ge 2400|\Delta_\alpha|^{-2}\varepsilon.$$

In fact, we can replace τ by τ_k and z' by z_k in Lemma 5.3.3, then

$$|x_k-\alpha_k|\ge \min\{e^{2.6\pi}\cdot 0.4\varepsilon, 5\cdot 10^{-7}, 800|\Delta_\alpha|^{-4}, 2400|\Delta_\alpha|^{-2}\varepsilon\}.$$

Notice that $|\Delta_\alpha|\ge 7$ and $\varepsilon<\min\{\frac{1}{3|\Delta_\alpha|^2}, 10^{-8}\}$, then

$$2400|\Delta_\alpha|^{-2}\varepsilon \le 800|\Delta_\alpha|^{-4},$$

$$2400|\Delta_\alpha|^{-2}\varepsilon \le \frac{2400}{49}\cdot 10^{-8} < 5\cdot 10^{-7},$$

$$2400|\Delta_\alpha|^{-2}\varepsilon \le \frac{2400}{49}\varepsilon \le 1410\varepsilon \le e^{2.6\pi}\cdot 0.4\varepsilon,$$

so we have our claim. Hence

$$\log^+|x_k-\alpha_k|^{-1} \le \log\left(\frac{|\Delta_\alpha|^2}{2400}\varepsilon^{-1}\right) \le \log(\varepsilon^{-1}) + 2\log|\Delta_\alpha| - 7.783,$$

$$\sum_{\substack{1\le k\le d\\ z_k\notin S_\varepsilon(\tau_k,\Delta)}} \log^+|x_k-\alpha_k|^{-1} \le d(\log(\varepsilon^{-1}) + 2\log|\Delta_\alpha| - 7.783). \tag{5.6}$$

Combine 5.5, 5.6 and the equality 5.4, we have the bound in (1).

For (2), the proof is similar as above. Since $j(\tau)=1728$, then $d=\mathcal{C}(\Delta)$ and

$$\sum_{k=1}^{\mathcal{C}(\Delta)} \log^+|x_k-1728|^{-1} = \sum_{\substack{1\le k\le \mathcal{C}(\Delta)\\ z_k\in S_\varepsilon(i,\Delta)}} \log^+|x_k-1728|^{-1} + \sum_{\substack{1\le k\le \mathcal{C}(\Delta)\\ z_k\notin S_\varepsilon(i,\Delta)}} \log^+|x_k-1728|^{-1}$$

For the first sum, by Lemma 5.3.5,

$$\log^+|x_k-1728|^{-1} \le \max\{0, 2\log|\Delta| - \log 2000\} \le 2\log|\Delta|,$$

$$\sum_{\substack{1\le k\le \mathcal{C}(\Delta)\\ z_k\in S_\varepsilon(i,\Delta)}} \log^+|x_k-1728|^{-1} \le 2\mathcal{C}_\varepsilon(i,\Delta)\log|\Delta|.$$

For the second sum, since $\varepsilon\le 7\cdot 10^{-3}$, $\varepsilon^{-2}>20000$ and $|z_k-i|\ge\varepsilon$, we have

$$|x_k-1728|^{-1} \le 20000^{-1}\min\{\varepsilon, 0.01\}^{-2} = 20000^{-1}\varepsilon^{-2},$$

$$\log^+|x_k-1728|^{-1} \le \max\{0, 2\log\varepsilon^{-1} - \log(20000)\} \le 2\log\varepsilon^{-1} - 9.9,$$

$$\sum_{\substack{1\le k\le \mathcal{C}(\Delta)\\ z_k\notin S_\varepsilon(i,\Delta)}} \log^+|x_k-1728|^{-1} \le \mathcal{C}(\Delta)(2\log\varepsilon^{-1} - 9.9).$$

Hence, as above, we have

$$\mathrm{h}(x-1728) \le 2\frac{\mathcal{C}_\varepsilon(i,\Delta)}{\mathcal{C}(\Delta)}\log|\Delta| + 2\log\varepsilon^{-1} - 9.9 + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)|.$$

5.3.2 Proof of Corollary 5.3.2

We will use the following lemmas from [7].

LEMMA 5.3.6 ([7] Lemma 3.5)**.** *Assume that* $|\Delta| \geq 10^{14}$*. Then* $F \geq |\Delta|^{0.34/\log\log(|\Delta|^{1/2})}$ *and* $F \geq 18.54 \log\log(|\Delta|^{1/2})$.

LEMMA 5.3.7 ([7] Lemma 3.6)**.** *For* $\Delta \neq -3, -4$*, we have*

$$\mathcal{C}(\Delta) \leq \pi^{-1}|\Delta|^{1/2}(2 + \log|\Delta|).$$

To prove (1), by Corollary 5.2.2, we have

$$\sum_{1\leq k\leq \mathcal{C}(\Delta_\alpha)} 4\frac{\mathcal{C}_\varepsilon(\tau_k,\Delta)}{d}\log\max\{|\Delta|,|\Delta_\alpha|\} \leq 4\frac{A\mathcal{C}(\Delta_\alpha)\left(46.488|\Delta|^{1/2}\varepsilon^2\log\log|\Delta|^{1/2} + 7.752|\Delta|^{1/2}\varepsilon + 2\right)}{d} \tag{5.7}$$

We can take $\varepsilon = 0.0003\frac{d}{A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}|\Delta_\alpha|^2}$, then $\varepsilon \leq \min\{\frac{1}{3|\Delta_\alpha|^2}, 10^{-8}\}$. Indeed, $F \geq 256$ if $|\Delta| \geq 10^{14}$, and by Lemma 5.3.6 and Lemma 5.3.7, we have

$$0.0003\frac{d}{A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}} \leq \frac{3\mathcal{C}(\Delta)}{10000F|\Delta|^{1/2}\log|\Delta|} \leq \frac{6+3\log(10^{14})}{10000\pi\log(10^{14})}\cdot\frac{1}{256} \leq \frac{1}{3},$$

$$0.0003\frac{d}{A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}|\Delta_\alpha|^2} \leq \frac{6+3\log(10^{14})}{490000\pi\log(10^{14})}\cdot\frac{1}{256} \leq 10^{-8}.$$

We estimate each term in the left of 5.7 with our ε:

$$\begin{aligned}
4\frac{46.488A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}\varepsilon^2\log\log|\Delta|^{1/2}}{d} &\leq 36\cdot10^{-8}\cdot46.488\frac{d\log\log|\Delta|^{1/2}}{A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}|\Delta_\alpha|^4}\\
&\leq \frac{36\cdot10^{-8}\cdot46.488}{|\Delta_\alpha|^4}\frac{\log\log|\Delta|^{1/2}}{F}\frac{\mathcal{C}(\Delta)}{|\Delta|^{1/2}\log|\Delta|}\\
&\leq \frac{36\cdot10^{-8}\cdot48.488\cdot(2+\log(10^{14}))}{18.54\cdot\pi\log(10^{14})}\cdot\frac{1}{7^4}\\
&\leq 0.0005;
\end{aligned}$$

$$\begin{aligned}
4\frac{7.752A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}\varepsilon}{d} &\leq 0.0003\cdot31.008|\Delta_\alpha|^{-2}\\
&\leq 0.0005.
\end{aligned}$$

With above, we have

$$\begin{aligned}
\mathrm{h}(x-\alpha) &\leq \frac{8A\mathcal{C}(\Delta_\alpha)}{d} + \log(\frac{A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}|\Delta_\alpha|^2}{d}) + 2\log|\Delta_\alpha| + 0.001 + \log(\frac{10000}{3}) - 7.783\\
&\quad + \frac{1}{d}\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|\\
&\leq \frac{8A\mathcal{C}(\Delta_\alpha)}{d} + \log(\frac{A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}}{d}) + 4\log|\Delta_\alpha| + 0.33 + \frac{1}{d}\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|.
\end{aligned}$$

For (2), the proof is similar. We set $\varepsilon = 0.3\frac{\mathcal{C}(\Delta)}{A|\Delta|^{1/2}}$, then $\varepsilon \leq 7\cdot10^{-3}$. Indeed, since $|\Delta| \geq 10^{14}$, so $F \geq 256$, hence

$$0.3\frac{\mathcal{C}(\Delta)}{A|\Delta|^{1/2}} = 0.3\frac{\mathcal{C}(\Delta)}{|\Delta|^{1/2}\log|\Delta|}\cdot\frac{1}{F} \leq 0.3\frac{2+\log(10^{14})}{\pi\log(10^{14})}\cdot\frac{1}{256} \leq 5\cdot10^{-4}.$$

By Corollary 5.2.2, Theorem 5.3.1(2), Lemma 5.3.6 and Lemma 5.3.7, we have

$$\begin{aligned}
\mathrm{h}(x-1728) &\le 2\frac{\mathcal{C}_\varepsilon(i,\Delta)}{\mathcal{C}(\Delta)}\log|\Delta| + 2\log\varepsilon^{-1} - 9.9 + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| \\
&\le 2\frac{A\left(46.488|\Delta|^{1/2}\varepsilon^2\log\log|\Delta|^{1/2} + 7.752|\Delta|^{1/2}\varepsilon + 2\right)}{\mathcal{C}(\Delta)} + 2\log\varepsilon^{-1} - 9.9 \\
&\quad + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| \\
&\le 2\cdot 46.488\cdot 0.3^2\frac{\log\log|\Delta|^{1/2}}{F}\frac{\mathcal{C}(\Delta)}{|\Delta|^{1/2}\log|\Delta|} + 2\cdot 0.3\cdot 7.752 + \frac{4A}{\mathcal{C}(\Delta)} \\
&\quad + 2\log\frac{A|\Delta|^{1/2}}{\mathcal{C}(\Delta)} - 2\log 0.3 - 9.9 + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| \\
&\le \frac{4A}{\mathcal{C}(\Delta)} + 2\log\frac{A|\Delta|^{1/2}}{\mathcal{C}(\Delta)} + 2\cdot 46.488\cdot 0.3^2\frac{2+\log(10^{14})}{18.54\cdot\pi\log(10^{14})} - 2.84 \\
&\quad + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| \\
&\le \frac{4A}{\mathcal{C}(\Delta)} + 2\log\frac{A|\Delta|^{1/2}}{\mathcal{C}(\Delta)} - 2.68 + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)|.
\end{aligned}$$

For (3), see [7, Corollary 3.2], without assuming that x is a singular unit, we add the term $\frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x)|$.

5.4 Lower Bounds for the Height of a Singular Modulus

We have these propositions from [7]:

Proposition 5.4.1 ([7] Proposition 4.1)**.** *Let x be a singular modulus of discriminant Δ. Assume that $|\Delta| \ge 16$. Then*

$$\mathrm{h}(x) \ge \frac{\pi|\Delta|^{1/2} - 0.01}{\mathcal{C}(\Delta)}.$$

Proposition 5.4.2. *Let x be a singular modulus of discriminant Δ. Then*

$$\mathrm{h}(x) \ge \frac{3}{\sqrt{5}}\log|\Delta| - 9.79;$$

$$\mathrm{h}(x) \ge \frac{1}{4\sqrt{5}}\log|\Delta| - 5.93.$$

Proof. The first one see [7, Proposition 4.3], the second one see [32, Lemma 14 (ii)] □

We can use the inequalilty $\mathrm{h}(x-\alpha) \ge \mathrm{h}(x) - \mathrm{h}(\alpha) - \log 2$ and the results above to give the lower bounds of $\mathrm{h}(x-\alpha)$ for an fixed α.

5.5 Proof of Theorem 5.1.1 (1)

As the set-up in section5.3, Proposition 5.4.1 and 5.4.2 allow us to give lower bounds of the height of $x-\alpha$:

$$\mathrm{h}(x-\alpha) \geq \mathrm{h}(x) - \mathrm{h}(\alpha) - \log 2 \geq \frac{\pi|\Delta|^{1/2} - 0.01}{\mathcal{C}(\Delta)} - \mathrm{h}(\alpha) - \log 2, \tag{5.8}$$

$$\mathrm{h}(x-\alpha) \geq \mathrm{h}(x) - \mathrm{h}(\alpha) - \log 2 \geq \frac{3}{\sqrt{5}} \log|\Delta| - \mathrm{h}(\alpha) - 9.79 - \log 2. \tag{5.9}$$

For (1), recall the upper bound of $x-\alpha$ in Corollary 5.3.2 (1) when $|\Delta| \geq 10^{14}$:

$$\mathrm{h}(x-\alpha) \leq \frac{8A\mathcal{C}(\Delta_\tau)}{d} + \log\left(\frac{A\mathcal{C}(\Delta_\alpha)|\Delta|^{1/2}}{d}\right) + 4\log|\Delta_\alpha| + 0.33 + \frac{1}{d}\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|, \tag{5.10}$$

Throughout the proof of (1), denote the discriminant of a singular modulus $x = j(z)$ by Δ, and we assume that $X = |\Delta| \geq \max\{e^{3.12}(\mathcal{C}(\Delta_\alpha)|\Delta_\alpha|^4 e^{\mathrm{h}(\alpha)})^3, 10^{15}\cdot\mathcal{C}(\Delta_\alpha)^6\}$. Hence $|\Delta| \geq |\Delta_\alpha|$, since $\mathrm{h}(\alpha) \geq 0$.

5.5.1 The main inequality

Recall that $A = F\max\{|\Delta|, |\Delta_\alpha|\} = F\log X$. Minding 0.01 in 5.8 we deduce from 5.10 the inequality

$$\frac{8A\mathcal{C}(\Delta_\alpha)}{d} + \log\left(\frac{AX^{1/2}}{d}\right) + C + \frac{1}{d}\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)| \geq Y$$

where

$$C = \log(\mathcal{C}(\Delta_\alpha)) + 4\log|\Delta_\alpha| + \mathrm{h}(\alpha) + 1.04,$$

$$Y = \max\left\{\frac{\pi X^{1/2}}{\mathcal{C}(\Delta)}, \frac{3}{\sqrt{5}}\log X - 9.78\right\}.$$

We rewrite this as

$$\frac{8A\mathcal{C}(\Delta_\alpha)/d}{Y} + \frac{\log A + C}{Y} + \frac{\log(X^{1/2}/d)}{Y} + \frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|/d}{Y} \geq 1. \tag{5.11}$$

Note that $C > 3.11 > 0$, $\log A \geq 0$ because $C \geq 4\log 7 + \frac{1}{4\sqrt{5}}\log 7 - 5.93 + 1.04 > 3.11$. Hence, we may replace Y by $\frac{3}{\sqrt{5}}\log X - 9.78$ in the second term of the left-hand side in 5.11. Similarly, in the 1st term and 4th term we may replace Y by $\pi X^{1/2}/\mathcal{C}(\Delta)$, and in the 3rd term we may replace $X^{1/2}\mathcal{C}(\Delta)$ by $\pi^{-1}Y$. Notice that $d \geq \mathcal{C}(\Delta)$, we obtain

$$\frac{8A\mathcal{C}(\Delta_\alpha)}{\pi X^{1/2}} + \frac{\log A + C}{\frac{3}{\sqrt{5}}\log X - 9.78} + \frac{\log(\pi^{-1}Y)}{Y} + \frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|}{\pi X^{1/2}} \geq 1. \tag{5.12}$$

To obtain a lower bound of $\log|\mathcal{N}_{H/\mathbb{Q}}(\alpha)|$, we will bound from above each of the three terms in its left-hand side.

From the results in [7, Section 5.2 and Section 5.3], we know that, when $X \geq 10^{15}$,

$$\log A \leq \frac{\log 2}{2}\frac{\log X}{\log\log X - c_1 - \log 2} + \log\log X, \tag{5.13}$$

where $c_1 < 1.1713142$.

5.5.2 Bound the first term in 5.12

From above, easy to know that when $X \geq 10^{15}$, we have

$$\frac{\log(AX^{-1/2})}{\log X} \leq u_0(X),$$

where

$$u_0(X) = \frac{\log 2}{2}\frac{1}{\log\log X - c_1 - \log 2} + \frac{\log\log X}{\log X} - \frac{1}{2}$$

which is decreasing for $X \geq 10^{15}$. Hence

$$\frac{\log(AX^{-1/2})}{\log X} \leq u_0(X) \leq u_0(10^{15}) \leq -0.1908,$$

so

$$\frac{8A\mathcal{C}(\Delta_\tau)}{\pi X^{1/2}} \leq \frac{8\mathcal{C}(\Delta_\tau)}{\pi}X^{-0.1908} \leq \frac{8}{\pi}\cdot 10^{15\cdot(-0.1908)} \leq 0.0035,$$

since $X \geq \mathcal{C}(\Delta_\tau)^6 \cdot 10^{15}$.

5.5.3 Bound the second term in 5.12

Obviously, by 5.13

$$\frac{\log A + C}{\frac{3}{\sqrt{5}}\log X - 9.78} \leq u_1(X)u_2(X),$$

where

$$u_1(X) = \frac{\log 2}{2}\frac{1}{\log\log X - c_1 - \log 2} + \frac{\log\log X + C}{\log X},$$

$$u_2(X) = (\frac{3}{\sqrt{5}} - \frac{9.78}{\log X})^{-1},$$

which are decreasing for $X \geq 10^{10}$.

Since $X \geq e^{3.12}(\mathcal{C}(\Delta_\alpha)|\Delta_\alpha|^4 e^{h(\alpha)})^3 = e^{3C}$, we have

$$\frac{\log\log X + C}{\log X} \leq 0.6.$$

Indeed, set $g(x) = \log x - 0.6x + C$, which is decreasing for $x > 5/3$. Let $x_0 = 3C > 9.33 \geq 5/3$, since $C > 3.11$. Hence

$$g(x) \leq g(x_0) = \log 3 + \log C - 0.8C \leq \log 3 + \log(3.11) - 0.8\cdot 3.11 < 0.$$

With this we have

$$u_1(X)u_2(X) \leq (\frac{\log 2}{2}\frac{1}{\log\log(10^{15}) - 1.1713142 - \log 2} + 0.6)\cdot u_2(10^{15}) < 0.7621.$$

5.5.4 Bound the third term in 5.12

For this term, we directly use the bound from [7, subsection 5.5]

$$\frac{\log(\pi^{-1}Y)}{Y} < 0.0672.$$

5.5.5 Summing up

We can combine the above estimates and bound $\frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|}{\pi X^{1/2}}$ by

$$\frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)|}{\pi X^{1/2}} > 1-(0.0035+0.7621+0.0672) = 0.1672,$$

so

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x-\alpha)| > \frac{|\Delta|^{1/2}}{2}.$$

5.6 Proof of Theorem 5.1.1 (2)

As in the last section, we assume that $X = |\Delta| \geq 10^{15}$. By inequality 5.8, 5.9 and Corollary 5.3.2 (2), we have

$$\frac{4A}{\mathcal{C}(\Delta)} + 2\log(\frac{AX^{1/2}}{\mathcal{C}(\Delta)}) + C + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| \geq Y$$

where

$$C = \mathrm{h}(1728) + \log 2 - 2.68 + 0.01 = \log(3456) - 2.67 > 0,$$

$$Y = \max\{\frac{\pi X^{1/2}}{\mathcal{C}(\Delta)}, \frac{3}{\sqrt{5}}\log X - 9.78\}.$$

We rewrite this as

$$\frac{4A/\mathcal{C}(\Delta)}{Y} + \frac{2\log A + C}{Y} + \frac{\log(X^{1/2}/\mathcal{C}(\Delta))}{Y} + \frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)|/\mathcal{C}(\Delta)}{Y} \geq 1.$$

Hence,

$$\frac{4A}{\pi X^{1/2}} + \frac{2\log A + C}{\frac{3}{\sqrt{5}}\log X - 9.78} + \frac{\log(\pi^{-1}Y)}{Y} + \frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)|}{\pi X^{1/2}} \geq 1. \tag{5.14}$$

Using the similar method to estimate each term when $X \geq 10^{15}$, we have

$$\frac{4A}{\pi X^{1/2}} < 0.0018,$$

$$\frac{2\log A + C}{\frac{3}{\sqrt{5}}\log X - 9.78} < 0.7337,$$

$$\frac{\log(\pi^{-1}Y)}{Y} < 0.0672,$$

$$\frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)|}{\pi X^{1/2}} \geq 1-(0.0018+0.7337+0.0672) = 0.1973.$$

Hence,

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x-1728)| \geq 0.1973\pi X^{1/2} \geq \frac{|\Delta|^{1/2}}{2}.$$

5.7 Proof of Theorem 5.1.1 (3)

As before, we assume that $X = |\Delta| \geq 10^{15}$. By Proposition 5.4.1, Proposition 5.4.2 and Corollary 5.3.2 (3), we have

$$\frac{12A}{\mathcal{C}(\Delta)} + 3\log\frac{AX^{1/2}}{\mathcal{C}(\Delta)} - 3.76 + \frac{1}{\mathcal{C}(\Delta)}\log|\mathcal{N}_{K/\mathbb{Q}}(x)| \geq Y,$$

where

$$Y = \max\{\frac{\pi X^{1/2}}{\mathcal{C}(\Delta)}, \frac{3}{\sqrt{5}}\log X - 9.78\},$$

We rewrite this as

$$\frac{12A/\mathcal{C}(\Delta)}{Y} + \frac{3\log A - 3.76}{Y} + \frac{3\log(X^{1/2}/\mathcal{C}(\Delta))}{Y} + \frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x)|/\mathcal{C}(\Delta)}{Y} \geq 1. \tag{5.15}$$

Noe that $3\log A - 3.76 > 0$ because $A \geq \log X \geq \log(10^{15}) > 30$. Hence, we obtain

$$\frac{12A}{\pi X^{1/2}} + \frac{3\log A - 3.76}{\frac{3}{\sqrt{5}}\log X - 9.78} + \frac{3\log(\pi^{-1}Y)}{Y} + \frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x)|}{\pi X^{1/2}} \geq 1.$$

From the results in [7, Page 23 to Page 25], we know that, when $X \geq 10^{15}$,

$$AX^{-1/2} < 0.0014,$$

$$\frac{3\log A - 3.76}{\frac{3}{\sqrt{5}}\log X - 9.78} < 0.7734,$$

$$\frac{\log(\pi^{-1}Y)}{Y} < 0.0672.$$

We can combine the above estimates and bound $\dfrac{\log|\mathcal{N}_{K/\mathbb{Q}}(x)|}{\pi X^{1/2}}$ by

$$\frac{\log|\mathcal{N}_{K/\mathbb{Q}}(x)|}{\pi X^{1/2}} > 1 - (12\pi^{-1}\cdot 0.0014 + 0.7734 + 3\cdot 0.0672) > 0.019,$$

so

$$\log|\mathcal{N}_{K/\mathbb{Q}}(x)| > \frac{|\Delta|^{1/2}}{20}.$$

Part III

The Artin Conductors and Discriminants of Hyperelliptic Curves

Chapter 6

Hyperelliptic Curves

This chapter provides sufficient background about hyperelliptic curves for our study. The main references are [40] and [41].

6.1 Basic Definition

This section comes from [41, Section 7.4]. In this section, k is a field, and we denote the function field of $\mathbb{P}^1_k$ by $k(x)$. For an integral projective curve C over k and a Cartier divisor D on C, we denote

$$L(D) = \{f \in k(C) \mid D + \operatorname{div}(f) \geq 0\},$$

$$\ell(D) = \dim_k L(D).$$

We also denote K_C the canonical divisor of C.

Definition 6.1.1. *Let C be a smooth, geometrically connected, projective curve over a field k of genus $g \geq 1$. We say C is a hyperelliptic curve if there exists a finite morphism $C \to \mathbb{P}^1_k$ of degree 2.*

Remark. *(1) The extension $k(C)/k(x)$ of fraction fields is Galois of* $\deg = 2$.

LEMMA 6.1.2 ([41], Lemma 7.4.8)**.** *Let C be a smooth, geometrically connected, projective curves over a field k of genus $g \geq 1$. Then C is hyperelliptic if and only if there exists a Cartier divisor D on C such that $\ell(D) = \deg D = 2$.*

Corollary 6.1.3 ([41], Proposition 7.4.9)**.** *Let C be a smooth, geometrically connected, projective curves over a field k. If C is elliptic or of genus $g = 2$. Then C is hyperelliptic.*

Proof. It suffices to find a Cartier divisor on C such that $\ell(D) = \deg D = 2$.

If C is elliptic, then $g = 1$ and there is a rational point $O \in C(k)$. We have $\deg(2O) = 2$ and

$$\ell(2O) = \deg(2O) + \chi(C) = 2$$

by Riemann-Roch Theorem.

If $g = 2$, then $\deg K_C = 2g - 2 = 2$ and $\ell(K_C) = g = 2$.. □

Definition 6.1.4. *Let C be a hyperelliptic curve over a field k with a separable morphism $f : C \to \mathbb{P}^1_k$ of degree 2. Let $\sigma \in \operatorname{Gal}(k(C)/k(x))$ be the generator. It induces an automorphism of order 2 of C, also denoted by σ. We will call it a hyperelliptic involution of C (associated to f).*

6.2 Hyperelliptic Equations

Proposition 6.2.1 ([41], Proposition 7.4.24)**.** *Let C be a hyperelliptic curve of genus $g \geq 1$ over a field k with a separable morphism $f : C \to \mathbb{P}^1_k$ of degree 2. Then*

(1) $k(C) = k(x)[y]$ with a relation

$$y^2 + Q(x) = P(x),\ Q(x), P(x) \in k[x]$$

with $2g+1 \leq \max\{2\deg Q(x), \deg P(x)\} \leq 2g+2$. We can take $Q(x) = 0$ if $\mathrm{Char}(k) \neq 2$.

(2) The curve C is the union of two affine open subschemes

$$U' = \mathrm{Spec}(k[x, Y])/(Y^2 + Q(x)Y - P(x)),$$

$$V' = \mathrm{Spec}(k[w, Z])/(Z^2 + Q_1(w)Z - P_1(w)),$$

where $Q_1(w) = Q(1/w)w^{g+1}, P_1(w) = P(1/w)w^{2g+2}$ and two open subschemes glue along $D(x) \simeq D(w)$ with relation $x = 1/w$ and $Y = x^{g+1}Z$.

(3) The ramification points of f are those $V(4P(x) + Q(x)^2) \subset U'$, plus the point $\{w = 0\} \in V'$ if $\deg(4P(x) + Q(x)^2) \leq 2g+1$.

Definition 6.2.2. *Let C be a hyperelliptic curve of genus $g \geq 1$ over a field k, with a separable morphism $f : C \to \mathbb{P}^1_k$ of degree 2. Let $x, y \in k(C)$ satisfying the following condition:*

(a) $k(\mathbb{P}^1_k) = k(x) \subset k(C)$;

(b) $y^2 + Q(x)y = P(x)$ with $P(x), Q(x) \in k[x]$ and $\deg Q(x) \leq g+1, \deg P(x) \leq 2g+2$;

(c) the equation above is normal.

We call $\{1, y\}$ a standard base of C and $y^2 + Q(x)y = P(x)$ a hyperelliptic equation of C. Such x,y exist due to Proposition 6.2.1.

When C is an elliptic curve, a hyperelliptic equation of C is called an elliptic equation of C if $\deg Q(x) \leq 1$ and $\deg P(x) \leq 3$.

Remark. *(1) From Proposition 6.2.1 (2), we know that C can be covered by two affine scheme of hyperelliptic equations.*

Corollary 6.2.3 ([41], Proposition 7.4.33)**.** *Let C be a hyperelliptic curve of genus $g \geq 1$ over a field k. Let*

$$(\mathcal{E}) : y^2 + Q(x)y = P(x),\ \ (\mathcal{E}') : v^2 + R(u)v = S(u)$$

be two hyperelliptic equations of C.

(1) Suppose that $g \geq 2$, then there exist $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(k)$, $e \in k^*$, $H(x) \in k[x]$, $\deg H \leq g+1$, such that

$$u = \frac{ax+b}{cx+d},\ \ v = \frac{H(x) + ey}{(cx+d)^{g+1}}.$$

(2) If C is an elliptic curve and the two equations are elliptic with the same origin. Then we have the same conclusion as above, with moreover $c = 0, d = 1$ and $\deg H(x) \leq 1$.

Definition 6.2.4. *Let C be a hyperelliptic curve of genus $g \geq 1$ over a field k. Let*

$$(\mathcal{E}) : y^2 + Q(x)y = P(x)$$

be a hyperelliptic equation of C. Let $R(x) := Q(x)^2 + 4P(x)$ with leading coefficient c. Then the discriminant of $(\mathcal{E})$ is defined as

$$\Delta(\mathcal{E}) := \begin{cases} 2^{-4(g+1)}\mathrm{disc}(R(x)) & \text{if } \deg R(x) = 2g+2, \\ 2^{-4(g+1)}c^2\mathrm{disc}(R(x)) & \text{if } \deg R(x) = 2g+1. \end{cases}$$

Remark. *(1) If $(\mathcal{E}), (\mathcal{E}')$ is as Corollary 6.2.3 with the change of coordinates, then*

$$\Delta(\mathcal{E}) = \Delta(\mathcal{E}')e^{-4(g+1)}(ad-bc)^{2(g+1)(2g+1)}.$$

6.3 Integral Models of Hyperelliptic Curves over a Discrete Valuation Field

LEMMA 6.3.1 ([40], Lemma 1)**.** *Let A be a PID, and $F/\mathrm{Frac}A[x]$ be a separable extension of $\deg = 2$ with integral closure B of $A[x]$ in F. Then we have the following properties:*

(1) The $A[x]$-module B is free and there exists $y \in B$ such that $\{1, y\}$ is a base of B.

(2) Suppose that F is the function field of a smooth projective curve of genus g over $\mathrm{Frac}A[x]$. We can choose a base $\{1, y\}$ such that $y^2 + Q(x)y = P(x)$ with $\deg Q(x) \leq g+1$ and $\deg P(x) \leq 2g+2$.

In the rest of this section R is a discrete valuation ring with valuation v, residue field k and fraction field K.

Definition 6.3.2. *Let C be a hyperelliptic curve of genus $g \geq 1$ over K. An integral equation of C is a hyperelliptic equation*

$$(\mathcal{E}) : y^2 + Q(x)y = P(x)$$

such that $\{1, y\}$ a base of the integral closure of $R[x] \subset K(x) = K(\mathbb{P}^1_K)$ in $K(C)$.

An integral equation $(\mathcal{E})$ of C is said to be minimal if $v(\Delta(\mathcal{E}))$ is minimal for all integral equations of C, where v is the discrete valuation on K. The integer $v(\Delta(\mathcal{E}))$ is called the minimal discriminant of C (in R), and denoted by $v(C)$.

Remark. *(1) Notice that a hyperelliptic equation $(\mathcal{E})$ is an integral equation if and only if $Q(x), P(x) \in R[x]$, and $R[x, Y]/(Y^2 + Q(x)Y - P(x))$ is normal.*

(2) The minimal (integral) equation of C always exists, but not unique in general.

LEMMA 6.3.3 ([40], Lemma 2)**.** *Let $B = R[x, Y]/(Y^2 + Q(x)Y - P(x))$ with $P(x), Q(x) \in R[x]$. Assume that $B \otimes_R K$ is normal. Then the following statements hold:*

(1) If $B \otimes_R k$ is reduced, then B is normal.

(2) The ring $B \otimes_R k$ is not reduced if and only if $4\overline{P}(x) + \overline{Q}(x) = 0$ and $-\overline{P}$ is not a square in $k[x]$.

(3) *If $B \otimes_R k$ is not reduced, then B is normal if and only if there exists $T(x) \in R[x]$ such that $v(P(x) + Q(x)T(x) - T(x)) = 1$.*

(4) *If $\mathrm{Char}(k) \neq 2$, then B is normal if and only if $v(4P(x) + Q(x)^2) \leq 1$.*

Next we talk about the Weierstrass models of a hyperelliptic curve.

Proposition 6.3.4. *We have a bijection:*

$$\{\text{generators of } K(\mathbb{P}^1_K) \text{ over } K\}/\sim \longleftrightarrow \{\text{smooth, proper, flat models of } \mathbb{P}^1_K \text{ over } R\},$$
$$x \mapsto \mathbb{P}^1_x (:= \mathrm{Spec}(R[x]) \cup \mathrm{Spec}(R[1/x])),$$

where $x \sim u$ for two generators x, u if there exists $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(R)$ such that $u = \dfrac{ax+b}{cx+d}$.

Proof. Obviously, this map is well-defined.

It is injective. Indeed, if $\mathbb{P}^1_x \simeq \mathbb{P}^1_u$ as models, we view $\mathbb{P}^1_u = \mathrm{Proj}(R[U_0, U_1]), u = U_1/U_0$. Then via this isomorphism, $\mathrm{Spec}(R[x]) = D_+(cU_1 + dU_0)$ for some $c, d \in R$. Hence there exists $a, b \in R$ such that $x = \frac{aU_1+bU_0}{cU_1+dU_0}$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(K)$. Since it induces isomorphism on special fibers, then $\begin{pmatrix} \overline{a} & \overline{b} \\ \overline{c} & \overline{d} \end{pmatrix} \in \mathrm{GL}_2(k)$, which implies $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(R)$.

It is surjective. Notice that every smooth, proper, flat model of $\mathbb{P}^1_K$ is isomorphic to $\mathbb{P}^1_R = \mathrm{Proj}(R[X_0, X_1])$, so we can take $x = \frac{X_1}{X_0}$. □

Definition 6.3.5. *Let C be a hyperelliptic curve over K with a fixed hyperelliptic involution $\sigma : C \to C$. A Weierstrass model W of C is a nomral, proper, flat model of C over R such that $W/\langle\sigma\rangle$ is smooth over R.*

Corollary 6.3.6. *Let C be a hyperelliptic curve over K with a fixed hyperelliptic involution $\sigma : C \to C$. Then we have the following bijections:*

$$\{\text{Integral equations of } C\}/\sim \longleftrightarrow \{\text{Weierstrass models of } C\} \longleftrightarrow$$
$$\{\text{smooth, proper, flat models of } \mathbb{P}^1_K \text{ over } R\}$$

where $(y^2 + Q(x)y = P(x)) \sim (w^2 + Q_1(z) = P_1(z))$ for two integral equations if there exists $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(R)$ such that $z = \dfrac{ax+b}{cx+d}$. In particular, every Weierstrass model of C is projective over R, and two Weierstrass models W_1, W_2 are isomorphic if and only if $W_1/\langle\sigma\rangle, W_2/\langle\sigma\rangle$ are isomorphic.

Proof. The map

$$\{\text{Integral equations of } C\}/\sim \to \{\text{smooth, proper, flat models of } \mathbb{P}^1_K \text{ over } R\},$$
$$(y^2 + Q(x)y = P(x)) \mapsto \mathbb{P}^1_x,$$

is well-defined and injective by Proposition 6.3.4. To show it is surjective, we should show that for any generator $x \in K(\mathbb{P}^1_K)$, there exists $\{1, y\} \subset K(C)$ such that $y^2 + Q(x)y = P(x)$ and the normalization of $R[X]$ in $K(C)$ is $R[x] + R[x]y$. This comes from Lemma 6.3.1 (1).

For a generator $x \in K(\mathbb{P}^1_K)$, we take the normalization $\pi : W_x \to \mathbb{P}^1_x$ of $\mathbb{P}^1_x$ in $K(C)$, which is finite. Hence W_x is a proper, normal, flat over R. Moreover, $W_{x,K} \simeq C$, since $K(W_{x,K}) = K(C)$ and $W_{x,K}, C$ are smooth projective over K. That implies that W_x is a model of C. To show that W_x is a Weierstrass model of C, it is sufficient to show that $W_x/\langle\sigma\rangle \simeq \mathbb{P}^1_x$. Indeed, $\pi^{-1}(\mathrm{Spec}(R_K[x])) = \mathrm{Spec}(B)$, where $B = R[x] + R[x]y$ for some $y \in K(C)$ such that $y^2 + Q(x)y = P(x)$. We have $\sigma(y) = -y - Q(x)$, so $\sigma(ax + by) = ax + by$ if and only if $b = 0$. Hence $(\mathrm{Spec}(B))/\langle\sigma\rangle\mathrm{Spec}(B^\sigma) = \mathrm{Spec}(R[x])$. It is similar for $\pi^{-1}(\mathrm{Spec}(R_K[1/x]))$.

On the other hand, for a Weierstrass model W of C, $W/\langle\sigma\rangle$ is a proper, smooth, flat model of $\mathbb{P}^1_K$ over R. Indeed, $W/\langle\sigma\rangle$ is smooth over R, and on generic fiber,

$$(W/\langle\sigma\rangle)_K = W_K/\langle\sigma\rangle = C/\langle\sigma\rangle \simeq \mathbb{P}^1_K.$$

It is proper, since for $\mathrm{Spec}(B) \subset C$ with $B = R[x,y]/(y^2 + Q(x)y - P(x)))$, $R \subset B^\sigma \subset B$, which implies that B^σ finite R-module and $\mathrm{Spec}(B)/\langle\sigma\rangle$ finite over $\mathrm{Spec}(R)$. By Proposition 6.3.4, $W/\langle\sigma\rangle \simeq \mathbb{P}^1_x$ for some generator $x \in K(\mathbb{P}^1_K)$, and $W \to \mathbb{P}^1_x$ is finite. Hence W is is the normalization of $\mathbb{P}^1_x$ in $K(C)$.

□

Remark. *(1) With the first bijection and the remark of Definition 6.2.4, we know that the discriminant of a Weierstrass model is well-defined.*

For our uses in Chapter 7, we will projective line have an equation similar to hyperelliptic equations, so we give the following definition.

Definition 6.3.7. *Assume that* $\mathrm{Char}(K) \neq 2$*, an integral (Weierstrass) equation over R is of the form*

$$y^2 = f(x)$$

with $f(x) \in R[x]$ such that $\deg(f) \geq 1$, $f(x) = 0$ has no multiple root and $v(f(x)) \leq 1$.

For such integral Weierstrass equation, let $x = 1/z, y = x^{g+1}w$, where $g = \left\lfloor \frac{\deg(f)-1}{2} \right\rfloor$, we get another integral Weierstrass equation $w^2 = z^{2g+2}f(1/z)$, the scheme over R glued by affine schemes defined by these two equations is called a Weierstrass model.

Remark. *(1) Compare to Definition 6.3.2 and Definition 6.3.5, Corollary 6.3.6, this definition just wants to include the case where $\deg(f) = 1$ or 2. By Lemma 6.3.3(d), for $y^2 = f(x)$, $\deg(f) \geq 3$, which defines a hyperelliptic curve, Definition 6.3.2 and Definition 6.3.7 coincide. It is similar for Weierstrass models.*

(2) Notice that if $\deg(f) = 1$, $v(f) = 1$, then $R[x,y]/(y^2 - f(x))$ is not normal.

Chapter 7

Artin Conductors of Hyperelliptic Curves

In this chapter, we are going to prove Theorem 7.2.1. It is the inductive process, Corollary 3.3, in Obus and Srinivasan's paper [47]. Our proof is different from theirs. In a word, they consider a specific regular scheme for an integral (Weierstrass) equation $y^2 = f(x)$ of a hyperelliptic curve, but we consider the minimial desingularization of the corresponding Weierstrass model.

7.1 Conductors of Arithmetic Curves

Definition 7.1.1. *Let (R, k) be a DVR with fraction field K and perfect residue field k, $S = \mathrm{Spec}(R)$. Let $X \to S$ be a regular, proper, flat S-scheme whose generic fiber X_η is a smooth, geometrically connected curve of genus $g \geq 1$. For a prime ℓ which is different to $p = \mathrm{Char}(k)$, we have a ℓ-adic Galois representation*

$$\rho : \mathrm{Gal}(K^{\mathrm{sep}}/K) \to \mathrm{Aut}_{\mathbb{Q}_\ell}(V_\ell),$$

where $V_\ell = \mathrm{H}^1_{\acute{e}t}(X_{\overline{\eta}}, \mathbb{Q}_\ell)$, $X_{\overline{\eta}} = X_\eta \times_{\mathrm{Spec}(k(\eta))} \mathrm{Spec}(\overline{k(\eta)})$. Set $\delta = \delta(\rho)$, the Swan conductor of ρ, see [15, Section 2]. We define the Artin conductor of X as

$$\mathrm{Art}(X/R) := \chi(X_\eta) - \chi(X_s) - \delta,$$

where X_s is the special fiber of X, and χ is the Euler's characteristic for étale topology. When there is no confusion, we simply denote $\mathrm{Art}(X/R)$ by $\mathrm{Art}(X)$.

The conductor of ρ in the definition is well-defined, since ρ is potentially semistable, see [30, Théorème 3.5].

We collect some results for Artin conductors.

Proposition 7.1.2 ([51], Theorem 3 and [41], Theorem 10.4.47)**.** *Keep the notations as Definition 7.1.1. Suppose that the genus of $X_\eta \geq 2$, and X is the relative minimal regular normal crossing divisor model of X_η. Then the follwoing conditions are quivalent.*

(1) The action of the wild inertia group P_K on $\mathrm{H}^1_{\acute{e}t}(X_{\overline{\eta}}, \mathbb{Q}_\ell)$ is trivial.

(2) Every irreducible component C of X_s whose multiplicity in X_s is divisible by p satisfies the following condition: C is isomorphic to $\mathbb{P}_k$ and intersects with other components of X_s at exactly two points and these components have p-prime multiplicities in X_s.

If R is strictly Henselian and $\mathrm{Char}(k) > 0$, they also equivalent to

(3) The curve X_η has stable reduction over a tamely ramified extension of K.

In particular, in this case, the Swan conductor $\delta(\rho) = 0$.

One of the important cases is the one where X is the minimal regular model of X_η. Let R^{sh} be the strict henselization of R, $K^{\mathrm{sh}} = \mathrm{Frac}(R^{\mathrm{sh}})$. Then $X_{R^{\mathrm{sh}}} := X \times_{\mathrm{Spec}(R)} \mathrm{Spec}(R^{\mathrm{sh}})$ is the minimal regular model of $X_\eta \times_{\mathrm{Spec}(K)} \mathrm{Spec}(K^{\mathrm{sh}})$ and $\mathrm{Art}(X_{R^{\mathrm{sh}}}/R^{\mathrm{sh}}) = \mathrm{Art}(X/R)$. Hence to study the conductors, we can suppose R strictly henselian.

Corollary 7.1.3. *Let (R, k) be a DVR with fraction field K and perfect residue field k. Let C be a smooth, projective, geometrically connected curve of genus $g \geq 2$ over K. We suppose that $\mathrm{Char}(k) > 2g + 1$. Then the Swan condutor $\delta(\rho) = 0$, where ρ is the ℓ-adic Galoi representation as Definition 7.1.1.*

Proof. We can assume that R is strictly henselian. Then the corrollary is a consequence of Proposition 7.1.2 and [41, Proposition 10.4.45]. □

Proposition 7.1.4 ([39], Proposition 1)**.** *Keep the notations as Definition 7.1.1. Suppose that the gcd of the multiplicities of irreducible component of X_s is 1. Then we have*

$$-\mathrm{Art}(X) = n - 1 + f,$$

where f is the Artin conductor of the ℓ-adic representation $\mathrm{H}^1_{\acute{e}t}(X_{\overline{\eta}}, \mathbb{Q}_\ell)$ as defined in Definition 7.1.1, n is the number of irreducible components of $X_{\bar{s}} = X_s \times_{\mathrm{Spec}(k(s))} \mathrm{Spec}(\overline{k(s)})$.

Remark. *(1) The condition on the gcd of multiplicities is satified if $X_\eta(K) \neq \emptyset$ or if the genus of X_η is 2.*

(2) If we denote the abelian rank and the unipotent rank of $X_{\bar{s}}$ by a and u respectively, then by [53, Lemma 1, Lemma 2], $f = 2u + t + \delta$, where δ is the Swan conductor of the ℓ-adic representation $\mathrm{H}^1_{\acute{e}t}(X_{\overline{\eta}}, \mathbb{Q}_\ell)$ as defined in Definition 7.1.1.

Theorem 7.1.5 (Ogg's formula)**.** *Keep the notations as Definition 7.1.1. If X_η is an elliptic curve and X is the minimal regular model of X_η. Then we have*

$$-\mathrm{Art}(X) = v(\Delta),$$

where $v(\Delta)$ is the minimal discriminant of of X_η.

This equality isn't true for general hyperelliptic curves. In [39], Liu proved that $-\mathrm{Art}(X) \leq v(\Delta)$ when X_η is a projective curve of genus 2. Unitil recently, Obus and Srinivasan [47] showed that this inequality holds for any hyperelliptic curve when $\mathrm{Char}(k) \neq 2$, i.e. the following theorem

Theorem 7.1.6 ([47], Theorem 1.1)**.** *Keep the notations as Definition 7.1.1. If k is perfect and $\mathrm{Char}(k) \neq 2$, X_η is a hyperelliptic curve of genus $g \geq 1$, X is the minimal regular model of X_η. Then we have*

$$-\mathrm{Art}(X) \leq v(\Delta),$$

where $v(\Delta)$ is the minimal discriminant of X_η.

7.2 Main Results

From now on, R is a strictly henselian, discrete valuation ring with valuation v, fraction field K and residue field k of $\mathrm{Char}(k) \neq 2$. Let π be a uniformizer of R, $S = \mathrm{Spec}(R)$. For a projective curve C over K (or k), we set

$\mathrm{n}(C)$: the number of irreducible components of C,
$\mathrm{p}_a(C)$: the arithmetic genus of C,
$\mathrm{a}(C)$: the abelian rank of C,
$\mathrm{t}(C)$: the toric rank of C,
$\mathrm{u}(C)$: the unipotent rank of C.

The definition of these quantities can be found in [41, Section 7.5]. Notice that, if C is smooth and geometrically connected, the arithmetic genus and geometric genus of C coincide. Hence we also denote $\mathrm{p}_a(C)$ by $g(C)$.

For a Noetherian scheme X, the set of regular points (resp. singular points) on X is denoted by $\mathrm{Reg}(X)$ (resp. $\mathrm{Sing}(X)$). If X is a regular fibered surface over R, i.e. regular, proper, flat scheme over R of dimension 2, the Artin conductor (resp. Swan conductor) of X is denoted by $\mathrm{Art}(X)$ (resp. $\delta(X)$). The special fiber and generic fiber of X is denoted by X_s and X_η respectively. For a Weierstrass model Y over R of a hyperelliptic curve, we denote $v(\Delta(Y))$ the valuation of the discriminant of Y. To simplify the notation, we denote this nonnegative integer $-\mathrm{Art}(X) - \delta(X)$ by $-\mathrm{Art}_{\mathrm{tame}}(X)$, which is $(\mathrm{a}(X_s) + 2\mathrm{u}(X_s)) + \mathrm{n}(X_s) - 1$.

We are going to prove the following inductive process for Artin Conductors and discriminant of hyperelliptic curves.

Theorem 7.2.1. *Let Y, Y_1 and Y_2 be the Weierstrass models over R of hyperelliptic curves. Suppose that they are defined by integral Weierstrass equations in one of the following cases:*

1. *$Y : y^2 = f_1(x)f_2(x)$, $Y_1 : y^2 = f_1(x)$ and $Y_2 : y^2 = f_2(x)$,*
2. *$Y : y^2 = \pi f_1(x)f_2(x)$, $Y_1 : y^2 = \pi f_1(x)$ and $Y_2 : y^2 = \pi f_2(x)$,*

where, in both cases, $\deg(f_i) = \deg(\overline{f}_i) \geq 1$ for $i = 1, 2$, and $\overline{f}_1, \overline{f}_2 \in k[x]$ are coprime. If for $i = 1, 2$,

$$-\mathrm{Art}_{\mathrm{tame}}(X_i) \leq v(\Delta(Y_i)),$$

then

$$-\mathrm{Art}_{\mathrm{tame}}(X) \leq v(\Delta(Y)),$$

where X, X_1 and X_2 are the minimal desingularizations of Y, Y_1 and Y_2 respectively. Moreover, if the equality holds for Y_1 and Y_2, it also holds for Y.

Remark. *(1) Notice that if $\deg(f) = 1$, we define its valuation of discriminant to be 0.*

(2) The assumption that $\deg f_i = \deg \overline{f}_i$ for $i = 1, 2$ can always be achieved for a Weierstrass model $Y : y^2 = f_1(x)f_2(x)$ after a suitable change of coordinates, e.g $x = 1/(x' - a)$ for some $a \in R$. Here, this assumption can make our result cleaner.

Theorem 7.2.1 comes directly from the following two theorems.

Theorem 7.2.2. *Let C be a hyperelliptic curve over K with an integral Weierstrass equation $y^2 = f_1(x)f_2(x)$, where $\deg(f_i) = \deg(\overline{f}_i) \geq 1$ for $i = 1, 2$, and $\overline{f}_1, \overline{f}_2 \in k[x]$ are coprime. Let C_1, C_2 be the hyperelliptic curves over K determined by Weierstrass equations $y^2 = f_1(x)$ and $y^2 = f_2(x)$ respectively. Then $y^2 = f_1(x)$ and $y^2 = f_2(x)$ define respective Weierstrass models Y_1, Y_2 of C_1, C_2, and*

$$v(\Delta(Y)) = v(\Delta(Y_1)) + v(\Delta(Y_2)),$$

$$\mathrm{n}(X_s) = \begin{cases} \mathrm{n}(X_{1,s}) + \mathrm{n}(X_{2,s}) - 1 & \text{if } \mathrm{n}(Y_{1,s}) = \mathrm{n}(Y_{2,s}) = 1, \\ \mathrm{n}(X_{1,s}) + \mathrm{n}(X_{2,s}) - 2 & \text{otherwise,} \end{cases}$$

$$\mathrm{a}(X_s) = \begin{cases} \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}) & \text{if } \mathrm{n}(Y_{1,s}) = \mathrm{n}(Y_{2,s}) = 1 \text{ and one of } \deg(f_1), \deg(f_2) \text{ is even,} \\ \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}) + 1 & \text{otherwise,} \end{cases}$$

$$\mathrm{t}(X_s) = \begin{cases} \mathrm{t}(X_{1,s}) + \mathrm{t}(X_{2,s}) & \text{if } \mathrm{n}(Y_{1,s}) = \mathrm{n}(Y_{2,s}) = 1, \\ \mathrm{t}(X_{1,s}) + \mathrm{t}(X_{2,s}) + 1 & \text{otherwise,} \end{cases}$$

$$\mathrm{u}(X_s) = \mathrm{u}(X_{1,s}) + \mathrm{u}(X_{2,s})$$

$$-\mathrm{Art}_{\mathrm{tame}}(X) = -\mathrm{Art}_{\mathrm{tame}}(X_1) - \mathrm{Art}_{\mathrm{tame}}(X_2)$$

where Y, Y_1 and Y_2 are the Weierstrass models defined by $y^2 = f_1(x)f_2(x)$, $y^2 = f_1(x)$ and $y^2 = f_2(x)$ with X/S, X_1/S and X_2/S as the minimal desingularizations respectively, and X_s, $X_{1,s}$ and $X_{2,s}$ are the special fibers of X, X_1 and X_2 respectively.

Theorem 7.2.3. *Let C be a hyperelliptic curve over K with an integral Weierstrass equation $y^2 = \pi f_1(x)f_2(x)$, where $\deg(f_i) = \deg(\overline{f}_i) \geq 1$ for $i = 1, 2$, and $\overline{f}_1, \overline{f}_2 \in k[x]$ are coprime. Let C_1, C_2 be the hyperelliptic curves over K determined by Weierstrass equations $y^2 = \pi f_1(x)$ and $y^2 = \pi f_2(x)$ respectively. Then $y^2 = \pi f_1(x)$ and $y^2 = \pi f_2(x)$ define respective Weierstrass models Y_1, Y_2 of C_1, C_2, and*

$$v(\Delta(Y)) = \begin{cases} v(\Delta(Y_1)) + v(\Delta(Y_2)) + 2 & \text{if } \deg(f_1) \text{ or } \deg(f_2) \text{ is even,} \\ v(\Delta(Y_1)) + v(\Delta(Y_2)) - 2 & \text{if } \deg(f_1), \deg(f_2) \text{ are both odd,} \end{cases}$$

$$\mathrm{n}(X_s) = \mathrm{n}(X_{1,s}) + \mathrm{n}(X_{2,s}) - 1,$$

$$\mathrm{a}(X_s) = \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}),$$

$$\mathrm{t}(X_s) = \mathrm{t}(X_{1,s}) + \mathrm{t}(X_{2,s}),$$

$$\mathrm{u}(X_s) = \mathrm{u}(X_{1,s}) + \mathrm{u}(X_{2,s})$$

$$-\mathrm{Art}_{\mathrm{tame}}(X) = \begin{cases} -\mathrm{Art}_{\mathrm{tame}}(X_1) - \mathrm{Art}_{\mathrm{tame}}(X_2) + 2 & \text{if } \deg(f_1) \text{ or } \deg(f_2) \text{ is even,} \\ -\mathrm{Art}_{\mathrm{tame}}(X_1) - \mathrm{Art}_{\mathrm{tame}}(X_2) - 2 & \text{if } \deg(f_1), \deg(f_2) \text{ are both odd,} \end{cases}$$

where Y, Y_1 and Y_2 are the Weierstrass models defined by $y^2 = f_1(x)f_2(x)$, $y^2 = f_1(x)$ and $y^2 = f_2(x)$ with X/S, X_1/S and X_2/S as the minimal desingularizations respectively.

With these theorems, we have the main result of [61, Theorem 1.2]:

Corollary 7.2.4. *Let Y be a Weierstrass model over R defined by $y^2 = f(x)$ or $y^2 = \pi f(x)$ with $f(x) = (x - b_1) \cdots (x - b_n)$ and $b_1, \cdots, b_n \in R$, $n \geq 1$. Then*

$$-\mathrm{Art}_{\mathrm{tame}}(X) \leq v(\Delta(Y)),$$

where X is the minimal desingularization of Y. If moreover, $\mathrm{Char}(k) \geq 2g(Y_K) + 1$, then the Conductor-discriminant inequality holds for corresponding hyperelliptic curve C. i.e.

$$-\mathrm{Art}(\mathcal{X}) \leq v(\Delta(C)),$$

where $\mathcal{X}$ is the minimal proper regular model of C, and $v(\Delta(C))$ is the valuation of the minimal discriminant of C.

7.3 Lemmas

In this section, we prove some lemmas that are going to be used later, they may have been proved somewhere, but I cannot find the proofs, so we prove them here.

We will use [38, Corollary 27.3] to prove the following lemmas.

LEMMA 7.3.1. *Let X be a two dimensional Noetherian scheme with finite closed singular points. Then a morphism $\tau : \tilde{X} \to X$ of finite type is a minimal desingularization if and only if $\tilde{X} \times_X \mathrm{Spec}(\mathcal{O}_{X,x}) \to \mathrm{Spec}(\mathcal{O}_{X,x})$ is a minimal desingularization for each $x \in X$. In particular, they are also equivalent to that $\tilde{X} \times_X U \to U$ is a minimal desingularization for any open subset U of X.*

Proof. Suppose that $x_1, \cdots, x_n \in X$ are all singular points of X. Denote $\mathcal{FP}_{X,\{x_1,\cdots,x_n\}}$ the category of morphisms $f : Y \to X$ of finite presentation which induce an isomorphism $f^{-1}(U) \to U$, where $U = X \setminus \{x_1, \cdots, x_n\}$. The morphisms in $\mathcal{FP}_{X,\{x_1,\cdots,x_n\}}$ are morphisms of schemes over X. For each i we set $X_i = \mathrm{Spec}(\mathcal{O}_{X,x_i})$ and $V_i = X_i \setminus \{x_i\}$. Similarly, we can define the category $\mathcal{FP}_{X_i,x_i}$. Then by [63, Lemma 51.6.1], the functor defined by base change

$$F : \mathcal{FP}_{X,\{x_1,\cdots,x_n\}} \to \mathcal{FP}_{X_1,x_1} \times \cdots \times \mathcal{FP}_{X_n,x_n}$$

is an equivalence of categories. Furthermore, if $f : Y \to X$ corresponds to $f_i : Y_i \to X_i$ under F, then f is proper if and only if f_i is proper for $i = 1, \cdots, n$, see in [63, Lemma 51.6.2]. It is obvious that Y is regular if and only if Y_i is regular for $i = 1, \cdots, n$ because of our choice of $x_1, \cdots, x_n$.

If $\tau_x : \tilde{X} \times_X \mathrm{Spec}(\mathcal{O}_{X,x}) \to \mathrm{Spec}(\mathcal{O}_{X,x})$ is a minimal desingularization for each $x \in X$, then for any $x \notin \{x_1, \cdots, x_n\}$, τ_x is an isomorphism, hence we can find an open neighborhood U_x of x such that $\tau^{-1}(U_x) \to U_x$ is an isomorphism, which means that $\tau \in \mathcal{FP}_{X,\{x_1,\cdots,x_n\}}$, and τ is a desingularization of X. To show that it is minimal, we take any integral exceptional curve E on $\tilde{X}$ relative to τ, and the image of E is a close point $x \in X$, it is also an integral exceptional curve on $\tilde{X} \times_X \mathrm{Spec}(\mathcal{O}_{X,x})$, and by [38, Corollary 27.3], we have the self-intersection number $(E, E) \leq -2\chi(E)$, where $\chi(E)$ is the Euler-Poincáre characteristic of $\mathcal{O}_E$. Hence τ is minimal by [38, Corollary 27.3] again.

The converse statement is similar by [38, Corollary 27.3]. □

LEMMA 7.3.2. *Let A be a two dimensional Noetherian local ring. Assume that the only singular point of $\mathrm{Spec}(A)$ is the closed point. Then a morphism $\tau : X \to \mathrm{Spec}(A)$ is a minimal desingularization if and only if $\widehat{\tau} : \widehat{X} = X \times_{\mathrm{Spec}(A)} \mathrm{Spec}(\widehat{A}) \to \mathrm{Spec}(\widehat{A})$ is a minimal desingularization, where $\widehat{A}$ is the completion of A.*

Proof. If $\tau : X \to \mathrm{Spec}(A)$ is a minimal desingularization, then by [63, Lemma 51.11.2], and because $A \to \widehat{A}$ faithfully flat, we know that X is regular if and only if $\widehat{X}$ is regular, and τ is proper if and only if $\widehat{\tau}$ is proper. Notice that X and $\widehat{X}$ have isomorphic fibers over closed points, so they have same integral exceptional curves.

By [38, Corollary 27.3], it suffices to show that the equality of intersection $(E, E) = (\widehat{E}, \widehat{E})$ for any integral exceptional curve of X over π, where $\widehat{E}$ is the corresponding exceptional curve on $\widehat{X}$. Indeed, if $\mathcal{I}$ is the sheaf of ideal of E, then $\widehat{\mathcal{I}} = \mathcal{I}\mathcal{O}_{\widehat{X}}$ is the correspondent sheaf of ideal, it's the pullback of $\mathcal{I}$ along $\widehat{X} \to X$, so $\mathcal{I}|_E \simeq \mathcal{I}\mathcal{O}_{\widehat{X}}|_{\widehat{E}}$ via $E \simeq \widehat{E}$, so we get the equality we want.

□

LEMMA 7.3.3. *Let A be a discrete valuation ring with an algebraically closed residue field, and $f : X \to Y$ be an étale morphism of integral curves over A, i.e. schemes X, Y are integral, of finite type over A and of relative dimension 1. Assume that Y is normal, and $\tilde{Y} \to Y$ is a minimal desingularization of Y, then $\tilde{X} := X \times_Y \tilde{Y} \to X$ is a minimal desingularization of X.*

Proof. Firstly, we claim that X, Y have finite singular point, which are closed points on their special fibers. Indeed, since X and Y are integral, of dimension two, flat and of finite type over A, and A is a discrete valuation ring, then the singular points on X, Y are closed, by [41, Corollary 8.2.38]. Furthermore, since Y is normal, hence we know that Y has finite many singular point, which are closed points on its special fibers. The same result is true for X, since $X \to Y$ is étale.

By Lemma7.3.1, we check the statement locally. If $y \in Y$ is regular, then $\tilde{Y} \times_Y \mathrm{Spec}(\mathcal{O}_{Y,y}) \simeq \mathrm{Spec}(\mathcal{O}_{Y,y})$, so $X \times_Y \tilde{Y} \times_Y \mathrm{Spec}(\mathcal{O}_{Y,y}) \simeq X \times_Y \mathrm{Spec}(\mathcal{O}_{Y,y})$. For each $x \in X$ above y, we have x is regular and $\mathrm{Spec}(\mathcal{O}_{X,x}) \times_Y \mathrm{Spec}(\mathcal{O}_{Y,y}) = \mathrm{Spec}(\mathcal{O}_{X,x})$, so $\mathrm{Spec}(\mathcal{O}_{X,x}) \times_X \widetilde{X} \simeq \mathrm{Spec}(\mathcal{O}_{X,x}) \times_X (X \times_Y \tilde{Y}) \simeq (\mathrm{Spec}(\mathcal{O}_{X,x}) \times_Y \mathrm{Spec}(\mathcal{O}_{Y,y})) \times_Y \tilde{Y} \simeq \mathrm{Spec}(\mathcal{O}_{X,x})$. If $y \in Y$ is singular, then $y \in Y_s$ is a closed point, and $\tilde{Y} \times_Y \mathrm{Spec}(\widehat{\mathcal{O}}_{Y,y}) \to \mathrm{Spec}(\widehat{\mathcal{O}}_{Y,y})$ is the minimal desingularization of $\mathrm{Spec}(\widehat{\mathcal{O}}_{Y,y})$. For each $x \in X$ above y, since the residue fields $k(x) = k(y) = k$, and f is étale, so $\widehat{\mathcal{O}}_{X,x} \simeq \widehat{\mathcal{O}}_{Y,y}$ by [41, Proposition 4.3.26]. Hence $\mathrm{Spec}(\widehat{\mathcal{O}}_{X,x}) \times_X (X \times_Y \tilde{Y}) \simeq \mathrm{Spec}(\widehat{\mathcal{O}}_{X,x}) \times_Y \tilde{Y} \to \mathrm{Spec}(\widehat{\mathcal{O}}_{X,x})$ is a minimal desingularization, so is $\mathrm{Spec}(\mathcal{O}_{X,x}) \times_X (X \times_Y \tilde{Y}) \simeq \mathrm{Spec}(\mathcal{O}_{X,x}) \times_Y \tilde{Y} \to \mathrm{Spec}(\mathcal{O}_{X,x})$ by Lemma7.3.2. □

Recall that, a morphism $f : X \to S$ is said to be a fibered surface, if with S is a Dedekind scheme and f is projective and flat, and X normal of dimension 2. As usual, the generic fiber is denoted by X_η.

LEMMA 7.3.4. *Let $f : X \to S$ be a normal fibered surface with S Dedekind scheme of dimension 1. Suppose that X_η is geometrically integral. Let $s \in S$ be a closed point such that one of the following conditions holds:*

(1) $\mathrm{Char}(k(s)) = 0$*;*

(2) d is prime to $\mathrm{Char}(k(s))$*, where d is the greatest common divisor of the multiplicities of the irreducible components of X_s.*

Then f is cohomologically flat at s. In particular, for any $i \geq 0$.

$$\dim \mathrm{H}^i(X_s, \mathcal{O}_{X_s}) = \dim \mathrm{H}^i(X_\eta, \mathcal{O}_{X_\eta}).$$

Moreover, if $S = \mathrm{Spec}(A)$ is affine, then the canonical map

$$\mathrm{H}^i(X, \mathcal{O}_X) \otimes_A k(s) \to \mathrm{H}^i(X_s, \mathcal{O}_{X_s})$$

is an isomorphism for each $i \geq 0$.

Proof. It is a statement in the introduction of [50]

We can suppose that $S = \mathrm{Spec}(A)$ is affine, furthermore, suppose that A is a discrete valuation ring. We have to check that $\mathcal{O}_X(X) = A$. Indeed, $\mathcal{O}_X(X)$ is finite over A, and since X_η is geometrically integral, so $\mathcal{O}_X(X) \subset \mathcal{O}_{X_\eta}(X_\eta) = \mathrm{Frac}(A)$, hence $\mathcal{O}_X(X) = A$. By the assumption on $k(s)$ and [50], we know that f is cohomologically flat.

The rest statements come from [28, Proposition 7.8.4 (e)] and [41, Theorem 5.3.20 (a)]. □

7.4 Proof of Theorem 7.2.2

We fix some notations during this section. Set $g = \mathrm{p}_a(Y_K)$, $g_i = \mathrm{p}_a(Y_{i,K})$, $i = 1, 2$, then we have

$$g = \begin{cases} g_1 + g_2 + 1 & \text{if } \deg(f_1) \text{ or } \deg(f_2) \text{ is even,} \\ g_1 + g_2 & \text{if } \deg(f_1), \deg(f_2) \text{ are both odd.} \end{cases} \tag{7.1}$$

Let

$$Y' = \mathrm{Spec}(\frac{R[x,y]}{(y^2 - f_1(x)f_2(x))}),$$

$$Y'' = \mathrm{Spec}(\frac{R[w,z]}{(z^2 - w^{2g+2}f_1(1/w)f_2(1/w))})$$

with $w = 1/x, z = y/x^{g+1}$. Then Y is covered by Y' and Y''. Similarly, for $i = 1$ or 2, Y_i is covered by Y_i' and Y_i'', where

$$Y_i' = \mathrm{Spec}(\frac{R[x,y]}{(y^2 - f_i(x))}),$$

$$Y_i'' = \mathrm{Spec}(\frac{R[w,z]}{(z^2 - w^{2g_i+2}f_i(1/w))})$$

with $w = 1/x, z = y/x^{g_i+1}$. We also set X', X'', X_i' and X_i'' the minimal desingularization of Y', Y'', Y_i' and Y_i'' respectively, $i = 1, 2$.

7.4.1 Discriminants

Since that $y^2 = f_1(x)f_2(x)$ is an integral Weierstrass equation of C, then $v(f_1(x)f_2(x)) = v(f_1(x)) + v(f_2(x)) = 0$, so $v(f_1(x)) = 0$ and $v(f_2(x)) = 0$, which means that $y^2 = f_1(x)$ and $y^2 = f_2(x)$ are integral Weierstrass equations, Definition 6.3.7.

The equality $v(\Delta(Y)) = v(\Delta(Y_1)) + v(\Delta(Y_2))$, follows directly from the fact that $\overline{f}_1, \overline{f}_2 \in k[x]$ are coprime and the leading coefficients of f_1 and f_2 are units.

7.4.2 Singular points

We analyze the singular points on Y. Since Y is normal of dimension 2 and $\mathrm{Reg}(Y)$ is open, so $\mathrm{Sing}(Y)$ is finite. Moreover, $\mathrm{Sing}(Y) \subset Y_s$.

Notice that R is strictly henselian with k perfect, so $k = \overline{k}$ and every singular closed point on $Y_s, Y_{1,s}$ and $Y_{2,s}$ is rational. We also know that $\mathrm{Sing}(Y) \subset \mathrm{Sing}(Y_s)$, then by Jacobian criteria for smoothness and the fact that $\deg(f_1) = \deg(\overline{f}_1), \deg(f_2) = \deg(\overline{f}_2)$, we have $\mathrm{Sing}(Y)$ is contained in Y', moreover, contained in

$$\left\{ (x - \overline{a}, y) \subset \frac{k[x,y]}{(y^2 - \overline{f}_1(x)\overline{f}_2(x))} \mid \overline{a} \in k \text{ with } \overline{f}_1(\overline{a}) = \overline{f}_1'(\overline{a}) = 0 \text{ or } \overline{f}_2'(\overline{a}) = \overline{f}_2(\overline{a}) = 0 \right\}.$$

It's similar for Y_1 and Y_2. We set

$$S_1 = \{(x - \overline{a}, y) \in \mathrm{Sing}(Y) \mid \overline{f}_1(\overline{a}) = \overline{f}_1'(\overline{a}) = 0\},$$

$$S_2 = \{(x - \overline{a}, y) \in \mathrm{Sing}(Y) \mid \overline{f}_2(\overline{a}) = \overline{f}_2'(\overline{a}) = 0\},$$

notice that S_1 an S_2 are disjoint, since $\overline{f}_1, \overline{f}_2 \in k[x]$ are coprime.

7.4.3 Étale coverings around singular points

In this subsection, we construct an étale covering around singular points on Y_1, which is also an étale covering near corresponding singular points on Y. It's similar for Y_2.

We set

$$W = \mathrm{Spec}(R[x,y,z]/(y^2 - f_1(x)f_2(x), z^2 - f_2(x)),$$

and we have an obvious morphism $W \to Y'$. We claim that the induced morphism $\mathrm{D}_W(z) \to \mathrm{D}_{Y'}(f_2(x))$ is étale. Indeed, $\frac{R[x,y,z]}{(y^2-f_1(x)f_2(x),z^2-f_2(x))}$ is a free $\frac{R[x,y]}{(y^2-f_1(x)f_2(x))}$-module of rank 2, and for each $P \in \mathrm{D}_{Y'}(f_2(x))$, we have an isomorphism of $k(P)$-algebras

$$\frac{k(P)[z]}{(z^2 - \overline{f_2(x)})} \simeq k(P) \times k(P),$$

since $\overline{f_2(x)} \neq 0$ in $k(P)$ and $\mathrm{Char}(k(P)) \neq 2$.

On the other hand, the morphism of R-algebras

$$\frac{R[x,y]}{(y^2 - f_1(x))} \to \left(\frac{R[x,y,z]}{(y^2 - f_1(x)f_2(x), z^2 - f_2(x))}\right)_z,\ x \mapsto x,\ y \mapsto y/z \tag{7.2}$$

induces an étale morphism $\mathrm{D}_W(z) \to \mathrm{D}_{Y_1'}(f_2(x))$. Indeed, notice that

$$\left(\frac{R[x,y,z]}{(y^2 - f_1(x)f_2(x), z^2 - f_2(x))}\right)_z \simeq \left(\frac{R[x,w,z]}{(w^2 - f_1(x), z^2 - f_2(x))}\right)_z, x \mapsto x,\ y \mapsto zw,\ z \mapsto z$$

via this isomorphism, the morphism in 7.2 is induced by

$$\frac{R[x,y]}{(y^2 - f_1(x))} \to \frac{R[x,w,z]}{(w^2 - f_1(x), z^2 - f_2(x))}, x \mapsto x, y \mapsto w.$$

We can see that $\frac{R[x,w,z]}{(w^2-f_1(x),z^2-f_2(x))}$ is a free $\frac{R[x,y]}{(y^2-f_1(x))}$-module of rank 2. We consider the scheme $W' = \mathrm{Spec}(\frac{R[x,w,z]}{(w^2-f_1(x),z^2-f_2(x))})$, then $\mathrm{D}_{W'}(z) = \mathrm{D}_{W'}(f_2(x)) \to \mathrm{D}_{Y_1'}(f_2(x))$ is flat, moreover, it is étale as before.

Set $V_1 = \mathrm{D}_{Y'}(f_2(x)), U_1 = \mathrm{D}_{Y_1'}(f_2(x))$ and $\tilde{V}_1 = X' \times_{Y'} V_1, \tilde{U}_1 = X_1' \times_{Y_1'} U_1$, and replace W by $\mathrm{D}_W(z)$. Notice that the singular poins of Y in S_1 are on V_1 and all singgular points of Y_1 are on U_1. Hence the étale morphisms $W \to V_1$ and $W \to U_1$ induce a bijection between $\mathrm{Sing}(Y_1)$ and S_1, and corresponding points have isomorphic completions of local rings. This is essential when we count the number of irreducible components of X_s.

On the other hand, let $\tilde{W}$ be the minimal desingularization of W. Then by Lemma 7.3.1 and Lemma 7.3.3 we have $\tilde{W} \simeq W \times_{V_1} \tilde{V}_1 \simeq W \times_{U_1} \tilde{U}_1$. Moreover, via $\tilde{W} \to \tilde{V}_1$, every point in $\tilde{V}_1$ has two preimages. This fact will be used afterwards.

It's similar for Y_2, and we can set V_2, U_2 similarly as we do to V_1, U_2.

Notice that $\mathrm{Sing(Y)} \subset V_1 \bigcup V_2$, $\mathrm{Sing(X_s)} \subset \tilde{V}_{1,s} \bigcup \tilde{V}_{2,s}$, and it is similar for Y_1, Y_2, we indeed have proved the following lemma:

LEMMA 7.4.1. *(1) There is a bijection*

$$\mathrm{Sing}(Y) \to \mathrm{Sing}(Y_1) \bigcup \mathrm{Sing}(Y_2)$$

such that the corresponding points have the same completion of local rings.

(2) There is a bijection

$$\mathrm{Sing}(X_s) \to \mathrm{Sing}(X_{1,s}) \bigcup \mathrm{Sing}(X_{2,s})$$

such that the corresponding points have the same completion of local rings.

7.4.4 Calculations

Irreducible components

There are inequalities between $\mathrm{n}(Y_s)$, $\mathrm{n}(Y_{1,s})$ and $\mathrm{n}(Y_{2,s})$. It is easy to see that $\mathrm{n}(Y_s) = 1$ if and only if $\overline{f}_1\overline{f}_2$ is not a square in $k[x]$. Moreover, it is also equivalent to that $\overline{f}_1$ or $\overline{f}_2$ is not a square, since $\overline{f}_1, \overline{f}_2 \in k[x]$ are coprime. Hence $\mathrm{n}(Y_s) = 1$ if and only if $\mathrm{n}(Y_{1,s}) = 1$ or $\mathrm{n}(Y_{2,s}) = 1$, and $\mathrm{n}(Y_s) = 2$ if and only if $\mathrm{n}(Y_{1,s}) = \mathrm{n}(Y_{2,s}) = 2$. We have the follwoing inequalities:

$$\mathrm{n}(Y_s) = \min\{\mathrm{n}(Y_{1,s}), \mathrm{n}(Y_{2,s})\} = \begin{cases} \mathrm{n}(Y_{1,s}) + \mathrm{n}(Y_{2,s}) - 1 & \text{if } \mathrm{n}(Y_{1,s}) = \mathrm{n}(Y_{2,s}) = 1, \\ \mathrm{n}(Y_{1,s}) + \mathrm{n}(Y_{2,s}) - 2 & \text{otherwise.} \end{cases} \tag{7.3}$$

We have that the genric points of these curves are reduced, hence they are regular and $\mathrm{Sing}(Y_s), \mathrm{Sing}(Y_{1,s}), \mathrm{Sing}(Y_{2,s})$ are finite sets of closed points. Indeed, if Y_s has a generic point that is not reduced, then $y^2 - \overline{f}_1(x)\overline{f}_2(x) = (y - \overline{g}(x))^2$ in $k(x)[y]$, so $\overline{f}_1(x)\overline{f}_2(x) = 0$, since $\mathrm{Char}(k) \neq 2$, but $\overline{f}_1(x)\overline{f}_2(x)$ are coprime in $k[x]$, hence we get a contradiction. Similar for $Y_{1,s}$ and $Y_{2,s}$.

For each singular point $y \in Y$, by Lemma7.3.2 and Lemma7.3.1, we have that $Z \to \mathrm{Spec}(\mathcal{O}_{Y,y})$ and $\widehat{Z} \to \mathrm{Spec}(\widehat{\mathcal{O}}_{Y,y})$ are the minimal desingularizations, where $Z = X \times_Y \mathrm{Spec}(\mathcal{O}_{Y,y})$, $\widehat{\mathcal{O}}_{Y,y}$ is the completion of $\mathcal{O}_{Y,y}$, and $\widehat{Z} = Z \times_{\mathrm{Spec}(\mathcal{O}_{Y,y})} \mathrm{Spec}(\widehat{\mathcal{O}}_{Y,y})$. Note that $\mathcal{O}_{Y,y}$ and $\widehat{\mathcal{O}}_{Y,y}$ have the same residue field, so the fibers over closed points are same, that is $\widehat{Z}_{\widehat{y}} = X_y$, where $\widehat{Z}_{\widehat{y}} = \widehat{Z} \times_{\mathrm{Spec}(\widehat{\mathcal{O}}_{Y,y})} \mathrm{Spec}(k(\widehat{y}))$, $\widehat{y}$ is the close point on $\mathrm{Spec}(\widehat{\mathcal{O}}_{Y,y})$ and X_y is the fiber on X over y. Similar results are true for Y_1 and Y_2. Hence, by Lemma 7.4.1, we have

$$\begin{aligned} \mathrm{n}(X_s) &= \mathrm{n}(Y_s) + \sum_{y \in \mathrm{Sing}(Y)} \mathrm{n}(X_y) \\ &= \mathrm{n}(Y_s) + \sum_{y \in \mathrm{Sing}(Y)} \mathrm{n}(\widehat{Z}_{\widehat{y}}) \\ &= \mathrm{n}(Y_s) + \sum_{y_1 \in \mathrm{Sing}(Y_1)} \mathrm{n}(\widehat{Z}_{\widehat{y}_1}) + \sum_{y_2 \in \mathrm{Sing}(Y_2)} \mathrm{n}(\widehat{Z}_{\widehat{y}_2}) \\ &= \mathrm{n}(Y_s) + \sum_{y_1 \in \mathrm{Sing}(Y_1)} \mathrm{n}(X_{1,y_1}) + \sum_{y_2 \in \mathrm{Sing}(Y_2)} \mathrm{n}(X_{2,y_2}) \end{aligned}$$

Combining this with 7.3, then

$$\mathrm{n}(X_s) = \begin{cases} \mathrm{n}(X_{1,s}) + \mathrm{n}(X_{2,s}) - 1 & \text{if } \mathrm{n}(Y_{1,s}) = \mathrm{n}(Y_{2,s}) = 1, \\ \mathrm{n}(X_{1,s}) + \mathrm{n}(X_{2,s}) - 2 & \text{otherwise.} \end{cases} \tag{7.4}$$

Ranks

We demonstrate the method to calculate the ranks firstly. From [41, Lemma 7.5.11, Lemma 7.5.18 and Theorem 7.5.19], for a connected projective curve C (not necessarily reduced) over an algebraically closed field k with the normalization $\sigma : C' = \coprod_{i=1}^{n} C'_i \to C$ and C'_i connected components of C', we have the following formulas

$$\mathrm{t}(C) = \mu(C) - \mathrm{n}(C) + 1, \tag{7.5}$$

$$\mathrm{a}(C) = \sum_{1 \leq i \leq n} \mathrm{p}_a(C'_i), \tag{7.6}$$

$$\mathrm{u}(C) = \dim_k \mathrm{H}^1(C, \mathcal{O}_C) - \mathrm{a}(C) - \mathrm{t}(C), \tag{7.7}$$

where $\mu(C) = \sum_{x \in C(k)} (m_x - 1)$, and $m_x = |\sigma^{-1}(x)|$. For m_x, by [41, Theorem 8.2.39(c)], we know that $\mathcal{O}_{C,x}$ is excellent, and m_x is the number of maximal ideals of $\mathcal{O}'_{C,x}$, where $\mathcal{O}'_{C,x}$ is the normalization of $\mathcal{O}_{C,x}$, which is finite over $\mathcal{O}_{C,x}$ by [41, Proposition 8.2.41(b)]. By (c) of this proposition, the number m_x also equals to the number of the irreducible components of $\mathrm{Spec}(\widehat{\mathcal{O}_{C,x}/\sqrt{0}}) = \mathrm{Spec}(\widehat{\mathcal{O}}_{C,x}/\sqrt{0}\widehat{\mathcal{O}}_{C,x})$, which is exactly the number of irreducible components of $\mathrm{Spec}(\widehat{\mathcal{O}}_{C,x})$, where $\widehat{\mathcal{O}}_{C,x}$ is the completion of $\mathcal{O}_{C,x}$.

Notice that R is a discrete valuation ring, X, X_1 and X_2 are projective over R with geometrically connected generic fibers, see [63, Lemma 51.16.11], then X_s, $X_{1,s}$ and $X_{2,s}$ are connected by Zariski's connectedness principle, see [41, Theorem 5.3.15].

By Lemma 7.4.1 (2), we have

$$\sum_{x \in X_s(k)} (m_x - 1) = \sum_{x \in X_{1,s}(k)} (m_x - 1) + \sum_{x \in X_{2,s}(k)} (m_x - 1),$$

and

$$\mathrm{t}(X_s) = \begin{cases} \mathrm{t}(X_{1,s}) + \mathrm{t}(X_{2,s}) & \text{if } \mathrm{n}(Y_{1,s}) = \mathrm{n}(Y_{2,s}) = 1, \\ \mathrm{t}(X_{1,s}) + \mathrm{t}(X_{2,s}) + 1 & \text{otherwise.} \end{cases} \tag{7.8}$$

Notice that Y_s is reduced, so the greatest common divisor of the multiplicities of the irreducible components of X_s is 1. Hence by Lemma 7.3.4, we have

$$\dim_k \mathrm{H}^1(X_s, \mathcal{O}_{X_s}) = \dim_K \mathrm{H}^1(X_K, \mathcal{O}_{X_K}) = \mathrm{p}_a(X_K) = g, \tag{7.9}$$

and it is similar for $X_{1,s}$ and $X_{2,s}$.

Next, we calculate the abelian ranks. We only consider the irreducible components of X_s that dominate the irreducible components of Y_s, since other components are the same as the corresponding ones of $X_{1,s}$ and $X_{2,s}$

If Y_s is irreducible, then there is only one irreducible component of X_s which dominant Y_s, denoted by C. Let $C' \to C$ be the normalization, hence $C' \to Y_s$ is also the normalization of Y_s. Suppose that

$$\overline{f}_1(x) = b_1 \prod_{i=1}^{d_1} (x - x_{1,i})^{e_{1,i}},$$

$$\overline{f}_2(x) = b_2 \prod_{j=1}^{d_2} (x - x_{2,j})^{e_{2,j}},$$

with $x_{1,i}$ and $x_{2,i}$ all distinct. Then C' will be defined by

$$w^2 = b_1 b_2 \prod_{e_{1,i}\text{odd}} (x - x_{1,i}) \prod_{e_{2,j}\text{odd}} (x - x_{2,j}).$$

Let $m_1 = \sum_{e_{1,i}\text{odd}} 1$ and $m_2 = \sum_{e_{2,j}\text{odd}} 1$, then $m_1 \equiv \deg(\overline{f}_1) \equiv \deg(f_1) \pmod 2$ and $m_2 \equiv \deg(\overline{f}_2) \equiv \deg(f_2) \pmod 2$. Since Y_s irreducible, we have $m_1 + m_2 > 0$ and $\mathrm{p}_a(C') = \left\lfloor \frac{m_1+m_2-1}{2} \right\rfloor$. If $\mathrm{n}(Y_{1,s}) = \mathrm{n}(Y_{2,s}) = 1$, then $\mathrm{p}_a(C'_1) = \left\lfloor \frac{m_1-1}{2} \right\rfloor$ and $\mathrm{p}_a(C'_2) = \left\lfloor \frac{m_2-1}{2} \right\rfloor$. In this case, with easy calculation and 7.1, 7.6, 7.7,7.8, 7.9, we have

$$\mathrm{p}_a(C') = \begin{cases} \mathrm{p}_a(C'_1) + \mathrm{p}_a(C'_2) + 1 & \text{if } \deg(f_1) \text{ or } \deg(f_2) \text{ is even,} \\ \mathrm{p}_a(C'_1) + \mathrm{p}_a(C'_2) & \text{if } \deg(f_1) \text{ and } \deg(f_2) \text{ are both odd,} \end{cases}$$

$$\mathrm{a}(X_s) = \begin{cases} \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}) + 1 & \text{if } \deg(f_1) \text{ or } \deg(f_2) \text{ is even,} \\ \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}) & \text{if } \deg(f_1) \text{ and } \deg(f_2) \text{ are both odd,} \end{cases}$$
$$\mathrm{u}(X_s) = \mathrm{u}(X_{1,s}) + \mathrm{u}(X_{2,s}).$$

If exactly one of $\mathrm{n}(Y_{1,s}), \mathrm{n}(Y_{2,s})$ equals to 2, we assume that $\mathrm{n}(Y_{1,s}) = 1$ and $\mathrm{n}(Y_{2,s}) = 2$. In this case, $m_2 = 0, \mathrm{p}_a(C_2') = 0$, so

$$\mathrm{p}_a(C') = \mathrm{p}_a(C_1') + \mathrm{p}_a(C_2'),$$
$$\mathrm{a}(X_s) = \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}),$$
$$\mathrm{u}(X_s) = \mathrm{u}(X_{1,s}) + \mathrm{u}(X_{2,s}).$$

If Y_s is not irreducible,we suppose that

$$\overline{f}_1(x) = \overline{h}_1(x)^2,$$
$$\overline{f}_2(x) = \overline{h}_2(x)^2.$$

Then $Z_1 : y = \overline{h}_1(x)\overline{h}_2(x)$ and $Z_2 : y = -\overline{h}_1(x)\overline{h}_2(x)$ are the irreducible components of Y_s, they are normal and have genus 0. It's similar for $Y_{1,s}$ and $Y_{2,s}$. Hence we have

$$\mathrm{a}(X_s) = \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}),$$

and

$$\mathrm{u}(X_s) = \mathrm{u}(X_{1,s}) + \mathrm{u}(X_{2,s})$$

Combining these two cases, we have the equalities in Theorem 7.2.2. Notice that the final inequality follows from Proposition 7.1.4:

$$\begin{aligned} -\mathrm{Art}_{\mathrm{tame}}(X) &= \mathrm{n}(X_s) - 1 + 2\mathrm{u}(X_s) + \mathrm{t}(X_s) \\ &= 2\mathrm{u}(X_s) + \mu(X_s). \end{aligned}$$

7.5 Proof of Theorem 7.2.3

The idea is almost the same as the proof of Theorem 7.2.2, but we will encounter more difficult situation. As before, we fix some nations. Set $g = \mathrm{p}_a(Y_K)$, $g_i = \mathrm{p}_a(Y_{i,K})$, $i = 1, 2$, then we have the following simple but important fact

$$g = \begin{cases} g_1 + g_2 + 1 & \text{if } \deg(f_1) \text{ or } \deg(f_2) \text{ is even,} \\ g_1 + g_2 & \text{if } \deg(f_1), \deg(f_2) \text{ are both odd.} \end{cases} \tag{7.10}$$

We call the case where $\deg(f_1)$ or $\deg(f_2)$ is even case 1, and the other case case 2. Let

$$Y' = \mathrm{Spec}\left(\frac{R[x,y]}{(y^2 - \pi f_1(x) f_2(x))}\right),$$
$$Y'' = \mathrm{Spec}\left(\frac{R[w,z]}{(z^2 - \pi w^{2g+2} f_1(1/w) f_2(1/w))}\right)$$

with $w = 1/x, z = y/x^{g+1}$. Then Y is covered by Y' and Y''. Similarly, for $i = 1$ or 2, Y_i is covered by Y_i' and Y_i'', where

$$Y_i' = \mathrm{Spec}\left(\frac{R[x,y]}{(y^2 - \pi f_i(x))}\right),$$
$$Y_i'' = \mathrm{Spec}\left(\frac{R[w,z]}{(z^2 - \pi w^{2g_i+2} f_i(1/w))}\right)$$

with $w = 1/x, z = y/x^{g_i+1}$. We also set X', X'', X_i' and X_i'' the minimal desingularization of Y', Y'', Y_i' and Y_i'' respectively, $i = 1, 2$.

For the convenience of readers, we mention some claims which are not exactly the same as section 7.4.

7.5.1 Discriminants

The proof of that $y^2 = \pi f_1(x)$ and $y^2 = \pi f_2(x)$ are integral Weierstrass equation is the same as the one in Subsection 7.4.1. By Definition 6.2.4, and the fact that $\overline{f}_1, \overline{f}_2$ are coprime, we have

$$\begin{aligned} v(\Delta(Y)) &= 1 - (-1)^{\deg(f_1 f_2)} + v(\mathrm{disc}(\pi f_1 f_2)) \\ &= 1 - (-1)^{\deg(f_1 f_2)} + 2(\deg(f_1) + \deg(f_2) - 1) + v(\mathrm{disc}(f_1)) + v(\mathrm{disc}(f_2)) \\ &= 1 - (-1)^{\deg(f_1 f_2)} + 2 + v(\mathrm{disc}(\pi f_1)) + v(\mathrm{disc}(\pi f_2)) \\ &= v(\Delta(Y_1)) + v(\Delta(Y_2)) + 2 + ((-1)^{\deg(f_1)} + (-1)^{\deg(f_2)} - (-1)^{\deg(f_1)+\deg(f_2)} - 1) \\ &= v(\Delta(Y_1)) + v(\Delta(Y_2)) + 2 - (1 - (-1)^{\deg(f_1)})((-1)^{\deg(f_2)} - 1), \end{aligned}$$

i.e.

$$v(\Delta(Y)) = \begin{cases} v(\Delta(Y_1)) + v(\Delta(Y_2)) + 2 & \text{if } \deg(f_1) \text{ or } \deg(f_2) \text{ is even;} \\ v(\Delta(Y_1)) + v(\Delta(Y_2)) - 2 & \text{if } \deg(f_1), \deg(f_2) \text{ are both odd.} \end{cases}$$

7.5.2 Singular points

Firstly, let $f(x) = f_1(x)f_2(x)$, then a close point $P \in Y'$ with corresponding maximal ideal $(x - a, y, \pi)$ is singular if and only if $\overline{f(a)} = 0$ in k. Indeed, a singular point must be on the special fibers, so the maximal ideal $\mathfrak{m}$ has the form $(x - a, y, \pi) \subset R[x, y]$. Set

$$I := \mathfrak{m}^2 + (y^2 - \pi f(x)) = ((x-a)^2, y^2, \pi^2, \pi(x-a), \pi y, y(x-a), y^2 - \pi f(x)).$$

Notice that $(x - a, \pi) + I \neq \mathfrak{m}$ and $(y, \pi) + I \neq \mathfrak{m}$, so P is singular if and only if $(x - a, y) + I \neq \mathfrak{m}$, that is $(x - a, y, \pi^2, \pi f(a)) \neq \mathfrak{m}$, which is equivalent to that $\overline{f(a)} = 0$ in k.

We set $Q \in Y''$ corresponding to the ideal (w, z, π) and $B = \{Q\}$ if $w = 0$ is a zero of $w^{2g+2}\overline{f}(1/w) = 0$, i.e. $\deg(f)$ is odd. Otherwise, we set $B = \emptyset$. Hence,

$$\mathrm{Sing}(Y) = \left\{ (x - a, y, \pi) \subset Y' \mid \overline{a} \in k \text{ with } \overline{f}_1(\overline{a}) = 0 \text{ or } \overline{f}_2(\overline{a}) = 0 \right\} \bigcup B.$$

It's similar for Y_1 and Y_2, and we can define Q_1, Q_2, B_1 and B_2 in the similar way.

7.5.3 Étale coverings around singular points

We replace W in Subection 7.4.3 to be

$$W = \mathrm{Spec}(R[x, y, z]/(y^2 - \pi f_1(x) f_2(x), z^2 - f_2(x)),$$

then we have étale morphisms. Similarly, we have the following lemma:

LEMMA 7.5.1. *(1) There is a bijection*

$$\mathrm{Sing}(Y') \to \mathrm{Sing}(Y_1') \bigcup \mathrm{Sing}(Y_2')$$

such that the corresponding points have the same completion of local rings.

(2) There is a bijection

$$\mathrm{Sing}(X_s') \to \mathrm{Sing}(X_{1,s}') \bigcup \mathrm{Sing}(X_{2,s}')$$

such that the corresponding points have the same completion of local rings.

With this lemma, we can deduce the similar lemma for Y.

LEMMA 7.5.2. *In case 1, the following statements hold:*

(1) There is a bijection

$$\mathrm{Sing}(Y) \to \mathrm{Sing}(Y_1) \bigcup \mathrm{Sing}(Y_2)$$

such that the corresponding points have the same completion of local rings.

(2) There is a bijection

$$\mathrm{Sing}(X_s) \to \mathrm{Sing}(X_{1,s}) \bigcup \mathrm{Sing}(X_{2,s})$$

such that the corresponding points have the same completion of local rings.

LEMMA 7.5.3. *In case 2, the following statements hold:*

(1) There is a bijection

$$\mathrm{Sing}(Y) \to \mathrm{Sing}(Y_1') \bigcup \mathrm{Sing}(Y_2')$$

such that the corresponding points have the same completion of local rings.

(2) There is a bijection

$$\mathrm{Sing}(X_s) \to \mathrm{Sing}(X_{1,s}') \bigcup \mathrm{Sing}(X_{2,s}')$$

such that the corresponding points have the same completion of local rings.

Hence, in case 2, there is no point corresponding to Q_1 and Q_2.

7.5.4 Calculations

A special case

For the calculation in case 2, we will need to know some quantities of the Weierstrass model defined by $y^2 = \pi x$ and its minimal desingularization. The results in this subsection are also useful in the proof of Corollary 7.2.4.

If $Y : y^2 = \pi x$, then Y is covered by the affine open subschemes $\mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi x)})$ and $\mathrm{Spec}(\frac{R[w,z]}{(z^2-\pi w)})$ with $w = 1/x, z = y/x$. We know that Y has two singular points (x, y, π) and (z, w, π), the blowup of $\mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi x)})$ at (x, y, π) is glued by 3 affine open subsets: $\mathrm{Spec}(\frac{R[x,y]}{(y^2-x)})$, $\mathrm{Spec}(\frac{R[x,y,z]}{(y^2-x,\pi-xz)})$ and $\mathrm{Spec}(\frac{R[x,y,z]}{(1-xy,\pi-yz)})$. Obviously, the first one is smooth over R, for the second and third, they are regular. Indeed, for each $\bar{b} \in k$, the point $(x, y, z - \bar{b})$ is singular on special fiber $\mathrm{Spec}(\frac{k[x,y,z]}{(y^2-x,xz)})$, so we consider $\mathfrak{m} = (x, y, z - b, \pi) \in \mathrm{Spec}(R[x, y, z])$, and set

$$I = \mathfrak{m}^2 + (y^2 - x, \pi - xz) = (x^2, y^2, (z-b)^2, \pi^2, \pi x, \pi y, \pi(z-b), xy, x(z-b), y(z-b), y^2 - x, \pi - xz).$$

With calculation, we have $I + (y, z - b) = \mathfrak{m}$, so $\mathfrak{m}$ is regular on $\mathrm{Spec}(\frac{R[x,y,z]}{(y^2-x,\pi-xz)})$. As for $\mathrm{Spec}(\frac{R[x,y,z]}{(1-xy,\pi-yz)})$, easy to check that it's smooth over R. Hence after blowing up Y at (x, y, π) and (z, w, π), we get a regular model X, which is minimal desingularization of Y. With calculation the exceptional curve over (x, y, π) is isomorphic to $\mathbb{P}^1_k$, hence we have $\mathrm{n}(X_s) = 3$, and the normalization of X_s is $\sigma : \mathbb{P}^1_k \sqcup \mathbb{P}^1_k \sqcup \mathbb{P}^1_k \to X_s$, and X_s has two singular points, and each has two preimage via σ, so $\mu(X_s) = 2$, $\mathrm{t}(X_s) = 0$, $\mathrm{a}(X_s) = 0$, $\mathrm{u}(X_s) = 0$ and $-\mathrm{Art}_{\mathrm{tame}}(X) = 2$ by Proposition 7.1.4.

Irreducible components

Since Y_s, $Y_{1,s}$ and $Y_{2,s}$ all have only one irreducible component, so we have

$$\mathrm{n}(Y_s) = \mathrm{n}(Y_{1,s}) + \mathrm{n}(Y_{2,s}) - 1.$$

Hence in case 1, as Subsection 4.4, we have

$$\begin{aligned}
\mathrm{n}(X_s) &= \mathrm{n}(Y_s) + \sum_{y \in \mathrm{Sing}(Y)} \mathrm{n}(X_y) \\
&= \mathrm{n}(Y_s) + \sum_{y_1 \in \mathrm{Sing}(Y_1)} \mathrm{n}(X_{1,y_1}) + \sum_{y_2 \in \mathrm{Sing}(Y_2)} \mathrm{n}(X_{2,y_2}) \\
&= (\mathrm{n}(Y_{1,s}) + \sum_{y_1 \in \mathrm{Sing}(Y_1)} \mathrm{n}(X_{1,y_1})) + (\mathrm{n}(Y_{2,s}) + \sum_{y_2 \in \mathrm{Sing}(Y_2)} \mathrm{n}(X_{2,y_2})) - 1 \\
&= \mathrm{n}(X_{1,s}) + \mathrm{n}(X_{1,s}) - 1.
\end{aligned}$$

In case 2, $Q_1 \in Y_1$ and $Q_2 \in Y_2$ are singular. We only consider Q_1, it is similar for Q_2. We know that Q_1 is on $Y_1'' : z^2 = w^{2g_1+2} f_1(1/w)$, and corresponds to the singular point of $\mathrm{Spec}(\frac{R[w,z]}{(z^2 - \pi w)})$ by Lemma 7.5.1. Hence the fiber X_{1,Q_1} on the minimal desingularization X_1 over Q_1 is isomorphic to $\mathbb{P}_k$. Then

$$\begin{aligned}
\mathrm{n}(X_s) &= \mathrm{n}(Y_s) + \sum_{y \in \mathrm{Sing}(Y)} \mathrm{n}(X_y) \\
&= \mathrm{n}(Y_s) + \sum_{y_1 \in \mathrm{Sing}(Y_1')} \mathrm{n}(X_{1,y_1}) + \sum_{y_2 \in \mathrm{Sing}(Y_2')} \mathrm{n}(X_{2,y_2}) \\
&= (\mathrm{n}(Y_{1,s}) + \sum_{y_1 \in \mathrm{Sing}(Y_1)} \mathrm{n}(X_{1,y_1})) + (\mathrm{n}(Y_{2,s}) + \sum_{y_2 \in \mathrm{Sing}(Y_2)} \mathrm{n}(X_{2,y_2})) - 3 \\
&= \mathrm{n}(X_{1,s}) + \mathrm{n}(X_{1,s}) - 3.
\end{aligned}$$

Ranks

We also consider case 1 at first. By Lemma 7.5.2,

$$\mu(X_s) = \sum_{x \in X_s(k)} (m_x - 1) = \sum_{x \in X_{1,s}(k)} (m_x - 1) + \sum_{x \in X_{2,s}(k)} (m_x - 1) = \mu(X_{1,s}) + \mu(X_{2,s}).$$

Hence

$$\begin{aligned}
\mathrm{t}(X_s) &= \mu(X_s) - \mathrm{n}(X_s) + 1 \\
&= (\mu(X_{1,s}) - \mathrm{n}(X_{1,s}) + 1) + (\mu(X_{2,s}) - \mathrm{n}(X_{2,s}) + 1) \\
&= \mathrm{t}(X_{1,s}) + \mathrm{t}(X_{2,s}).
\end{aligned}$$

The normalizations of Y_s, $Y_{1,s}$ and $Y_{2,s}$ are the projective line which has genus 0, so

$$\mathrm{a}(X_s) = \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}).$$

Since the multiplicities of the irreducible components of Y_s is 2, then the greatest common divisor of the multiplicities of the irreducible components of X_s is 1 or 2, and $\mathrm{Char}(k) \neq 2$, so by Lemma 7.3.4, we also have

$$\dim_k \mathrm{H}^1(X_s, \mathcal{O}_{X_s}) = \mathrm{p}_a(Y_K) = g,$$

$$\dim_k \mathrm{H}^1(X_{1,s}, \mathcal{O}_{X_{1,s}}) = g_1,$$

$$\dim_k \mathrm{H}^1(X_{2,s}, \mathcal{O}_{X_{2,s}}) = g_2,$$

$$g = g_1 + g_2 + 1.$$

Hence

$$\begin{aligned} \mathrm{u}(X_s) &= g - \mathrm{a}(X_s) - \mathrm{t}(X_s) \\ &= (g_1 - \mathrm{a}(X_{1,s}) - \mathrm{t}(X_{1,s})) + (g_2 - \mathrm{a}(X_{2,s}) - \mathrm{t}(X_{2,s})) + 1 \\ &= \mathrm{u}(X_{1,s}) + \mathrm{u}(X_{2,s}) + 1, \end{aligned}$$

$$\begin{aligned} -\mathrm{Art}(X) - \delta(X) &= \mathrm{n}(X_s) - 1 + 2\mathrm{u}(X_s) + \mathrm{t}(X_s) \\ &= 2\mathrm{u}(X_s) + \mu(X_s) \\ &= (-\mathrm{Art}(X_1) - \delta(X_1)) + (-\mathrm{Art}(X_2) - \delta(X_2)) + 2. \end{aligned}$$

For case 2, by the discussion in the end of the case $y^2 = \pi x$ and Lemma 7.5.3, we know that X_{1,Q_1} intersects with the other irreducible components of$X_{1,s}$ at only one point x_1 and $m_{x_1} = 2$. It is similar for X_2. Hence we have

$$\begin{aligned} \mu(X_s) &= \sum_{x \in X_s(k)} (m_x - 1) \\ &= \sum_{x \in X'_{1,s}(k)} (m_x - 1) + \sum_{x \in X'_{2,s}(k)} (m_x - 1) \\ &= (\sum_{x \in X_{1,s}(k)} (m_x - 1) - 1) + (\sum_{x \in X_{2,s}(k)} (m_x - 1) - 1) \\ &= \mu(X_{1,s}) + \mu(X_{2,s}) - 2, \end{aligned}$$

$$\begin{aligned} \mathrm{t}(X_s) &= \mu(X_s) - \mathrm{n}(X_s) + 1 \\ &= (\mu(X_{1,s}) - \mathrm{n}(X_{1,s}) + 1) + (\mu(X_{2,s}) - \mathrm{n}(X_{2,s}) + 1) \\ &= \mathrm{t}(X_{1,s}) + \mathrm{t}(X_{2,s}). \end{aligned}$$

Since the normalization of $Y_s, Y_{1,s}, Y_{2,s}$ are the projective line which has genus 0, and X_{1,Q_1}, X_{2,Q_2} are both isomorphic to $\mathbb{P}^1_K$, so

$$\mathrm{a}(X_s) = \mathrm{a}(X_{1,s}) + \mathrm{a}(X_{2,s}).$$

As the discussion in case 1, we have

$$\dim_k \mathrm{H}^1(X_s, \mathcal{O}_{X_s}) = \mathrm{p}_a(Y_K) = g,$$

$$\dim_k \mathrm{H}^1(X_{1,s}, \mathcal{O}_{X_{1,s}}) = g_1,$$

$$\dim_k \mathrm{H}^1(X_{2,s}, \mathcal{O}_{X_{2,s}}) = g_2,$$

$$g = g_1 + g_2.$$

Hence

$$\begin{aligned} \mathrm{u}(X_s) &= g - \mathrm{a}(X_s) - \mathrm{t}(X_s) \\ &= (g_1 - \mathrm{a}(X_{1,s}) - \mathrm{t}(X_{1,s})) + (g_2 - \mathrm{a}(X_{2,s}) - \mathrm{t}(X_{2,s})) \\ &= \mathrm{u}(X_{1,s}) + \mathrm{u}(X_{2,s}), \end{aligned}$$

$$\begin{aligned} -\mathrm{Art}_{\mathrm{tame}}(X) &= \mathrm{n}(X_s) - 1 + 2\mathrm{u}(X_s) + \mathrm{t}(X_s) \\ &= 2\mathrm{u}(X_s) + \mu(X_s) \\ &= -\mathrm{Art}_{\mathrm{tame}}(X_1) - \mathrm{Art}_{\mathrm{tame}}(X_2) - 2. \end{aligned}$$

7.6 Proof of Corollary 7.2.4

For a Weierstrass model Y, we set

$$h(Y) := v(\Delta(Y)) + \mathrm{Art}_{\mathrm{tame}}(X),$$

where X is minimal desingularization of Y. Then with the notations in Theorem 7.2.1, we have

$$h(Y) = h(Y_1) + h(Y_2)$$

Hence, after a change of coordinates, sufficient to prove the inequality that

$$h(Y) \geq 0, \tag{7.11}$$

for $Y : y^2 = \pi^\epsilon f(x)$ with $\overline{f}(x) = ux^n \in k[x]$, where $f(x) = x(x-b_2)\cdots(x-b_n)$ with $n \geq 4$, $v(b_i) \geq 1$ for each i, and $\epsilon = 0$ or 1. We can suppose that $u = 1$ since R is strictly henselian. We will prove this by induction on n, so we consider all cases for $n \geq 1$. Before that, let us determine the singular points on Y.

Notice that $Y = \mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi^\epsilon f(x))}) \bigcup \mathrm{Spec}(\frac{R[z,w]}{(w^2-\pi^\epsilon z^{2\lfloor (n+1)/2\rfloor} f(\frac{1}{z}))})$, where $x = \frac{1}{z}$, $y = x^{\lfloor (n+1)/2\rfloor} w$.

If $\epsilon = 0, n \geq 2$ or $\epsilon = 1, n$ is even, the Y has only one singular point $(x, y, \pi) \in \mathrm{Spec}(\frac{R[x,y]}{(y^2-f(x))})$. Firstly, a point of the form $P = (z, w-b, \pi)$ on Y is regular: if $\epsilon = 0$, and P is singular, then $\overline{b} = 0$, and easy to check that $z = 0$ not a multiple root of $z^{2\lfloor (n+1)/2\rfloor} f(\frac{1}{z}) = 0$, so it's not singular; if $\epsilon = 1, n$ is even, then $z = 0$ is not a root of $z^n f(\frac{1}{z}) = 0$, and any point of the form $(z, w-b, \pi)$ is regular, see Section 7.5.2. Hence, in both cases, the possible singular point is $\mathfrak{m} = (x, y, \pi) \in \mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi^\epsilon f(x))})$. Next, we prove that m is singular. If $\epsilon = 1, n$ is even, see Section 7.5.2; if $\epsilon = 0$, the proof is similar: consider $\mathfrak{m} = (x, y, \pi) \in \mathrm{Spec}(R[x, y])$, and set

$$I = \mathfrak{m}^2 + (y^2 - f(x)) = (x^2, y^2, \pi^2, \pi x, \pi y, xy, y^2 - f(x)),$$

then $I + (x, \pi) = (x, y^2, \pi) \neq \mathfrak{m}$, $I + (x, y) = (x, y, \pi^2) \neq \mathfrak{m}$, so $\mathfrak{m}$ is singular if and only if $I + (y, \pi) = (x^2, y, \pi, x^n) \neq \mathfrak{m}$, that is $n \geq 2$. It also shows that If $\epsilon = 0, n = 1$, then Y is regular.

If $\epsilon = 1, n$ is odd, then Y has two singular points $(x, y, \pi) \in \mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi f(x))})$ and $(z, w, \pi) \in \mathrm{Spec}(\frac{R[z,w]}{(w^2-\pi z^{n+1} f(\frac{1}{z}))})$, see Section 7.5.2.

We prove the inequality 7.11 by induction on n. For $n = 1$, if $Y : y^2 = x$, we have that Y is smooth, hence the minimal desingularization is exactly Y, and easy to know that $Y_s \simeq \mathbb{P}_k$, so $\mu(Y_s) = \mathrm{p}_a(Y_s) = \mathrm{t}(Y_s) = \mathrm{a}(Y_s) = \mathrm{u}(Y_s) = 0$, so $-\mathrm{Art}_{\mathrm{tame}}(Y) = 0$, i.e. $h(Y) = 0$.

For $Y : y^2 = \pi x$, we know that $h(Y) = 0$ in the proof of Theorem 7.2.3.

If $n \geq 2$, suppose that $v(b_n) = \max\limits_{2\leq i\leq n}\{v(b_i)\}$, then after change of coordinates

$$\begin{cases} x = \pi^m \tilde{x} \\ y = \pi^{\lfloor (nm+\epsilon)/2\rfloor} \tilde{y}, \end{cases}$$

we get another Weierstrass model $\tilde{Y} : \tilde{y}^2 = \pi^{\tilde{\epsilon}} \tilde{x}(\tilde{x} - \frac{b_2}{\pi^m}) \cdots (\tilde{x} - \frac{b_n}{\pi^m})$, where $\tilde{\epsilon} = 0$ or 1. We claim that

$$h(Y) \geq h(\tilde{Y}) + 2m(n-2). \tag{7.12}$$

If so, then by inductive hypothesis and Theorem 7.2.1, we have $h(Y) \geq h(\tilde{Y}) \geq 0$. Hence, to show the first statement of Corollary 7.2.4 hold, it is sufficient to prove following lemma:

LEMMA 7.6.1. *Let $Y : y^2 = \pi^\epsilon x(x-b_2)\cdots(x-b_n)$ be a Weierstrass model of a hyperelliptic curve with $n \geq 2$, $v(b_i) \geq 1$, $\epsilon = 0$ or 1, via*

$$\begin{cases} x = \pi\tilde{x} \\ y = \pi^{\lfloor (n+\epsilon)/2 \rfloor}\tilde{y}, \end{cases}$$

we have another Weierstrass model $\tilde{Y} : \tilde{y}^2 = \pi^{\tilde{\epsilon}}\tilde{x}(\tilde{x}-\frac{b_2}{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi})$ with $\tilde{\epsilon} = n+\epsilon-2\lfloor (n+\epsilon)/2 \rfloor$ Then

$$h(Y) \geq h(\tilde{Y}) + 2(n-2)$$

Proof. Obviously, we have $v(\Delta(Y)) - v(\Delta(\tilde{Y})) = (2n-2)\epsilon - (2n-2)\tilde{\epsilon} + 2\binom{n}{2}$, and $\delta(X) = \delta(\tilde{X})$, wheren X (resp.$\tilde{X}$) are the minimal desingularization of Y (resp.$\tilde{Y}$).

Set $U = \mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi^\epsilon x(x-b_2)\cdots(x-b_n))}) \subset Y$.

If n is even, then $\epsilon = \tilde{\epsilon}$, and Y has only one singular point $(x,y,\pi) \in \mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi^\epsilon x(x-b_2)\cdots(x-b_n))})$, and the all singular points of $\tilde{Y}$ are on $\mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi^{\tilde{\epsilon}} x(x-\frac{b_2}{\pi})\cdots(x-\frac{b_n}{\pi}))})$. Suppose that W is the blowup of Y at (π,x,y), it's glued by some affine open subsets.

If n is even and $\epsilon = \tilde{\epsilon} = 0$, we can know that W is glued by

$$W_1 = \mathrm{Spec}(\frac{R[\tilde{x},\tilde{y}]}{(\tilde{y}^2 - \pi^{n-2}\tilde{x}(\tilde{x}-\frac{b_2}{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi}))}),$$

$$W_2 = \mathrm{Spec}(\frac{R[x,\tilde{\pi},\tilde{y}]}{(\tilde{y}^2 - x^{n-2}(1-\frac{b_2}{\pi}\tilde{\pi})\cdots(1-\frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}x)}),$$

$$W_3 = \mathrm{Spec}(\frac{R[y,\tilde{\pi},\tilde{x}]}{(1 - \tilde{x}y^{n-2}(\tilde{x}-\frac{b_2}{\pi}\tilde{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}y)}),$$

and take the normalization $\tilde{W}$ of W. we can check that W_3 is regular by using the standard method as before. We know that the normalization $\tilde{W}_1$ of W_1 is $\mathrm{Spec}(\frac{R[\tilde{x},\tilde{y}]}{(\tilde{y}^2-\tilde{x}(\tilde{x}-\frac{b_2}{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi}))})$, which is an open subset of $\tilde{Y}$, and the normalization $\tilde{W}_2$ of W_2 is $\mathrm{Spec}(\frac{R[x,\tilde{\pi},\tilde{y}]}{(\tilde{y}^2-(1-\frac{b_2}{\pi}\tilde{\pi})\cdots(1-\frac{b_n}{\pi}\tilde{\pi}),\pi-x\tilde{\pi}))})$. We consider the closed points on $\mathrm{V}(\tilde{\pi}) \subset \tilde{W}_2$, each of them are regular on $\tilde{W}_2$: by Jacobian criteria, the possible singular points on $V(\tilde{\pi})$ are corresponding ideals $(\pi,\tilde{\pi},\tilde{y}\pm 1,x)$, set $\mathfrak{m} = (\pi,\tilde{\pi},\tilde{y}-1,x) \subset R[x,\tilde{\pi},\tilde{y}]$ and

$$I = \mathfrak{m}^2 + (\tilde{y}^2 - (1-\frac{b_2}{\pi}\tilde{\pi})\cdots(1-\frac{b_n}{\pi}\tilde{\pi}), \pi - x\tilde{\pi}),$$

easy to check that $I + (\tilde{\pi},\tilde{y}-1) = \mathfrak{m}$, so the point corresponding to $\mathfrak{m}$ is regular on $\tilde{W}_2$, and similar for another point. We use this process to check if a point is regular or not in the following proof. Hence the singular points of $\tilde{W}$ are on $\tilde{W}_1$. With calculation, we know that the preimage of $(\pi,x,y) \in Y$ on $\tilde{W}$ is isomorphic to $\tilde{Y}_s$, then

$$\mathrm{n}(\tilde{W}_s) = \mathrm{n}(Y_s) + \mathrm{n}(\tilde{Y}_s),$$

i.e. $\mathrm{n}(\tilde{W}_s) - \mathrm{n}(\tilde{Y}_s) = \mathrm{n}(Y_s) = 2$. Since the minimal desingularization can be obtained by a series of normalized blowups at singular points, and $\tilde{W}$, $\tilde{Y}$ have the same singular points, so we have $\mathrm{n}(X_s) - \mathrm{n}(\tilde{X}_s) = \mathrm{n}(\tilde{W}_s) - \mathrm{n}(\tilde{Y}_s) = 2$, and with the fact that Y and $\tilde{Y}$ have the isomorphic generic fiber, we have that $\mathrm{Art}(\widetilde{X}) - \mathrm{Art}(X) = 2$. Hence

$$h(Y) - h(\tilde{Y}) - 2(n-2) = n(n-1) - 2 - 2(n-2) = (n-1)(n-2) \geq 0.$$

If n is even and $\epsilon = \tilde{\epsilon} = 1$, then W is glued by

$$W_1 = \mathrm{Spec}(\frac{R[\tilde{x},\tilde{y}]}{(\tilde{y}^2 - \pi^{n-1}\tilde{x}(\tilde{x}-\frac{b_2}{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi}))}),$$

$$W_2 = \mathrm{Spec}(\frac{R[x,\tilde{\pi},\tilde{y}]}{(\tilde{y}^2 - \tilde{\pi}x^{n-1}(1-\frac{b_2}{\pi}\tilde{\pi})\cdots(1-\frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}x)}),$$

$$W_3 = \mathrm{Spec}(\frac{R[y,\tilde{\pi},\tilde{x}]}{(1 - \tilde{\pi}\tilde{x}y^{n-1}(\tilde{x}-\frac{b_2}{\pi}\tilde{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}y)}),$$

and $\tilde{W}$ is the normalization of W. Similarly, we have that W_3 is regular, the normalization $\tilde{W}_1$ of W_1 is $\mathrm{Spec}(\frac{R[\tilde{x},\tilde{y}]}{(\tilde{y}^2-\pi\tilde{x}(\tilde{x}-\frac{b_2}{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi}))})$, which is an open subset of $\tilde{Y}$, and the normalization $\tilde{W}_2$ of W_2 is $\mathrm{Spec}(\frac{R[x,\tilde{\pi},\tilde{y}]}{(\tilde{y}^2-\tilde{\pi}x(1-\frac{b_2}{\pi}\tilde{\pi})\cdots(1-\frac{b_n}{\pi}\tilde{\pi}),\pi-x\tilde{\pi})})$. With calculation as before, we know that $(\pi, x, \tilde{\pi}, \tilde{y})$ is the only point on $\mathrm{V}(\tilde{\pi})$ and singular on $\tilde{W}_2$. To resolve it, it's sufficient to consider the completion of its local ring, that is

$$\frac{\hat{R}[[x,\tilde{\pi},\tilde{y}]]}{(\tilde{y}^2 - \tilde{\pi}x(1-\frac{b_2}{\pi}\tilde{\pi})\cdots(1-\frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}x)},$$

where $\hat{R}$ is the completion of R. Notice that this ring is isomorphic to the completion of local ring of $T = \mathrm{Spec}(\frac{R[x,\tilde{\pi},y]}{(\tilde{y}^2-x\tilde{\pi},\pi-x\tilde{\pi})})$ at $(\pi, x, \tilde{\pi}, \tilde{y})$, which is the only singular point on T. Take the blowup $\tilde{T}$ of T at $(\pi, x, \tilde{\pi}, \tilde{y})$, and with calculation, we know that $\tilde{T}$ is regular and the fiber on $\tilde{T}$ over $(\pi, x, \tilde{\pi}, \tilde{y})$ is isomorphic to $\mathbb{P}^1_k$, which means that

$$\mathrm{n}(\tilde{T}_s) = \mathrm{n}(T_s) + 1.$$

Except $(\pi, x, \tilde{\pi}, \tilde{y}) \in \tilde{W}_2$, the rest of singular points on $\tilde{W}$ are on $\tilde{W}_1$. With calculation, we know that the preimage of (π, x, y) on $\tilde{W}$ is exactly isomorphic to $\tilde{Y}_s$, then

$$\mathrm{n}(\tilde{W}_s) = \mathrm{n}(Y_s) + \mathrm{n}(\tilde{Y}_s),$$

i.e. $\mathrm{n}(\tilde{W}_s) - \mathrm{n}(\tilde{Y}_s) = \mathrm{n}(Y_s) = 1$. Hence $\mathrm{n}(X_s) - \mathrm{n}(\tilde{X}_s) = \mathrm{n}(\tilde{W}_s) - \mathrm{n}(\tilde{Y}_s) + (\mathrm{n}(\tilde{T}_s) - \mathrm{n}(T_s)) = 2$, and $\mathrm{Art}(\widetilde{X}) - \mathrm{Art}(X) = 2$. It is similar as above, we have $h(Y) - h(\tilde{Y}) \geq 2(n-2)$.

If n is odd, $\epsilon = 0$, then $\tilde{\epsilon} = 1$, and Y has only one singular point $(\pi, x, y) \in \mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi^\epsilon x(x-b_2)\cdots(x-b_n))})$, but $\tilde{Y}$ has a singular point that is not on affine open subset $\mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi x(x-\frac{b_2}{\pi})\cdots(x-\frac{b_n}{\pi}))})$, that is $(\pi, z, w) \in \mathrm{Spec}(\frac{R[z,w]}{(w^2-\pi z(1-\frac{b_2}{\pi}z)\cdots(1-\frac{b_n}{\pi}z))})$. As above, suppose that the normalized blowup of Y at (π, x, y) is $\tilde{W}$, which is glued by

$$\tilde{W}_1 = \mathrm{Spec}(\frac{R[\tilde{x},\tilde{y}]}{(\tilde{y}^2 - \pi\tilde{x}(\tilde{x}-\frac{b_2}{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi}))}),$$

$$\tilde{W}_2 = \mathrm{Spec}(\frac{R[x,\tilde{y},\tilde{\pi}]}{(\tilde{y}^2 - x(1-\frac{b_2}{\pi}\tilde{\pi})\cdots(1-\frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}x)}),$$

$$\tilde{W}_3 = \mathrm{Spec}(\frac{R[y,\tilde{\pi},\tilde{x}]}{(1 - \tilde{x}y^{n-2}(\tilde{x}-\frac{b_2}{\pi}\tilde{\pi})\cdots(\tilde{x}-\frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}y)}),$$

and we can check that every point on $\mathrm{V}(\tilde{\pi}) \subset \tilde{W}_2$ is regular on $\tilde{W}_2$, similar for $\tilde{W}_3$. Hence, singular points of $\tilde{W}$ are on $\tilde{W}_1$, which is isomorphic to an open subset of $\tilde{Y}$. With calculation, we know that the preimage of (π, x, y) on $\tilde{W}$ is exactly $\tilde{Y}_s$, then

$$\mathrm{n}(\tilde{W}_s) = \mathrm{n}(Y_s) + \mathrm{n}(\tilde{Y}_s),$$

i.e. $\mathrm{n}(\tilde{W}_s) - \mathrm{n}(\tilde{Y}_s) = \mathrm{n}(Y_s) = 1$. Since $\tilde{Y}$ has to resolve (π, z, w), so $\mathrm{n}(X_s) - \mathrm{n}(\tilde{X}_s) \leq \mathrm{n}(\tilde{W}_s) - \mathrm{n}(\tilde{Y}_s) - 1 = 0$, and $\mathrm{Art}(\widetilde{X}) - \mathrm{Art}(X) \leq 0$. Hence

$$h(Y) - h(\tilde{Y}) - 2(n-2) \geq -4(n-1) + n(n-1) - 2(n-2) = (n-3)(n-2) \geq 0.$$

If n is odd, $\epsilon = 1$, then $\tilde{\epsilon} = 0$, and all singular points of $\tilde{Y}$ are on $\mathrm{Spec}(\frac{R[x,y]}{(y^2 - x(x-\frac{b_2}{\pi})\cdots(x-\frac{b_n}{\pi}))})$, but Y has two singular point, one of which is not on affine open subset $\mathrm{Spec}(\frac{R[x,y]}{(y^2-\pi x(x-b_2)\cdots(x-b_n))})$, that is (π, z, w) on $\mathrm{Spec}(\frac{R[z,w]}{(w^2-\pi z(1-b_2z)\cdots(1-b_nz))})$. Take the blowup Y' of Y at $P = (\pi, z, w)$, then Y' will resolve this singular point. Indeed, take the completion $\widehat{R[z,w]_P} = \hat{R}[[z,w]]$ of $R[z,w]_P$, by Hensel's Lemma, there exists $T \in (\widehat{R[z,w]_P})^*$ such that $T^2 = (1-b_2z)\cdots(1-b_nz)$, so $\widehat{\mathcal{O}_{Y,P}} \simeq \frac{\widehat{R[z,w]_P}}{(w^2-\pi z)}$, and we have seen that the minimal desingularization of $\mathrm{Spec}(\frac{R[z,w]_P}{w^2-\pi z})$ is obtain by blowing up at P, with fiber $\mathbb{P}^1$ over P, hence it is same for $\widehat{\mathcal{O}_{Y,P}}$ and Y. Suppose that $\tilde{W}$ is the normalized blowup of U at (x, y, π), then $\tilde{W}$ is glued by

$$\tilde{W}_1 = \mathrm{Spec}(\frac{R[\tilde{x}, \tilde{y}]}{(\tilde{y}^2 - \tilde{x}(\tilde{x} - \frac{b_2}{\pi})\cdots(\tilde{x} - \frac{b_n}{\pi}))}),$$

$$\tilde{W}_2 = \mathrm{Spec}(\frac{R[x, \tilde{y}, \tilde{\pi}]}{(\tilde{y}^2 - \tilde{\pi}(1 - \frac{b_2}{\pi}\tilde{\pi})\cdots(1 - \frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}x)}),$$

$$\tilde{W}_3 = Spec(\frac{R[y, \tilde{\pi}, \tilde{x}]}{(1 - \tilde{\pi}\tilde{x}y^{n-2}(\tilde{x} - \frac{b_2}{\pi}\tilde{\pi})\cdots(\tilde{x} - \frac{b_n}{\pi}\tilde{\pi}), \pi - \tilde{\pi}y)}),$$

and we can check that $\tilde{W}_3$ is regular, and every point on $\mathrm{V}(\tilde{\pi}) \subset \tilde{W}_2$ is regular on $\tilde{W}_2$. Hence, singular points of $\tilde{W}$ are on $\tilde{W}_1$, which is isomorphic to an open subset of $\tilde{Y}$. With calculation, we know that the preimage of (π, x, y) on $\tilde{W}$ is exactly $\tilde{Y}_s$, then

$$\mathrm{n}(\tilde{W}_s) = \mathrm{n}(Y_s) + \mathrm{n}(\tilde{Y}_s),$$

i.e. $\mathrm{n}(\tilde{W}_s) - \mathrm{n}(\tilde{Y}_s) = \mathrm{n}(Y_s) = 1$. Hence $\mathrm{n}(X_s) - \mathrm{n}(\widetilde{X}_s) = 1 + \mathrm{n}(\tilde{W}_s) - \mathrm{n}(\tilde{Y}_s) = 2$, so $\mathrm{Art}(\widetilde{X}) - \mathrm{Art}(X) = 2$. Hence, we have

$$h(Y) - h(\tilde{Y}) = 2(n-1) + n(n-1) - 2 - 2(n-2) = n(n-1) \geq 0.$$

□

The second statement comes from the first statement and Corollary 7.1.3.

Bibliography

[1] Bajolet, A., Bilu, Yu., and Matschke, B. (2012). Computing integral points on x_ns^+(p). *arXiv preprint arXiv:1212.0665*.

[2] Baker, A. and Wüstholz, G. (2007). *Logarithmic forms and Diophantine geometry*, volume 9 of *New Mathematical Monographs*. Cambridge University Press, Cambridge.

[3] Bilu, Yu. (1993). *Effective Analysis of Integral Points on Algebraic Curves*. PhD thesis, Beer Sheva.

[4] Bilu, Yu. (1995). Effective analysis of integral points on algebraic curves. *Israel J. Math.*, 90(1-3):235–252.

[5] Bilu, Yu. (2002). Baker's method and modular curves. In *A panorama of number theory or the view from Baker's garden (Zürich, 1999)*, pages 73–88. Cambridge Univ. Press, Cambridge.

[6] Bilu, Yu., Faye, B., and Zhu, H. (2019). Separating singular moduli and the primitive element problem. *arXiv preprint arXiv:1903.07126*.

[7] Bilu, Yu., Habegger, P., and Kühne, L. (2018). No Singular Modulus Is a Unit. *Int. Math. Res. Not. IMRN*. rny274.

[8] Bilu, Yu. and Illengo, M. (2011). Effective Siegel's theorem for modular curves. *Bull. Lond. Math. Soc.*, 43(4):673–688.

[9] Bilu, Yu. and Parent, P. (2011a). Runge's method and modular curves. *Int. Math. Res. Not. IMRN*, (9):1997–2027.

[10] Bilu, Yu. and Parent, P. (2011b). Serre's uniformity problem in the split Cartan case. *Ann. of Math. (2)*, 173(1):569–584.

[11] Bilu, Yu. and Strambi, M. (2010). Quantitative Riemann existence theorem over a number field. *Acta Arith.*, 145.

[12] Bilu, Yu., Strambi, M., and Surroca, A. (2013). Quantitative Chevalley-Weil theorem for curves. *Monatsh. Math.*, 171(1):1–32.

[13] Bloch, S. (1987). De rham cohomology and conductors of curves. *Duke Math. J.*, 54(2):295–308.

[14] Bombieri, E. and Gubler, W. (2006). *Heights in Diophantine geometry*, volume 4 of *New Mathematical Monographs*. Cambridge University Press, Cambridge.

[15] Brumer, A. and Kramer, K. (1994). The conductor of an abelian variety. *Compositio Mathematica*, 92(2):227–248.

[16] Bugeaud, Y. and Győry, K. (1996). Bounds for the solutions of unit equations. *Acta Arith.*, 74(1):67–80.

[17] Cassels, J. W. S. (1997). *An introduction to the geometry of numbers*. Classics in Mathematics. Springer-Verlag, Berlin. Corrected reprint of the 1971 edition.

[18] Childress, N. (2009). *Class field theory*. Universitext. Springer, New York.

[19] Conrad, B. (2007). Arithmetic moduli of generalized elliptic curves. *J. Inst. Math. Jussieu*, 6(2):209–278.

[20] Cox, D. A. (2013). *Primes of the form $x^2 + ny^2$*. Pure and Applied Mathematics (Hoboken). John Wiley & Sons, Inc., Hoboken, NJ, second edition. Fermat, class field theory, and complex multiplication.

[21] Deligne, P. and Rapoport, M. (1973). Les schémas de modules de courbes elliptiques. In *Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, pages 143–316. Lecture Notes in Math., Vol. 349.

[22] Diamond, F. and Shurman, J. (2005). *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.

[23] Evertse, J.-H. and Győry, K. (2015). *Unit equations in Diophantine number theory*, volume 146 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge.

[24] Faltings, G. (1983). Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73(3):349–366.

[25] Faye, B. and Riffaut, A. (2018). Fields generated by sums and products of singular moduli. *J. Number Theory*, 192:37–46.

[26] Friedman, E. (1989). Analytic formulas for the regulator of a number field. *Invent. Math.*, 98(3):599–622.

[27] Gross, B. H. and Zagier, D. B. (1985). On singular moduli. *J. Reine Angew. Math.*, 355:191–220.

[28] Grothendieck, A. (1961). Éléments de géométrie algébrique. III. Étude cohomologique des faisceaux cohérents. I. *Inst. Hautes Études Sci. Publ. Math.*, (11):167.

[29] Grothendieck, A. (1971). *Revêtements étales et groupe fondamental (SGA 1)*, volume 224 of *Lecture notes in mathematics*. Springer-Verlag.

[30] Grothendieck, A. (1972). *Groupes de Monodromie en géométrie algébrique. (SGA 7I)*, volume 288 of *Lecture Notes in Mathematics*. Springer-Verlag.

[31] Habegger, P. (2015). Singular moduli that are algebraic units. *Algebra Number Theory*, 9(7):1515–1524.

[32] Habegger, P., Jones, G., and Masser, D. (2017). Six unlikely intersection problems in search of effectivity. *Math. Proc. Cambridge Philos. Soc.*, 162(3):447–477.

[33] Hindry, M. and Silverman, J. H. (2000). *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics*. Springer-Verlag, New York. An introduction.

[34] Katz, N. M. and Mazur, B. (1985). *Arithmetic moduli of elliptic curves*, volume 108 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ.

[35] Kubert, D. S. and Lang, S. (1981). *Modular units*, volume 244 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York-Berlin.

[36] Lang, S. (1987). *Elliptic functions*, volume 112 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition. With an appendix by J. Tate.

[37] Li, Y. (2018). Singular units and isogenies between cm elliptic curves. *arXiv preprint arXiv:1810.13214*.

[38] Lipman, J. (1969). Rational singularities, with applications to algebraic surfaces and unique factorization. *Inst. Hautes Études Sci. Publ. Math.*, (36):195–279.

[39] Liu, Q. (1994). Conducteur et discriminant minimal de courbes de genre 2. *Compositio Math.*, 94(1):51–79.

[40] Liu, Q. (1996). Modèles entiers des courbes hyperelliptiques sur un corps de valuation discrète. *Trans. Amer. Math. Soc.*, 348(11):4577–4610.

[41] Liu, Q. (2002). *Algebraic geometry and arithmetic curves*, volume 6 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford. Translated from the French by Reinie Erné, Oxford Science Publications.

[42] Liu, Q. and Lorenzini, D. (1999). Models of curves and finite covers. *Compositio Math.*, 118(1):61–102.

[43] Loher, T. and Masser, D. (2004). Uniformly counting points of bounded height. *Acta Arith.*, 111(3):277–297.

[44] Louboutin, S. (2000). Explicit bounds for residues of Dedekind zeta functions, values of L-functions at $s = 1$, and relative class numbers. *J. Number Theory*, 85(2):263–282.

[45] Matveev, E. M. (2000). An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. II. *Izv. Math.*, 64(6):1217–1269.

[46] Neukirch, J. (1999). *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[47] Obus, A. and Srinivasan, P. (2019). Conductor-discriminant inequality for hyperelliptic curves in odd residue characteristic.

[48] Ogg, A. P. (1967). Elliptic curves and wild ramification. *Amer. J. Math.*, 89:1–21.

[49] Pethö, A. and de Weger, B. M. M. (1986). Products of prime powers in binary recurrence sequences. I. The hyperbolic case, with an application to the generalized Ramanujan-Nagell equation. *Math. Comp.*, 47(176):713–727.

[50] Raynaud, M. (1970). Spécialisation du foncteur de Picard. *Inst. Hautes Études Sci. Publ. Math.*, (38):27–76.

[51] Saito, T. (1987). Vanishing cycles and geometry of curves over a discrete valuation ring. *American Journal of Mathematics*, 109(6):1043–1085.

[52] Serre, J.-P. (1970). Facteurs locaux des fonctions zêta des varietés algébriques (définitions et conjectures). In *Séminaire Delange-Pisot-Poitou. 11e année: 1969/70. Théorie des nombres. Fasc. 1: Exposés 1 à 15; Fasc. 2: Exposés 16 à 24*, page 15. Secrétariat Math., Paris.

[53] Serre, J.-P. and Tate, J. (1968). Good reduction of abelian varieties. *Ann. of Math. (2)*, 88:492–517.

[54] Sha, M. (2014a). Bounding the j-invariant of integral points on certain modular curves. *Int. J. Number Theory*, 10(6):1545–1551.

[55] Sha, M. (2014b). Bounding the j-invariant of integral points on modular curves. *Int. Math. Res. Not. IMRN*, (16):4492–4520.

[56] Shimura, G. (1994). *Introduction to the arithmetic theory of automorphic functions*, volume 11 of *Publications of the Mathematical Society of Japan*. Princeton University Press, Princeton, NJ. Reprint of the 1971 original, Kanô Memorial Lectures, 1.

[57] Siegel, C. L. (1929). Über einige Anwendungen diophantischer Approximationen. *Abh. Preuß. Akad. Wiss., Phys.-Math. Kl.*, 1929(1):70 s.

[58] Siegel, C. L. (1969). Abschätzung von Einheiten. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*, 1969:71–86.

[59] Silverman, J. H. (1994). *Advanced topics in the arithmetic of elliptic curves*, volume 151 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.

[60] Silverman, J. H. (2009). *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer, Dordrecht, second edition.

[61] Srinivasan, P. (2015). Conductors and minimal discriminants of hyperelliptic curves with rational weierstrass points.

[62] Srinivasan, P. (2019). Conductors and minimal discriminants of hyperelliptic curves: A comparison in the tame case.

[63] Stacks project authors, T. (2019). The stacks project. `https://stacks.math.columbia.edu`.

[64] Viola, C. (2016). *An introduction to special functions*, volume 102 of *Unitext*. Springer, [Cham]. La Matematica per il 3+2.

[65] Voutier, P. (1996). An effective lower bound for the height of algebraic numbers. *Acta Arith.*, 74(1):81–95.

[66] Waldschmidt, M. (2000). *Diophantine approximation on linear algebraic groups*, volume 326 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin. Transcendence properties of the exponential function in several variables.

[67] Yu, K. (2007). p-adic logarithmic forms and group varieties. III. *Forum Math.*, 19(2):187–280.