Higher-dimensional modular equations, applications to isogeny computations and point counting

Jean Kieffer

Thèse présentée pour obtenir le grade de

Docteur de l'Université de Bordeaux

École doctorale de Mathématiques et Informatique
Spécialité Mathématiques pures

par Jean KIEFFER

Équations modulaires en dimension supérieure, applications au calcul d'isogénies et au comptage de points

Sous la direction de : Damien ROBERT
Co-encadrant : Aurel PAGE

Soutenue le 13 juillet 2021

Membres du jury :

M. Jean-Marc COUVEIGNES	Professeur, Université de Bordeaux	Président
M. Kamal KHURI-MAKDISI	Professeur, American University of Beirut	Rapporteur
M. David KOHEL	Professeur, Aix-Marseille Université	Rapporteur (exc.)
Mme Christine BACHOC	Professeure, Université de Bordeaux	Examinatrice
Mme Sorina IONICA	Maîtresse de conférences, Université de Picardie	Examinatrice
M. David LUBICZ	Ingénieur de recherche, DGA	Examinateur
M. Damien ROBERT	Chargé de recherche, INRIA	Directeur
M. Aurel PAGE	Chargé de recherche, INRIA	Invité

Titre: Équations modulaires en dimension supérieure, applications au calcul d'isogénies et au comptage de points

Résumé: L'objectif de cette thèse est de généraliser la méthode d'Elkies, un ingrédient fondamental de l'algorithme SEA pour le comptage de points d'une courbe elliptique sur un corps fini, au cas des variétés abéliennes polarisées de dimension supérieure. Les *équations modulaires* jouent un rôle central dans cette étude. Premièrement, nous donnons un algorithme de calcul d'isogénies entre surfaces abéliennes à partir d'équations modulaires. Deuxièmement, nous obtenons des bornes de degré et de hauteur pour les équations modulaires en fonction de leur niveau. Troisièmement, nous décrivons un algorithme rigoureux permettant d'évaluer des équations modulaires pour les surfaces abéliennes via des approximations complexes. Combiner ces résultats permet d'obtenir un algorithme de comptage de points de meilleure complexité pour les surfaces abéliennes principalement polarisées avec multiplication réelle.

Mots-clés: Variétés abéliennes; isogénies; équations modulaires; algorithmes.

Title: Higher-dimensional modular equations, applications to isogeny computations and point counting

Abstract: This thesis aims to generalize Elkies's method, a fundamental ingredient in the SEA algorithm for counting points on elliptic curves over finite fields, to the case of polarized abelian varieties of higher dimensions. *Modular equations* play a central role in this study. First, we design an algorithm using these modular equations to compute isogenies between abelian surfaces. Second, we give degree and height bounds for modular equations in terms of their level. Third, we describe a rigorous algorithm to evaluate modular equations for abelian surfaces via complex approximations. Combining these results yields an asymptotically faster point counting algorithm for principally polarized abelian surfaces with fixed real multiplication.

Keywords: Abelian varieties; isogenies; modular equations; algorithms.

Unité de recherche: Institut de Mathématiques de Bordeaux, UMR 5251, 351 cours de la Libération, 33400 Talence, France.

Remerciements

Durant ma thèse, j'ai eu la chance de côtoyer un grand nombre de personnes qui ont contribué à rendre ces trois années aussi agréables qu'enrichissantes. C'est un grand plaisir de pouvoir les remercier ici.

Mes remerciements vont tout d'abord à mes directeurs de thèse, Damien Robert et Aurel Page, pour m'avoir proposé un sujet fascinant et m'avoir guidé avec bienveillance chaque fois que j'en avais besoin – d'autant plus dans un contexte de Covid, de soutenance d'HDR et d'enfants en bas âge. Je remercie les membres du jury pour le temps qu'ils auront consacré à l'évaluation de la thèse, et particulièrement Kamal Khuri-Makdisi et David Kohel pour leurs rapports détaillés et élogieux. Merci à Luca De Feo, Ben Smith, Renaud Dubois et Jean-Marc Couveignes pour leurs encouragements. Merci à Craig Costello, Renate Scheidler, Michael Jacobson et Noam Elkies, ainsi qu'aux membres des différentes équipes au sein desquelles j'ai donné des exposés, pour leur intérêt pour mes travaux.

Je voudrais également remercier tous ceux qui, à différents niveaux, font de l'IMB un fantastique environnement de travail (mais pas seulement). Merci à tous les membres de l'équipe LFANT pour les discussions animées du séminaire et des pauses café. Merci à Ida Sirben, Muriel Hachemi, Sabrina Blondel–Duthil, Sylvaine Granier et à la cellule informatique sans qui la machine ne tournerait pas aussi bien. Merci à mes collègues, anciens et futurs, de Lambda – espérons que ce pot de rentrée pourra s'organiser un jour ! Merci à mes co-bureaux du deuxième et du troisième étage pour leur calme, c'est important parfois. Merci aux aînés, Alexandre, Corentin, Lara et les autres, pour m'avoir chaleureusement accueilli lorsque j'étais en première année. Merci à Thomas pour le café, Adrien pour le thé, Paul pour les pommes, et Bianca et Nicoletta pour les *curgete*.

À Bordeaux, il me reste à remercier tous ceux que j'ai rencontrés en dehors du labo ; les revoir après les confinements a été un vrai plaisir. Merci tout d'abord à toute la famille Cervantès dont j'ai vraiment eu le sentiment de faire partie, notamment pour la logistique de la soutenance. Merci à Marine Zubeldia pour la pédagogie et la flexibilité des cours de

hautbois, et merci au tarif jeunes de l'ONBA. Merci aux membres du club de go Kitani, Paul en particulier, pour toutes ces victoires et défaites – on taira lesquelles ont été les plus fréquentes – et à Christine et Guillaume pour leur accueil toujours chaleureux au Bistroquet.

Entre les week-ends et les confinements, j'ai dû passer autant de ces trois ans à Paris qu'à Bordeaux ; merci, en passant, à Tgvmax pour les réveils à 5h du matin. De ce fait j'ai pu garder contact avec une brochette d'amis formidables que je garderai probablement toute ma vie. Merci à "en face", Pierre-Louis, Malo et Gwenaël, pour la coinche, le tarot et les repas clandestins ; mention spéciale à PL pour accepter de courir à ma vitesse. Merci aussi à Alexandre et Edwige, à Benoît et Juliette, aux deux Margaux, à Amaury, Quentin, Laura, Thomas, Théo, Victor, Alban, Marine, Brieuc, Camille et tous les autres, pour les discussions autour d'un apéro, les jeux, et les expériences partagées de ski, de vélo sous la pluie, et de crapahutage dans les montagnes de Samoëns.

Je termine par la famille, que je ne peux que remercier d'être toujours présente. Merci à Martin et Simon – ils terminent leurs études également, meilleurs vœux pour la suite ! Merci à mes grands-parents pour l'invitation constante à venir prendre l'air breton. Merci à mes parents pour les prêts de voiture, les dîners d'anniversaire et le premier confinement cinq étoiles, en plus du reste. Merci aux Williart pour leur accueil pendant les six dernières années déjà, et merci à mes cousin(e)s, oncles et tantes – que je devrais voir plus souvent. Enfin, merci à Alice, bien sûr : profite bien du chat, et j'ai hâte d'être l'année prochaine !

Résumé étendu en français

Le problème du *comptage de points* pour les courbes elliptiques consiste, étant donnée une courbe elliptique E sur un corps fini $\mathbb{F}_q$, à calculer le nombre de points de E définis sur $\mathbb{F}_q$. Ses analogues en dimension supérieure demandent, étant donnée une variété abélienne polarisée A sur $\mathbb{F}_q$, de calculer le polynôme caractéristique de l'endomorphisme de Frobenius π_A de A. Une des motivations principales pour étudier le problème du comptage de points provient de la cryptographie à clés publiques, qui utilise des courbe elliptiques et des Jacobiennes de dimension 2 dont le nombre de points est (quasiment) premier [Mil85, Kob87, Kob89]. Au-delà de cette application, le problème du comptage a un intérêt intrinsèque en théorie algorithmique des nombres.

L'algorithme de Schoof [Sch85, Sch95] pour les courbes elliptiques, et ses généralisations en dimensions supérieures [Pil90, GH00, GKS11, GS12, Abe18], permettent de résoudre le problème du comptage de points en temps polynomial en $\log(q)$. Pour une série de petits premiers ℓ, cet algorithme calcule le polynôme caractéristique de π_A modulo ℓ en étudiant l'action de π_A sur le sous-groupe de ℓ-torsion $A[\ell]$. On conclut ensuite à l'aide du théorème chinois et des bornes de Hasse–Weil [Sil09, §V.1, Thm. 1.1], [Mil86a, Thm. 19.1]. La méthode de Schoof est l'état de l'art pour le comptage de points en grande caractéristique, *sauf* dans le cas des courbes elliptiques où une amélioration importante existe.

Pour calculer le nombre de points d'une courbe elliptique E sur $\mathbb{F}_q$ modulo ℓ, Elkies [Elk98] montre que l'on peut, environ une fois sur deux, remplacer $E[\ell]$, noyau de l'*endomorphisme* $[\ell]$, par le noyau d'une *ℓ-isogénie* de source E. Celle-ci est calculée à l'aide du *polynôme modulaire elliptique* Φ_ℓ de niveau ℓ, une équation bivariée qui relie les j-invariants de courbes elliptiques ℓ-isogènes [Cox13, §11.C]. On peut résumer les prérequis de la méthode d'Elkies en trois points :

- Un algorithme de *calcul d'isogénies* permettant, à partir de Φ_ℓ et de deux courbes elliptiques ℓ-isogènes E, E', de calculer le noyau de l'isogénie correspondante [Elk98, BMSS08] ;

- Des *bornes de complexité* (degré et taille des coefficients) pour le polynôme Φ_ℓ, fournies par un résultat de Cohen [Coh84] explicité par la suite [BS10, Paz19] ;
- Un moyen d'*évaluer efficacement* le polynôme Φ_ℓ et ses dérivées, soit en précalculant ces polynômes [Eng09b, BLS12], soit de manière directe [Sut13].

La méthode d'Elkies permet de réduire le degré des polynômes manipulés au cours de l'algorithme de comptage. L'algorithme obtenu est de meilleure complexité asymptotique en moyenne que la méthode de Schoof [SS14, SS15], et le gain en pratique pour des données de taille cryptographique est clair. La méthode d'Elkies entre, avec d'autres idées dues à Atkin, dans la composition de l'*algorithme SEA* pour le comptage de points des courbes elliptiques [Sch95], implanté dans de nombreux logiciels de calcul formel et théorie des nombres comme Pari/GP ou Magma. Les polynômes modulaires elliptiques et l'algorithme d'Elkies apparaissent également dans les algorithmes fondés sur les *graphes d'isogénies* de courbes elliptiques [Koh96].

Le but de cette thèse est de généraliser la méthode d'Elkies aux variétés abéliennes de dimension supérieure, notamment dans deux cas où la méthode de Schoof est bien étudiée : les surfaces abéliennes principalement polarisées, avec ou sans multiplication réelle. Le rôle du polynôme modulaire elliptique Φ_ℓ est alors joué par les *équations modulaires* de Hilbert et de Siegel, respectivement, dont des exemples sont connus [BL09, Mil15, Mar20, MR20]. D'un point de vue géométrique, ces équations modulaires définissent le lieu, dans les espaces de modules correspondants, des surfaces abéliennes liées par une isogénie de type fixé. Le langage des variétés de Shimura PEL [Mil05] permet à la fois d'unifier ces deux exemples et d'en considérer de vastes généralisations.

Les contributions de cette thèse remplissent les trois prérequis de la méthode d'Elkies évoqués précédemment :

- Le chapitre 3 présente un algorithme de calcul d'isogénies entre surfaces abéliennes principalement polarisées, à partir des équations modulaires de Hilbert ou Siegel ;
- Le chapitre 5 présente des bornes de degré et de hauteur (i.e. de taille de coefficients) pour les équations modulaires dans le cadre général des variétés de Shimura PEL ;
- Le chapitre 6 présente un algorithme d'évaluation des équations modulaires de Hilbert et de Siegel pour les surfaces abéliennes via des approximations complexes.

Les chapitre 2 présente les espaces de modules de variétés abéliennes et les équations modulaires, et est nécessaire pour la suite de la thèse. Le chapitre 4 donne des résultats reliant la hauteur d'une fraction rationnelle à celles de ses évaluations, et est nécessaire pour le chapitre 5.

Enfin, le chapitre 7 combine les résultats précédents et les applique au problème du comptage de points. Nous obtenons en particulier un algorithme de comptage pour les surfaces abéliennes principalement polarisées à multiplication réelle (fixée) qui est, sous heuristiques et à facteurs constants près, de même complexité asymptotique que l'algorithme SEA pour les courbes elliptiques.

Ces résultats ont donné lieu à des prépublications : [KPR19, Kie20c, Kie20a] et [Kie20b] correspondent approximativement aux chapitres 3, 4, 5 et 6 respectivement. Le chapitre 6 reprend également une partie des résultats l'article [Kie21] à paraître. La suite de ce résumé présente les résultats principaux des différents chapitres.

Chapitre 3 : Calcul d'isogénies entre surfaces abéliennes

Ce chapitre étend l'algorithme de calcul d'isogénies à partir des polynômes modulaires, connu dans le cas des courbes elliptiques, au cas des surfaces abéliennes principalement polarisées avec ou sans multiplication réelle. Présentons ici le second cas : on s'intéresse, pour un premier ℓ, à calculer une *ℓ-isogénie* φ (de degré ℓ^2) entre deux surfaces abéliennes principalement polarisées A et A'. Soit $\mathcal{A}_2$ l'espace de modules correspondant ; il est birationnel à $\mathbb{P}^3$, et son corps des fonctions est engendré par les trois *invariants d'Igusa* notés j_1, j_2, j_3. L'*application de Kodaira–Spencer* relie les déformations possibles de l'isogénie φ, qui se calculent à l'aide des équations modulaires de Siegel et de leurs dérivées, à l'application $d\varphi\colon T_0(A) \to T_0(A')$ induite par φ sur les espaces tangents à A et A' en zéro.

Nous explicitons cette application de Kodaira–Spencer dans le cas des Jacobiennes de courbes de genre 2 : concrètement, une formule explicite permet d'évaluer certaines fonctions modulaires de Siegel vectorielles, les dérivées des invariants d'Igusa, en termes des coefficients d'une équation de courbe hyperelliptique de genre 2. Le résultat suivant en découle.

Théorème 1. *Soit ℓ un nombre premier, et soit k un corps de caractéristique 0 ou $> 8\ell + 7$. Soit $\mathcal{U}$ l'ouvert de $\mathcal{A}_2$ formé des surfaces abéliennes A sans automorphismes supplémentaires et telles que $j_3(A) \neq 0$. On suppose qu'il existe un algorithme permettant d'évaluer les dérivées des équations modulaires de Siegel de niveau ℓ en un point de $\mathcal{U} \times \mathcal{U}$ sur k où elles sont*

définies, utilisant $C_{\text{eval}}(\ell)$ opérations dans k. Soit $\varphi\colon A \to A'$ une ℓ-isogénie définie sur k telle que A, A' soient dans $\mathcal{U}$ et telle que la sous-variété de $\mathcal{A}_2 \times \mathcal{A}_2$ définie par les équations modulaires soit normale en (A, A'). Alors, étant donnés les invariants d'Igusa de A et A', on peut calculer une représentation explicite de l'isogénie φ en utilisant $O(C_{\text{eval}}(\ell)) + \widetilde{O}(\ell)$ opérations dans k.

Chapitre 4 : Hauteurs de fractions rationnelles et interpolation

Ce chapitre présente tout d'abord la notion de *hauteur* pour les polynômes et fractions rationnelles à coefficients dans un corps de nombres L. Cette quantité positive reflète la taille des coefficients. Par exemple, lorsque $L = \mathbb{Q}$ et $F \in \mathbb{Q}(X)$, on écrit $F = P/Q$ avec $P, Q \in \mathbb{Z}[X]$ premiers entre eux ; alors la hauteur $h(F)$ de F est donnée par le maximum des $\log|c|$, lorsque c parcourt les coefficients non nuls de P et Q.

Ce chapitre présente ensuite des résultats, nécessaires au chapitre 5, reliant la hauteur d'un polynôme ou d'une fraction rationnelle à la hauteur de ses *évaluations* en certains points ; cela permet de borner $h(F)$ lorsque F est une fraction rationnelle obtenue par interpolation, comme les équations modulaires. On considère ici des points d'évaluation x_i entiers et (presque) consécutifs, on se donne H telle que $h(F(x_i)) \leq H$ pour tout i, et l'on recherche en particulier des énoncés où $h(F)$ est bornée en $O(H)$.

Théorème 2. *Soit L un corps de nombres de degré d_L sur $\mathbb{Q}$, soit $[\![A, B]\!]$ un intervalle de $\mathbb{Z}$, et notons $D = B - A$ et $M = \max\{|A|, |B|\}$. Soit $F \in L(X)$ une fraction rationnelle de degré au plus $d \geq 1$. Soit $S \subset [\![A, B]\!]$ ne contenant pas de pôle de F, soit $\eta \geq 1$, et soit $H \geq \max\{4, \log(2M)\}$. Supposons de plus que*

1. *$h(F(x)) \leq H$ pour tout $x \in S$.*
2. *S contient au moins D/η éléments.*
3. *$D \geq \max\{\eta d^3 H, 4\eta d d_L\}$.*

Alors

$$h(F) \leq H + C_L \eta d \log(\eta d H) + d \log(2M) + \log(d+1)$$

où C_L désigne une constante qui dépend uniquement du corps de nombres L. On peut prendre $C_{\mathbb{Q}} = 960$.

Chapitre 5 : Bornes de degré et de hauteur pour les équations modulaires

Ce chapitre fournit des bornes de degré et de hauteur pour les équations modulaires dans le cadre général des variétés de Shimura PEL. On considère la variété de Shimura associée à un certain groupe réductif G et un choix de niveau, des composantes connexes $\mathcal{S}, \mathcal{T}$ de cette variété, et un choix d'invariants sur ces composantes connexes. Les variétés algébriques $\mathcal{S}$ et $\mathcal{T}$ sont des espaces de modules pour les variétés abéliennes complexes munies d'une polarisation, d'endomorphismes, et d'une structure de niveau de types fixés.

Pour un élément adélique δ de G, les *équations modulaires de niveau* δ sont des fractions rationnelles multivariées définissant la correspondance de Hecke de niveau δ dans le produit $\mathcal{S} \times \mathcal{T}$. Cette correspondance de Hecke décrit, au niveau des variétés abéliennes, un certain nombre $d(\delta)$ d'isogénies de degré noté $l(\delta)$. Nous montrons que le degré et la hauteur des équations modulaires de niveau δ est bornée en termes de $d(\delta)$ et $l(\delta)$.

Théorème 3. *Avec les notations ci-dessus, il existe des constantes C_1 et C_2 (indépendantes de δ) vérifiant la propriété suivante. Soit δ un élément adélique de G, et F une fraction rationnelle obtenue comme coefficient de l'une des équations modulaires de niveau δ. Alors le degré total de F est borné par $C_1\, d(\delta)$, et la hauteur de F est bornée par $C_2\, d(\delta) \log l(\delta)$.*

Dans le cas des équations modulaires de Siegel de niveau ℓ, nous avons $d(\delta) = \ell^3 + \ell^2 + \ell + 1$ et $l(\delta) = \ell^2$. Ainsi, le degré des équations modulaires de Siegel de niveau ℓ est $O(\ell^3)$, et leur hauteur est $O(\ell^3 \log \ell)$. Dans le cas du polynôme modulaire elliptique Φ_ℓ, on retrouve, aux constantes près, les bornes de degré et de hauteur précédemment connues.

Chapitre 6 : Évaluation des équations modulaires pour les surfaces abéliennes

Ce chapitre présente un algorithme d'évaluation des équations modulaires de Siegel et de Hilbert pour les surfaces abéliennes, ainsi que de leurs dérivées, en un point défini sur un corps de nombres (ou un corps fini, par relèvement). Cet algorithme est fondé sur des approximations complexes, et fait notamment intervenir le calcul rapide de certaines formes modulaires, les *thêta-constantes* en dimension deux.

Le calcul rapide de ces thêta-constantes repose classiquement sur deux heuristiques, à savoir la correction de certains choix de signes dans les

formules de duplication, et la bonne convergence d'itérations de Newton. Un résultat auxiliaire de ce chapitre permet de lever la première heuristique : les choix de signes habitullement utilisés dans l'algorithme de calcul de thêta-constantes en dimension deux sont en effet corrects. La seconde ne pose pas de problèmes expérimentalement.

En dehors de ces heuristiques, l'utilisation de l'arithmétique d'intervalles permet de certifier les résultats obtenus par l'algorithme d'évaluation des équations modulaires. Un cas particulier s'énonce comme suit.

Théorème 4 (Sous heuristiques relatives au calcul rapide de thêta-constantes en dimension deux)**.** *Soit ℓ un nombre premier, et $(j_1, j_2, j_3) \in \mathbb{Q}^3$ un triplet de hauteur $O(1)$ où les dénominateurs des équations modulaires de Siegel de niveau ℓ ne s'annulent pas. Alors on peut évaluer ces équations modulaires et leurs dérivées en (j_1, j_2, j_3) en utilisant $\widetilde{O}(\ell^6)$ opérations binaires.*

Ce coût est à comparer au coût de stockage des équations modulaires sous forme de fractions rationnelles, que l'on estime de $\Theta(\ell^{15} \log \ell)$ au vu des résultats du chapitre 5 : lorsque la dimension de l'espace de modules associé est au moins deux, le précalcul des équations modulaires ne semble pas la stratégie à adopter.

Dans le cas des équations modulaires de Hilbert, le coût de l'algorithme d'évaluation peut décroître jusqu'à $\widetilde{O}(\ell^2)$ opérations binaires seulement. Dans ce cas, la manipulation des équations modulaires devient asymptotiquement négligeable devant le reste de l'algorithme de comptage de points.

Contents

Notation and conventions

Here we gather notations appearing in the thesis. When appropriate, we indicate the page where a more precise definition can be found.

We also adopt the following conventions concerning the words *variety, degree*, and *lattice*. Let k be a field.

- An *algebraic variety* over k is an integral (i.e. reduced and irreducible) separated scheme of finite type over k.
- The *degree* of a rational fraction $F \in k(X)$ is $\max\{\deg(P), \deg(Q)\}$, where $P, Q \in k(X)$ are coprime and $F = P/Q$. The *degree* of a polynomial $P \in k[X_1, \ldots, X_n]$ is its total degree. The *degree* of a rational fraction $F \in k(X_1, \ldots, X_n)$ is the maximum of the total degrees of its numerator and denominator.
- Let V be a finite-dimensional vector space over $\mathbb{Q}$. A *lattice* in V is a subgroup of $V(\mathbb{Q})$ generated by a basis of V over $\mathbb{Q}$; a *lattice* in $V(\mathbb{R})$ is a subgroup of $V(\mathbb{R})$ generated by a basis of $V(\mathbb{R})$ over $\mathbb{R}$; if p is a prime number, a *lattice* in $V(\mathbb{Q}_p)$ is a subgroup of the form $\bigoplus_{i\in I} \mathbb{Z}_p e_i$ where $(e_i)_{i\in I}$ is a basis of $V(\mathbb{Q}_p)$ over $\mathbb{Q}_p$; and finally, a *lattice* in $V(\mathbb{A}_f)$ (see notation below) is a product of lattices in $V(\mathbb{Q}_p)$ for each p that are equal to $V(\mathbb{Z}_p)$ for all p but finitely many.

General notation

$[\![A, B]\!]$	Integer interval $\{n \in \mathbb{Z} \colon A \leq n \leq B\}$
$\# S$	Cardinality of S
$\mathbf{1}_P$	Function with value 1 when property P holds, and 0 otherwise
$\mathrm{Mat}_{n\times m}(A)$	Module of $n \times m$ matrices over the ring A
Sp_{2g}	Symplectic group of order $2g$ (p. 66)
GSp_{2g}	General symplectic group of order $2g$
m^t	Transpose of the matrix m
m^{-t}	Inverse of m^t
I_n	Identity matrix of size $n \times n$

$\mathrm{Diag}(a_1, \ldots, a_n)$	Diagonal $n \times n$ matrix with entries $a_1, \ldots, a_n$
$(A\ B)$	Concatenation of matrices A and B
$\mathbb{P}^n$	Projective space of dimension $n \geq 1$
$\mathbb{A}^n$	Affine space of dimension $n \geq 1$
$\mathbb{G}_m$	Multiplicative group
$\mathbb{Z}_F$	Ring of integers of the number field F
∂_F^{-1}	Inverse different of the number field F (p. 76)
(x)	Ideal generated by x
$\mathrm{Tr}_{F/\mathbb{Q}}$, $N_{F/\mathbb{Q}}$	Trace and norm in the finite extension $F/\mathbb{Q}$
$\mathcal{P}_F$	Set of finite primes in the number field F
$\bar{k}$	Algebraic closure of the field k
$V(k)$	Set of k-points of the algebraic variety V
$\Omega^1(V)$	Vector space of global differential forms on V
$\log^+(x)$	$\log(\max\{1, x\})$, where $x \in \mathbb{R}$
$f = O_X\big(g(n_1, n_2, \ldots)\big)$	There exists $C > 0$ depending on X such that $f(n_1, n_2, \ldots) \leq C \max\{1, g(n_1, n_2, \ldots)\}$ for all values of $n_1, n_2, \ldots \geq 0$
$f = \widetilde{O}_X\big(g(n_1, n_2, \ldots)\big)$	There exist $C, k > 0$ depending on X such that $f \leq C \max\big\{1, g(n_1, \ldots) \log^+(g(n_1, \ldots))^k\big\}$ for all values of $n_1, n_2, \ldots \geq 0$

Chapter 1

Φ_ℓ	Elliptic modular polynomial of level ℓ (p. 24)
$\mathbb{H}_1$	Poincaré upper half plane (p. 22)
$j(E)$, $j(\tau)$	j-invariant (p. 23)
$\Lambda(\tau)$, where $\tau \in \mathbb{H}_1$	Lattice $\mathbb{Z} \oplus \tau\mathbb{Z} \subset \mathbb{C}$
$E(\tau)$, where $\tau \in \mathbb{H}_1$	Complex elliptic curve $\mathbb{C}/\Lambda(\tau)$
$X(1)$	Modular curve of level one (p. 22)
$\theta_0(\tau), \theta_1(\tau)$ for $\tau \in \mathbb{H}_1$	Theta constants on $\mathbb{H}_1$ (p. 22)
$E[n]$	n-torsion subgroup of the elliptic curve E
$T_\ell(E)$	ℓ-adic Tate module of the elliptic curve E (p. 27)
$T_0(E)$	Tangent space of E at the neutral point
E_4, E_6	Normalized Eisenstein series on $\mathbb{H}_1$
$\mathcal{G}(\ell, \mathbb{F}_q)$ and others	Isogeny graphs of elliptic curves (p. 30)
$\mathrm{Ell}(\mathcal{O}, k)$	Set of elliptic curves with endomorphism ring $\mathcal{O}$ over k up to isomorphism (p. 31)
$\mathcal{A}_2$	Siegel moduli space for abelian surfaces (p. 36)
$\mathcal{A}_{2,F}$	Hilbert moduli space for abelian surfaces with real multiplication by $\mathbb{Z}_F$, where F is a real quadratic field (p. 36)

j_1, j_2, j_3	Igusa invariants on $\mathcal{A}_2$
$\Psi_{\ell,1}, \Psi_{\ell,2}, \Psi_{\ell,3}$	Siegel modular equations of level ℓ for principally polarized abelian surfaces

Chapter 2

$\widehat{A}$	Dual of the abelian variety A (p. 44)
t_x	Translation by x (p. 43)
$A[n]$	n-torsion subgroup of the abelian variety A
$\mathrm{Pic}^0(A)$	Group of line bundles algebraically equivalent to zero on A, up to linear equivalence (p. 44)
$\mathrm{NS}(A)$	Néron–Severi group of A (p. 44)
$\mathrm{Jac}(\mathcal{C})$	Jacobian of the smooth curve $\mathcal{C}$ (p. 45)
η_P	Embedding $\mathcal{C} \hookrightarrow \mathrm{Jac}(\mathcal{C}),\ Q \mapsto [Q-P]$
$\mathrm{End}_k(A)$	Ring of endomorphisms of A defined over k
$\mathrm{End}^0(A)$	Endomorphism algebra of A over $\bar{k}$ (p. 47)
$\dagger$	Rosati involution (p. 47)
$\mathrm{End}(A)^\dagger$	Set of real endomorphisms of A (p. 47)
$\mathrm{End}^0(A)^\dagger$	Subspace of invariants under $\dagger$ in $\mathrm{End}^0(A)$
$\mathbb{A}_f$	Ring of finite adeles of $\mathbb{Q}$ (p. 50)
$\widehat{\mathbb{Z}}$	Profinite completion of $\mathbb{Z}$ (p. 51)
G^{der}, Z, G^{ad}, T, ν, $G(\mathbb{R})_+$, $G(\mathbb{Q})_+$	See p. 53
(G, X_+)	PEL datum (p. 55)
$\mathrm{Sh}_K(G, X_+)$ (or Sh_K)	PEL Shimura variety of level K attached to the PEL datum (G, X_+) (p. 55)
$\mathcal{C} \subset G(\mathbb{A}_f)$	Set of representatives for connected components of $\mathrm{Sh}_K(\mathbb{C})$ (p. 56)
$\mathcal{S}_c = \Gamma_c \backslash X_+$	Connected component of $\mathrm{Sh}_K(\mathbb{C})$ associated with $c \in \mathcal{C}$ (p. 56)
$E(G, X_+)$	Reflex field of the PEL datum (G, X_+) (p. 56)
(Λ_c, ψ_c)	Standard lattice attached to $c \in \mathcal{C}$ (p. 58)
$T_p(A)$	p-adic Tate module of A (p. 60)
$\widehat{T}(A)$	Global Tate module of A (p. 60)
$a_n(f)$	Fourier coefficient of f (p. 70)
q_1, q_2, q_3	Variables in Fourier expansions on $\mathbb{H}_2$ (p. 70)
$\theta_{a,b}$ for $a, b \in \{0,1\}^g$	Theta constants on $\mathbb{H}_g$ (p. 71)
θ_j for $0 \le j \le 15$	Theta constants on $\mathbb{H}_2$ (p. 71)
$h_4, h_6, h_{10}, h_{12}, \psi_4, \psi_6, \chi_{10}, \chi_{12}$	Siegel modular forms on $\mathbb{H}_2$ (p. 72)
j_1, j_2, j_3	Igusa invariants on $\mathbb{H}_2$ (p. 73)

$\det^k \mathrm{Sym}^n$	Irreducible representations of $\mathrm{GL}_2(\mathbb{C})$ (p. 73)
$\frac{df}{d\tau}$	Vector-valued derivative of a Siegel modular form f on $\mathbb{H}_2$ (p. 74)
$f_{8,6}$	Siegel modular form of weight $\det^8 \mathrm{Sym}^6$ (p. 75)
H_R	Hilbert embedding (p. 77)
G_2, F_6, F_{10}	Hilbert modular forms for $F = \mathbb{Q}(\sqrt{5})$ (p. 80)
g_1, g_2	Gundlach invariants for $F = \mathbb{Q}(\sqrt{5})$ (p. 81)
b_j for $0 \leq j \leq 6$	Coefficients of the pullback of $f_{8,6}$ on $\mathbb{H}_1^2$ (p. 81)
H_δ	Hecke correspondence of level $\delta \in G(\mathbb{A}_f)$ (p. 84)
$d(\delta)$	Degree of the Hecke correspondence H_δ (p. 85)
$l(\delta)$	Degree of the isogenies described by H_δ (p. 86)
$\Psi_{\delta,m}$	Modular equations of level δ (p. 89)

Chapter 3

$\mathcal{A}_g$	Moduli space of principally polarized (p.p.) abelian varieties of dimension g
$\mathcal{A}_{g,\ell}$	Moduli space of p.p. abelian varieties of dimension g endowed with the kernel of an ℓ-isogeny
$H_\ell = (H_{\ell,1}, H_{\ell,2})$	Map $\mathcal{A}_{g,\ell} \to \mathcal{A}_g \times \mathcal{A}_g$, $(A, K) \mapsto (A, A/K)$ (p. 96)
$d\varphi$	Tangent map of the isogeny φ (p. 96)
$\omega(\mathcal{C})$	Basis of differential forms on $\mathrm{Jac}(\mathcal{C})$ associated with the hyperelliptic equation $\mathcal{C}$ (p. 98)
$A(\tau)$	P.p. abelian surface $\mathbb{C}^2/\Lambda(\tau)$ attached to $\tau \in \mathbb{H}_2$ (p. 99)
$A_F(\tau)$	P.p. abelian surface $\mathbb{C}^2/\Lambda_F(\tau)$ with real multiplication by $\mathbb{Z}_F$ attached to $\tau \in \mathbb{H}_1^2$ (p. 100)
$\omega(\tau)$	Standard basis of differential forms on $A(\tau)$ or $A_F(\tau)$ (p. 99)
$\mathcal{C}(\tau), \mathcal{C}_F(\tau)$	Standard genus 2 hyperelliptic equations attached to $\tau \in \mathbb{H}_2$ (resp. $\tau \in \mathbb{H}_1^2$) (p. 100)
$\mathrm{Cov}(f)$	Covariant attached to the Siegel modular form f (p. 101)
$I_2, I_4, I_6, I_6', I_{10}, R$	Scalar-valued covariants (p. 103)
y_1, y_2, y_3, X	Vector-valued covariants (p. 103)
$\partial j(\tau)$	3×3 matrix of derivatives of Igusa invariants at $\tau \in \mathbb{H}_2$ (p. 108)
$D\Psi_{\ell,L}$, $D\Psi_{\ell,R}$	3×3 matrices of derivatives of Siegel modular equations of level ℓ (p. 109)
$\partial j^{(F)}(\tau)$	3×2 matrix of derivatives of Igusa invatiants at $\tau \in \mathbb{H}_1^2$ (p. 111)

$D\Psi_{\beta,L},\ D\Psi_{\beta,R}$	3×3 matrices of derivatives of Hilbert modular equations of level β (p. 111)
$W(k)$	Witt ring of the field k
$\mathscr{X}_g \to \mathscr{A}_g$	Moduli stack of abelian varieties of dimension g with its universal family (p. 115)
$\mathscr{A}_g(\ell), \mathscr{A}_{g,\ell}, \mathscr{A}_{g,F}, \mathscr{A}_{g,\beta}$	Moduli stacks of abelian varieties of dimension g with suitable structure (pp. 116 and 117)
$\mathcal{C}^{<2>}$	Symmetric square of the curve $\mathcal{C}$
φ^*	Pullback by the isogeny φ on line bundles or differential forms
Θ	Theta divisor (p. 126)
$\mathcal{L}_A^\alpha$	Line bundle on A (up to algebraic equivalence) associated with $\alpha \in \mathrm{End}(A)^\dagger$ (p. 127)
$(\mathcal{L} \cdot \mathcal{L}')$	Intersection pairing for line bundles on a smooth algebraic surface
$C_{\mathrm{eval}}(\ell),\ C_{\mathrm{eval}}(\beta)$	Cost of evaluation for modular equations of Siegel type and level ℓ (resp. Hilbert type and level β) (p. 130)

Chapter 4

$h(F)$	Height of F (p. 139)
d_L	Degree of the number field L over $\mathbb{Q}$
$\mathcal{V}_L, \mathcal{V}_L^0, \mathcal{V}_L^\infty$	Set of all (resp. all non-archimedean, resp. all archimedean) places of L
$v_{\mathfrak{p}}$	$\mathfrak{p}$-adic valuation, where $\mathfrak{p} \in \mathcal{P}_L$ (p. 139)
$\lvert\cdot\rvert_v$	Normalized absolute value associated with the place v (p. 139)
$D!$	Factorial of $D \geq 1$
$\widetilde{h}(x)$	$\frac{1}{d_L} \log N_{L/\mathbb{Q}}(x)$ (p. 146)

Chapter 5

$j_1, \ldots, j_{n+1}$	Modular functions on $\mathcal{S}, \mathcal{T}$ (p. 158)
$J_1, \ldots, J_{n+1}, Y_1, \ldots$	Indeterminates
E	Equation satisfied by $j_1, \ldots, j_{n+1}$ on $\mathcal{S}$ (p. 158)
Σ, K_0, K', K''	See p. 159
g_δ, m	Denominator of $\Psi_{\delta,m}$ (p. 160)
$\mathrm{SGC}(j_1, \ldots, j_{n+1})$	Symmetric geometric complexity (p. 161)
e	Degree of E in J_{n+1}
d_E	Total degree of E in $J_1, \ldots, J_n$

$\mathrm{Res}_X(P,Q)$	Resultant of P, Q with respect to the variable X
$\mathrm{GC}(j_1,\dots,j_{n+1})$	Geometric complexity (p. 167)
$\Gamma^0(\ell)$	Congruence subgroup of $\mathrm{Sp}_4(\mathbb{Z})$ (p. 168)
C_ℓ	Set of representatives for $\Gamma^0(\ell)\backslash \mathrm{Sp}_4(\mathbb{Z})$ (p. 168)
g_ℓ	Classical Siegel modular form; denominator of Siegel modular equations of level ℓ (p. 168)
w_ℓ	Weight of g_ℓ
$f^{(i)}_{\ell,m}(\tau)$	Coefficient of Y^i in $g_\ell(\tau)\Psi_{\ell,m}\big(j_1(\tau), j_2(\tau), j_3(\tau)\big)$
D_ℓ	Complete denominator of Siegel modular equations of level ℓ (p. 170)
$\Gamma^0_F(\beta)$	Congruence subgroup of $\mathrm{SL}(\mathbb{Z}_F \oplus \partial_F^{-1})$ (p. 171)
Δ	Discriminant of F
C_β, C^σ_β	Set of representatives modulo $\Gamma^0_F(\beta)$ (p. 171)
g_β	Classical Hilbert modular form; denominator of Hilbert modular equations of level β (p. 171)
w_β	Weight of g_β
$f^{(i)}_{\beta,m}(\tau)$	Coefficient of Y^i in $g_\beta(\tau)\Psi_{\beta,m}\big(g_1(\tau), g_2(\tau)\big)$
D_β	Complete denominator of Hilbert modular equations of level β (p. 172)
$h_F(A)$	Faltings height of the abelian variety A
$h_{\Theta,r}(A,\mathcal{L})$	Theta height of level r of the principally polarized abelian variety $(A,\mathcal{L})$
$h_j(A)$	j-height of A (p. 175)
$\overline{h}_F, \overline{h}_{\Theta,r}, \overline{h}_j$	$\max\{1, h_F\}$, etc.
$C(g,r)$	See p. 176
$\mathcal{A}_{\Theta,r}$	Moduli space of principally polarized abelian varieties of dimension g with level r theta structure (p. 176)
$\mathcal{U}$	Zariski open set of $\mathcal{S}$ (p. 179)
$\mathcal{V}_\delta$	Zariski open set of $\mathcal{S}$; in particular the modular equations $\Psi_{\delta,m}$ have no poles on V_δ (p. 180)
$C_{X.Y.Z}$	Constant ≥ 1 so that prop. $X.Y.Z$ holds with this value of C
$C_{\log}$	We have $\log d(\delta) \leq C_{\log} \max\{1, \log l(\delta)\}$ (p. 182)
$\mathcal{I}_k(T)$	Evaluation set of the tree T (p. 182)

Chapter 6

$\log_2^+(x)$	$\log_2 \max\{1, x\}$
$\lvert P\rvert$	Largest absolute value among the coefficient of P (p. 194)

α_d	Hermite's constant (p. 198)
m_F	Generating matrix of $\mathbb{Z}_F$ in the complex embeddings of F (p. 129)
$B(\tau)$	Borchardt sequence of theta constants at $2^n\tau$ for $n \geq 0$ (pp. 200 and 202)
γ_k for $0 \leq k \leq 3$	Matrices in $\mathrm{Sp}_4(\mathbb{Z})$ (p. 202)
$\mathcal{F}_2$	Fundamental domain in $\mathbb{H}_2$ (p. 203)
$\mathfrak{S} \subset \mathrm{Sp}_4(\mathbb{Z})$	Finite set defining the boundary of $\mathcal{F}_2$ (p. 203)
$z_j(\tau) = x_j(\tau) + iy_j(\tau)$ for $1 \leq j \leq 3$	Entries of $\tau \in \mathbb{H}_2$ (p. 204)
$q_j(\tau), \lambda_1(\tau), r(\tau)$	See p. 204
$\mathcal{F}_2'$	Subset of $\mathbb{H}_2$ containing $\mathcal{F}_2$ (p. 204)
$\mathcal{F}_2^\varepsilon$	Neighborhood of $\mathcal{F}_2$ (p. 216)
$\Lambda(\tau)$	$\log \max\{2, \lvert\tau\rvert, \det(\mathrm{Im}\,\tau)^{-1}\}$ (p. 217)
$m_1(\tau), m_2(\tau)$	Successive minima of $\mathrm{Im}(\tau)$ on $\mathbb{Z}^2$ (p. 217)
Θ	Finite recipe of algebraic extensions (p. 223)
$B_{\Theta,\mu}$	Upper bound on the absolute values of the algebraic numbers in $\Theta(j_1, j_2, j_3)$ in the complex embedding μ of L (p. 223)

Chapter 7

π_A	Frobenius endomorphism of A
χ_A	Characteristic polynomial of π_A (p. 234)
ξ_A	Real endomorphism $\pi_A + \pi_A^\dagger$ (p. 234)
$\lambda_1, \lambda_2, r_1, r_2, s_1, s_2$	See p. 234
$\langle \cdot, \cdot \rangle$	Weil pairing on $A[\ell]$ (p. 235)
$\widetilde{P}$	q-reciprocal of the polynomial P (p. 235).

Chapter 1

Introduction

This thesis is about *higher-dimensional modular equations*. These equations encode isogenies between polarized abelian varieties; from a geometrical point of view, they define subvarieties of certain moduli spaces consisting of pairs of isogenous abelian varieties.

The first examples of modular equations are the *elliptic modular polynomials*, which encode isogenies between abelian varieties of dimension 1, i.e. elliptic curves. Let $\ell \geq 1$ be a prime number. The elliptic modular polynomial of level ℓ, denoted by $\Phi_\ell \in \mathbb{Z}[X, Y]$, satisfies the following property: if k is an algebraically closed field of characteristic distinct from ℓ, and E, E' are elliptic curves over k, then there exists an isogeny $\varphi\colon E \to E'$ of degree ℓ if and only if the equality

$$\Phi_\ell(j(E), j(E')) = 0$$

holds, where j denotes the usual j-invariant.

This introductory chapter presents the construction of elliptic modular polynomials (§1.1) before reviewing existing applications in algorithmic number theory and cryptography (§1.2), in particular Elkies's method for counting points on elliptic curves over finite fields. Other examples of modular equations in dimension 2, i.e. for abelian surfaces, have also been computed, and we give an overview of the state of the art on this subject (§1.3). Finally we present the plan of the thesis (§1.4). Our contributions aim to generalize properties and algorithms from the case of elliptic modular polynomials to higher dimensions. Combining our results yields an asymptotically faster point-counting algorithm based on Elkies's method for principally polarized abelian surfaces with real multiplication.

1.1 Elliptic modular polynomials

Throughout this introduction, we assume familiarity with elementary properties of elliptic curves, as presented in Silverman's book [Sil09]. In particular, elliptic curves are algebraic curves, and can be defined over any field, such as finite fields. Over $\mathbb{C}$, elliptic curves are identified with their sets of $\mathbb{C}$-points, which are complex tori. We also assume familiarity with classical modular forms; see for instance [DS05, §1–2].

Perhaps the easiest way to define elliptic modular polynomials is to work over $\mathbb{C}$. We recall the necessary facts about the moduli space of complex elliptic curves in §1.1.1. Then we define the elliptic modular polynomials and list some of their properties in §1.1.2.

1.1.1 The moduli space of complex elliptic curves

Denote by $\mathbb{H}_1$ the Poincaré upper half plane,

$$\mathbb{H}_1 = \{z \in \mathbb{C} \colon \operatorname{Im}(z) > 0\}.$$

The elliptic modular group $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$ acts on $\mathbb{H}_1$ on the left in the following way: if $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(1)$ and $\tau \in \mathbb{H}_1$, then

$$\gamma\tau = \frac{a\tau + b}{c\tau + d}. \tag{1.1}$$

With every $\tau \in \mathbb{H}_1$, one associates the lattice $\Lambda(\tau) = \mathbb{Z} \oplus \tau\mathbb{Z} \subset \mathbb{C}$ and the complex elliptic curve $E(\tau) = \mathbb{C}/\Lambda(\tau)$. The map $\tau \mapsto E(\tau)$ induces a bijection between the quotient $\Gamma(1)\backslash\mathbb{H}_1$ and the set of isomorphism classes of complex elliptic curves. We say that $\Gamma(1)\backslash\mathbb{H}_1$ is a *moduli space* for complex elliptic curves.

The quotient $\Gamma(1)\backslash\mathbb{H}_1$ can be compactified by adding one point called the *cusp*. The resulting compactification, denoted by $X(1)$, is the *modular curve of level one* over $\mathbb{C}$. As every compact Riemann surface, $X(1)$ is an algebraic curve over $\mathbb{C}$. Rational functions $X(1) \to \mathbb{C}$ are identified with modular functions of weight zero and level $\Gamma(1)$ on $\mathbb{H}_1$.

An isomorphism $X(1) \simeq \mathbb{P}^1$ is given by the *j-invariant* $\tau \mapsto j(\tau)$. Consider the *theta constants* $\theta_0, \theta_1 \colon \mathbb{H}_1 \to \mathbb{C}$ defined as follows:

$$\theta_0(\tau) = \sum_{n\in\mathbb{Z}} \exp(\pi i \tau n^2),$$
$$\theta_1(\tau) = \sum_{n\in\mathbb{Z}} (-1)^n \exp(\pi i \tau n^2).$$

Then the j-invariant is defined by

$$j = 256\frac{(\theta_0^8 - \theta_0^4\theta_1^4 + \theta_1^8)^3}{\theta_0^8\theta_1^8(\theta_0^4 - \theta_1^4)^2}. \tag{1.2}$$

In particular, the j-invariant generates the function field of $X(1)$ over $\mathbb{C}$. Therefore every modular function of weight zero and level $\Gamma(1)$ can be expressed as a rational function in j with coefficients in $\mathbb{C}$. On $X(1)$, the j-invariant has only one pole located at the cusp; therefore one can consider the j-invariant as a quantity attached to elliptic curves, with the property that two complex elliptic curves have equal j-invariants if and only if they are isomorphic.

Far-reaching generalizations of the modular curve $X(1)$ are given by the *PEL Shimura varieties* of finite level (see chapter 2). PEL Shimura varieties are moduli spaces for abelian varieties with certain polarization, endomorphism, and level structures. They provide a convenient unified language to study higher-dimensional modular equations.

1.1.2 Elliptic modular polynomials

Recall that an *isogeny* $\varphi\colon E \to E'$ between complex elliptic curves is a surjective morphism with finite kernel. The *degree* of φ is its degree as a rational map; and the *kernel* of φ is its kernel as a group morphism from $E(\mathbb{C})$ to $E'(\mathbb{C})$, which has cardinality $\deg(\varphi)$.

Let us restrict to isogenies of prime degree; every isogeny is a composition of these. If an isogeny φ has prime degree ℓ, then the kernel of φ is a cyclic group isomorphic to $\mathbb{Z}/\ell\mathbb{Z}$; we say that φ is an *ℓ-isogeny*. For every $\tau \in \mathbb{H}_1$, the identity map on $\mathbb{C}$ induces an ℓ-isogeny $E(\tau) \to E(\tau/\ell)$. Moreover every ℓ-isogeny over $\mathbb{C}$ is isomorphic to an isogeny of this form.

One can check that the function $\tau \mapsto j(\tau/\ell)$ is a modular function of weight zero for the congruence subgroup

$$\Gamma^0(\ell) = \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})\colon b = 0 \bmod \ell\} \tag{1.3}$$

of $\Gamma(1)$. Therefore, for every $\tau \in \mathbb{H}_1$, the polynomial

$$P_\ell(\tau) = \prod_{\gamma \in \Gamma^0(\ell)\backslash\Gamma(1)} \left(Y - j\left(\tfrac{1}{\ell}\gamma\tau\right)\right) \quad \in \mathbb{C}[Y]$$

is well-defined. The roots of $P_\ell(\tau)$ are the j-invariants of all elliptic curves that are ℓ-isogenous to $E(\tau)$; more precisely, the roots of $P_\ell(\tau)$ counted with multiplicities are the j-invariants of the elliptic curves of the form $E(\tau)/K$,

where K runs through the $\ell+1$ cyclic subgroups of the ℓ-torsion subgroup $E(\tau)[\ell]$.

One checks that the coefficients of P_ℓ are modular functions of weight zero and level $\Gamma(1)$ with poles at the cusp only. Therefore, there exists a polynomial $\Phi_\ell \in \mathbb{C}[X,Y]$ such that for every $\tau \in \mathbb{H}_1$, we have

$$\Phi_\ell(j(\tau), Y) = P_\ell(\tau) = \prod_{\gamma \in \Gamma^0(\ell)\backslash\Gamma(1)} \Big(Y - j\left(\tfrac{1}{\ell}\gamma\tau\right)\Big). \tag{1.4}$$

This bivariate polynomial Φ_ℓ is the *elliptic modular polynomial of level* ℓ. Let us list some of its properties.

1. The coefficients of Φ_ℓ lie in fact in $\mathbb{Z}$: one way to see this is to show that the Fourier expansions of the coefficients of P_ℓ have integer coefficients [Cox13, Thm. 11.18 (i)].
2. The polynomial Φ_ℓ is symmetric in X and Y. This comes from the existence of the *dual isogeny*: for every ℓ-isogeny $\varphi\colon E \to E'$, the dual of φ is an ℓ-isogeny $E' \to E$ [Cox13, Thm. 11.18 (iii)].
3. The degree of Φ_ℓ is $\ell+1$ in both variables. Indeed, its degree in Y is $[\Gamma(1) : \Gamma^0(\ell)] = \ell+1$, and Φ_ℓ is symmetric.
4. Define the height of Φ_ℓ, denoted by $h(\Phi_\ell)$, as follows: $h(\Phi_\ell)$ is the maximum value of $\log|c|$, where c runs through the nonzero coefficients of Φ_ℓ. Then, as ℓ grows to infinity, we have

   $$h(\Phi_\ell) \sim 6\ell \log \ell.$$

   This asymptotic result was first proved by Cohen [Coh84], and explicit bounds from above for $h(\Phi_\ell)$ are also known [BS10, Paz19].
5. Let k be an algebraically closed field of characteristic distinct from ℓ, and let E, E' be two elliptic curves over k. Then the equality

   $$\Phi_\ell(j(E), j(E')) = 0$$

   holds if and only if E and E' are ℓ-isogenous over k. By construction, this property holds true over $\mathbb{C}$, and extends to every characteristic zero field by Lefschetz's principle (every field k that is finitely generated over $\mathbb{Q}$ embeds in $\mathbb{C}$) [Sil09, §VI.6]. The general argument uses the language of moduli schemes: one invokes the smoothness of the moduli space of elliptic curves with level ℓ structure over $\mathbb{Z}[1/\ell]$ [DR73, §IV, Thm. 2.5] to lift ℓ-isogenies from k to characteristic zero.

Given the above bounds on the degree and height of Φ_ℓ, the cost of storing Φ_ℓ on a computer is $O(\ell^3 \log \ell)$. This becomes quickly too large to fit Φ_ℓ on a page, but is reasonable enough to store large databases of these polynomials: for instance, Sutherland [Sut] has a database of all elliptic modular polynomials of prime levels up to 1000. As an example, we have

$$\begin{aligned}\Phi_2(X,Y) = X^3 - X^2Y^2 + 1488X^2Y - 162000X^2 + 1488XY^2 \\ + 40773375XY + 8748000000X + Y^3 - 162000Y^2 \\ + 8748000000Y - 157464000000000.\end{aligned}$$

Several methods are available to compute elliptic modular polynomials in quasi-linear time in their size. One method consists in computing complex approximations of its coefficients using eq. (1.4) [Eng09b]. An essential subroutine of this algorithm is to compute theta constants, and hence the j-invariant by eq. (1.2), in quasi-linear time: for every precision $N \geq 1$ and every $\tau \in \mathbb{H}_1$ lying in a certain fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$, the theta-constants $\theta_0(\tau), \theta_1(\tau)$ can be computed up to an error of 2^{-N} in quasi-linear time in N [Dup11]. Another method to compute Φ_ℓ consists in using the Chinese remainder theorem (CRT), and computing Φ_ℓ modulo several prime numbers p by generating sufficiently many pairs of ℓ-isogenous curves over $\mathbb{F}_p$ [BLS12]. Both methods are suitable to evaluate $\Phi_\ell(x_0, Y)$, for a given x_0 in a finite field or number field; the CRT method even achieves memory improvements in this case [Sut13].

Other kinds of modular polynomials for elliptic curves can be defined, using other modular functions than the j-invariant, for example the Weber f-function, whose congruence subgroup has index 24 in $\mathrm{SL}_2(\mathbb{Z})$. Their heights are smaller by a constant factor (see [EM02] for an explanation of this phenomenon in the case of class polynomials of imaginary quadratic fields), but the asymptotic complexity of manipulating them is similar.

1.2 Algorithmic applications

This section presents several algorithms using elliptic modular polynomials. We focus on elliptic curves over finite fields, although some of the algorithms also apply to other base fields or rings. Let p be a prime number, and let $q = p^r$ be a power of p.

Let E be an elliptic curve over the finite field $\mathbb{F}_q$, and let $\ell \neq p$ be a prime. A direct use of the elliptic modular polynomial of level ℓ is to compute all elliptic curves (up to isomorphism) that are ℓ-isogenous to E: one computes the roots of $\Phi_\ell(j(E), Y) \in \mathbb{F}_q[Y]$, using for instance the

Cantor–Zassenhaus algorithm [vG13, §14]. The polynomial Φ_ℓ can also be used to *compute* isogenies: given two ℓ-isogenous elliptic curves E, E' over $\mathbb{F}_q$, an algorithm of Elkies [Elk98] recovers an ℓ-isogeny $\varphi\colon E \to E'$ as an explicit rational map, at least when p is large with respect to ℓ and E, E' are sufficiently generic. Elkies's algorithm is an essential part of the SEA point counting algorithm for elliptic curves over finite fields [Sch95]. Note that Elkies's algorithm has a converse: given a cyclic subgroup $K \subset E(\overline{\mathbb{F}}_q)$ of order ℓ, one can compute an equation for the elliptic curve E/K as well as rational fractions describing the ℓ-isogeny $E \to E/K$ using Vélu's formulæ [Vé71].

The first part of this section presents the point counting problem and the SEA algorithm, starting with Schoof's algorithm (§1.2.1), then presenting Elkies's improvement (§1.2.2). The second part discusses other applications of modular polynomials in algorithmic number theory (§1.2.3) and cryptography (§1.2.4).

1.2.1 Schoof's point counting algorithm

The *point counting problem* for elliptic curves is the following: given an elliptic curve E over $\mathbb{F}_q$, compute $\#E(\mathbb{F}_q)$. Counting points has applications in cryptography (see §1.2.4), and also in algorithmic number theory, for instance to compute Euler factors of L-functions attached to elliptic curves. Schoof's algorithm solves the point counting problem in polynomial time in $\log(q)$.

Let π_E be the Frobenius endomorphism of E, and consider its *characteristic polynomial* [Sil09, §V, Thm. 2.3.1]:

$$\chi_E = X^2 - t_E X + q \in \mathbb{Z}[X],$$

where $t_E \in \mathbb{Z}$ is the *trace of Frobenius*. If we consider $\mathbb{Z}$ as a subset of $\mathrm{End}(E)$, the polynomial χ_E is equal to $(X - \pi_E)(X - \widehat{\pi}_E)$ where $\widehat{\pi}_E$ denotes the dual of Frobenius; hence $\chi_E(\pi_E) = 0$. Computing χ_E is equivalent to computing $\#E(\mathbb{F}_q)$ since

$$\#E(\mathbb{F}_q) = q + 1 - t_E.$$

The Weil conjectures, in this case Hasse's theorem [Sil09, §V, Thm. 2.4], state that the roots of χ_E in $\mathbb{C}$ all have absolute value $\sqrt{q}$. Therefore the *Hasse bound* holds:

$$|t_E| \leq 2\sqrt{q}. \tag{1.5}$$

Schoof's algorithm [Sch85] computes $t_E \bmod \ell$ for a series of distinct prime numbers $\ell \neq p$, then recovers $t_E \in \mathbb{Z}$ using the Chinese remainder

theorem and eq. (1.5). The trace t_E is recovered uniquely as soon as the product of ℓ's is greater than $4\sqrt{q}$. By the prime number theorem, this bound can be reached using $O(\log q)$ prime numbers $\ell \in O(\log q)$.

Let us now turn to the computation of $t_E \bmod \ell$, where $\ell \neq p$ is a prime number. Let

$$T_\ell(E) = \varprojlim_{n\to\infty} E[\ell^n]$$

be the ℓ-adic Tate module of E, which is a free $\mathbb{Z}_\ell$-module of rank 2. Then χ_E, seen as a polynomial with coefficients in $\mathbb{Z}_\ell$, is the characteristic polynomial of π_E seen as an automorphism of $T_\ell(E)$ [Sil09, §III, Prop. 8.6]. In particular, $t_E \bmod \ell$ is the unique element $t \in \mathbb{Z}/\ell\mathbb{Z}$ satisfying the relation

$$t\,\pi_E(P) = \pi_E^2(P) + qP \tag{1.6}$$

for every ℓ-torsion point $P \in E[\ell](\overline{\mathbb{F}}_q)$.

One usually avoids computing all ℓ-torsion points on E, and rather tests eq. (1.6) on the formal ℓ-torsion point P_0 with coefficients in the coordinate ring of $E[\ell]\backslash\{0\}$. One computes $\pi_E(P_0)$, $\pi_E^2(P_0)$ and qP_0 using square-and-multiply algorithms, the latter using the addition law on E, for instance [Sil09, §II, Alg. 2.3]. Using asymptotically fast algorithms for polynomial multiplication [vG13, §8–9], the cost of finding $t_E \bmod \ell$ is then $\widetilde{O}(\ell^2 \log q)$ operations in $\mathbb{F}_q$. The ℓ^2 factor comes from the degree of defining equations for $E[\ell]\backslash\{0\}$; in other words this ℓ^2 stands for the *order* of the subgroup of E used to compute $t_E \bmod \ell$. Overall, the complexity of Schoof's algorithm is $\widetilde{O}(\log^5(q))$ binary operations.

1.2.2 Elkies's method and the SEA algorithm

Let $\ell \neq p$ be an odd prime number. Given Weierstrass models of two ℓ-isogenous elliptic curves E, E' over $\mathbb{F}_q$,

$$E\colon x^2 = y^3 + ax + b, \qquad E'\colon y^2 = x^3 + a'x + b',$$

Elkies's algorithm [Elk98] attemps to compute an ℓ-isogeny $\varphi\colon E \to E'$ as an explicit rational map. For simplicity, assume that the polynomial $\Phi_\ell(j(E), Y)$ has simple roots over $\overline{\mathbb{F}}_q$, and that E and E' have no extra automorphisms. Then Elkies's algorithm succeeds as soon as $p > 4\ell - 1$. The result yields an equation for the subgroup $\ker(\varphi)$ of $E[\ell]$, which is defined over $\mathbb{F}_q$ and cyclic of order ℓ.

In the context of Schoof's algorithm, assume that the discriminant of χ_E is a nonzero square modulo ℓ. Then the endomorphism π_E restricted to $E[\ell]$

diagonalizes over $\mathbb{Z}/\ell\mathbb{Z}$: in some basis of $E[\ell]$, the matrix of π_E is of the form

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

for some distinct eigenvalues $\lambda, \mu \in (\mathbb{Z}/\ell\mathbb{Z})^\times$ such that $\lambda\mu = q$. There are exactly two cyclic subgroups of $E[\ell]$ which are stable under π_E, i.e. defined over $\mathbb{F}_q$. Assume, as above, that $\Phi_\ell(j(E), Y)$ has only simple roots over $\overline{\mathbb{F}}_q$. Then *Elkies's method* to compute $t_E \bmod \ell$ is the following:

1. Compute the two roots α, β of $\Phi_\ell(j(E), Y)$ belonging to $\mathbb{F}_q$;
2. Construct an elliptic curve E' over $\mathbb{F}_q$ such that $j(E') = \alpha$;
3. Compute the kernel of an ℓ-isogeny $\varphi\colon E \to E'$ using Elkies's algorithm;
4. Compute the eigenvalue of π_E on $\ker(\varphi)$, say λ;
5. Output $t_E = \lambda + q/\lambda \bmod \ell$.

Step 4 uses the formal point of $\ker(\varphi)\backslash\{0\}$ in a similar way to Schoof's algorithm; it costs $\widetilde{O}(\ell \log q)$ operations in $\mathbb{F}_q$ since $\ker(\varphi)$ has order ℓ, and dominates the rest of the algorithm.

Heuristically, Elkies's method applies for about half of the auxiliary primes ℓ, called *Elkies primes*. If enough Elkies primes can be found, then the resulting point-counting algorithm has a complexity of $\widetilde{O}(\log^4(q))$ binary operations. It is known that enough Elkies primes can be found *on average*, when considering either all elliptic curves over $\mathbb{F}_q$ [SS14], or (under the generalized Riemann hypothesis) the reductions of a given elliptic curve E defined over $\mathbb{Q}$ modulo primes [SS15].

The Schoof–Elkies–Atkin (SEA) point counting algorithm [Sch95] adds a number of practical improvements to Elkies's method, and has the same asymptotic complexity. In particular, *Atkin's method* consists in studying the factorization patterns of $\Phi_\ell(j(E), Y)$ to obtain partial information about $t_E \bmod \ell$. This information is used in an exponential-time sieve which reduces the number of Elkies primes to consider. Other improvements consist in using chains of ℓ-isogenies [FM02], or even Schoof's method, to compute t_E modulo powers of ℓ when ℓ is small.

Let us now describe Elkies's isogeny algorithm. Fixing Weierstrass models specifies differential forms ω, ω' on E, E' respectively, defined by the formula $dx/2y$ [Sil09, §III.1]. By duality, the Weierstrass models define bases of the tangent spaces $T_0(E), T_0(E')$ of E and E' at the neutral point. The tangent map

$$d\varphi\colon T_0(E) \to T_0(E'),$$

when written in these bases, is a scalar $c \in \mathbb{F}_q^\times$ called the *normalization factor* of φ. We have $\varphi^*\omega' = c\omega$.

If c is known, then one can write differential equations satisfied by the rational functions defining φ [BMSS08]. These equations can be solved in power series using Newton iterations: computing n terms of the solutions costs $\widetilde{O}(n)$ operations in $\mathbb{F}_q$. Elkies's algorithm concludes with a rational reconstruction step using the extended Euclidean algorithm [vG13, §12.3]. The degrees of the rational fractions describing φ are in $O(\ell)$, hence the cost of Elkies's algorithm *once c is determined* is $\widetilde{O}(\ell)$ operations in $\mathbb{F}_q$. The assumption that $p > 4\ell - 1$ appears since one has to perform divisions by $2, 3, \ldots, 4\ell - 1$ during the Newton iterations.

Computing the normalization factor c involves explicit formulæ about modular functions, in particular the derivative

$$j' := \frac{1}{2i\pi}\frac{dj}{d\tau}$$

of the j-invariant, which is a modular function of weight 2. Assume, for the moment, that E, E' are defined over $\mathbb{C}$. Then, given a modular function f of weight k, the quantity $f(E, \omega)$ is well-defined: choose $\tau \in \mathbb{H}_1$ and an isomorphism $\eta\colon E \xrightarrow{\sim} E(\tau)$, let $\omega(\tau)$ be the differential form given by $2\pi i\, dz$ on $E(\tau) = \mathbb{C}/\Lambda(\tau)$, let $r \in \mathbb{C}^\times$ be such that $\eta^*\omega(\tau) = r\omega$, and set

$$f(E, \omega) = r^k f(\tau),$$

which is indeed independent of τ and η. The classical theory of complex elliptic curves [Sil09, §VI.1] gives the equalities

$$E_4(E, \omega) = -48a, \qquad E_6(E, \omega) = 864b,$$

where E_4 and E_6 denote the normalized Eisenstein series of weights 4 and 6 respectively. Moreover

$$j'(E, \omega) = -\frac{E_6(E, \omega)}{E_4(E, \omega)} j(E) = \frac{2^9 3^5 a^2 b}{4a^3 + 27b^2}. \tag{1.7}$$

Since E, E' are ℓ-isogenous over $\mathbb{C}$, there exists a $\tau \in \mathbb{H}_1$ and a commutative diagram

$$\begin{array}{ccc} E & \xrightarrow{\varphi} & E' \\ \downarrow{\scriptstyle \eta} & & \downarrow{\scriptstyle \eta'} \\ E(\tau) & \xrightarrow{z \mapsto z} & E(\tau/\ell). \end{array}$$

where η, η' are isomorphisms. Let $r, r' \in \mathbb{C}^\times$ such that

$$\eta^*\omega(\tau) = r\omega, \quad \eta'^*\omega(\tau/\ell) = r'\omega(\tau).$$

Then the normalization factor of φ is $c = r'r^{-1}$. The crucial equality comes from differentiating the relation $\Phi_\ell\big(j(\tau), j(\tau/\ell)\big) = 0$ with respect to τ:

$$\partial_X \Phi_\ell\big(j(E), j(E')\big)\, j'(\tau) + \frac{1}{\ell} \partial_Y \Phi_\ell\big(j(E), j(E')\big)\, j'(\tau/\ell) = 0,$$

where ∂_X and ∂_Y denote derivation with respect to X and Y in $\mathbb{Z}[X, Y]$, respectively. Hence

$$c^2 \partial_X \Phi_\ell\big(j(E), j(E')\big)\, j'(E, \omega) + \frac{1}{\ell} \partial_Y \Phi_\ell\big(j(E), j(E')\big)\, j'(E', \omega') = 0. \quad (1.8)$$

Equation (1.8) is entirely algebraic, using eq. (1.7) to evaluate the modular function j'. It is also valid when E, E' are defined over $\mathbb{F}_q$; a simple way to see this is to lift the isogeny φ to characteristic zero. Under our assumptions, the derivative of Φ_ℓ with respect to Y does not vanish; moreover $j'(E', \omega') \neq 0$ because E' has no extra automorphisms. Therefore eq. (1.8) allows us to compute the normalization factor of φ up to sign, which is enough for the rest of Elkies's algorithm.

Recall that the modular polynomial Φ_ℓ has degree $O(\ell)$ in both variables X and Y, and height $O(\ell \log \ell)$. Recall also that $\ell \in O(\log q)$. Therefore, assuming that Φ_ℓ has been precomputed, evaluating the derivatives of Φ_ℓ at $(j(E), j(E'))$ costs $\widetilde{O}(\ell^2)$ operations in $\mathbb{F}_q$; it is the most costly step in Elkies's isogeny algorithm, but is still dominated by root finding and Frobenius computations in the context of the SEA algorithm.

1.2.3 Other applications of modular polynomials

Besides Elkies's algorithm, elliptic modular polynomials have other algorithmic applications. To present them, it is convenient to introduce the *ℓ-isogeny graph* $\mathcal{G}(\ell, \mathbb{F}_q)$, where $\ell \neq p$ is a prime. The vertices of $\mathcal{G}(\ell, \mathbb{F}_q)$ are elliptic curves defined over $\mathbb{F}_q$ up to $\overline{\mathbb{F}}_q$-isomorphism, and can be labeled by j-invariants in $\mathbb{F}_q$. Two vertices v_1, v_2 are linked by an edge if and only if there exists an ℓ-isogeny between them. The graph $\mathcal{G}(\ell, \mathbb{F}_q)$ is undirected and without multiple edges. More generally, if L is a list of primes not containing p, we define $\mathcal{G}(L, \mathbb{F}_q)$ as the superposition of the graphs $\mathcal{G}(\ell, \mathbb{F}_q)$ for $\ell \in L$. A subscript E denotes the connected components of these graphs containing the elliptic curve E.

Kohel's thesis [Koh96] shows that isogeny graphs of elliptic curves have a very particular shape. Let E be an elliptic curve over $\mathbb{F}_q$, and denote by $\mathcal{O} = \mathrm{End}_{\mathbb{F}_q}(E)$ the ring of endomorphisms of E defined over $\mathbb{F}_q$. There are two possibilities, depending on the trace of Frobenius t_E [Wat69, Thm. 4.1]:

1. If $|t_E| < 2\sqrt{q}$, then $\mathcal{O}$ is isomorphic to an order in an imaginary quadratic field F; this case includes all ordinary elliptic curves (for which t_E is prime to p), and also all supersingular elliptic curves over $\mathbb{F}_q$ when q is not a square.

2. If $|t_E| = 2\sqrt{q}$, then $\mathcal{O}$ is isomorphic to a maximal order in a certain quaternion algebra.

In case 1, let Δ be the discriminant of $F = \mathcal{O} \otimes \mathbb{Q}$, and let $\left(\frac{\Delta}{\ell}\right)$ be the Legendre symbol. Denote by $\mathbb{Z}_F$ the ring of integers in F. Each vertex of $\mathcal{G}(\ell, \mathbb{F}_q)_E$ can be given a *depth*, which is an integer between 0 and $d_{\max}$, the ℓ-adic valuation of $[\mathbb{Z}_F : \mathbb{Z}[\pi_E]]$. Each vertex of depth d is connected to

- 1 vertex of depth $d-1$ and ℓ vertices of depth $d+1$, if $0 < d < d_{\max}$;
- 1 vertex of depth $d-1$, if $0 < d = d_{\max}$;
- $1 + \left(\frac{\Delta}{\ell}\right)$ vertices of depth 0, if $0 = d = d_{\max}$;
- $1 + \left(\frac{\Delta}{\ell}\right)$ vertices of depth 0 and $\ell - \left(\frac{\Delta}{\ell}\right)$ of depth 1, if $0 = d < d_{\max}$.

We say that $\mathcal{G}(\ell, \mathbb{F}_q)_E$ is an *ℓ-volcano*; the vertices of depth zero form the *crater* or *surface* of the volcano, and vertices of depth $d_{\max}$ form the *floor*. A typical example is depicted on fig. 1.1.

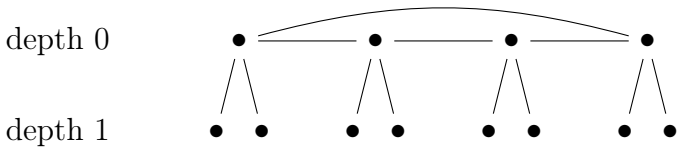

Fig. 1.1. An ℓ-volcano with $\ell = 3$, $d_{\max} = 1$, and $\left(\frac{\Delta}{\ell}\right) = 1$.

In case 2, the graphs $\mathcal{G}(\ell, \mathbb{F}_q)_E$ do not have a volcano structure. Instead, they have the Ramanujan property [Piz90]: they are "optimal" expander graphs, therefore random walks in these graphs mix rapidly. Under the generalized Riemann hypothesis, one can also obtain expander graphs out of isogeny volcanoes, by considering the surface of $\mathcal{G}(L, \mathbb{F}_q)_E$ when L is sufficiently large [JMV09].

The elliptic modular polynomial Φ_ℓ allows us to navigate ℓ-isogeny graphs by computing all vertices adjacent to a given one. Various number-theoretic algorithms rely on this fact. From now on, when k is a field and $\mathcal{O}$ is an imaginary quadratic order, we denote by $\mathrm{Ell}(\mathcal{O}, k)$ the set of $\bar{k}$-isomorphism classes of elliptic curves over k whose ring of endomorphisms over k is isomorphic to $\mathcal{O}$.

1. *Computing depths* [Koh96]. Given the rigid structure of ℓ-volcanoes, one can compute the depth of a given vertex v in $\mathcal{G}(\ell, \mathbb{F}_q)$ by following three paths from v and stopping when one of them reaches the floor. If E is an elliptic curve over $\mathbb{F}_q$ representing v, this depth equals the ℓ-adic valuation of $\left[\mathbb{Z}_F : \mathrm{End}_{\mathbb{F}_q}(E)\right]$.

2. *Finding isogeny chains* [FM02]. In the SEA algorithm, if ℓ is a very small Elkies prime (say $\ell = 2$ or 3), it is interesting to use a chain of ℓ-isogenies to compute the trace of Frobenius modulo ℓ^r, where r is the length of the chain. Walking along the surface of the corresponding isogeny volcano allows us to compute chains of arbitrary lengths.

3. *Explicit CM theory.* Let $\mathcal{O}$ be an imaginary quadratic order, and assume that $\mathrm{Ell}(\mathcal{O}, \mathbb{F}_q)$ is nonempty. The theory of *complex multiplication* (CM) defines a simply transitive action of the class group $\mathrm{Cl}(\mathcal{O})$ on $\mathrm{Ell}(\mathcal{O}, \mathbb{F}_q)$ [Wat69, Thm. 4.5]. Let $\ell \neq p$ be a prime that splits in two distinct invertible $\mathcal{O}$-ideals, $\ell = \mathfrak{l}_1\mathfrak{l}_2$. Then, for every elliptic curve $E \in \mathrm{Ell}(\mathcal{O}, \mathbb{F}_q)$, the elliptic curves $[\mathfrak{l}_1] \cdot E$ and $[\mathfrak{l}_2] \cdot E$ are the images of two ℓ-isogenies with domain E; all three elliptic curves E, $[\mathfrak{l}_1] \cdot E$ and $[\mathfrak{l}_2] \cdot E$ lie on the surface of $\mathcal{G}(\ell, \mathbb{F}_q)_E$. Moreover, the actions of $[\mathfrak{l}_1]$ and $[\mathfrak{l}_2]$ can be distinguished by computing the Frobenius eigenvalue on the kernel of the corresponding ℓ-isogeny. Therefore the action of $\mathrm{Cl}(\mathcal{O})$ on $\mathrm{Ell}(\mathcal{O}, \mathbb{F}_q)$ can be computed explicitly, by walking along the surface of ℓ-isogeny volcanoes (using depth computations) and using Elkies's algorithm described in §1.2.2.

4. *Computing Hilbert class polynomials* [Sut11]. Let $\mathcal{O}$ be an imaginary quadratic order. The *Hilbert class polynomial* of $\mathcal{O}$, denoted by $H_{\mathcal{O}}$, is a monic polynomial with coefficients in $\mathbb{Z}$ whose complex roots are the j-invariants of all the elements of $\mathrm{Ell}(\mathcal{O}, \mathbb{C})$. It is a defining polynomial for the ring class field of $\mathcal{O}$ over $F = \mathcal{O} \otimes \mathbb{Q}$. A possible strategy to compute $H_{\mathcal{O}}$ is to compute it modulo several primes p and using the Chinese remainder theorem. The roots of $H_{\mathcal{O}}$ mod p are given by the j-invariants of elements in $\mathrm{Ell}(\mathcal{O}, \overline{\mathbb{F}}_p)$, and can be enumerated using the explicit CM action.

5. *Computing elliptic modular polynomials* [BLS12, Sut13]. Isogeny volcanoes can also be used to compute modular polynomials via interpolation and Chinese remainders, as mentioned in §1.1.2. For instance, knowing all the vertices of a graph $\mathcal{G}(3, \mathbb{F}_q)_E$ of the shape given in fig. 1.1 would allow us to reconstruct $\Phi_3(X, Y)$ mod p, by interpolating $\Phi_3(x, Y)$ for each j-invariant x on the surface, then $\Phi_3(X, Y)$. In this algorithm, the vertices of $\mathcal{G}(\ell, \mathbb{F}_q)_E$ are enumerated using the

explicit CM action of ideals of small norms.

6. *Computing endomorphism rings of elliptic curves.* Let E be an ordinary elliptic curve over $\mathbb{F}_q$. Computing $\#E(\mathbb{F}_q)$ determines the endomorphism algebra of E; afterwards, walking in isogeny graphs gives additional information about the precise endomorphism *ring* of E over $\mathbb{F}_q$. For instance, if ℓ is a small prime, computing the depth of E in $\mathcal{G}(\ell, \mathbb{F}_q)_E$ is an efficient way of determining the ℓ-adic valuation of the conductor of $\mathrm{End}_{\mathbb{F}_q}(E)$. For larger primes, one can use the explicit CM theory to rule out candidates for $\mathrm{End}_{\mathbb{F}_q}(E)$. Kohel [Koh96] uses these observations to design an exponential-time algorithm (under the generalized Riemann hypothesis) for computing $\mathrm{End}_{\mathbb{F}_q}(E)$. A probabilistic subexponential time algorithm (under heuristics), using so-called *smooth relations* in ideal class groups, also exists [BS11].

1.2.4 Relevance to cryptography

The number-theoretic algorithms involving elliptic modular polynomials have applications in two areas of public-key cryptography. The first area is *elliptic curve cryptography*, which relies on the hardness of the discrete logarithm problem in the group $E(\mathbb{F}_q)$, where E is an elliptic curve over $\mathbb{F}_q$. The second area is *isogeny-based cryptography*, which relies on the hardness of finding paths in isogeny graphs. Elliptic curve cryptography is widespread in real life; isogeny-based cryptography is younger and less used in industry, but attracts interest because of its conjectured resistance to quantum attacks.

Elliptic curve cryptography. Let G be a finite group of order n which is effective, in the sense that its elements can be encoded in a computer and the basic operations (multiplication, equality tests) can be performed algorithmically. Let $g \in G$ and $0 \leq a \leq n-1$. Then computing g^a is easy: using a square-and-multiply algorithm, it can be done in $O(\log n)$ operations in G. Conversely, the *discrete logarithm problem* asks, given $g \in G$ and $h \in \langle g \rangle$, to compute a such that $g^a = h$. For a generic group G, the best known classical algorithms to do so have complexity $O(\sqrt{n})$, which is exponential in $\log n$ [Pol75]. This exponential complexity gap makes the discrete logarithm problem suitable for use in public-key cryptography.

Another computational problem closely related to discrete logarithms is the *Diffie–Hellman problem*: given $g, g^a, g^b \in G$, compute g^{ab}. For a generic G, the best known classical algorithms to solve the Diffie–Hellman problem also are in $O(\sqrt{n})$; but the problem becomes easy if one can com-

pute discrete logarithms in G. The well-known Diffie–Hellman key exchange protocol [DH76], the first published example of a public-key cryptographic protocol, relies on the hardness of the Diffie–Hellman problem.

In concrete implementations, one has to choose a group G. In general, one should choose $n = \#G$ to be either prime or "almost" prime, in order to avoid generic attacks based on Chinese remainders [PH78]. The idea of elliptic curve cryptography, introduced by Koblitz [Kob87] and Miller [Mil85], is to take $G = E(\mathbb{F}_q)$ where E is an elliptic curve over $\mathbb{F}_q$. For a generic E, the best known classical algorithms to compute discrete logarithms in $E(\mathbb{F}_q)$ are in $O(\sqrt{q})$. This compares favorably with the case of multiplicative groups of finite fields, where subexponential attacks are known [Pom87]. We refer to [GM17] for a review of the discrete logarithm problem, both in the generic and finite field cases.

Several algorithms are available to generate suitable elliptic curves, say of prime order. One can run the SEA algorithm (recall §1.2.2) on a family of candidates, for instance chosen at random, until finding one of prime order. Since the SEA algorithm computes $\#E(\mathbb{F}_q)$ modulo small primes, one can abort the algorithm early if a small factor of $\#E(\mathbb{F}_q)$ is detected. Another strategy is to use the *CM method*, first described in [AM93], which constructs elliptic curves such that $\#E(\mathbb{F}_q)$ is completely determined, at the cost of some loss of control on q: if $\mathcal{O}$ is an imaginary quadratic order whose Hilbert class polynomial $H_{\mathcal{O}}$ is known, and E is an elliptic curve over $\mathbb{F}_q$ such that $j(E)$ is a root of $H_{\mathcal{O}}$, then the trace of Frobenius of E is determined up to sign.

On the destructive side, isogenies can be used to attack the discrete logarithm problem on a given elliptic curve E: one tries to find an isogeny chain from E to another elliptic curve where the discrete logarithm problem is easier to solve [GHS02].

The discrete logarithm problem, even in generic groups, can be solved in polynomial time on a quantum computer by Shor's algorithm [Sho94]. Therefore elliptic curve cryptography may become insecure as quantum computers develop.

Isogeny-based cryptography. Let $\mathcal{G}$ be a finite connected graph, let v_0 be a vertex of $\mathcal{G}$, and assume that there exists an algorithm to compute the neighbors of a given vertex of $\mathcal{G}$. If $\mathcal{G}$ has good expansion properties, then one can efficiently sample vertices in $\mathcal{G}$ with a close to uniform distribution by following random walks from v_0. Conversely, given a random vertex v of $\mathcal{G}$, finding a path from v_0 to v is a hard problem in general. Informally, isogeny-based cryptography relies on the hardness of this problem when $\mathcal{G}$ is an isogeny graph over a finite field: $\mathcal{G}$ can be either a supersingular

graph of ℓ-isogenies for some fixed ℓ, or a superposition of isogeny cycles of different prime degrees. Isogeny-based primitives are one of the families of cryptographic proposals (along with, for instance, lattice-based or code-based primitives) that are conjectured to be quantum-resistant.

Examples of isogeny-based protocols are given by the Charles–Goren–Lauter hash function [CLG09], which uses a supersingular graph of 2-isogenies; the SIDH key exchange protocol [dJP11], which uses supersingular graphs of 2- and 3-isogenies; and the CSIDH key exchange protocol [dKS18, CLM+18], which uses a superposition of isogeny cycles for supersingular curves over a prime field $\mathbb{F}_p$. Isogeny-based cryptography is a very active research area, with recent cryptographic proposals such as verifiable delay functions [dMPS19] and digital signatures [dG19].

For performance reasons, these protocols rely on computing isogenies from their kernels using Vélu's formulæ to navigate isogeny graphs, instead of finding roots of modular polynomials.

1.3 State of the art in higher dimensions

After elliptic curves, it is natural to study *abelian varieties*, which are by definition complete group varieties over a field; elliptic curves correspond to abelian varieties of dimension one. Abelian varieties are fundamental objects in algebraic geometry and number theory. They appear for instance in the study of algebraic curves: much information about a smooth algebraic curve of genus g can be obtained from its Jacobian, which is an abelian variety of dimension g. Abelian varieties have a very rich theory, that we do not develop here: we refer to [Mum70, Mil86a, Mil86b], and [BL04] for the complex theory.

1.3.1 Modular equations of Siegel and Hibert type

Just as elliptic modular polynomials encode ℓ-isogenies between elliptic curves for a fixed prime ℓ, higher-dimensional modular equations encode isogenies of a fixed type between abelian varieties. To define these modular equations, one works with moduli spaces of abelian varieties, for instance over $\mathbb{C}$, as we did for elliptic curves in §1.1. In order to obtain well-behaved moduli spaces, one considers isomorphism classes of abelian varieties endowed with an additional structure called a *polarization*. A polarization on an abelian variety A is a certain equivalence class of ample line bundles on A; it can also be seen as an isogeny from A to its dual abelian variety $\widehat{A}$, and the polarization is called *principal* if this isogeny is an isomorphism.

In the dimension 2 case, i.e. for abelian surfaces, two moduli spaces are of particular interest.

1. The *Siegel moduli space* $\mathcal{A}_2$ classifies principally polarized abelian surfaces. As an algebraic variety, $\mathcal{A}_2$ can be defined over $\mathbb{Q}$, and is birational to $\mathbb{P}^3$, by results of Igusa [Igu62]: in other words, isomorphism classes of principally polarized abelian surfaces are generically classified by three invariants. For each prime ℓ, the *Siegel modular equations* of level ℓ encode the presence of ℓ-isogenies between principally polarized abelian surfaces [BL09, Mil15]; the degree of these isogenies is ℓ^2.

2. Let F be a fixed quadratic field. The *Hilbert moduli space* $\mathcal{A}_{2,F}$ classifies principally polarized abelian surfaces with real multiplication by $\mathbb{Z}_F$, the ring of integers of F. As an algebraic variety, $\mathcal{A}_{2,F}$ has dimension 2. For each totally positive prime $\beta \in \mathbb{Z}_F$, the *Hilbert modular equations* of level β encode the presence of β-isogenies [Mar20, MR20]; the degree of these isogenies is $N_{F/\mathbb{Q}}(\beta)$.

In practice, one makes a choice of invariants on $\mathcal{A}_2$ or $\mathcal{A}_{2,F}$; then the modular equations can be seen as explicit multivariate rational fractions. For instance, a standard choice of invariants on the Siegel moduli space $\mathcal{A}_2$ is given by the three *Igusa invariants*, denoted by j_1, j_2, j_3. The modular equations of Siegel type and level ℓ are given by three polynomials $\Psi_{\ell,m} \in \mathbb{Q}(j_1, j_2, j_3)[Y]$ for $1 \leq m \leq 3$, with the following property: there exist an open subvariety $\mathcal{U}$ of $\mathcal{A}_2$ such that for every two points z, z' of $\mathcal{U}(\mathbb{C})$, the equalities

$$\begin{cases} \Psi_{\ell,1}\big(j_1(z), j_2(z), j_3(z), j_1(z')\big) = 0 \\ \partial_X \Psi_{\ell,1}\big(j_1(z), j_2(z), j_3(z), j_1(z')\big) \cdot j_2(z') = \Psi_{\ell,2}\big(j_1(z), j_2(z), j_3(z), j_1(z')\big) \\ \partial_X \Psi_{\ell,1}\big(j_1(z), j_2(z), j_3(z), j_1(z')\big) \cdot j_3(z') = \Psi_{\ell,3}\big(j_1(z), j_2(z), j_3(z), j_1(z')\big) \end{cases}$$

hold if and only if the principally polarized abelian surfaces associated with z and z' are ℓ-isogenous.

Several examples of modular equations for abelian surfaces have already been computed, but they tend to have frightening sizes due to their large number of variables, degrees, and heights. Writing down the Siegel modular equations uses approximately 1 MB for $\ell = 2$, and approximately 410 MB for $\ell = 3$; the computations for $\ell = 5$ have not been carried out to our knowledge. As in the case of elliptic modular polynomials, one can consider variants of modular equations using invariants of abelian surfaces with level structure. For instance, Siegel modular equations in terms of

theta constants in dimension 2 (instead of Igusa invariants) have smaller coefficients, and have been computed up to $\ell = 7$ [Mil].

The most efficient algorithms to compute Siegel or Hilbert modular equations currently use complex approximations [Mil15, MR20]. As in the elliptic case, an essential subroutine in these algorithms is to compute theta constants in quasi-linear time in the precision [Dup06, LT16].

1.3.2 Applications

In terms of applications, the literature on modular polynomials for abelian surfaces is much less developed than for their elliptic counterparts.

The point counting problem. Prior to this thesis, no generalization of Elkies's isogeny algorithm was known for abelian surfaces. In the literature, the asymptotically fastest point counting algorithms for abelian surfaces over finite fields are based on Schoof's method, which admits generalizations to abelian varieties of any dimension [Pil90]. Schoof's method uses the action of Frobenius on kernels of *endomorphisms* of abelian varieties: multiplication by ℓ in the generic case [GH00, GS12], and real multiplication if available [GKS11].

Atkin's point counting method, which studies the factorization patterns of modular equations, has also been generalized to dimension 2 [BGLG$^+$16], but this result does not affect the asymptotic complexity of point counting algorithms. Moreover, the complexity of this method has not been worked out, as no bounds from above for the degree and heights of modular equations for abelian surfaces appear in the literature.

Isogeny graphs. The structure of isogeny graphs of ℓ-isogenies between principally polarized abelian surfaces over $\mathbb{F}_q$ can be quite intricate, but isogeny graphs of β-isogenies between ordinary principally polarized abelian surfaces with real multiplication are still volcanoes [BJW17, Mar18]. However, the chosen method to navigate isogeny graphs in dimension 2 is often to compute isogenies from their kernels using higher-dimensional analogues of Vélu's formulæ instead of using modular equations; these formulæ are available both for ℓ- and β-isogenies [LR12, CR15, CE15, DJRV17].

Given the current state of knowledge, there are several reasons for preferring to compute isogenies from their kernels: the complexity of manipulating modular equations is unclear, finding directions in isogeny graphs is hard due to the lack of explicit kernels, and the databases of precomputed modular equations are small. The drawback is that the kernel has to be

computed in the first place, and computing the full ℓ-torsion subgroup of an abelian surface is costly.

1.4 Results and plan of the thesis

The contributions of this thesis aim to better understand modular equations in higher dimensions and algorithms using them. A particular focus is placed on modular equations of Hilbert and Siegel type for principally polarized abelian surfaces. In short, this thesis provides the three necessary ingredients to generalize Elkies's point counting method to the case of abelian surfaces:

1. An *isogeny algorithm* which, given isogenous abelian surfaces, computes an isogeny between them as an explicit rational map;
2. General *complexity bounds* for modular equations, i.e. height and degree bounds, in terms of their level;
3. An *evaluation algorithm* for modular equations and their derivatives.

Combining these results also shows that one can efficiently navigate isogeny graphs for abelian surfaces using modular equations. Let us now detail the contents of each chapter.

Chapter 2. We give an introduction to *PEL Shimura varieties*, which are moduli spaces for complex abelian varieties with specified polarization, endomorphism, and level structures. This formalism encompasses the different moduli spaces mentioned in chapter 1, and we detail its connection with the classical language of Siegel half spaces and modular forms. Adopting the language of Shimura varieties allows us to give a unified definition of modular equations, of which elliptic modular polynomials, and Siegel or Hilbert modular equations for abelian surfaces are special cases.

Chapter 3. This chapter presents an analogue of Elkies's isogeny algorithm in dimension 2 using modular equations of either Hilbert or Siegel type, and corresponds to the preprint [KPR19]. Let us present the Siegel case here: we want to compute an ℓ-isogeny φ between principally polarized abelian surfaces A and A'. The *Kodaira–Spencer map* relates the possible deformations of φ, which can be computed from derivatives of Siegel modular equations, with the tangent map $d\varphi\colon T_0(A) \to T_0(A')$ induced by φ on the tangent spaces of A and A' at zero. We make the Kodaira–Spencer map explicit in the case of Jacobians of genus 2 curves: concretely, an explicit

formula yields the value of certain vector-valued Siegel modular functions, namely derivatives of Igusa invariants, in terms of the coefficients of a hyperelliptic curve equation. This explicit formula is the dimension 2 analogue of eq. (1.7) for the derivative of the j-invariant in the case of elliptic curves. The result is the following.

Theorem 1.4.1. *Let ℓ be a prime, and let k be a field of characteristic zero or $p > 8\ell + 7$. Let $\mathcal{U}$ be the open subvariety of $\mathcal{A}_2$ consisting of abelian surfaces A without extra automorphisms and such that $j_3(A) \neq 0$. Assume that there exists an algorithm to evaluate derivatives of Siegel modular equations of level ℓ at a given point of $\mathcal{U} \times \mathcal{U}$ over k, using $C_{\text{eval}}(\ell)$ operations in k. Let $\varphi\colon A \to A'$ be an ℓ-isogeny defined over k such that A, A' lie in $\mathcal{U}$ and such that the subvariety of $\mathcal{A}_2 \times \mathcal{A}_2$ cut out by the Siegel modular equations of level ℓ is normal at (A, A'). Then, given the Igusa invariants of A and A', one can compute an explicit representation of the isogeny φ using $O(C_{\text{eval}}(\ell)) + \widetilde{O}(\ell)$ operations in k.*

Chapter 4. This chapters contains prerequisites to the results of chapter 5, and corresponds to the preprint [Kie20c]. The first part presents the notion of *heights* of polynomials and rational fractions over a number field. This nonnegative quantity reflects the size of the coefficients. For instance, if $F \in \mathbb{Q}(X)$, one can write $F = P/Q$ where $P, Q \in \mathbb{Z}[X]$ are coprime; then the height $h(F)$ of F is the maximum of $\log|c|$, where c runs through the nonzero coefficients of P and Q.

The second part presents results relating the height of a polynomial or rational fraction with the heights of its evaluations at certain points; this allows us to bound $h(F)$ from above when F is a rational fraction obtained from an interpolation process, and will be applied to modular equations in chapter 5. We take evaluation points to be (almost) consecutive integers x_i, we are given $H \geq 0$ such that $h(F(x_i)) \leq H$ for every i, and we are particularly looking for results where $h(F)$ is in $O(H)$.

Theorem 1.4.2. *Let L be a number field of degree d_L over $\mathbb{Q}$, let $[\![A, B]\!]$ be an interval in $\mathbb{Z}$, and write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L(X)$ be a rational fraction of degree at most $d \geq 1$. Let S be a subset of $[\![A, B]\!]$ containing no pole of F, let $\eta \geq 1$, and let $H \geq \max\{4, \log(2M)\}$. Moreover, assume that*

1. *$h(F(x)) \leq H$ for every $x \in S$.*
2. *S contains at least D/η elements.*
3. *$D \geq \max\{\eta d^3 H, 4\eta d d_L\}$.*

Then

$$h(F) \leq H + C_L \eta d \log(\eta d H) + d \log(2M) + \log(d+1),$$

where C_L is a constant depending only on L. One can take $C_{\mathbb{Q}} = 960$.

Chapter 5. This chapter gives degree and height bounds for modular equations in the general setting of PEL Shimura varieties in terms of their level. The associated preprint is [Kie20a]. Let us use the notation of chapter 2: we consider the PEL Shimura variety associated with a certain reductive group G and a choice of level, connected components $\mathcal{S}, \mathcal{T}$ of this Shimura variety, and a choice of invariants on these connected components. The algebraic varieties $\mathcal{S}$ and $\mathcal{T}$ are moduli spaces for complex abelian varieties with PEL structure.

Let δ be an adelic element of G. The modular equations of level δ are multivariate rational fractions defining the Hecke correspondence of level δ in the product $\mathcal{S} \times \mathcal{T}$. In the modular interpretation, this Hecke correspondence describes a certain number $d(\delta)$ of isogenies of degree denoted by $l(\delta)$ between abelian varieties with PEL structure. We show that the degrees and heights of modular equations of level δ are bounded above in terms of $d(\delta)$ and $l(\delta)$.

Theorem 1.4.3. *With the notation above, there exist constants C_1 and C_2 (independent of δ) with the following property. Let δ be an adelic element of G, and let F be a rational fraction obtained as a coefficient of one of the modular equations of level δ. Then the total degree of F is bounded above by $C_1\, d(\delta)$, and the height of F is bounded above by $C_2\, d(\delta) \log l(\delta)$.*

In the case of Siegel modular equations of level ℓ for principally polarized abelian surfaces, we have $d(\delta) = \ell^3 + \ell^2 + \ell + 1$ and $l(\delta) = \ell^2$. Thus the degree of Siegel modular equations of level ℓ is $O(\ell^3)$, and their height is $O(\ell^3 \log \ell)$. In the case of elliptic modular polynomials, theorem 1.4.3 recovers Cohen's bounds [Coh84] up to constant factors.

Chapter 6. This chapter corresponds to the preprint [Kie20b]. We present an evaluation algorithm for Siegel and Hilbert modular equations in dimension 2 and their derivatives at a given point defined over a number field. The algorithm also applies to finite fields, by considering lifts to characteristic zero. We use complex approximations, and rely on fast algorithms to compute theta constants in dimension 2.

Computing theta constants in dimension 2 in quasi-linear time in the demanded precision relies on two heuristics, namely the correctness of certain sign choices in duplication formulæ, and the convergence of certain

Newton iterations. An auxiliary result of this chapter removes the first heuristic: the sign choices usually made in the algorithm to compute theta constants are indeed correct. The complete proof appears in the accepted paper [Kie21]. The second heuristic is valid in practice.

Apart from the computation of theta constants, our evaluation algorithm for modular equations is made provably correct by a rigorous analysis of precision losses. In particular we obtain the following result.

Theorem 1.4.4 (Under heuristics relative to the computation of theta constants in dimension 2)**.** *Let ℓ be a prime, and let $(j_1, j_2, j_3) \in \mathbb{Q}^3$ be a tuple of height $O(1)$ where the denominators of Siegel modular equations of level ℓ do not vanish. Then one can evaluate these modular equations and their derivatives at (j_1, j_2, j_3) using $\widetilde{O}(\ell^6)$ binary operations.*

One should compare this cost to the cost of storing Siegel modular equations in full, which we estimate to be $\Theta(\ell^{15} \log \ell)$ given the results of chapter 5: when the dimension of the associated moduli space is greater than one, evaluating modular equations on the fly seems superior to precomputing them.

In the case of Hilbert modular equations encoding β-isogenies where β has norm ℓ over $\mathbb{Q}$, the cost of the evaluation algorithm can decrease to $\widetilde{O}(\ell^2)$ binary operations only. If this happens in the context of Elkies's method, then manipulating modular equations becomes asymptotically negligible when compared to the rest of the point counting algorithm.

Chapter 7. We conclude this thesis by applying the results of all preceding chapters to the point counting problem for abelian surfaces. In particular, we obtain a heuristic point counting algorithm for principally polarized abelian surfaces over finite fields with real multiplication by a fixed real quadratic field F which, under heuristics and up to constant factors depending on F, has the same asymptotic complexity as the SEA algorithm for elliptic curves.

Chapter 2

Moduli spaces of abelian varieties and modular equations

Modular equations, such as the elliptic modular polynomials presented in §1.1, can be defined as equations relating certain modular functions defined on moduli spaces of abelian varieties. This chapter starts with reminders on abelian varieties (§2.1), then presents the language of *PEL Shimura varieties*, which are moduli spaces for abelian varieties with certain polarization, endomorphism, and level structures (§2.2). The moduli spaces that are used to define modular equations of Siegel and Hilbert type for abelian surfaces, namely the Siegel threefold and Hilbert surfaces, are special cases of PEL Shimura varieties (§2.3). Using this language, we define modular equations in the general setting of PEL Shimura varieties as equations defining Hecke correspondences (§2.4). The modular polynomials mentioned in chapter 1 are all special cases of this definition. This unified language will be useful in chapter 5, where we prove size bounds for all modular equations in terms of their level.

2.1 Reminders on abelian varieties

By definition, an *abelian variety* over a field k is a complete group variety over k. In order to study abelian varieties, we endow them with *polarizations*. Depending on the context, a polarization on A can be seen either as an isogeny from A to its dual abelian variety $\widehat{A}$, or as an equivalence class of ample line bundles on A. An important special case is that of *principal polarizations*, when the isogeny $A \to \widehat{A}$ is an isomorphism. Elliptic curves, as well as Jacobians of smooth curves in general, are canonically endowed with a principal polarization.

When $k = \mathbb{C}$, it is possible to describe abelian varieties as complex tori of the form $\mathbb{C}^g/\Lambda$, where $g \geq 1$ is the dimension and $\Lambda \subset \mathbb{C}^g$ is a lattice. In this case we have an additional way of describing polarizations, namely as a symplectic form taking integer values on Λ.

Endowing an abelian variety A with a polarization yields a "positivity structure" on its endomorphism algebra, given by the Rosati involution. For instance, it makes sense to talk about *real endomorphisms* of A, i.e. endomorphisms that are invariant under the Rosati involution. Polarizations also allow us to classify isogenies: the notion of *ℓ-isogenies* between polarized abelian varieties is defined in terms of pullbacks of polarizations. Similarly, if $\mathcal{O}$ is an order in a totally real field F of degree g over $\mathbb{Q}$, and $\beta \in \mathcal{O}$ is totally positive, then we can define the notion of *β-isogenies* between abelian varieties of dimension g with real multiplication by $\mathcal{O}$.

References for this section are [Mum70, Mil86a, Mil86b], and also [BL04] for abelian varieties over $\mathbb{C}$.

2.1.1 Abelian varieties and polarizations

Let $g \geq 1$ be an integer, let k be a field, and let $\bar{k}$ be an algebraic closure of k. Let A, B be abelian varieties of the same dimension g over k.

Recall that a *group variety* over k is, by definition, an algebraic variety over k endowed with a k-rational neutral point and suitable morphisms defined over k encoding the group operations. Abelian varieties are smooth varieties [Mil86a, §1], and also abelian groups [Mil86a, Cor. 2.4]; we will use the additive notation. If $x \in A(\bar{k})$, we denote by t_x the translation by x:

$$t_x \colon A \to A, \quad y \mapsto y + x.$$

If $x \in A(k)$, then this morphism is defined over k.

An *isogeny* $A \to B$ is a group morphism that has finite kernel (or equivalently, is surjective [Mil86a, Prop. 8.1]). For every integer $n \neq 0$, the multiplication-by-n map on A, denoted by $[n]_A$, is an isogeny of degree n^{2g} [Mil86a, Thm. 8.2]. If n is prime to the characteristic of k, then the *n-torsion subgroup* $A[n] := \ker([n]_A)$ of A is an étale group scheme, and the group of its $\bar{k}$-points is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^{2g}$.

Since A is smooth, there is a correspondence between line bundles and divisors on A. One can define three different notions of equivalence for line bundles on A defined over $\bar{k}$.

1. Two line bundles $\mathcal{L}_1$ and $\mathcal{L}_2$ on A, corresponding to divisors D_1 and D_2, are *linearly equivalent* if $D_1 - D_2$ is the divisor of a function;

equivalently, if $\mathcal{L}_1 \otimes \mathcal{L}_2^{-1}$ is globally trivial. We often identify line bundles with their linear equivalence classes.

2. The line bundles $\mathcal{L}_1$ and $\mathcal{L}_2$ are *algebraically equivalent* if there exists a smooth variety T, a line bundle $\mathcal{M}$ on $A \times T$, and closed points t_1, t_2 of T such that $\mathcal{L}_1 \simeq \mathcal{M}_{|A\times\{t_1\}}$ and $\mathcal{L}_2 \simeq \mathcal{M}_{|A\times\{t_2\}}$. We say that $\mathcal{L}_1$ is *algebraically equivalent to zero* if it is algebraically equivalent to the trivial line bundle; the group of line bundles algebraically equivalent to zero on A is denoted by $\mathrm{Pic}^0(A)$. Linearly equivalent line bundles are algebraically equivalent. Line bundles up to algebraic equivalence on A form a group under the tensor product, called the *Néron–Severi group* $\mathrm{NS}(A)$ of A.

3. Two algebraically equivalent line bundles on A are always *numerically equivalent*, meaning that one can be replaced by the other in intersection pairings.

Recall that a line bundle $\mathcal{L}$ on A is *very ample* if the global sections of $\mathcal{L}$ define a projective embedding of A, and that $\mathcal{L}$ is *ample* if some positive power of $\mathcal{L}$ is very ample. Lefschetz's theorem [Mum70, §III.17] states that for every ample line bundle $\mathcal{L}$ on A, its third power $\mathcal{L}^{\otimes 3}$ is very ample.

The *dual* of A, denoted by $\widehat{A}$, is also an abelian variety of dimension g over k [Mil86a, §9]. There is a canonical bijection between $\widehat{A}(\bar{k})$ and $\mathrm{Pic}^0(A)$. For every line bundle $\mathcal{L}$ on A, the map

$$\phi_{\mathcal{L}} \colon A \to \widehat{A}, \quad x \mapsto (t_x^*\mathcal{L}) \otimes \mathcal{L}^{-1}$$

is a group morphism by the theorem of the square [Mil86a, §6]. Two line bundles on A are algebraically equivalent if and only if they give rise to the same map $\phi_{\mathcal{L}}$. The map $\phi_{\mathcal{L}}$ is zero if and only if $\mathcal{L}$ is algebraically equivalent to zero, and $\phi_{\mathcal{L}}$ is an isogeny if and only $\mathcal{L}$ is ample [Mil86a, Prop. 10.1].

For every isogeny $\varphi \colon A \to B$, we have a *dual isogeny* $\widehat{\varphi} \colon \widehat{B} \to \widehat{A}$; and the double dual of A is identified with A. Under this identification, isogenies of the form $\phi_{\mathcal{L}}$ are self-dual.

A *polarization* on A is an isogeny $\varphi \colon A \to \widehat{A}$ such that there exists an ample line bundle $\mathcal{L}$ over A defined over $\bar{k}$ such that $\varphi = \phi_{\mathcal{L}}$. A polarization is called *principal* if it is an isomorphism. In other words, a polarization on A is an algebraic equivalence class of ample line bundles on A, and we often make a choice of such a line bundle. Yet another way of looking at polarizations is to view φ as a symmetric correspondence on $A \times A$.

2.1.2 Jacobians

If $\mathcal{C}$ is a smooth curve of genus g over k, then its *Jacobian*, denoted by $\mathrm{Jac}(\mathcal{C})$, is an abelian variety of dimension g defined over k [Mil86b]. The $\bar{k}$-points of $\mathrm{Jac}(\mathcal{C})$ are identified with linear equivalence classes of degree zero divisors on $\mathcal{C}$. Let K be a divisor of degree g on $\mathcal{C}$ (for instance the canonical divisor, when $g = 2$). Then a generic divisor of degree zero on $\mathcal{C}$ can be written as

$$[P_1 + \cdots + P_g - K]$$

where $P_1, \ldots, P_g \in \mathcal{C}(\bar{k})$ are uniquely determined up to permutation. This defines a birational map between the g-th symmetric power $\mathcal{C}^{<g>}$ of $\mathcal{C}$, which is also a smooth variety of dimension g, and the Jacobian variety $\mathrm{Jac}(\mathcal{C})$ [Mil86b, Thm. 5.1].

If $P \in \mathcal{C}(\bar{k})$, then we have an embedding

$$\eta_P \colon \mathcal{C} \hookrightarrow \mathrm{Jac}(\mathcal{C}), \quad Q \mapsto [Q - P]. \tag{2.1}$$

The morphism η_P induces an isomorphism

$$\eta_P^* \colon \Omega^1(\mathrm{Jac}(\mathcal{C})) \to \Omega^1(\mathcal{C})$$

between the spaces of global differential forms on $\mathcal{C}$ and $\mathrm{Jac}(\mathcal{C})$. The isomorphism η_P^* does not depend on the choice of P, hence η_P^* is defined over k [Mil86b, Prop. 2.2]. The image $\eta_P(\mathcal{C})$ is a divisor on $\mathrm{Jac}(\mathcal{C})$, called the *theta divisor*; its algebraic equivalence class does not depend on P. The associated line bundle is ample, and defines a canonical principal polarization on $\mathrm{Jac}(\mathcal{C})$ [Mil86b, Thm. 6.6].

Example 2.1.1. Recall that an elliptic curve over k is a smooth curve of genus 1 endowed with a base point. An elliptic curve E is canonically isomorphic to its Jacobian, and therefore E is a principally polarized abelian variety of dimension 1. If 0_E denotes the neutral point of E, then $\mathcal{L} := \mathcal{L}([0_E])$ (in the usual Riemann–Roch notation) is an ample line bundle such that $\phi_{\mathcal{L}}$ is the principal polarization of E. Since $\mathcal{L}^{\otimes 3}$ is very ample, its global sections induce a projective embedding of E; we obtain the existence of Weierstrass models of E in this way.

In the case $g = 2$, every principally polarized abelian surface A over $\bar{k}$ is either isomorphic to the product of two elliptic curves with their canonical principal polarizations, or isomorphic to the Jacobian of a smooth genus 2 *hyperelliptic curve* over $\bar{k}$ [OU73]. A hyperelliptic curve of genus g over $\bar{k}$ is the normalization of a plane curve with equation

$$y^2 = P(x)$$

where P is a polynomial of degree $2g+1$ or $2g+2$ with simple roots. If $g = 3$, then the threefold A is either the Jacobian of a plane quartic curve, the Jacobian of a genus 3 hyperelliptic curve, or a product of principally polarized abelian varieties of lower dimensions.

2.1.3 Abelian varieties over $\mathbb{C}$

Every abelian variety A of dimension g over $\mathbb{C}$ is isomorphic to a complex torus of the form $\mathbb{C}^g/\Lambda$, where $\Lambda \subset \mathbb{C}^g$ is a lattice, i.e. a discrete free abelian group of rank $2g$ [Mum70, §I.1]. A matrix whose columns generate such a lattice Λ is called a *period matrix* of A.

Conversely, every *algebraizable* complex torus is an abelian variety; this condition is not automatic, and holds if and only if the generating matrix of Λ satisfies certain symmetry and positivity conditions called the *Riemann relations* [BL04, §4.2]. One can deduce from these relations that every abelian variety of dimension g over $\mathbb{C}$ admits a period matrix of the form $(I_g\ \tau)$ where τ belongs to the *Siegel upper half space* $\mathbb{H}_g$, i.e. $\tau \in \mathrm{Mat}_{g\times g}(\mathbb{C})$ is symmetric and $\mathrm{Im}(\tau)$ is positive definite.

Let $A = \mathbb{C}^g/\Lambda$ be a complex abelian variety, and let $\mathcal{L}$ be a line bundle on A. Since the pullback of $\mathcal{L}$ to $\mathbb{C}^g$ is trivial, the line bundle $\mathcal{L}$ can be described in terms of *factors of automorphy* for the lattice Λ; we refer to [BL04, §2.1–2.2] for the precise definition. Starting from any factor of automorphy of $\mathcal{L}$, one can construct a real-valued alternating form $\psi_{\mathcal{L}}$ on $\mathbb{C}^g$ taking integer values on Λ. In fact, the alternating form $\psi_{\mathcal{L}}$ can be identified with the first Chern class of $\mathcal{L}$ [BL04, Thm. 2.1.2], so that two line bundles on A are algebraically equivalent if and only if their associated alternating forms are equal. The form $\psi_{\mathcal{L}}$ has the property that $\psi_{\mathcal{L}}(iu, iv) = \psi_{\mathcal{L}}(u, v)$ for every $u, v \in \mathbb{C}^g$, and the line bundle $\mathcal{L}$ is a polarization if and only if the form $(u, v) \mapsto \psi_{\mathcal{L}}(u, iv)$ is symmetric and positive definite [BL04, §4.1]. Conversely, if ψ is an alternating form on $\mathbb{C}^g$ such that $\psi(\Lambda, \Lambda) \subset \mathbb{Z}$ and $\psi(iu, iv) = \psi(u, v)$ for every $u, v \in \mathbb{C}^g$, then ψ is the first Chern class of a line bundle on A.

Let ψ be a polarization on A, seen as an alternating form. Then there exists a $\mathbb{Z}$-basis of Λ in which the matrix of ψ takes the form

$$\begin{pmatrix} & & & D_1 & \\ & (0) & & & \ddots & \\ & & & & & D_g \\ -D_1 & & & & & \\ & \ddots & & & (0) & \\ & & -D_g & & & \end{pmatrix}$$

for some positive integers $D_1 | \cdots | D_g$, uniquely determined by ψ. The tuple $(D_1, \ldots, D_g)$ is called the *type* of the polarization ψ. The type of ψ can

also be seen as the isomorphism class of the finite group $\Lambda^\perp/\Lambda$, where

$$\Lambda^\perp = \{u \in \mathbb{C}^g \colon \forall v \in \Lambda,\ \psi(u,v) \in \mathbb{Z}\}\,.$$

The notion of polarization type can be extended to abelian varieties over any field k, at least when the degree of the polarization is prime to the characteristic of k.

2.1.4 Endomorphisms and isogenies

Let A be an abelian variety of dimension g over a field k endowed with a polarization $\phi\colon A \to \widehat{A}$. We denote by $\mathrm{End}(A)$ (resp. $\mathrm{End}_k(A)$) the ring of endomorphisms of A defined over $\overline{k}$ (resp. k). Scalar multiplications define an embedding $\mathbb{Z} \hookrightarrow \mathrm{End}_k(A)$, and the ring $\mathrm{End}(A)$ is torsion-free. The endomorphism algebra of A is $\mathrm{End}^0(A) := \mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q}$.

The Rosati involution. The polarization ϕ defines a *Rosati involution* on the algebra $\mathrm{End}^0(A)$ [Mil86a, §17], defined formally as

$$f \mapsto f^\dagger := \phi^{-1} \circ \widehat{f} \circ \phi.$$

An endomorphism f of A is called *real* if $f^\dagger = f$. The map

$$\mathcal{L} \mapsto \phi^{-1} \circ \phi_{\mathcal{L}} \tag{2.2}$$

is a bijection between $\mathrm{NS}(A) \otimes \mathbb{Q}$ and the subspace $\mathrm{End}^0(A)^\dagger$ of real endomorphisms of A [Mil86a, Prop. 17.2], carrying tensor product to addition. If ϕ is a principal polarization, then the Rosati involution leaves $\mathrm{End}(A)$ stable, and the map (2.2) is a bijection between $\mathrm{NS}(A)$ and $\mathrm{End}(A)^\dagger$.

In the case $k = \mathbb{C}$, if we write $A = \mathbb{C}^g/\Lambda$, then endomorphisms of A are identified with endomorphisms of $\mathbb{C}^g$ as a $\mathbb{C}$-vector space (i.e. $g \times g$ matrices) leaving Λ stable; and the Rosati involution is adjunction with respect to the symplectic form ψ attached to the polarization [BL04, §5].

The Rosati involution can be used to classify the possible endomorphism algebras of abelian varieties. The involution $\dagger$ is always a *positive involution* [Mum70, §21, Thm. 1]. Recall that if B is a finite-dimensional $\mathbb{Q}$-algebra, a positive involution on B is an "anti-involution" $*\colon B \to B$ (i.e. $b^{**} = b$ and $(bb')^* = b'^* b^*$ for all $b, b' \in B$) such that the bilinear form $(b, b') \mapsto \mathrm{Tr}_{B/\mathbb{Q}}(bb'^*)$ is symmetric and positive definite. If A is *simple* (i.e. A is not isogenous to products of abelian varieties of smaller dimensions), then $\mathrm{End}^0(A)$ is a division algebra [Mum70, §19, Cor. 2]. Further restrictions on $\mathrm{End}^0(A)$ are given in [Mum70, §21, Thm. 2].

Isogeny types. Let $(A, \mathcal{L})$ and $(A', \mathcal{L}')$ be polarized abelian varieties; here we consider polarizations as line bundles. Let $\ell \geq 1$. We say that an isogeny $\varphi\colon A \to A'$ is an *ℓ-isogeny* if $\varphi^*\mathcal{L}'$ is algebraically equivalent to $\mathcal{L}^{\otimes \ell}$; equivalently, φ is an ℓ-isogeny if $\varphi^*\mathcal{L}' \in \mathrm{NS}(A)$ corresponds to the real endomorphism $\ell \in \mathrm{End}^0(A)^\dagger$ under the bijection (2.2). An ℓ-isogeny has degree ℓ^{2g}. If A and A' have dimension 1, we recover the usual notion of ℓ-isogenies between elliptic curves.

When more endomorphisms are present, there are other isogeny types to consider. Let F be a totally real field of degree g over $\mathbb{Q}$, and let $\mathcal{O}$ be an order in F. Let $(A, \mathcal{L})$ and $(A', \mathcal{L}')$ be principally polarized abelian varieties of dimension g, and assume that we are given *real multiplication embeddings*

$$\iota\colon \mathcal{O} \hookrightarrow \mathrm{End}(A)^\dagger, \quad \iota'\colon \mathcal{O} \hookrightarrow \mathrm{End}(A')^\dagger.$$

We say that A, A' have *real multiplication* by $\mathcal{O}$. Let $\beta \in \mathcal{O}$. We say that an isogeny $\varphi\colon A \to A'$ is a *β-isogeny* if

- the line bundle $\varphi^*\mathcal{L}' \in \mathrm{NS}(A)$ corresponds to the real endomorphism $\iota(\beta) \in \mathrm{End}(A)^\dagger$, and
- the isogeny φ is compatible with the real multiplication embeddings, in other words we have $\varphi \circ \iota(\alpha) = \iota'(\alpha) \circ \varphi$ for every $\alpha \in \mathcal{O}$.

A necessary condition for a β-isogeny to exist is that $\beta \in \mathcal{O}$ is totally positive [DJRV17, Thm. 1.1].

2.2 The theory of PEL Shimura varieties

2.2.1 Motivating the definition

Consider the modular curve $X(1)$, introduced in §1.1.1. It is a compactification of the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}_1$, where the modular group $\mathrm{SL}_2(\mathbb{Z})$ acts on the upper half plane $\mathbb{H}_1$ as in eq. (1.1). More generally, modular curves of higher level are obtained as compactifications of quotients of the form $\Gamma\backslash\mathbb{H}_1$, where $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup. Recall that a subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is *congruence* if it contains a subgroup of the form

$$\Gamma(N) = \left\{\gamma \in \mathrm{SL}_2(\mathbb{Z})\colon \gamma = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod N\right\}, \tag{2.3}$$

for some $N \geq 1$. An example is given by the group $\Gamma = \Gamma^0(\ell)$ where ℓ is a prime, defined by eq. (1.3).

A drawback of this classical definition is that it is not immediately clear what the generalizations of $\mathbb{H}_1$ and $\mathrm{SL}_2(\mathbb{Z})$ should be in order to

obtain moduli spaces of abelian varieties of higher dimensions. The goal of this section is to present another way of defining modular curves which is better suited for generalization. Our presentation is heavily inspired by Milne [Mil05, §4-7].

Another interpretation of $\mathbb{H}_1$. The action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}_1$ is the restriction of a "larger" action by the group $\mathrm{GL}_2(\mathbb{R})_+$ consisting of all real 2×2 matrices with positive determinant. The action of $\mathrm{GL}_2(\mathbb{R})_+$ on $\mathbb{H}_1$ is transitive. Let us fix $i \in \mathbb{H}_1$ as a base point. The stabilizer of i in $\mathrm{GL}_2(\mathbb{R})_+$ is the subgroup

$$K_\infty = \left\{ \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix} \right) : a, b \in \mathbb{R},\ (a,b) \neq (0,0) \right\},$$

and we have a bijection

$$\mathrm{GL}_2(\mathbb{R})_+ / K_\infty \simeq \mathbb{H}_1, \qquad \gamma \mapsto \gamma i.$$

As a quotient of Lie groups, $\mathrm{GL}_2(\mathbb{R})_+ / K_\infty$ has a natural structure of a real differentiable manifold which corresponds to that of $\mathbb{H}_1$.

In turn, we can give an "intrinsic" definition of the subgroup K_∞. By definition, the group $\mathrm{GL}_2(\mathbb{R})_+$ acts on the real vector space $V(\mathbb{R}) = \mathbb{R}^2$ (we think of V as the algebraic variety $\mathbb{A}^2$ over $\mathbb{Q}$, hence the notation). Moreover $V(\mathbb{R})$ has a natural *complex structure*, meaning an endomorphism x_0 of $V(\mathbb{R})$ such that $x_0 \circ x_0 = -1$: take

$$x_0 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right)$$

in the canonical basis (e_1, e_2) of $V(\mathbb{R})$. We check the following facts.

1. The group $\mathrm{GL}_2(\mathbb{R})_+$ acts on the set of complex structures on $V(\mathbb{R})$ by conjugation: if x is a complex structure on $V(\mathbb{R})$ and $\gamma \in G(\mathbb{R})_+$, we set $\gamma x := \gamma \circ x \circ \gamma^{-1}$.
2. The stabilizer of x_0 under this action is K_∞; this is the intrisic definition of K_∞ we were looking for.
3. The orbit of x_0 under the action of $\mathrm{GL}_2(\mathbb{R})_+$ consists of all complex structures $x = \left(\begin{smallmatrix} s & t \\ u & v \end{smallmatrix} \right)$ such that $t > 0$.

To explain point 3, we define a symplectic form ψ on V by setting

$$\psi(e_1, e_2) = -1.$$

Then the complex structure x_0 is *positive* for ψ, meaning that the bilinear form defined by $(u, v) \mapsto \psi(u, x_0(v))$ on $V(\mathbb{R}) \times V(\mathbb{R})$ is symmetric and

positive definite; indeed, its matrix in the canonical basis is the identity. The action of $\mathrm{GL}_2(\mathbb{R})$ on V is *compatible* with ψ, in the sense that for every $\gamma \in \mathrm{GL}_2(\mathbb{R})$, there exists $\lambda \in \mathbb{R}^\times$ such that

$$\forall u, v \in V(\mathbb{R}),\ \psi(\gamma u, \gamma v) = \lambda \psi(u, v).$$

Therefore, if we set $x = \gamma x_0$, the bilinear form $(u, v) \mapsto \psi(u, x(v))$ is still symmetric. Moreover, $\mathrm{GL}_2(\mathbb{R})_+$ is connected: therefore, if $\gamma \in \mathrm{GL}_2(\mathbb{R})_+$, we have $\lambda > 0$ and hence the symmetric bilinear form $\psi(u, x(v))$ is positive definite. The point 3 above states that the orbit of x_0 under $\mathrm{GL}_2(\mathbb{R})_+$ consists of all complex structures on $V(\mathbb{R})$ that are positive for ψ.

Turning things around, we start from the vector space V endowed with the symplectic form ψ; we define $G = \mathrm{GL}_2$ as the algebraic group of automorphisms of V that are compatible with ψ; and we fix a complex structure x_0 on $V(\mathbb{R})$ that is positive for ψ. Then the upper half plane $\mathbb{H}_1$, with its usual action of $\mathrm{GL}_2(\mathbb{R})_+$, is identified with the orbit of x_0 under the action of $\mathrm{GL}_2(\mathbb{R})_+$ by conjugation.

Another interpretation of congruence subgroups. The definition of congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$ also seems rather arbitrary: why consider the subgroups $\Gamma(N)$ in the first place? In fact, these subgroups arise naturally when working with adeles.

Let $\mathcal{P}_\mathbb{Q}$ denote the set of prime numbers. Recall that the ring of finite adeles $\mathbb{A}_f$ of $\mathbb{Q}$ is defined as a restricted product over $\mathcal{P}_\mathbb{Q}$:

$$\mathbb{A}_f = \prod_{p \in \mathcal{P}_\mathbb{Q}}{}' (\mathbb{Q}_p\,;\mathbb{Z}_p), \tag{2.4}$$

where $\mathbb{Q}_p$ and $\mathbb{Z}_p$ denote the field of p-adic numbers and the ring of p-adic integers, respectively. The elements of $\mathbb{A}_f$ are tuples $(x_p)_{p \in \mathcal{P}_\mathbb{Q}}$ such that $x_p \in \mathbb{Q}_p$ for all p, and $x_p \in \mathbb{Z}_p$ for all p but finitely many. The ring $\mathbb{A}_f$ is endowed with the adelic topology: a basis of neighborhoods of $0 \in \mathbb{A}_f$ consists of open sets of the form $\prod_p U_p$, where $U_p \subset \mathbb{Q}_p$ is an open set containing 0, and $U_p = \mathbb{Z}_p$ for every p but finitely many.

If G is any affine algebraic group over $\mathbb{Q}$ (for instance $G = \mathrm{GL}_2$), we can define $G(\mathbb{A}_f)$ as a topological group. To do so, choose a closed immersion η of G in an affine space $\mathbb{A}^n_\mathbb{Q}$, for some $n \geq 1$. This choice allows us to define $G(\mathbb{Z})_\eta$ and $G(\mathbb{Z}_p)_\eta$ for every prime p, as the sets of solutions in $\mathbb{Z}$ (resp. $\mathbb{Z}_p$) of the equations defining G; these sets depend on η, hence the subscript. Then we form the restricted product

$$G(\mathbb{A}_f)_\eta = \prod_{p \in \mathcal{P}_\mathbb{Q}}{}' \big(G(\mathbb{Q}_p)\,;G(\mathbb{Z}_p)_\eta\big).$$

and we declare that a basis of neigborhoods of $1 \in G(\mathbb{A}_f)_\eta$ consists of products of the form $\prod_p U_p$, where $U_p \subset G(\mathbb{Q}_p)$ is an open set containing 1, and $U_p = G(\mathbb{Z}_p)_\eta$ for every p but finitely many.

If η' is another choice of closed immersion of G into an affine space, then $\eta(G)$ and $\eta'(G)$ are isomorphic over $\mathbb{Q}$. Therefore they are isomorphic over $\mathbb{Z}[1/D]$, for some $D \geq 1$. The subsets $G(\mathbb{Z}_p)_\eta$ and $G(\mathbb{Z}_p)_{\eta'}$ of $G(\mathbb{Q}_p)$ are the same for every prime p not dividing D, therefore $G(\mathbb{A}_f)_\eta = G(\mathbb{A}_f)_{\eta'}$ as topological spaces. Thus $G(\mathbb{A}_f)$ is well defined.

We return to the special case $G = \mathrm{GL}_2$, and denote by $\mathrm{GL}_2(\mathbb{Q})_+$ the subset of $\mathrm{GL}_2(\mathbb{Q})$ consisting of matrices with positive determinant. Recall the usual notation

$$\widehat{\mathbb{Z}} = \prod_{p \in \mathcal{P}_\mathbb{Q}} \mathbb{Z}_p = \varprojlim_{n \to \infty} \mathbb{Z}/n\mathbb{Z}.$$

Proposition 2.2.1. *For every congruence subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$, there exists a compact open subgroup K of $\mathrm{GL}_2(\mathbb{A}_f)$ contained in $\mathrm{GL}_2(\widehat{\mathbb{Z}})$ such that $\Gamma = \mathrm{GL}_2(\mathbb{Q})_+ \cap K$. Conversely, for every compact open subgroup K of $\mathrm{GL}_2(\mathbb{A}_f)$ that is contained in $\mathrm{GL}_2(\widehat{\mathbb{Z}})$, the intersection $\mathrm{GL}_2(\mathbb{Q})_+ \cap K$ is a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$.*

Proof. This is [Mil05, Prop. 4.1], with the additional remark that

$$\mathrm{GL}_2(\mathbb{Z}) \cap \mathrm{GL}_2(\mathbb{Q})_+ = \mathrm{SL}_2(\mathbb{Z}). \qquad \square$$

In particular, we have $\mathrm{SL}_2(\mathbb{Z}) = \mathrm{GL}_2(\mathbb{Q})_+ \cap K_0$, with $K_0 = \mathrm{GL}_2(\widehat{\mathbb{Z}})$. We already know that the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}_1$ can be rewritten as the double quotient $\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{GL}_2(\mathbb{R})_+/K_\infty$, where K_∞ is the stabilizer of the complex structure x_0 on $V(\mathbb{R})$.

Proposition 2.2.2. *The map*

$$\begin{array}{ccc} \mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{GL}_2(\mathbb{R})_+/K_\infty & \to & \mathrm{GL}_2(\mathbb{Q})_+\backslash\big(\mathrm{GL}_2(\mathbb{R})_+ \times \mathrm{GL}_2(\mathbb{A}_f)\big)/\big(K_\infty \times K_0\big) \\ \gamma & \mapsto & \big(\gamma, \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\big) \end{array}$$

is a bijection.

Proof. First, we have to explain how the double quotient on the right is formed: K_∞ acts on $\mathrm{GL}_2(\mathbb{R})_+$ by multiplication on the right, K_0 acts on $\mathrm{GL}_2(\mathbb{A}_f)$ by multiplication on the right, and $\mathrm{GL}_2(\mathbb{Q})_+$ acts by multiplication on the left simultaneously on $\mathrm{GL}_2(\mathbb{R})_+$ and $\mathrm{GL}_2(\mathbb{A}_f)$.

By [Mil05, Thm. 5.17 and Lem. 5.20] (the assumption that K is sufficiently small is not actually needed there; moreover SL_2 is simply connected

as an algebraic group [Mil17, Ex. 18.20]), the determinant induces a bijection

$$\begin{aligned}\mathrm{GL}_2(\mathbb{Q})_+ \backslash \mathrm{GL}_2(\mathbb{A}_f)/K_0 &\to \det\big(\mathrm{GL}_2(\mathbb{Q})_+\big)\backslash \det\big(\mathrm{GL}_2(\mathbb{A}_f)\big)/\det(K_0)\\ &= \mathbb{Q}_+^\times \backslash \mathbb{A}_f^\times / \widehat{\mathbb{Z}}^\times.\end{aligned}$$

From description of $\mathbb{A}_f$ given in eq. (2.4), we see that $\mathbb{A}_f^\times = \mathbb{Q}_+^\times \widehat{\mathbb{Z}}^\times$. Therefore the double quotient $\mathbb{Q}_+^\times \backslash \mathbb{A}_f^\times / \widehat{\mathbb{Z}}^\times$ is trivial, and every element of $\mathrm{GL}_2(\mathbb{R})_+ \times \mathrm{GL}_2(\mathbb{A}_f)$ has a representative of the form $\big(\gamma, \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)\big)$ for some $\gamma \in \mathrm{GL}_2(\mathbb{R})_+$ modulo the action of $\mathrm{GL}_2(\mathbb{Q})_+$ and K_0.

To conclude, we only have to show that for every $\gamma, \gamma' \in \mathrm{GL}_2(\mathbb{R})_+$, the pairs $\big(\gamma, \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)\big)$ and $\big(\gamma', \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)\big)$ are equal in the double quotient if and only if γ and γ' are equal modulo $\mathrm{SL}_2(\mathbb{Z})$. This is straightforward given that $\mathrm{SL}_2(\mathbb{Z}) = \mathrm{GL}_2(\mathbb{Q})_+ \cap K_0$. □

Note that the double quotient on the right hand side of proposition 2.2.2 would make sense for algebraic groups other than GL_2: we have succeeded in finding another definition of the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}_1$ that is suitable for generalization. It remains to find out which algebraic groups give rise to moduli spaces of polarized abelian varieties.

The modular interpretation of $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}_1$. Finally, we take another look at the well-known fact that $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}_1$ parametrizes isomorphism classes of elliptic curves over $\mathbb{C}$. The elliptic curve associated with $\tau \in \mathbb{H}_1$ is $E(\tau) = \mathbb{C}/\Lambda(\tau)$, where $\Lambda(\tau) = \mathbb{Z} \oplus \tau\mathbb{Z}$. Consider the bijection

$$\eta_\tau \colon V(\mathbb{R})/\mathbb{Z}^2 \to \mathbb{C}/\Lambda(\tau), \quad e_1 \mapsto 1,\ e_2 \mapsto -\tau.$$

Then, after simple computations with 2×2 matrices, we obtain:

Proposition 2.2.3. *Let $\gamma \in \mathrm{GL}_2(\mathbb{R})_+$, and let $\tau = \gamma i \in \mathbb{H}_1$. Then the pullback under η_τ of the multiplication-by-i map on $\mathbb{C}$ is the complex structure γx_0 on $V(\mathbb{R})$.*

In other words, we have a distiguished lattice $\Lambda_0 = \mathbb{Z}^2 \subset V(\mathbb{R})$; and for every complex structure $x \in \mathrm{GL}_2(\mathbb{R})_+/K_\infty$, the elliptic curve associated with x is the quotient $V(\mathbb{R})/\Lambda_0$, with x as a complex structure.

We can also describe the natural principal polarization of $E(\tau)$. Recall that V is endowed with the symplectic form ψ such that $\psi(e_1, e_2) = -1$. When restricted to $\Lambda_0 \times \Lambda_0$, the form ψ takes values in $\mathbb{Z}$, and corresponds via η_τ to the polarization of $E(\tau)$.

In order to construct moduli spaces of polarized abelian varieties of higher dimensions, it is promising to start from a $\mathbb{Q}$-vector space V of higher dimension, endowed with a symplectic form ψ and a complex structure x_0 that is positive for ψ, and to consider the algebraic group G of automorphisms of V compatible with this data (and possibly additional endomorphisms of V, if we want to classify abelian varieties with an endomorphism structure). Then we construct double quotients of the form

$$G(\mathbb{Q})_+ \backslash \big(G(\mathbb{R})_+ \times G(\mathbb{A}_f)\big) / \big(K_\infty \times K\big)$$

where K_∞ is the stabiliser of x_0, the subgroup $K \subset G(\mathbb{A}_f)$ is compact and open, and the "plus" subscripts encode some connectedness condition. We arrive at the definition of a PEL Shimura variety.

2.2.2 Simple PEL Shimura varieties of type (A) or (C)

We continue to follow Milne's exposition of the theory of Shimura varieties [Mil05], which is itself inspired from Deligne's presentation of Shimura's works [Del70].

Notation. If G is a connected affine algebraic group over $\mathbb{Q}$, then

- G^{der} is the derived group of G,
- Z is the center of G,
- $G^{\mathrm{ad}} = G/Z$ is the adjoint group of G,
- $T = G/G^{\mathrm{der}}$ is the largest abelian quotient of G,
- $\nu \colon G \to T$ is the natural quotient map,
- $G^{\mathrm{ad}}(\mathbb{R})_+$ is the connected component of the identity in $G^{\mathrm{ad}}(\mathbb{R})$ for the real topology,
- $G(\mathbb{R})_+$ is the preimage of $G^{\mathrm{ad}}(\mathbb{R})_+$ in $G(\mathbb{R})$, and finally
- $G(\mathbb{Q})_+ = G(\mathbb{Q}) \cap G(\mathbb{R})_+$.

Note that Z, G^{der}, T and G^{ad} are algebraic groups, i.e. exist as schemes. Defining them is not as direct as in the case of "abstract" groups. First, there exists a subgroup Z of G such that for every $\mathbb{Q}$-algebra R, the group $Z(R)$ is the center of the group $G(R)$ [Mil17, Prop. 1.92]. Then the quotient G^{ad} exists by [Mil17, Thm. 5.14], and has following property: if k is an algebraically closed field containing $\mathbb{Q}$, then $G^{\mathrm{ad}}(k) = G(k)/Z(k)$ [Mil17, Prop. 5.47]. The group G^{der} is defined as the smallest normal subgroup N of G such that G/N is commutative; G^{der} can also be characterized

as the subgroup of G generated by commutators [Mil17, Prop. 6.18]. Finally, T is also defined as a scheme-theoretic quotient.

For instance, if $G = \mathrm{GL}_2$, we have $G^{\mathrm{der}} = \mathrm{SL}_2$; $Z = \mathbb{G}_m$; $G^{\mathrm{ad}} = \mathrm{PGL}_2$; $T = \mathbb{G}_m$; $\nu = \det$; and $G(\mathbb{R})_+$, $G(\mathbb{Q})_+$ are the sets of real (resp. rational) 2×2 matrices with positive determinant. This coincides with the notation used in §2.2.1.

PEL data. Let us introduce the data necessary to define a PEL Shimura variety. As indicated in §2.2.1, it consists of a $\mathbb{Q}$-vector space V equipped with a symplectic form, certain endomorphisms, and a class of complex structures, that are "compatible" in a way. Following Milne [Mil05, §8], we start with the endomorphism algebra of V; it is a finite-dimensional $\mathbb{Q}$-algebra B endowed with a positive involution denoted by $*$. We also assume that B is *simple*, meaning that its only two-sided ideals are 0 and B.

The center F of B is a number field; let $F_0 \subset F$ be the subfield of invariants under $*$. We make the technical assumption that B is either of type (A) or (C) [Mil05, Prop. 8.3]: this means that for every embedding θ of F_0 in an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$, the algebra with positive involution $(B \otimes_{F_0,\theta} \overline{\mathbb{Q}}, *)$ is isomorphic to a product of factors of the form, respectively,

(A) $\mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}}) \times \mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}})$ with $(a, b)^* = (b^t, a^t)$, or

(C) $\mathrm{Mat}_{n\times n}(\overline{\mathbb{Q}})$ with $a^* = a^t$.

Let (V, ψ) be a faithful symplectic $(B, *)$-module. In other words, V is a finite-dimensional $\mathbb{Q}$-vector space equipped with a faithful B-module structure and a nondegenerate alternating $\mathbb{Q}$-bilinear form ψ such that for all $b \in B$ and for all $u, v \in V$,

$$\psi(b^* u, v) = \psi(u, bv). \tag{2.5}$$

Let $\mathrm{GL}_B(V)$ denote the group of automorphisms of V respecting the action of B, and let G be its (reduced) algebraic subgroup defined by

$$G(\mathbb{Q}) = \big\{ g \in \mathrm{GL}_B(V) \colon \psi(gx, gy) = \psi(\mu(g)x, y) \text{ for some } \mu(g) \in F_0^\times \big\}.$$

The group G is affine, connected, and reductive, and its derived group is $G^{\mathrm{der}} = \ker(\mu) \cap \ker(\det)$ [Mil05, Prop. 8.7]. We warn the reader that our G is denoted by G_1 in [Mil05, §8 of the 2017 version]. In Milne's terminology, our G will define a Shimura variety (so that the results of [Mil05, §5] apply), but not strictly speaking a PEL Shimura variety. This choice of reductive group will allow us to consider more Hecke correspondences later on.

Let x be a complex structure on $V(\mathbb{R})$, meaning an endomorphism of $V(\mathbb{R})$ such that $x^2 = -1$. We say that x is *positive for* ψ if it commutes with the action of B and if the bilinear form $(u, v) \mapsto \psi\big(u, x(v)\big)$ on $V(\mathbb{R})$ is symmetric and positive definite. In particular, $x \in G(\mathbb{R})$ and $\mu(x) = 1$. Such a complex structure x_0 exists [Mil05, Prop. 8.14]. Define X_+ to be the orbit of x_0 under the action of $G(\mathbb{R})_+$ by conjugation. We call the tuple $(B, *, V, \psi, G, X_+)$ a *simple PEL Shimura datum of type (A) or (C)*, or simply a *PEL datum*. To simplify notations, we abbreviate PEL data as pairs (G, X_+), the underlying data (V, ψ) and $(B, *)$ being implicit.

Example 2.2.4. In the example $V = \mathbb{Q}^2$ studied in §2.2.1, there were no endomorphisms of V to consider. This corresponds to taking $B = \mathbb{Q}$ with the involution $*$ being the identity map; in this case $(B, *)$ is indeed a simple $\mathbb{Q}$-algebra with positive involution of type (C). The bilinear form ψ satisfies eq. (2.5) by $\mathbb{Q}$-linearity. All automorphisms of V respect the action of B, and the space of symplectic forms on $\mathbb{Q}^2$ is one-dimensional, hence $G = \mathrm{GL}_2$. The rest of the definitions coincide with those of §2.2.1.

Remark 2.2.5. Let (G, X_+) be a PEL datum. One can show that the space X_+ is a *hermitian symmetric domain* [Mil05, Cor. 5.8]. By definition, a connected hermitian manifold Y is a hermitian symmetric domain if

- The group of holomorphic isometries of Y acts transitively on Y,
- For each $y \in Y$, there exists an involution of Y having y as an isolated fixed point.

If one gives $\mathbb{H}_1$ the usual hermitian metric, then the group $\mathrm{GL}_2(\mathbb{R})_+$ acts on $\mathbb{H}_1$ by holomorphic isometries, and $z \mapsto -1/z$ is the required symmetry at $i \in \mathbb{H}_1$ [Mil05, Ex. 1.1]. This shows that $\mathbb{H}_1$ is indeed a hermitian symmetric domain.

PEL Shimura varieties. Let (G, X_+) be a PEL datum as above, let K be a compact open subgroup of $G(\mathbb{A}_f)$, and let K_∞ be the stabilizer of x_0 in $G(\mathbb{R})_+$. The *PEL Shimura variety* associated with (G, X_+) of level K is the double quotient

$$\begin{aligned} \mathrm{Sh}_K(G, X_+)(\mathbb{C}) &= G(\mathbb{Q})_+ \backslash (X_+ \times G(\mathbb{A}_f))/K \\ &= G(\mathbb{Q})_+ \backslash (G(\mathbb{R})_+ \times G(\mathbb{A}_f))/K_\infty \times K. \end{aligned} \tag{2.6}$$

Actually, this quotient will be the set of $\mathbb{C}$-points of the Shimura variety, hence the notation. In the first line of (2.6), the group $G(\mathbb{Q})_+$ acts on X_+ and $G(\mathbb{A}_f)$ by conjugation and left multiplication respectively, and K acts on $G(\mathbb{A}_f)$ by right multiplication. When the context is clear, we

omit (G, X_+) from the notation. The set $\mathrm{Sh}_K(\mathbb{C})$ is given the quotient topology obtained from the real topology on $G(\mathbb{R})_+$ and the adelic topology on $G(\mathbb{A}_f)$.

In order to describe $\mathrm{Sh}_K(\mathbb{C})$ more explicitly, we study its connected components. The projection to the second factor induces a map with connected fibers from $\mathrm{Sh}_K(\mathbb{C})$ to the double quotient $G(\mathbb{Q})_+\backslash G(\mathbb{A}_f)/K$, which is finite [Mil05, Lem. 5.12]. Let $\mathcal{C}$ be a set of representatives in $G(\mathbb{A}_f)$ for this double quotient. The connected component $\mathcal{S}_c$ of $\mathrm{Sh}_K(\mathbb{C})$ indexed by $c \in \mathcal{C}$ can be identified with $\Gamma_c\backslash X_+$, where $\Gamma_c = G(\mathbb{Q})_+ \cap cKc^{-1}$ is an arithmetic subgroup of $\mathrm{Aut}(X_+)$ [Mil05, Lem. 5.13]. Thus, the Shimura variety $\mathrm{Sh}_K(\mathbb{C})$ has a natural structure of a complex analytic space, and is an algebraic variety by the theorem of Baily and Borel [Mil05, Thm. 3.12].

Since G^{der} is simply connected, by [Mil05, Thm. 5.17 and Lem. 5.20] (as above, the assumption that K is sufficiently small is not necessary), the map ν induces an isomorphism

$$G(\mathbb{Q})_+\backslash G(\mathbb{A}_f)/K \simeq \nu(G(\mathbb{Q})_+)\backslash T(\mathbb{A}_f)/\nu(K).$$

Therefore the set of connected components of $\mathrm{Sh}_K(\mathbb{C})$ is a finite abelian group. Moreover, each connected component is itself a Shimura variety with underlying group G^{der} [Mil05, Rem. 5.23].

A fundamental theorem states that the Shimura variety $\mathrm{Sh}_K(G, X_+)$ exists as an algebraic variety defined over the *reflex field* $E(G, X_+)$, which is a number field contained in $\mathbb{C}$, depending only on the PEL datum [Mil05, §12-14]. The field of definition of the individual connected components of $\mathrm{Sh}_K(\mathbb{C})$ depends on K, and is a finite abelian extension of $E(G, X_+)$.

Example 2.2.6. By §2.2.1, the modular curve $X(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}_1$ is the PEL Shimura variety associated with $(G, X_+) = (\mathrm{GL}_2, \mathbb{H}_1)$ and congruence subgroup $K_0 = \mathrm{GL}_2(\widehat{\mathbb{Z}})$, choosing the identity matrix as the unique element of $\mathcal{C}$. Let us also recover the classical modular curves of higher level as special cases of PEL Shimura varieties.

Choose an integer $N \geq 1$, and for every $p \in \mathcal{P}_{\mathbb{Q}}$, denote by $v_p(N)$ the p-adic valuation of N. Let K_N be the compact open subgroup of $\mathrm{GL}_2(\mathbb{A}_f)$ defined by

$$K_N = \left\{\gamma \in G(\mathbb{A}_f) \colon \forall\, p \in \mathcal{P}_{\mathbb{Q}},\ \gamma_p \in \mathrm{GL}_2(\mathbb{Z}_p) \text{ and } \gamma_p = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \bmod p^{v_p(N)}\right\}.$$

It is a subgroup of $\mathrm{GL}_2(\widehat{\mathbb{Z}})$ defined by congruence conditions modulo the prime divisors of N.

If $N > 1$, then the group of connected components of $\mathrm{Sh}_{K_N}(\mathbb{C})$ is not trivial. Still, choosing the identity matrix as an element of $\mathcal{C}$ identifies one of

the connected components of $\mathrm{Sh}_{K_N}(\mathbb{C})$ with the quotient $\Gamma(N)\backslash\mathbb{H}_1$: indeed, by (2.3), we have $\Gamma(N) = \mathrm{GL}_2(\mathbb{Q})_+ \cap K_N$. This quotient is the classical modular curve $X(N)$ of level N; one can check that $X(N)$ is defined over the cyclotomic field $\mathbb{Q}(\zeta_N)$.

2.2.3 Modular interpretation of PEL Shimura varieties

Our motivation in constructing PEL Shimura varieties is to obtain moduli spaces of complex abelian varieties with polarization, endomorphism, and level structures. This *modular interpretation* of PEL Shimura varieties is usually formulated in terms of isogeny classes of abelian varieties (see for instance [Mil05, Thm. 8.17]). In order to obtain a modular intepretation in terms of *isomorphism* classes of abelian varieties in the spirit of [Car86, §2.6], as in the case of elliptic curves, we fix

- A PEL datum (G, X_+),
- A lattice $\Lambda_0 \subset V$ (see p. 14 for our conventions on lattices),
- A compact open subgroup $K \subset G(\mathbb{A}_f)$ which is assumed to stabilize the adelic lattice $\widehat{\Lambda}_0 = \Lambda_0 \otimes \widehat{\mathbb{Z}} \subset V(\mathbb{A}_f)$,
- A set $\mathcal{C} \subset G(\mathbb{A}_f)$ of representatives for the finite double quotient $G(\mathbb{Q})_+\backslash G(\mathbb{A}_f)/K$.

Finally, let $\mathcal{O}$ be the largest order in B stabilizing Λ_0. As in §2.2.2, for every $c \in \mathcal{C}$, we write $\Gamma_c = G(\mathbb{Q})_+ \cap cKc^{-1}$. We also denote by $\mathcal{S}_c = \Gamma_c\backslash X_+$ the connected component of $\mathrm{Sh}_K(G, X_+)(\mathbb{C})$ associated with c.

The local-global principle for lattices holds: the map $\Lambda \mapsto \widehat{\Lambda} = \Lambda \otimes \widehat{\mathbb{Z}}$ is a bijection between lattices in V and lattices in $V(\mathbb{A}_f)$, and its inverse is intersection with $V(\mathbb{Q})$. The assumption that K stabilizes $\widehat{\Lambda}_0$ does not imply a loss of generality, because every compact open subgroup of $G(\mathbb{A}_f)$ stabilizes some lattice in $V(\mathbb{A}_f)$.

We define a *polarized lattice* to be a pair (Λ, ϕ) where Λ is a free $\mathbb{Z}$-module of finite rank and $\phi\colon \Lambda \times \Lambda \to \mathbb{Z}$ is a nondegenerate alternating form. Given a polarized lattice (Λ, ϕ), we can extend ϕ to the $\mathbb{Q}$-vector space $\Lambda \otimes \mathbb{Q}$, and we define

$$\Lambda^\perp = \{v \in \Lambda \otimes \mathbb{Q}\colon \forall w \in \Lambda,\ \phi(v, w) \in \mathbb{Z}\}.$$

Then $\Lambda^\perp/\Lambda$ is a finite abelian group called the *polarization type* of (Λ, ϕ). We say that ϕ is a *principal polarization* on Λ if $\Lambda^\perp = \Lambda$.

A modular interpretation in terms of lattices. Using the data above, we define a *standard polarized lattice* for each connected component of the Shimura variety $\mathrm{Sh}_K(\mathbb{C})$, as follows.

Definition 2.2.7. For each $c \in \mathcal{C}$, we define

$$\widehat{\Lambda}_c = c(\widehat{\Lambda}_0), \quad \text{and} \quad \Lambda_c = \widehat{\Lambda}_c \cap V(\mathbb{Q}).$$

The action of c, or any other element of $G(\mathbb{A}_f)$, on adelic lattices is easily defined locally at each prime. Since c respects the action of B on $V(\mathbb{A}_f)$, the order $\mathcal{O}$ is again the stabilizer of $\widehat{\Lambda}_c$, and thus of Λ_c. Let $\lambda_c \in \mathbb{Q}_+^\times$ be such that $\psi_c := \lambda_c \psi$ satisfies $\psi_c(\Lambda_c \times \Lambda_c) = \mathbb{Z}$. We call (Λ_c, ψ_c) with its structure of $\mathcal{O}$-module the *standard polarized lattice* associated with (Λ_0, c).

Choose $c \in \mathcal{C}$, and let (Λ_c, ψ_c) be the standard polarized lattice associated with (Λ_0, c). We consider tuples $(\Lambda, x, \iota, \phi, \eta K)$ where

- Λ is a free $\mathbb{Z}$-module of rank $\dim V$,
- $x \in \mathrm{End}(\Lambda \otimes \mathbb{R})$ is a complex structure on $\Lambda \otimes \mathbb{R}$,
- ι is an embedding $\mathcal{O} \hookrightarrow \mathrm{End}_{\mathbb{Z}}(\Lambda)$,
- $\phi\colon \Lambda \times \Lambda \to \mathbb{Z}$ is a nondegenerate alternating $\mathbb{Z}$-bilinear form on Λ,
- ηK is a K-orbit of $\widehat{\mathbb{Z}}$-linear isomorphisms of $\mathcal{O}$-modules $\widehat{\Lambda}_0 \to \Lambda \otimes \widehat{\mathbb{Z}}$,

satisfying the following condition of compatibility with (Λ_c, ψ_c):

($\star$) There exists an isomorphism of $\mathcal{O}$-modules $a\colon \Lambda \to \Lambda_c$, carrying ηK to cK and x to an element of X_+, such that

$$\exists \zeta \in \mu(\Gamma_c),\ \forall u, v \in \Lambda,\ \phi(u,v) = \psi_c\big(\zeta a(u), a(v)\big).$$

For short, we call such a tuple a *lattice with PEL structure defined by* (Λ_0, c), or simply a *lattice with PEL structure* when the dependency on (Λ_0, c) is understood.

By definition, an *isomorphism* between two lattices with PEL structure denoted by $(\Lambda, x, \iota, \phi, \eta K)$ and $(\Lambda', x', \iota', \phi', \eta' K)$ is an isomorphism of $\mathcal{O}$-modules $f\colon \Lambda \to \Lambda'$ that sends x to x', sends ηK to $\eta' K$, and such that $\phi(u,v) = \phi'\big(\zeta f(u), f(v)\big)$ for some $\zeta \in \mu(\Gamma_c)$.

For every lattice with PEL structure $(\Lambda, x, \iota, \phi, \eta K)$, the compatibility condition ($\star$) implies in particular that the complex structure x is positive for ϕ, the adjunction involution defined by ϕ coincides with $*$ on B, the action of B on $\Lambda \otimes \mathbb{Q}$ leaves the complex structure x invariant, and the polarized lattices (Λ, ϕ) and (Λ_c, ψ_c) have the same polarization type.

Proposition 2.2.8. *Let $c \in \mathcal{C}$, and let $\mathcal{Z}_c$ be the set of isomorphism classes of lattices with PEL structure defined by (Λ_0, c). Then the map*

$$\begin{array}{ccc} \mathcal{Z}_c & \longrightarrow & \mathcal{S}_c \\ (\Lambda, x, \iota, \phi, \eta K) & \longmapsto & [axa^{-1},\, c] \quad \text{where } a \text{ is as in } (\star) \end{array} \tag{2.7}$$

is well-defined and bijective. The inverse map is

$$[x, c] \mapsto (\Lambda_c, x, \iota, \psi_c, cK).$$

where ι is the natural action of $\mathcal{O}$ on Λ_c.

Proof. We check that

1. *The map (2.7) is well defined.* The complex structure axa^{-1} belongs to X_+ because of condition $(\star)$. We show that, up to action of Γ_c by conjugation, it does not depend on the choice of a. Indeed, if a' is another choice of isomorphism, then $a' = qa$ where q is an automorphism of Λ_c leaving ψ_c invariant (modulo multiplication by some $\zeta \in \mu(\Gamma_c)$) and respecting the action of B; so q is an element of $G(\mathbb{Q})$. Also, the automorphism q sends cK to cK, and carries X_+ to X_+, hence $q \in G(\mathbb{Q})_+ \cap cKc^{-1} = \Gamma_c$ by [Mil05, Prop. 5.7].
2. *Isomorphic lattices with PEL structure map to the same point.* Let $(\Lambda, x, \iota, \phi, \eta K)$ and $(\Lambda', x', \iota', \phi', \eta' K)$ be two such lattices, and let f be an isomorphism between them. If $a\colon \Lambda \to \Lambda_c$ is as in $(\star)$, then $af\colon \Lambda' \to \Lambda_c$ is also a valid isomorphism. The complex structures on $V(\mathbb{R})$ constructed from (Λ, x, a) and (Λ', x', af) are the same.
3. *If two lattices with PEL structure map to the same class, then they are isomorphic.* Let $(\Lambda, x, \iota, \phi, \eta K)$ and $(\Lambda', x', \iota', \phi', \eta' K)$ be two lattices with PEL structure, and choose isomorphisms a, a' as in $(\star)$. By assumption, the complex structures axa^{-1} and $a'x'a'^{-1}$ on $V(\mathbb{R})$ differ by conjugation by an element $q \in \Gamma_c$. Then $a^{-1}qa'\colon \Lambda' \to \Lambda$ is the required isomorphism.
4. *The map (2.7) is surjective.* Let $[x, c] \in \mathcal{S}_c$. Then $[x, c]$ is the image of the lattice with PEL structure given by $(\Lambda_c, x, \iota, \psi_c, cK)$ where ι is given by the natural action of $\mathcal{O}$ on Λ_c: indeed, the identity $\Lambda_c \to \Lambda_c$ satisfies condition $(\star)$. □

A modular interpretation in terms of abelian varieties. Recall from §2.1.3 that every complex abelian variety is isomorphic to a complex torus of the form $\mathbb{C}^g/\Lambda$ for some lattice $\Lambda \subset \mathbb{C}^g$; in fact $\Lambda = H_1(A, \mathbb{Z})$. The inclusion $\Lambda \subset \mathbb{C}^g$ defines a natural complex structure on $\Lambda \otimes \mathbb{R}$. Under

this identification, endomorphisms of A correspond to endomorphisms of Λ that respect the complex structure, and we have seen that a polarization of A can be seen as a certain symplectic form taking integer values on Λ.

Let A be a complex abelian variety, and let $\Lambda = H_1(A, \mathbb{Z})$ be the lattice attached to A under the above equivalence of categories. Recall that for every prime number p, the *Tate module* $T_p(A)$ is defined as the projective limit of the torsion subgroups $A[p^n]$ as n tends to infinity:

$$T_p(A) = \varprojlim_{n\to\infty} A[p^n] = \varprojlim_{n\to\infty} \Lambda/p^n\Lambda = \Lambda \otimes \mathbb{Z}_p.$$

Therefore $\Lambda \otimes \widehat{\mathbb{Z}}$ is canonically isomorphic to the global Tate module $\widehat{T}(A)$ of A, defined as

$$\widehat{T}(A) = \prod_{p \text{ prime}} T_p(A).$$

Fix $c \in \mathcal{C}$, and let (Λ_c, ψ_c) be the standard polarized lattice associated with (Λ_0, c). We define a *complex abelian variety with PEL structure defined by* (Λ_0, c) to be a tuple $(A, \phi, \iota, \eta K)$ where

- (A, ϕ) is a complex polarized abelian variety of dimension $\dim V$,
- ι is an embedding $\mathcal{O} \hookrightarrow \mathrm{End}(A)$,
- ηK is a K-orbit of $\widehat{\mathbb{Z}}$-linear isomorphisms of $\mathcal{O}$-modules $\widehat{\Lambda}_0 \to \widehat{T}(A)$,

satisfying the following condition of compatibility with (Λ_c, ψ_c):

($\star\star$) There exists an isomorphism of $\mathcal{O}$-modules $a\colon H_1(A, \mathbb{Z}) \to \Lambda_c$, carrying ϕ to ψ_c, carrying ηK to cK, and such that the complex structure induced by a on $V(\mathbb{R})$ belongs to X_+.

If $(A, \phi, \iota, \eta K)$ is a complex abelian variety with PEL structure defined by (Λ_0, c), then condition ($\star\star$) implies that A and (Λ_c, ψ_c) have the same polarization type, and that the Rosati involution on $\mathrm{End}(A) \otimes \mathbb{Q}$ (which is adjunction with respect to ϕ) restricts to $*$ on B. In particular, ι takes values in $\mathrm{End}(A)^\dagger$.

An *isomorphism* between complex abelian varieties with PEL structure $(A, \phi, \iota, \eta K)$ and $(A', \phi', \iota', \eta' K)$ is an isomorphism of complex polarized abelian varieties $f\colon (A, \phi) \to (A,' \phi')$ respecting the action of $\mathcal{O}$ and sending ηK to $\eta' K$.

The difference with the setting of proposition 2.2.8 is that isomorphisms of complex abelian varieties with PEL structure must respect the polarizations exactly, rather than up to an element of $\mu(\Gamma_c)$. In general, $\mu(\Gamma_c) \neq \{1\}$, but there is the following workaround. If $\varepsilon \in F^\times$ lies in

the center of B, then multiplication by ε defines an element in the center of $G(\mathbb{Q})$. Therefore it makes sense to define

$$\mathcal{E}_K = \{\varepsilon \in F^\times : \varepsilon \in K\} = \{\varepsilon \in F^\times : \varepsilon \in \Gamma_c\}, \quad \text{for every } c \in G(\mathbb{A}_f).$$

Proposition 2.2.9. *Let $c \in \mathcal{C}$, and let (Λ_c, ψ_c) be the standard polarized lattice associated with (Λ_0, c). If $\mu(\mathcal{E}_K) = \mu(\Gamma_c)$, then the map*

$$[x, c] \mapsto \big(V(\mathbb{R})/\Lambda_c, \psi_c, \iota, cK\big),$$

where $V(\mathbb{R})$ is seen as a complex vector space via x, and ι is the action of $\mathcal{O}$ on $V(\mathbb{R})/\Lambda_c$ induced by the action of B on $V(\mathbb{R})$, is a bijection between $\mathcal{S}_c$ and the set of isomorphism classes of complex abelian varieties with PEL structure defined by (Λ_0, c).

Proof. When defining $\mathcal{Z}_c$ as in proposition 2.2.8, we can impose $\zeta = 1$ in condition $(\star)$ and strengthen the notion of isomorphism between lattices with PEL structure to respect the polarizations exactly. Indeed, multiplying the isomorphism a of $(\star)$ by $\varepsilon \in \mathcal{E}_K$ leaves everything invariant except the alternating form, which is multiplied by $\mu(\varepsilon)$. The result follows then from the equivalence of categories between lattices and complex abelian varieties outlined above. □

Remark 2.2.10. The group $\mu(\mathcal{E}_K)$ always has finite index in $\mu(\Gamma_c)$. Indeed, if $\mathbb{Z}_{F_0}^\times$ denotes the unit group of F_0, then we have

$$\mu(\mathcal{E}_K) \subset \mu(\Gamma_c) \subset \mathbb{Z}_{F_0}^\times$$

and $\mu(\mathcal{E}_K)$ contains a subgroup of finite index in $\mathbb{Z}_{F_0}^\times$, namely all the squares of elements in $\mathbb{Z}_{F_0}^\times \cap K$. By [Che51, Thm. 1], there exists a compact open subgroup M of $\mu(K)$ such that $\mathbb{Z}_{F_0}^\times \cap M = \mu(\mathcal{E}_K)$. Define $K' = K \cap \mu^{-1}(M)$. Then $\mathcal{E}_{K'} = \mathcal{E}_K$, and for every $c \in G(\mathbb{A}_f)$ we have

$$G(\mathbb{Q})_+ \cap cK'c^{-1} = \{\gamma \in \Gamma_c : \mu(\gamma) \in \mu(\mathcal{E}_K)\}.$$

Therefore the hypothesis of proposition 2.2.9 will be satisfied for the smaller congruence subgroup K'.

Example 2.2.11. Let us detail the modular interpretation of the classical modular curve of level $N \geq 1$ introduced in example 2.2.6, keeping the notation used there. We take $\Lambda_0 = \mathbb{Z}^2$, so that $\mathcal{O} = \mathbb{Z}$, and $K_N \subset \mathrm{GL}_2(\widehat{\mathbb{Z}})$ indeed stabilizes $\widehat{\Lambda}_0$.

We focus on the connected component associated with $c = I_2$. Then the standard polarized lattice defined by (Λ_0, c) is $(\Lambda_c, \psi_c) = (\Lambda_0, \psi)$, and

the polarization ψ on Λ_0 is principal. Since $\Gamma_c = \Gamma(N) \subset \mathrm{SL}_2(\mathbb{Z})$, we have $\mu(\Gamma_c) = \{1\}$, so proposition 2.2.9 applies.

In the definition of a complex abelian variety with PEL structure, we can ignore the action of $\mathcal{O} = \mathbb{Z}$ which is always present. Giving a $\widehat{\mathbb{Z}}$-linear isomorphism $\widehat{\Lambda}_0 \to \widehat{T}$ is the same as specifying a $\mathbb{Z}_p$-basis of $T_p(A)$ for every prime p, and two such isomorphisms differ by an element of K_N if and only if the induced maps $(\mathbb{Z}/N\mathbb{Z})^2 \to A[N]$ are equal. Therefore, a level structure for K_N is simply an isomorphism $\eta\colon (\mathbb{Z}/N\mathbb{Z})^2 \to A[N]$.

Finally, let (A, ϕ) be a principally polarized abelian variety of dimension 1 over $\mathbb{C}$ (i.e. a complex elliptic curve) endowed with an isomorphism $\eta\colon (\mathbb{Z}/N\mathbb{Z})^2 \to A[N]$, and let $\Lambda = H_1(A, \mathbb{Z})$. Let $\overline{u}_1, \overline{u}_2 \in A[N] = \Lambda/N\Lambda$ be the images of $e_1 = (1, 0)$ and $e_2 = (0, 1)$ by η. We claim that condition $(\star\star)$ holds if and only if

$$\phi(\overline{u}_1, \overline{u}_2) = -1 \bmod N \tag{2.8}$$

Since $(\star\star)$ imples $\psi(e_1, e_2) = -1$, eq. (2.8) is indeed satisfied. Conversely, if (2.8) holds, then the isomorphism $a\colon \Lambda \to \Lambda_0$ that we want to construct will be given by a $\mathbb{Z}$-basis (u_1, u_2) of Λ such that $\phi(u_1, u_2) = -1$ and (u_1, u_2) reduces to $\overline{u}_1, \overline{u}_2 \bmod N$. Such a basis exists because ϕ is principal and the reduction map $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ is surjective. The complex structure induced by a on $\Lambda_0 \otimes \mathbb{R}$ belongs to X_+ because X_+ is the set of all complex structures on $\Lambda_0 \otimes \mathbb{R}$ that are positive for ψ, as we computed in §2.2.1.

Therefore, by proposition 2.2.9, the connected component $\mathcal{S}_c = \Gamma(N)\backslash\mathbb{H}_1$ of $\mathrm{Sh}_{K_N}(\mathbb{C})$ classifies isomorphism classes of elliptic curves (A, ϕ) over $\mathbb{C}$ endowed with an isomorphism $\eta\colon (\mathbb{Z}/N\mathbb{Z})^2 \to A[N]$ such that

$$\phi\big(\eta(0, 1), \eta(1, 0)\big) = 1 \bmod N.$$

This is indeed the usual modular interpretation of the classical modular curve of level N.

2.2.4 Modular forms on PEL Shimura varieties

The definition of elliptic modular polynomials given in §1.1.2 uses a particular function, the j-invariant, on the modular curve $X(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}_1$. This invariant has an expression in terms of modular forms: we have

$$j = 1728\frac{E_4^3}{E_4^3 - E_6^2}$$

where E_4, E_6 are the normalized Eisenstein series of weight 4 and 6 on $\mathbb{H}_1$, respectively. The fact that j is an isomorphism between $X(1)$ and $\mathbb{P}^1(\mathbb{C})$

is intimately related with the fact that the graded $\mathbb{C}$-algebra of modular forms for $\mathrm{SL}_2(\mathbb{Z})$ is $\mathbb{C}[E_4, E_6]$.

The notion of modular forms can be generalized to PEL Shimura varieties, and this section attempts to present this generalization without going into technical details. Modular forms can be used to construct coordinate functions: by the fundamental theorem of Baily and Borel [BB66], modular forms on a given PEL Shimura variety realize projective embeddings of its connected components.

Let (G, X_+) be a PEL datum, and let $K_\infty \subset G(\mathbb{R})_+$ be the stabilizer of a fixed complex structure $x_0 \in X_+$. Attached to this data is a certain canonical character of K_∞ [BB66, §1.8], denoted by $\rho\colon K_\infty \to \mathbb{C}^\times$. Let K be a compact open subgroup of $G(\mathbb{A}_f)$. A *modular form* of weight $w \in \mathbb{Z}$ on $\mathrm{Sh}_K(G, X_+)(\mathbb{C})$ is a function

$$f\colon G(\mathbb{Q})_+ \backslash \big(G(\mathbb{R})_+ \times G(\mathbb{A}_f)\big)/K \to \mathbb{C}$$

that satisfies suitable growth and holomorphy conditions [Mil90, Prop. 3.2], and such that

$$\forall x \in G(\mathbb{R})_+,\ \forall g \in G(\mathbb{A}_f),\ \forall k_\infty \in K_\infty,\ f([xk_\infty, g]) = \rho(k_\infty)^w f([x, g]).$$

The weight of f is denoted by $\mathrm{wt}(f)$. We also say that f is *of level* K.

Let $\mathcal{S}$ be a connected component of $\mathrm{Sh}_K(\mathbb{C})$, or a union of these, and let L be its field of definition. A *modular form of weight w on $\mathcal{S}$* is the restriction of a modular form of weight w on $\mathrm{Sh}_K(\mathbb{C})$ to the preimage of $\mathcal{S}$ in $G(\mathbb{Q})_+ \backslash \big(G(\mathbb{R}_+) \times G(\mathbb{A}_f)\big)/K$ by the natural projection. There is a canonical notion of modular forms on $\mathcal{S}$ being defined over L [Mil90, Chap. III]. A *modular function* on $\mathcal{S}$ is the quotient of two modular forms of the same weight, the denominator being nonzero on each connected component of $\mathcal{S}$.

The following result is well-known; as we could not find a precise reference, we present a short proof.

Theorem 2.2.12. *Let $\mathcal{S}$ be a connected component of $\mathrm{Sh}_K(G, X_+)(\mathbb{C})$, and let L be its field of definition. Then the graded L-algebra of modular forms on $\mathcal{S}$ defined over L is finitely generated. Moreover there exists a weight $w \geq 1$ such that modular forms of weight w defined over L realize a projective embedding of $\mathcal{S}$. Every element of the function field $L(\mathcal{S})$ is a quotient of two modular forms of the same weight defined over L.*

Proof. Choose an element $c \in \mathcal{C} \subset G(\mathbb{A}_f)$ defining the connected component $\mathcal{S}$, so that $\mathcal{S} = \Gamma_c \backslash X_+$ where $\Gamma_c = G(\mathbb{Q})_+ \cap cKc^{-1}$. Assume first that the congruence subgroup K of $G(\mathbb{A}_f)$ is sufficiently small, so that Γ_c is torsion free. By the Baily–Borel theorem [BB66, Thm. 10.11], there exists

an ample line bundle $\mathcal{M}_{\mathbb{C}}$ on $\mathcal{S}$ such that for every $w \geq 1$, the algebraic sections of $\mathcal{M}_{\mathbb{C}}^{\otimes w}$ are exactly the modular forms of weight w on $\mathcal{S}$.

In fact, $\mathcal{M}_{\mathbb{C}}$ is the inverse determinant of the tangent bundle on $\mathcal{S}$ [BB66, §7.3]. Since $\mathcal{S}$ has a model over L, there exists a line bundle $\mathcal{M}$ on $\mathcal{S}$ defined over L such that $\mathcal{M} \otimes_L \mathbb{C} = \mathcal{M}_{\mathbb{C}}$. This is a special case of a general result on the rationality of automorphic vector bundles [Mil90, Chap. III, Thm. 4.3]. For every $w \geq 1$, the L-vector space of modular forms of weight w on $\mathcal{S}$ defined over L is $H^0(\mathcal{S}, \mathcal{M}^{\otimes w})$. Since $\mathcal{M} \otimes_L \mathbb{C}$ is ample, $\mathcal{M}$ is ample too, and this implies the conclusions of the theorem.

In general, we can always find a congruence subgroup K' of finite index in K such that the arithmetic subgroups $G(\mathbb{Q})_+ \cap cK'c^{-1}$ for $c \in G(\mathbb{A}_f)$ are torsion free [Mil05, Prop. 3.5], and we can assume that K' is normal in K. Let $\mathcal{S}'$ be a connected component of $\mathrm{Sh}_{K'}(\mathbb{C})$ lying over $\mathcal{S}$, and let L' be its field of definition. Then the conclusions of the theorem hold for $\mathcal{S}'$. We can identify the modular forms on $\mathcal{S}$ defined over L with the modular forms on $\mathcal{S}'$ defined over L' that are invariant under the action of a subgroup of K/K' (we refer to §2.4.3 for the definition of this action). Therefore the conclusions of the theorem also hold for $\mathcal{S}$ by Noether's theorem [Noe15] on invariants under finite groups. □

We can also consider modular forms that are symmetric under certain automorphisms of Sh_K. Let Σ be a finite group of automorphisms of V as a $\mathbb{Q}$-vector space that leaves the symplectic form ψ invariant, and also acts on B in such a way that

$$\forall u \in V,\ \forall b \in B,\ \forall \sigma \in \Sigma,\ \sigma(bu) = \sigma(b)\sigma(u).$$

This implies that the elements of Σ commute with the involution $*$, and hence leave F_0 stable. Under these assumptions, each element $\sigma \in \Sigma$ induces an automorphism of G defined over $\mathbb{Q}$, also denoted by σ. Assume further that these automorphisms leave $G(\mathbb{R})_+$, X_+, K, K_∞, ν and the character ρ invariant. Then Σ can be seen as a finite group of automorphisms of $\mathcal{S}$, and one can check as in [Mil05, Thm. 13.6] that these automorphisms are defined over L. For every modular form f of weight w on $\mathcal{S}$ defined over L, and every $\sigma \in \Sigma$, the function

$$\sigma \cdot f \colon [x, g] \mapsto f\big([\sigma^{-1}(x), \sigma^{-1}(g)]\big)$$

is a modular form of weight w on $\mathcal{S}$ defined over L. We say that f is *symmetric* under Σ if $\sigma \cdot f = f$ for every $\sigma \in \Sigma$.

Proposition 2.2.13. *Let Σ be a finite group of automorphisms of G as above. Then the graded L-algebra of symmetric modular forms on $\mathcal{S}$ defined*

over L is finitely generated, and every symmetric modular function on $\mathcal{S}$ defined over L is the quotient of two symmetric modular forms of the same weight defined over L.

Proof. This results from theorem 2.2.12 and Noether's theorem. □

Example 2.2.14. When $\mathcal{S}$ is the classical modular curve of level $N \geq 1$, we can identify the modular forms on $\mathcal{S}$ with the classical modular forms of level N. Our PEL datum is $(G, X_+) = (\mathrm{GL}_2, \mathbb{H}_1)$, and we fix $i \in \mathbb{H}_1$ as a base point, so that $K_\infty = \mathbb{R}^\times \cdot \mathrm{SO}_2(\mathbb{R})$. Recall that a classical modular form of level N and weight w on $\mathbb{H}_1$ is a holomorphic function $g\colon \mathbb{H}_1 \to \mathbb{C}$ that is holomorphic at infinity and such that for every $\tau \in \mathbb{H}_1$ and every matrix $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(N)$, we have

$$g(\gamma\tau) = (\gamma^*\tau)^w g(\tau)$$

where

$$\gamma\tau := \frac{a\tau + b}{c\tau + d} \quad \text{and} \quad \gamma^*\tau := c\tau + d.$$

We have seen in example 2.2.6 that $\mathcal{S}$ is the connected component of $\mathrm{Sh}_{K_N}(G, X_+)(\mathbb{C})$ associated with the identity matrix $I_2 \in G(\mathbb{A}_f)$. Every element of $G(\mathbb{Q})_+ \backslash G(\mathbb{R})_+ \times G(\mathbb{A}_f)/K$ above $\mathcal{S}$ has a representative of the form $[\gamma, I_2]$ where $\gamma \in G(\mathbb{R})_+$.

Given a classical modular form g of level $\Gamma(N)$, we can check that the function

$$f\colon [\gamma, I_2] \mapsto (\gamma^* i)^{-w} g(\gamma i)$$

is a modular form of weight w on $\mathcal{S}$; here the character $\rho\colon K_\infty \to \mathbb{C}^\times$ is given by

$$\rho\left(\lambda \left(\begin{smallmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{smallmatrix}\right)\right) = \lambda^{-1} e^{i\theta}$$

for every $\lambda \in \mathbb{R}^\times$ and $\theta \in \mathbb{R}$. Conversely, let f be a modular form on $\mathcal{S}$. Let $\tau \in \mathbb{H}_1$, and choose a matrix $\gamma \in G(\mathbb{R})_+$ such that $\tau = \gamma i$. Then the quantity

$$g(\tau) = (\gamma^* i)^w f([\gamma, I_2])$$

does not depend on the choice of γ, and g is a classical modular form of level N. Under this identification, the modular form f is defined over $\mathbb{Q}(\zeta_N)$ if and only if the Fourier coefficients of g belong to this number field; this is a special case of very general results relating different notions of "rationality" for modular forms [Mil05, Chap. III, §2], [Har85, Har86].

2.3 Examples of PEL Shimura varieties

In this section, we present two examples of PEL Shimura varieties playing a central role in the rest of this thesis, namely the Siegel and Hilbert moduli spaces. For correct choices of congruence subgroups and connected components, Siegel moduli spaces classify isomorphism classes of principally polarized abelian varieties over $\mathbb{C}$; and the Hilbert moduli space associated with a totally real number field F of degree g classifies isomorphism classes of principally polarized abelian varieties of dimension g over $\mathbb{C}$ with real multiplication by $\mathbb{Z}_F$. Both types of moduli spaces admit classical descriptions as quotients of hermitian symmetric domains. We link these classical descriptions to their definitions as Shimura varieties as we did in §2.2.1 in the case of modular curves.

In the special case of moduli spaces for abelian varieties of dimension $g = 2$ (i.e. abelian surfaces), explicit generators for the graded algebra of Siegel modular forms are known by works of Igusa [Igu60, Igu79]. Similar results are available for a few examples of Hilbert surfaces.

References for this section are [van08], [Mil05, §6], [BL04, §8] for Siegel moduli spaces, and [van88, Bru08], [BL04, §9] for Hilbert moduli spaces.

2.3.1 Siegel moduli spaces

Let $g \geq 1$ be an integer. We want to build a moduli space for principally polarized abelian varieties of dimension g over $\mathbb{C}$. We start from the vector space $V = \mathbb{Q}^{2g}$ endowed with the symplectic form ψ defined by

$$\forall u, v \in V,\ \psi(u,v) = u^t \begin{pmatrix} 0 & -I_g \\ I_g & 0 \end{pmatrix} v.$$

The symplectic form ψ is a principal polarization on the lattice $\Lambda_0 = \mathbb{Z}^{2g}$. The endomorphism algebra is chosen to be $B = \mathbb{Q}$, with $*$ being the trivial involution; $(B, *)$ is simple of type (C). Finally, the complex structure

$$x_0 = \begin{pmatrix} 0 & I_g \\ -I_g & 0 \end{pmatrix}$$

is positive for ψ.

The algebraic group associated with this data is the general symplectic group $G = \mathrm{GSp}_{2g}$. Let $\mu \colon G \to \mathbb{G}_m$ be the morphism such that

$$\forall g \in \mathrm{GSp}_{2g}(\mathbb{Q}),\ \forall u, v \in V,\ \psi(gu, gv) = \mu(g)\psi(u,v).$$

The symplectic group Sp_{2g} is, by definition, the kernel of μ. Since $\mathrm{Sp}_{2g}(\mathbb{R})$ is connected [Mil05, Thm. 5.2 and Prop. 8.7], the group $\mathrm{GSp}_{2g}(\mathbb{R})_+$ consists

of all elements $g \in \mathrm{GSp}_{2g}(\mathbb{R})$ such that $\mu(g) > 0$. The orbit X_+ of x_0 under the action of $\mathrm{GSp}_{2g}(\mathbb{R})$ by conjugation consists of all complex structures on $V(\mathbb{R})$ that are positive for ψ [Mil05, §6]. Finally, one can show that the reflex field of Siegel spaces is $\mathbb{Q}$ [Mil05, §14]. When $g = 1$, we recover the PEL data from §2.2.1 associated with classical modular curves.

The classical description of Siegel spaces. Reasoning as in the case of elliptic curves (§2.2.1), we attach to each complex structure $x \in X_+$ the unique matrix $\tau(x) \in \mathrm{Mat}_{g\times g}(\mathbb{C})$ with invertible imaginary part satisfying the following property: under the isomorphism $\eta_x \colon V(\mathbb{R}) \to \mathbb{C}^g$ whose matrix in the canonical bases is $(I_g \ -\tau(x))$, the pullback of the natural complex structure on $\mathbb{C}^g$ is x. In particular, the matrix attached to x_0 is $\tau(x_0) = iI_g$.

Lemma 2.3.1. *For each $x \in X_+$, the matrix $\tau(x)$ is symmetric and its imaginary part is positive definite; in other words $\tau(x)$ belongs to the Siegel upper half space $\mathbb{H}_g$.*

Proof. Let $\eta_{x,*}\psi$ denote the pushforward of ψ under η_x; it is an $\mathbb{R}$-bilinear form on $\mathbb{C}^g$. Let $(e_1, \dots, e_g)$ be the canonical basis of $\mathbb{C}^g$. Then the matrix of $\eta_{x,*}\psi$ in the $\mathbb{R}$-basis of $\mathbb{C}^g$ given by $(e_1, \dots, e_g, ie_1, \dots, ie_g)$ is (in $g \times g$ blocks)

$$\begin{pmatrix} 0 & \mathrm{Im}(\tau)^{-1} \\ \mathrm{Im}(\tau)^{-t} & -\mathrm{Im}(\tau)^{-t}(\mathrm{Re}(\tau) - \mathrm{Re}(\tau)^t)\,\mathrm{Im}(\tau)^{-1} \end{pmatrix}.$$

The relation $\eta_{x,*}\psi(iu, iv) = \eta_{x,*}\psi(u, v)$ implies that $\mathrm{Re}(\tau)$ is symmetric, and the fact that the bilinear form $\psi(u, iv)$ is symmetric and positive definite implies that $\mathrm{Im}(\tau)$ is symmetric and positive definite. □

Lemma 2.3.2. *Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GSp}_{2g}(\mathbb{R})_+$ (written in $g \times g$ blocks), and let $x \in X_+$. Then we have*

$$\tau(\gamma x) = (a\tau(x) + b)(c\tau(x) + d)^{-1}.$$

Proof. We have a commutative diagram

$$\begin{array}{ccccc} V(\mathbb{R}) & \xrightarrow{\gamma^{-1}} & V(\mathbb{R}) & \xrightarrow{\eta_x} & \mathbb{C}^g \\ \downarrow{\scriptstyle \gamma x} & & \downarrow{\scriptstyle x} & & \downarrow{\scriptstyle i} \\ V(\mathbb{R}) & \xrightarrow{\gamma^{-1}} & V(\mathbb{R}) & \xrightarrow{\eta_x} & \mathbb{C}^g. \end{array}$$

Therefore, γx is the pullback of the multiplication by i under the isomorphism $V(\mathbb{R}) \to \mathbb{C}^g$ whose matrix is $(I_g \ -\tau(x))\,\gamma^{-1}$. Since $\gamma \in \mathrm{GSp}_{2g}(\mathbb{R})$,

this matrix is, up to a scalar in $\mathbb{R}^\times$, equal to

$$(I_g \ -\tau(x)) \begin{pmatrix} d^t & -c^t \\ -b^t & a^t \end{pmatrix} = (\tau(x)^t c^t + d^t, -\tau(x)^t a^t - b^t). \tag{2.9}$$

By lemma 2.3.1, we know that $\tau(x) \in \mathbb{H}_g$. By [BL04, Prop. 8.2.2] the matrix $c\tau(x)+d$ is invertible, and moreover $(a\tau(x)+b)(c\tau(x)+d)^{-1} \in \mathbb{H}_g$ is also symmetric. Starting from eq. (2.9), we act on $\mathbb{C}^g$ by $(\tau(x)^t c^t + d^t)^{-1}$, and we obtain:

$$\forall \gamma \in \mathrm{GSp}_{2g}(\mathbb{R})_+, \ \tau(\gamma x) = (a\tau(x)+b)(c\tau(x)+d)^{-1}. \qquad \square$$

The modular interpretation of Siegel spaces. The maximal compact open subgroup of $G(\mathbb{A}_f)$ leaving the lattice $\Lambda_0 \otimes \widehat{\mathbb{Z}}$ stable is $K_0 = \mathrm{GSp}_{2g}(\widehat{\mathbb{Z}})$. The PEL Shimura variety $\mathrm{Sh}_{K_0}(\mathbb{C})$ has only one connected component identified with $\Gamma \backslash \mathbb{H}_g$, where

$$\Gamma = \mathrm{GSp}_{2g}(\mathbb{Q})_+ \cap K_0 = \mathrm{Sp}_{2g}(\mathbb{Z}).$$

The action of Γ on $\mathbb{H}_g$ is given by the formula from lemma 2.3.2. By proposition 2.2.9, the quotient $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g$ is a moduli space for principally polarized abelian varieties of dimension g over $\mathbb{C}$. This modular interpretation coincides with the classical one [BL04, §8.1].

More generally, choose positive integers $D_1 | \cdots | D_g$ such that $D_1 = 1$, and let $\Lambda \subset V(\mathbb{Q})$ be the lattice generated by the vectors

$$e_1, \ldots, e_g, D_1 e_{g+1}, \ldots, D_g e_{2g},$$

where $e_1, \ldots, e_{2g}$ denote the canonical basis of $\mathbb{Q}^{2g}$. Then the polarization ψ has type $(D_1, \ldots, D_g)$ on Λ. Let K be a compact open subgroup of $G(\mathbb{A}_f)$ that stabilizes $\Lambda \otimes \widehat{\mathbb{Z}}$, and let $\mathcal{S}$ denote the connected component of $\mathrm{Sh}_K(\mathbb{C})$ defined by the identity matrix in $G(\mathbb{A}_f)$. Then $\mathcal{S}$ is identified with $\Gamma \backslash \mathbb{H}_g$, where

$$\Gamma = \mathrm{GSp}_{2g}(\mathbb{Q})_+ \cap K = \mathrm{Sp}_{2g}(\mathbb{Q}) \cap K.$$

By proposition 2.2.9, the variety $\mathcal{S}$ is a moduli space for polarized abelian varieties A of dimension g over $\mathbb{C}$ with polarization type $(D_1, \ldots, D_g)$ and level K structure, such that $H_1(A, \mathbb{Z})$ is isomorphic to the standard polarized lattice (Λ_0, ψ) with its natural level K structure.

2.3.2 Siegel modular forms in dimension 2

Given the classical description of Siegel spaces from §2.3.1, modular forms on Siegel spaces can be identified with certain holomorphic functions

on the Siegel half space $\mathbb{H}_g$. This identification is the higher-dimensional analogue of example 2.2.14.

Let us focus on the case of the modular group $\mathrm{Sp}_{2g}(\mathbb{Z})$, where $g \geq 2$ is an integer. For every $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GSp}_{2g}(\mathbb{R})_+$ and every $\tau \in \mathbb{H}_g$, we write

$$\gamma\tau = (a\tau + b)(c\tau + d)^{-1} \quad \text{and} \quad \gamma^*\tau = c\tau + d.$$

Let W be a finite-dimensional vector space over $\mathbb{C}$, and let

$$\rho\colon \mathrm{GL}_g(\mathbb{C}) \to \mathrm{GL}(W)$$

be an irreducible holomorphic representation of $\mathrm{GL}_g(\mathbb{C})$ on W. A *Siegel modular form* of dimension g and weight ρ is a holomorphic function $f\colon \mathbb{H}_g \to W$ satisfying the transformation rule

$$\forall \gamma \in \mathrm{Sp}_{2g}(\mathbb{Z}),\ \forall \tau \in \mathbb{H}_g, \quad f(\gamma\tau) = \rho(\gamma^*\tau) f(\tau).$$

We say that f is *scalar-valued* if $\dim W = 1$, and *vector-valued* otherwise. A *Siegel modular function* is only required to be meromorphic instead of holomorphic. Since $g \geq 2$, there is no need to enforce a holomorphy condition at the cusps: Koecher's principle [van08, Thm. 4.4] asserts that it is automatically satisfied.

If $w \geq 0$ is an integer, $W = \mathbb{C}$, and $\rho = \det^w$, then the Siegel modular forms of weight ρ can be identified with the modular forms on $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g$ of weight w in the sense of §2.2.4. We also call these functions (scalar-valued) Siegel modular forms of weight w. More precisely, we fix $\tau_0 = iI_g \in \mathbb{H}_g$ as a base point. Then, if h is a classical Siegel modular form of weight w for $\mathrm{Sp}_{2g}(\mathbb{Z})$, the function

$$f\colon [\gamma, I_{2g}] \mapsto (\gamma^*\tau_0)^{-w} h(\gamma\tau_0)$$

is a modular form in the sense of Shimura varieties. Conversely, given a modular form f of weight w on $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g$ in the sense of §2.2.4, the function $h\colon \mathbb{H}_g \to \mathbb{C}$ given by

$$\forall \gamma \in \mathrm{GSp}_{2g}(\mathbb{R})_+,\ h(\gamma\tau_0) = (\gamma^*\tau_0)^w f([\gamma, I_{2g}])$$

is well defined, and is a classical Siegel modular form of weight w for $\mathrm{Sp}_{2g}(\mathbb{Z})$.

By theorem 2.2.12, scalar-valued Siegel modular forms realize a projective embedding of the quotient $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g$.

Remark 2.3.3. We have seen in the proof of theorem 2.2.12 that modular forms on a PEL Shimura variety $\mathcal{S}$, at least when the congruence subgroup is small enough, have a geometric interpretation as the algebraic sections of

a certain line bundle, namely the inverse determinant of the tangent bundle of $\mathcal{S}$. Siegel modular forms, scalar-valued or not, also have a geometric interpretation as sections of vector bundles constructed from the *Hodge bundle* $\mathbb{E}$ on $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g$ [van08, §10 and §13] (strictly speaking, $\mathbb{E}$ should be considered as a vector bundle on $\Gamma\backslash\mathbb{H}_g$ for some suitable normal subgroup Γ of $\mathrm{Sp}_{2g}(\mathbb{Z})$ of finite index, endowed with an action of $\mathrm{Sp}_{2g}(\mathbb{Z})/\Gamma$). The fiber of $\mathbb{E}$ at $\tau \in \mathbb{H}_g$ is identified with the vector space $\Omega^1(A(\tau))$ of global differential forms on $A(\tau)$, where $A(\tau) = \mathbb{C}^g/(\mathbb{Z}^g \oplus \tau\mathbb{Z}^g)$ denotes the principally polarized abelian variety of dimension g attached to τ in the modular interpretation.

Concretely, if A is a principally polarized abelian variety of dimension g over $\mathbb{C}$, if ω is a basis of $\Omega^1(A)$, and if f is a Siegel modular form of weight ρ on $\mathbb{H}_g$, then the quantity $f(A, \omega)$ makes sense. This fact will play an important role in chapter 3.

A convenient way to manipulate Siegel modular forms is to write down Fourier expansions [van08, §4]. Let f be a Siegel modular form on $\mathbb{H}_g$ of any weight ρ, with underlying vector space W. If $s \in \mathrm{Mat}_{g\times g}(\mathbb{Z})$ is symmetric, then $f(\tau + s) = f(\tau)$ for every $\tau \in \mathbb{H}_g$. Hence $f(\tau)$ has a Fourier expansion in terms of the quantities $\exp(2\pi i \tau_{k,l})$ where $\tau_{k,l}$ is an entry of the matrix τ. This Fourier expansion is often written in the form

$$f(\tau) = \sum_{\substack{n \in \mathrm{Mat}_{g\times g}(\mathbb{Q}) \\ \text{symmetric}}} \exp\big(2\pi i \,\mathrm{Tr}(n\tau)\big)\, a_n(f).$$

The Fourier coefficient $a_n(f) \in W$ can be nonzero only when n is half-integral (meaning that the coefficients of n lie in $\frac{1}{2}\mathbb{Z}$, and its diagonal coefficients lie in $\mathbb{Z}$), and moreover n is positive as a symmetric matrix.

In the special case $W = \mathbb{C}$, the modular form f is defined over $\mathbb{Q}$ if and only if all its Fourier coefficients $a_n(f)$ are rational numbers; see for instance [Har86, Thm. 6.4]. We say that f is *defined over* $\mathbb{Z}$ if its Fourier coefficients are all integers.

In the special case $g = 2$, we write for $\tau \in \mathbb{H}_2$:

$$\tau = \begin{pmatrix} z_1(\tau) & z_3(\tau) \\ z_3(\tau) & z_2(\tau) \end{pmatrix}, \quad \text{and} \quad q_j = \exp\big(2\pi i z_j(\tau)\big) \quad \text{for } 1 \le j \le 3.$$

The terms appearing in Fourier expansions of Siegel modular forms in dimension 2 are of the form $a\, q_1^{n_1} q_2^{n_2} q_3^{n_3}$ with $a \in \mathbb{C}$, $n_1 \ge 0$, $n_2 \ge 0$, and $n_3^2 \le 4n_1n_2$ (note that n_3 can still be negative). Hence the Fourier expansion of a Siegel modular form in dimension 2 is an element of the power series ring $\mathbb{C}[q_3, q_3^{-1}][[q_1, q_2]]$.

Explicit generators for $g = 2$. We start by defining the *theta constants* on $\mathbb{H}_g$ for any $g \geq 1$ [van08, §7]. Given $a, b \in \{0,1\}^g$, we write

$$\theta_{a,b}(\tau) = \sum_{m \in \mathbb{Z}^g} \exp\left(\pi i \left(\left(m + \tfrac{a}{2}\right)^t \tau \left(m + \tfrac{a}{2}\right) + \left(m + \tfrac{a}{2}\right)^t \tfrac{b}{2}\right)\right). \tag{2.10}$$

The pair (a, b) is called the *theta characteristic* of the theta constant $\theta_{a,b}$. Theta constants are holomorphic functions on $\mathbb{H}_g$, and are identically zero unless they are *even*, i.e. $a^t b \in \mathbb{Z}$ is even.

Even theta constants are not modular forms for the full Siegel modular group $\mathrm{Sp}_{2g}(\mathbb{Z})$, but they satisfy certain transformation properties under $\mathrm{Sp}_{2g}(\mathbb{Z})$ which make them convenient building blocks. Every scalar-valued Siegel modular function of any weight on $\mathbb{H}_g$ has an expression in terms of quotients of theta constants [Igu72, Thm. 9 p. 222]. Moreover, for each $1 \leq g \leq 3$, the stronger result that every scalar-valued Siegel modular form is a polynomial in the theta constants holds [Igu64, Igu66, FSM19]. In the case $g = 1$, we recover the "genus 1" theta constants θ_0 and θ_1 defined in §1.1.1 as the theta constants with characteristics $(0,0)$ and $(0,1)$.

In the case $g = 2$, we introduce the following indexation of theta constants [Dup06, §6.2]: write $a = (a_1, a_2) \in \{0,1\}^2$ and $b = (b_1, b_2) \in \{0,1\}^2$. Then we define

$$\theta_{8a_2+4a_1+2b_2+b_1} := \theta_{a,b}.$$

There are ten even genus 2 theta constants, denoted with this convention by $\tau \mapsto \theta_j(\tau)$ for $j \in \{0,1,2,3,4,6,8,9,12,15\}$. Let us present the construction of explicit generators of the graded algebra of scalar-valued Siegel modular forms for $g = 2$ in terms of theta constants.

Definition 2.3.4 ([Str10, §II.7.1])**.** Let $T \subset \{0,1\}^4$ be the set of ten even theta characteristics for $g = 2$.

- A *Göpel quadruple* is a subset $C \subset T$ of four characteristics such that $\sum_{c \in C} c \in 2\mathbb{Z}^4$; there are 15 Göpel quadruples.
- A *syzygous triple* is a set of three theta characteristics $\{b, c, d\} \subset T$ that is a subset of a Göpel quadruple; there are 60 syzygous triples.
- If $\{b, c, d\}$ is a syzygous triple, we write $b = (b_1, b_2, b_3, b_4)$ and similarly for c, d; then we define $\varepsilon(b, c, d) \in \{\pm 1\}$ as -1 to the power

$$\begin{aligned} &b_1 + b_2 + c_1 + c_2 + d_1 + d_2 + b_1c_1 + b_2c_2 + b_4c_2 + b_1c_3 - b_2c_4 \\ &+ b_1d_1 - b_3d_1 + c_1d_1 + b_2d_2 + c_2d_2 + c_4d_2 + c_1d_3 - b_2b_3c_1 \\ &- b_2b_4c_2 - b_1b_2c_3 - b_2b_3d_1 - b_3c_1d_1 - b_1c_3d_1 - b_2c_3d_1 - b_2b_4d_2 \\ &- b_4c_2d_2 - b_1b_2d_3 - b_1c_1d_3 - b_2c_1d_3. \end{aligned}$$

Finally, we define the holomorphic functions h_j on $\mathbb{H}_2$ for $j \in \{4, 6, 10, 12\}$ as follows:

$$\begin{aligned} h_4 &= \sum_{c \in T} \theta_c^8, \\ h_6 &= \sum_{\substack{\{b,c,d\} \\ \text{syzygous}}} \varepsilon(b,c,d)(\theta_b \theta_c \theta_d)^4, \\ h_{10} &= \prod_{c \in T} \theta_c^2, \qquad \text{and} \\ h_{12} &= \sum_{\substack{C \text{ Göpel} \\ \text{quadruple}}} \prod_{c \in T \setminus C} \theta_c^4. \end{aligned}$$

For each $j \in \{4, 6, 10, 12\}$, the function h_j is a Siegel modular form of weight j for $\mathrm{Sp}_4(\mathbb{Z})$, and is defined over $\mathbb{Q}$.

Theorem 2.3.5 ([Igu60])**.** *The modular forms h_4, h_6, h_{10}, h_{12} are algebraically independent, and generate the graded $\mathbb{Q}$-algebra of scalar-valued Siegel modular forms of even weight in dimension 2.*

We refer to [Bol87] for the correspondence between Igusa's language in [Igu60] and the modular functions h_j from definition 2.3.4.

Remark 2.3.6. Other normalizations of these generators are present in the literature. For instance, in order to normalize Fourier expansions, one defines

$$\psi_4 := 2^{-2} h_4, \ \psi_6 := 2^{-2} h_6, \ \chi_{10} := -2^{-12} h_{10}, \text{ and } \chi_{12} := 2^{-15} h_{12}. \quad (2.11)$$

We warn the reader that different definitions appear in the literature: our χ_{10} is -4 times the modular form χ_{10} appearing in Igusa's paper [Igu79], and our χ_{12} is 12 times Igusa's χ_{12}. The Fourier expansions of the modular forms $\psi_4, \psi_6, \chi_{10}$ and χ_{12} are

$$\begin{aligned} \psi_4(\tau) &= 1 + 240(q_1 + q_2) \\ &\quad + \big(240 q_3^2 + 13440 q_3 + 30240 + 13340 q_3^{-1} + 240 q_3^{-2}\big) q_1 q_2 \\ &\quad + O\big(q_1^2, q_2^2\big), \\ \psi_6(\tau) &= 1 - 504(q_1 + q_2) \\ &\quad + \big(-504 q_3^2 + 44352 q_3 + 166320 + 44352 q_3^{-1} - 504 q_3^{-2}\big) q_1 q_2 \\ &\quad + O\big(q_1^2, q_2^2\big), \\ \chi_{10}(\tau) &= \big(q_3 - 2 + q_3^{-1}\big) q_1 q_2 + O(q_1^2, q_2^2), \qquad \text{and} \\ \chi_{12}(\tau) &= \big(q_3 + 10 + q_3^{-1}\big) q_1 q_2 + O\big(q_1^2, q_2^2\big). \end{aligned}$$

The absence of constant terms in the Fourier expansions of χ_{10} and χ_{12} indicates that they are *cusp forms*, hence the different notation.

The Siegel modular form χ_{10} satisfies the following important property: the abelian surface $A(\tau) = \mathbb{C}^2/(\mathbb{Z}^2 \oplus \tau\mathbb{Z}^2)$ attached to $\tau \in \mathbb{H}_2$ is the product of two elliptic curves over $\mathbb{C}$ if and only if $\chi_{10}(\tau) = 0$ [Igu62]. Otherwise, this abelian surface is isomorphic to the Jacobian of a genus 2 hyperelliptic curve over $\mathbb{C}$.

The *Igusa invariants*, defined in [Str10, §II.2.1] as

$$\begin{aligned} j_1 &:= \frac{h_4 h_6}{h_{10}} = -2^{-8}\frac{\psi_4\psi_6}{\chi_{10}}, \\ j_2 &:= \frac{h_{12}h_4^2}{h_{10}^2} = 2^{-5}\frac{\psi_4^2\chi_{12}}{\chi_{10}^2}, \qquad \text{and} \\ j_3 &:= \frac{h_4^5}{h_{10}^2} = 2^{-14}\frac{\psi_4^5}{\chi_{10}^2} \end{aligned}$$

are Siegel modular functions of weight zero. By theorems 2.2.12 and 2.3.5, they realize a birational morphism defined over $\mathbb{Q}$ from the Siegel threefold $\mathrm{Sp}_4(\mathbb{Z})\backslash\mathbb{H}_2$, viewed as an algebraic variety, to the projective space $\mathbb{P}^3$. Igusa invariants are well-defined for all Jacobians of genus 2 curves over $\mathbb{C}$.

A stronger result than theorem 2.3.5 is to describe the structure of the graded ring of Siegel modular forms in dimension 2 that are defined over $\mathbb{Z}$. This is also done by Igusa [Igu79], who gives a set of fourteen generators, the first ones being $\psi_4, \psi_6, \chi_{10}$ and χ_{12}. We will not need the explicit list in this thesis, only the following consequence of Igusa's result.

Proposition 2.3.7. *Let f be a scalar-valued Siegel modular form of even weight k defined over $\mathbb{Z}$. Then $12^k f \in \mathbb{Z}[h_4, h_6, h_{10}, h_{12}]$.*

Vector-valued Siegel modular forms in dimension 2. In order to describe vector-valued modular forms explicitly in dimension 2, we recall the usual classification of irreducible representations of $\mathrm{GL}_2(\mathbb{C})$.

Let $n \geq 0$ be an integer. We denote by Sym^n the n-th symmetric power of the standard representation of $\mathrm{GL}_2(\mathbb{C})$ on $\mathbb{C}^2$. Explicitly, Sym^n is a representation on the vector space $\mathbb{C}_n[x]$ of polynomials of degree at most n, with

$$\mathrm{Sym}^n\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right)P(x) = (bx+d)^n P\left(\frac{ax+c}{bx+d}\right).$$

We take $(x^n, \dots, x, 1)$ as the standard basis of $\mathbb{C}_n[x]$. Then we can write endomorphisms of $\mathbb{C}_n[x]$ as matrices, and we have

$$\operatorname{Sym}^2\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right) = \begin{pmatrix} a^2 & ab & b^2 \\ 2ac & ad+bc & 2bd \\ c^2 & cd & d^2 \end{pmatrix}.$$

Proposition 2.3.8. *The irreducible finite-dimensional holomorphic representations of the group* $\mathrm{GL}_2(\mathbb{C})$ *are exactly the representations* $\det^k \operatorname{Sym}^n$*, for* $k \in \mathbb{Z}$ *and* $n \in \mathbb{N}$*.*

Proof. Since $\mathrm{SL}_2(\mathbb{C})$ is a simply connected Lie group, there is an equivalence between holomorphic finite-dimensional representations of $\mathrm{SL}_2(\mathbb{C})$ and representations of its Lie algebra $\mathfrak{sl}_2(\mathbb{C})$ [Bou72, Ch. III, §6.1, Th. 1]. By [Bou75, Ch. VIII, §1.3, Th. 1], irreducible representations of $\mathfrak{sl}_2(\mathbb{C})$ are classified by their higher weight; on the Lie group side, this shows that the holomorphic finite-dimensional irreducible representations of $\mathrm{SL}_2(\mathbb{C})$ are exactly the representations Sym^n for $n \in \mathbb{N}$. The case of $\mathrm{GL}_2(\mathbb{C})$ follows easily. □

Writing Sym^n as a representation on $\mathbb{C}_n[x]$ allows us to multiply Siegel modular forms; hence, they naturally generate a graded $\mathbb{C}$-algebra. In contrast with the case of scalar-valued modular forms, this graded algebra is not finitely generated [van08, §25].

We conclude this section with two examples of vector-valued Siegel modular forms in dimension 2. The first example is given by the derivatives of modular functions of weight zero, for instance Igusa invariants.

Proposition 2.3.9. *Let* f *be a Siegel modular function on* $\mathbb{H}_2$ *of weight zero. Then the meromorphic function*

$$\frac{df}{d\tau} := \frac{1}{2\pi i}\left(\frac{\partial f}{\partial z_1}x^2 + \frac{\partial f}{\partial z_3}x + \frac{\partial f}{\partial z_2}\right)$$

is a Siegel modular function of weight Sym^2*.*

This result can be proved by differentiating the relation $f(\gamma\tau) = f(\tau)$ with respect to τ, for every $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$; but the proof is not very enlighening. We refer to chapter 3 for a geometric interpretation of proposition 2.3.9, which is fundamental in the isogeny algorithm.

The second example is that of a Siegel modular form of weight $\det^8 \operatorname{Sym}^6$ which will also play a fundamental role in chapter 3.

Example 2.3.10. Following Ibukiyama [Ibu12], let $E_8 \subset \mathbb{R}^8$ denote the lattice of half-integer vectors $v = (v_1, \dots, v_8)$ subject to the conditions

$$\sum_{k=1}^{8} v_k \in 2\mathbb{Z} \quad \text{and} \quad \forall\, 1 \le k, l \le 8,\ v_k - v_l \in \mathbb{Z}.$$

Set $a = (2, 1, i, i, i, i, i, 0)$ and $b = (1, -1, i, i, 1, -1, -i, i) \in \mathbb{C}^8$. Using the notation

$$\langle v, w \rangle := \sum_{k=1}^{8} v_k w_k,$$

we write for $0 \le j \le 6$ and $\tau \in \mathbb{H}_2$:

$$\Theta_j(\tau) := \sum_{v,v' \in E_8} \langle v, a\rangle^j \cdot \langle v', a\rangle^{6-j} \cdot \begin{vmatrix} \langle v, a\rangle & \langle v', a\rangle \\ \langle v, b\rangle & \langle v', b\rangle \end{vmatrix}^4 \cdot \exp\Big(\pi i\big(\langle v, v\rangle\, z_1(\tau) + 2\,\langle v, v'\rangle\, z_3(\tau) + \langle v', v'\rangle\, z_2(\tau)\big)\Big).$$

Finally, we define

$$f_{8,6}(\tau) := \frac{1}{111456000} \sum_{j=0}^{6} \tbinom{6}{j} \Theta_j(\tau)\, x^j.$$

Then $f_{8,6}$ is a nonzero Siegel modular form of weight $\det^8 \mathrm{Sym}^6$. This definition provides an explicit, but slow, method to compute the first coefficients of the q-expansion; using the expression of $f_{8,6}$ in terms of theta series as in [CFv17] would be faster. We have

$$\begin{aligned} f_{8,6}(\tau) = \quad & \big((4q_3^2 - 16q_3 + 24 - 16q_3^{-1} + 4q_3^{-2})q_1^2 q_2 + \cdots\big)\, x^6 \\ & + \big((12q_3^2 - 24q_3 + 24q_3^{-1} - 12q_3^{-2})q_1^2 q_2 + \cdots\big)\, x^5 \\ & + \big((-q_3 + 2 - q_3^{-1})q_1 q_2 + \cdots\big)\, x^4 \\ & + \big((-2q_3 + 2q_3^{-1})q_1 q_2 + \cdots\big)\, x^3 \\ & + \big((-q_3 + 2 - q_3^{-1})q_1 q_2 + \cdots\big)\, x^2 \\ & + \big((12q_3^2 - 24q_3 + 24q_3^{-1} - 12q_3^{-2})q_1 q_2^2 + \cdots\big)\, x \\ & + \big((4q_3^2 - 16q_3 + 24 - 16q_3^{-1} + 4q_3^{-2})q_1 q_2^2 + \cdots\big)\,. \end{aligned}$$

2.3.3 Hilbert moduli spaces

Let F be a fixed totally real number field of degree g over $\mathbb{Q}$, and let $\mathcal{O}$ be an order in the ring of integers $\mathbb{Z}_F$. Recall that a principally polarized

abelian variety A of dimension g has *real multiplication by* $\mathcal{O}$ if A is endowed with an embedding $\eta\colon \mathcal{O} \to \mathrm{End}(A)^\dagger$.

Our goal is to construct a moduli space of principally polarized abelian varieties of dimension g with real multiplication by $\mathbb{Z}_F$. Let ∂_F^{-1} denote the inverse different of F, i.e. the dual of $\mathbb{Z}_F$ for the trace form:

$$\partial_F^{-1} = \left\{x \in F\colon \forall y \in \mathbb{Z}_F,\ \mathrm{Tr}_{F/\mathbb{Q}}(xy) \in \mathbb{Z}\right\}.$$

Consider the lattice

$$\Lambda_0 = \mathbb{Z}_F \oplus \partial_F^{-1} \subset F^2.$$

The symplectic form ψ on F^2 defined by

$$\forall a, b, c, d \in F,\ \psi\big((a,b),(c,d)\big) = \mathrm{Tr}_{F/\mathbb{Q}}(ad - bc)$$

is a principal polarization on Λ_0; the fractional ideal ∂_F^{-1} was chosen to ensure precisely this. We let $B = F$, with $*$ the trivial involution (which is indeed positive since F is totally real), act on F^2 by multiplication. The $\mathbb{Q}$-algebra $(B, *)$ is simple of type (C). Let $V = F^2$; then (V, ψ) is a symplectic $(B,*)$-module, and the order of B fixing Λ_0 is precisely $\mathbb{Z}_F$.

The algebraic group associated with this data is $G = \mathrm{GL}_2(F)$. The g real embeddings $\sigma_1, \dots, \sigma_g$ of F induce an identification

$$G(\mathbb{R}) = \prod_{i=1}^{g} \mathrm{GL}_2(\mathbb{R}),$$

and the subgroup $G(\mathbb{R})_+$ consists of matrices with totally positive determinant; in other words $G(\mathbb{R})_+ = \prod_{i=1}^g \mathrm{GL}_2(\mathbb{R})_+$.

Finally, we consider the complex structure $x_0 \in G(\mathbb{R})$ on $V(\mathbb{R}) \simeq (\mathbb{R}^2)^g$ given by

$$x_0 = \left(\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\right)_{1 \le i \le g},$$

which is positive for ψ. Let X_+ be the $G(\mathbb{R})_+$-conjugacy class of x_0. The pair (G, X_+) is called a *Hilbert Shimura datum*. Its reflex field is $\mathbb{Q}$: see [van88, §X.4] in the case $g = 2$, and [Mil05, Ex. 12.4] in general.

The classical description of Hilbert moduli spaces. Given the description of $G(\mathbb{R})_+$ above, and the identification of the orbit of $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ under $\mathrm{GL}_2(\mathbb{R})_+$ with the upper half plane $\mathbb{H}_1$ (recall §2.2.1), the domain X_+ can be identified with $\mathbb{H}_1^g$ endowed with the action of $\mathrm{GL}_2(\mathbb{R})_+$ on each coordinate. Explicitly, the action of $\mathrm{GL}_2(F)$ on $\mathbb{H}_1^g$ is the following: for every $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(F)$ and $(\tau_1, \dots, \tau_g) \in \mathbb{H}_1^g$, we have

$$\gamma \cdot (\tau_1, \dots, \tau_g) = \left(\frac{\sigma_j(a)\tau_j + \sigma_j(b)}{\sigma_j(c)\tau_j + \sigma_j(d)}\right)_{1 \le j \le g}.$$

In the notation of §2.2.2, the morphism μ in the case of the algebraic group $\mathrm{GL}_2(F)$ is the determinant. As in remark 2.2.10, we can make a choice of compact open subgroup $K \subset G(\mathbb{A}_f)$ such that $G(\mathbb{Q})_+ \cap K$ contains exactly those matrices $\gamma \in \mathrm{GL}(\mathbb{Z}_F \oplus \partial_F^{-1})$ (i.e. automorphisms of Λ_0, whose determinant is a priori a unit of $\mathbb{Z}_F$) whose determinant is a square. Let $\mathcal{S}$ be the connected component of $\mathrm{Sh}_K(G, X_+)(\mathbb{C})$ associated with the identity matrix in $G(\mathbb{A}_f)$. Then we have a natural bijection

$$\mathcal{S} \simeq \mathrm{SL}(\mathbb{Z}_F \oplus \partial_F^{-1}) \backslash \mathbb{H}_1^g.$$

The group $\Gamma(1)_F := \mathrm{SL}(\mathbb{Z}_F \oplus \partial_F^{-1})$ is called the *Hilbert modular group*. By proposition 2.2.9, the connected component $\mathcal{S}$ classifies isomorphism classes of principally polarized abelian varieties A of dimension g over $\mathbb{C}$, with real multiplication by $\mathbb{Z}_F$, satisfying the compatibility condition $(\star\star)$. In this case, the compatibility condition is the following: there exists an isomorphism of $\mathbb{Z}_F$-modules $H_1(A, \mathbb{Z}) \to \Lambda_0$ carrying the principal polarization to ψ [van88, §IX.1].

The Hilbert embedding. Given the modular interpretation of Hilbert moduli spaces, there should exist a forgetful map, called the *Hilbert embedding*, from $\mathrm{SL}(\mathbb{Z}_F \oplus \partial_F^{-1}) \backslash \mathbb{H}_1^g$ to the Siegel moduli space $\mathrm{Sp}_{2g}(\mathbb{Z}) \backslash \mathbb{H}_g$. This map is constructed by choosing an isomorphism between the lattices $\mathbb{Z}_F \oplus \delta_F^{-1} \subset F^2$ and $\mathbb{Z}^{2g} \subset \mathbb{Q}^{2g}$ that respects the symplectic forms, or in other words a symplectic basis of $\mathbb{Z}_F \oplus \partial_F^{-1}$. The map $\mathbb{H}_1^g \to \mathbb{H}_g$ between the sets of acceptable complex structures depends on this choice of basis, but the induced map on the moduli spaces does not.

The concrete matrix calculations are as follows. Choose a $\mathbb{Z}$-basis $(e_1, \ldots, e_g)$ of $\mathbb{Z}_F$, and embed F in $\mathbb{R}^g$ via $(\sigma_1, \ldots, \sigma_g)$. Let $R \in \mathrm{GL}_g(\mathbb{R})$ be the matrix whose (k, l)-th entry is $\sigma_k(e_l)$; then we have $\mathbb{Z}_F = R\,\mathbb{Z}^g$, and

$$\mathbb{Z}_F \oplus \partial_F^{-1} = \begin{pmatrix} R & 0 \\ 0 & R^{-t} \end{pmatrix} \mathbb{Z}^{2g}.$$

Moreover, the map $\left(\begin{smallmatrix} R & 0 \\ 0 & R^{-t} \end{smallmatrix}\right) : \mathbb{Z}^{2g} \to \mathbb{Z}_F \oplus \partial_F^{-1}$ respects the symplectic forms. The associated Hilbert embedding is

$$\begin{array}{rccc} H_R\colon & \mathbb{H}_1^g & \to & \mathbb{H}_g \\ & (\tau_1, \ldots, \tau_g) & \mapsto & R^t \,\mathrm{Diag}(\tau_1, \ldots, \tau_g) R, \end{array} \tag{2.12}$$

and the associated map on modular groups, also denoted by H_R, is

$$\begin{array}{rccc} H_R\colon & \mathrm{SL}(\mathbb{Z}_F \oplus \partial_F^{-1}) & \to & \mathrm{Sp}_{2g}(\mathbb{Z}) \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \mapsto & \begin{pmatrix} R^t & 0 \\ 0 & R^{-1} \end{pmatrix} \begin{pmatrix} a^\sigma & b^\sigma \\ c^\sigma & d^\sigma \end{pmatrix} \begin{pmatrix} R^{-t} & 0 \\ 0 & R^{-1} \end{pmatrix}. \end{array}$$

with the notation $y^\sigma = \mathrm{Diag}\big(\sigma_1(y), \dots, \sigma_g(y)\big) \in \mathrm{Mat}_{g\times g}(\mathbb{R})$ when $y \in F$.

Remark 2.3.11. The Hilbert embedding is not always a true embedding. For instance, when $g = 2$ (so that F is a real quadratic field), the induced map between moduli spaces is generically 2-1: if (A, η) is a principally polarized abelian surface with real multiplication by $\mathbb{Z}_F$, and $\overline{\eta} := \eta \circ \sigma$ where σ denotes real conjugation in F, then (A, η) and $(A, \overline{\eta})$ are not isomorphic in general but map to the same point in the Siegel threefold.

If $g = 2$, the image of the Hilbert embedding is a surface contained in the Siegel threefold, called the *Humbert surface* associated with the real quadratic field F (or rather $\mathbb{Z}_F$). Examples of equations of Humbert surfaces in terms of Igusa invariants can be computed when the discriminant of F is small [Gru10].

2.3.4 Hilbert modular forms in dimension 2

Keep the notation of §2.3.3. For all $\gamma \in \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(1)_F$ and $\tau = (\tau_1, \dots, \tau_g) \in \mathbb{H}_g$, we write

$$\gamma\tau := \left(\frac{\sigma_j(a)\tau_j + \sigma_j(b)}{\sigma_j(c)\tau_j + \sigma_j(d)}\right)_{1\le j\le g} \quad \text{and} \quad \gamma^*\tau := \prod_{i=1}^{g}\big(\sigma_j(c)\tau_j + \sigma_j(d)\big).$$

Let $w_1, \dots, w_g$ be integers. A *Hilbert modular form* of weight $(w_1, \dots, w_g)$ for F is a holomorphic function $f\colon \mathbb{H}_1^g \to \mathbb{C}$ such that for every $\gamma \in \Gamma(1)_F$ and every $\tau \in \mathbb{H}_1^g$, we have

$$f(\gamma\tau) = \prod_{j=1}^{g}\big(\sigma_j(c)\tau_j + \sigma_j(d)\big)^{w_j} f(\tau).$$

As in the case of Siegel modular forms of degree $g \ge 2$, Koecher's principle makes enforcing holomorphy conditions at the cusps unnecessary [Bru08, Thm. 1.20]. Since every irreducible (finite-dimensional, holomorphic) representation of $\mathrm{GL}_1(\mathbb{C})^g$ has dimension one, there is no need to consider vector-valued modular forms in the Hilbert case.

For every $w \ge 1$, Hilbert modular forms for F of *parallel weight* w, i.e. weight $(w, \dots, w)$, are identified with modular forms of weight w on the Shimura component $\mathcal{S} = \Gamma(1)_F\backslash\mathbb{H}_1^g$ described in §2.3.3. Explicitly, we choose $\tau_0 = (i, \dots, i) \in \mathbb{H}_1^g$ as a base point; if h is a classical Hilbert modular form of weight $(w, \dots, w)$ for F, then the function

$$f\colon [\gamma, (I_2)^g] \mapsto (\gamma^*\tau_0)^{-w} h(\gamma\tau_0)$$

is a modular function of weight w on $\mathcal{S}$ in the sense of §2.2.4.

The goal of this section is to describe certain generators of graded rings of Hilbert modular forms explicitly, as we did in §2.3.2 in the case of Siegel modular forms. Such generators depend on F, and they are known for a few examples of real quadratic fields F of small discriminants; we focus on the example of $F = \mathbb{Q}(\sqrt{5})$. Explicitly, Hilbert modular forms can be described in terms of Fourier expansions, and also through their relations to Siegel modular forms via the Hilbert embedding.

For the moment, let F be a general real quadratic field. We fix an embedding of F in $\mathbb{R}$, so that the real embeddings σ_1 and σ_2 are the identity and the real conjugation (denoted by σ), respectively. Let $\Sigma = \{1, \sigma\}$ be the Galois group of F. Then Σ acts on the $\mathbb{Q}$-vector space $V = F^2$, and satisfies the hypotheses listed in §2.2.4; hence it makes sense to consider Hilbert modular forms that are symmetric under Σ. We check that a Hilbert modular form f is symmetric if and only if $f(\tau_1, \tau_2) = f(\tau_2, \tau_1)$ for every $\tau_1, \tau_2 \in \mathbb{H}_1$.

Let (e_1, e_2) be a $\mathbb{Z}$-basis of $\mathbb{Z}_F$ and let $R = \left(\begin{smallmatrix} e_1 & e_2 \\ \bar{e}_1 & \bar{e}_2 \end{smallmatrix}\right)$ be the matrix defining the associated Hilbert embedding H_R, as in §2.3.3.

Proposition 2.3.12. *Let $k \in \mathbb{Z}$, $n \in \mathbb{N}$, and let $f\colon \mathbb{H}_2 \to \mathbb{C}_n[x]$ be a Siegel modular form of weight $\rho = \det^k \mathrm{Sym}^n$. Define the functions $g_i\colon \mathbb{H}_1^2 \to \mathbb{C}$ for $0 \leq i \leq n$ by the following equality:*

$$\forall t \in \mathbb{H}_1^2,\ \sum_{i=0}^{n} g_i(t)\, x^i = \rho(R) f\big(H_R(t)\big).$$

Then each g_i is a Hilbert modular form of weight $(k{+}i,\ k{+}n{-}i)$. Moreover, if $n = 0$, then the Hilbert modular form $g_0 = \det(R)^k H_R^* f$ is symmetric.

Proof. It is straightforward to check the transformation rule using the formula for H_R given in §2.3.3. The heart of the computation is that on diagonal matrices $\left(\begin{smallmatrix} r_1 & 0 \\ 0 & r_2 \end{smallmatrix}\right)$, the representation $\det^k \mathrm{Sym}^n$ splits: the coefficient before x^i is multiplied by $(r_1 r_2)^k\, r_1^i\, r_2^{n-i}$.

In the case $n = 0$, we observe that real conjugation in F induces a symplectic automorphism of the lattice $\mathbb{Z}_F \oplus \partial_F^{-1}$; this implies that there exists a matrix $\gamma_\sigma \in \mathrm{Sp}_4(\mathbb{Z})$ such that $H_R(\tau_2, \tau_1) = \gamma_\sigma H_R(\tau_1, \tau_2)$ for every $\tau_1, \tau_2 \in \mathbb{H}_1$. Symmetry follows from the fact that $\det(\gamma_\sigma) = 1$. □

As a consequence, the pullbacks of Igusa invariants to $\mathbb{H}_1^2$ under H_R are symmetric Hilbert modular functions of weight $(0, 0)$. They define a birational map from the Humbert surface to the subvariety of $\mathbb{C}^3$ cut out by the Humbert equation.

Remark 2.3.13. If h is a Siegel modular form of weight Sym^2 obtained as the derivative of a Siegel modular form of weight zero as in proposition 2.3.9, say $f(\tau)$, we can check that for every $(\tau_1, \tau_2) \in \mathbb{H}_1^2$,

$$\rho(R)h\big(H_R(\tau_1,\tau_2)\big) = \frac{1}{2\pi i}\left(\frac{\partial}{\partial\tau_1}f\big(H_R(\tau_1,\tau_2)\big)x^2 + \frac{\partial}{\partial\tau_2}f\big(H_R(\tau_1,\tau_2)\big)\right).$$

In general, if f is a Hilbert modular form of weight $(0,0)$, then its partial derivatives with respect to τ_1 and τ_2 are Hilbert modular forms of weight $(2,0)$ and $(0,2)$ respectively; this is consistent with the result of proposition 2.3.12.

In the special case $F = \mathbb{Q}(\sqrt{5})$, we take $e_1 = 1$ and $e_2 = (1-\sqrt{5})/2$, and we still take

$$R = \begin{pmatrix} e_1 & e_2 \\ \overline{e}_1 & \overline{e}_2 \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}). \tag{2.13}$$

Hilbert modular forms for F have Fourier expansions in terms of

$$w_1 = \exp\big(2\pi i(e_1\tau_1 + \overline{e}_1\tau_2)\big) \quad \text{and} \quad w_2 = \exp\big(2\pi i(e_2\tau_1 + \overline{e}_2\tau_2)\big).$$

We use this notation and the term *w-expansions* to avoid confusion with q-expansions of Siegel modular forms. A Hilbert modular form is *defined over* $\mathbb{Z}$ if its Fourier coefficients are all integers.

If $(a,b) \in \mathbb{Z}^2 \setminus \{(0,0)\}$, a nonzero coefficient before $w_1^a w_2^b$ can appear in the w-expansion of a Hilbert modular form only when $ae_1 + be_2$ is a totally positive element of $\mathbb{Z}_F$. Since $e_1 = 1$ and e_2 has negative norm, for a given a, only finitely many b's can appear. Therefore we can consider of w-expansions as elements of the power series ring $\mathbb{C}[w_2, w_2^{-1}][[w_1]]$.

Theorem 2.3.14 ([Nag83])**.** *The graded $\mathbb{C}$-algebra of symmetric Hilbert modular forms of even parallel weight for $F = \mathbb{Q}(\sqrt{5})$ is generated by three elements G_2, F_6, F_{10} of respective weights 2, 6 and 10, with w-expansions*

$$\begin{aligned}
G_2(\tau_1,\tau_2) &= 1 + (120w_2 + 120)w_1 \\
&\quad + \big(120w_2^3 + 600w_2^2 + 720w_2 + 600 + 120w_2^{-1}\big)w_1^2 + O(w_1^3), \\
F_6(\tau_1,\tau_2) &= (w_2+1)w_1 + \big(w_2^3 + 20w_2^2 - 90w_2 + 20 + w_2^{-1}\big)w_1^2 + O(w_1^3), \\
F_{10}(\tau_1,\tau_2) &= (w_2^2 - 2w_2 + 1)w_1^2 + O(w_1^3).
\end{aligned}$$

The ring of Hilbert modular forms of even parallel weight defined over $\mathbb{Z}$ is generated by G_2, F_6, F_{10}, and

$$F_{12} = \frac{1}{4}(F_6^2 - G_2F_{10}).$$

As a consequence, we can give an analogue of proposition 2.3.7 in the case of Hilbert modular forms for $\mathbb{Q}(\sqrt{5})$, whose proof is immediate given theorem 2.3.14.

Proposition 2.3.15. *Let f be a symmetric Hilbert modular form of even weight k for the real quadratic field $\mathbb{Q}(\sqrt{5})$. Then $2^k f \in \mathbb{Z}[G_2, G_6, F_{10}]$.*

The *Gundlach invariants* for $F = \mathbb{Q}(\sqrt{5})$ are

$$g_1 = \frac{G_2^5}{F_{10}} \quad \text{and} \quad g_2 = \frac{G_2^2 F_6}{F_{10}}.$$

They define a birational map from the associated Humbert surface to $\mathbb{C}^2$.

By proposition 2.3.12, the pullbacks of the Siegel modular forms ψ_4, ψ_6, χ_{10} and χ_{12} via the Hilbert embedding H_R are symmetric Hilbert modular forms of even weight. Their expressions in terms of G_2, F_6, F_{10} can be computed using linear algebra on Fourier expansions [LY11, Prop. 3.2]: in our case we have

$$q_1 = w_1, \quad q_2 = w_1 w_2, \quad q_3 = w_2.$$

In the following result, we take into account the fact that our choice of normalization for the Siegel modular forms χ_{10} and χ_{12} (recall §2.3.2) differs from that of [LY11].

Proposition 2.3.16 ([LY11, Thm. 4.4])**.** *In the case $F = \mathbb{Q}(\sqrt{5})$, with R as in (2.13), we have*

$$\begin{aligned} H_R^* \psi_4 &= G_2^2, \\ H_R^* \psi_6 &= G_2^3 - 864 F_6, \\ H_R^* \chi_{10} &= -F_{10}, \\ H_R^* \chi_{12} &= -12(3F_6^2 - 2G_2 F_{10}), \end{aligned}$$

As another example, we study the pullback of the modular form $f_{8,6}$ from example 2.3.10 via H_R.

Proposition 2.3.17. *Let R be as in eq. (2.13), and define the functions $b_i(t)$ on $\mathbb{H}_1^2$ for $0 \le i \le 6$ by*

$$\forall t \in \mathbb{H}_1^2, \ \det{}^8 \operatorname{Sym}^6(R) f_{8,6}\big(H_R(t)\big) = \sum_{i=0}^{6} b_i(t)\, x^i.$$

Then b_2 and b_4 are identically zero, and

$$\begin{aligned} b_3^2 &= 4F_{10}F_6^2, \\ b_1b_5 &= \frac{36}{25}F_{10}F_6^2 - \frac{4}{5}F_{10}^2G_2, \\ b_0b_6 &= \frac{-4}{25}F_{10}F_6^2 + \frac{1}{5}F_{10}^2G_2, \\ b_3\left(b_0^2b_5^3 + b_1^3b_6^2\right) &= 123F_{10}^3F_6 - \frac{32}{25}F_{10}^2F_6^2G_2^2 + \frac{288}{125}F_{10}F_6^4G_2 - \frac{3456}{3125}F_6^6. \end{aligned}$$

Proof. By proposition 2.3.12, each coefficient b_i is a Hilbert modular form of weight $(8+i, 14-i)$. We can check that σ exchanges b_i and b_{6-i}. From the q-expansion of $f_{8,6}$, we can compute the w-expansions of the b_i's; then, we use linear algebra to identify symmetric combinations of the b_i's of parallel even weight in terms of the generators G_2, F_6, F_{10} from theorem 2.3.14. In particular we find that $b_2b_4 = 0$; since b_2 and b_4 are exchanged by σ, both are identically zero. □

2.4 Modular equations

In this section, we define modular equations in the setting of PEL Shimura varieties as equations describing Hecke correspondences. All the types of modular equations described in chapter 1, namely elliptic modular polynomials and modular equations of Siegel and Hilbert type for abelian surfaces, can be recovered as special cases of the general definition.

2.4.1 The example of elliptic modular polynomials

In chapter 1, we defined the classical modular polynomial Φ_ℓ, where ℓ is a prime, as follows. We start with the modular invariant j, which generates the function field of the moduli space $\mathcal{S} = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}_1$ over $\mathbb{Q}$. Then, we look at the function

$$\tau \mapsto j(\tau/\ell),$$

which is invariant under the action of the congruence subgroup

$$\Gamma^0(\ell) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z}) \colon b = 0 \bmod \ell \right\}.$$

Then the coefficients of the polynomial

$$P_\ell(\tau) = \prod_{\gamma \in \Gamma^0(\ell)\backslash \mathrm{SL}_2(\mathbb{Z})} \left(Y - j\left(\tfrac{1}{\ell}\gamma\tau\right)\right) \quad \in \mathbb{C}[Y]$$

are modular functions for $\mathrm{SL}_2(\mathbb{Z})$. The polynomial Φ_ℓ is the unique polynomial in $\mathbb{C}[X,Y]$ satisfying the relation $\Phi_\ell(j(\tau),Y) = P_\ell(\tau)$ for every $\tau \in \mathbb{H}_1$.

In order to introduce the definition of modular equations in a broader sense, we translate the definition of Φ_ℓ in the language of PEL Shimura varieties. Recall that the underlying algebraic group in the case of modular curves is $G = \mathrm{GL}_2$. Note that for every $\tau \in \mathbb{H}_1$, we have

$$\tau/\ell = \delta\tau \quad \text{with } \delta = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \ell \end{smallmatrix}\right) \in G(\mathbb{Q})_+,$$

and moreover

$$\Gamma^0(\ell) = \mathrm{SL}_2(\mathbb{Z}) \cap \left(\delta^{-1}\,\mathrm{SL}_2(\mathbb{Z})\delta\right).$$

Geometrically, we consider the diagram

$$\begin{array}{ccc} \Gamma^0(\ell)\backslash\mathbb{H}_1 & \xrightarrow{\tau\mapsto\delta\tau} & \Gamma_0(\ell)\backslash\mathbb{H}_1 \\ \downarrow & & \downarrow \\ \mathcal{S} & & \mathcal{S} \end{array} \tag{2.14}$$

where $\Gamma_0(\ell) = \left(\delta\,\mathrm{SL}_2(\mathbb{Z})\delta^{-1}\right) \cap \mathrm{SL}_2(\mathbb{Z})$. It gives a morphism

$$\Gamma^0(\ell)\backslash\mathbb{H}_1 \to \mathcal{S}\times\mathcal{S}, \quad \tau \mapsto (\tau, \delta\tau).$$

The product $\mathcal{S}\times\mathcal{S}$ is then birational to $\mathbb{P}^1\times\mathbb{P}^1$ via (j,j). The modular curve $\Gamma^0(\ell)\backslash\mathbb{H}_1$ is birational to its image in $\mathbb{P}^1\times\mathbb{P}^1$, and the elliptic modular polynomial of level ℓ is as an equation for this image.

In the modular interpretation, the lattice $\mathbb{Z}\oplus\delta\tau\mathbb{Z}$ contains the lattice $\mathbb{Z}\oplus\tau\mathbb{Z}$ as a sublattice of index ℓ, so the elliptic curves attached to τ and $\delta\tau$ are indeed ℓ-isogenous.

Recall from example 2.2.14 that if g is a classical modular form, then the function

$$f\colon [\gamma, I_2] \mapsto (\gamma^* i)^{-w} g(\gamma i)$$

is a modular form on $\mathcal{S}$ in the sense of §2.2.4: the function f is left-invariant under $G(\mathbb{Q})_+$, right-invariant under a compact subgroup K of $G(\mathbb{A}_f)$ (with $K = K_0 = \mathrm{GL}_2(\widehat{\mathbb{Z}})$ if g was modular for $\mathrm{SL}_2(\mathbb{Z})$), and satisfies a certain transformation rule under right action by the subgroup $K_\infty \subset G(\mathbb{R})_+$. Under this correspondence, we can consider j as a modular form on $\mathcal{S}$ in the sense of §2.2.4. Then, if $\tau = \gamma i \in \mathbb{H}_1$, we have

$$j(\delta\tau) = j\left([\delta\gamma, I_2]\right) = j\left([\gamma, \delta^{-1}]\right).$$

Therefore, the classical modular function $\tau \mapsto j(\delta\tau)$ for $\Gamma^0(\ell)$ corresponds to the modular function

$$\begin{array}{rccc} j_{\delta^{-1}}\colon & G(\mathbb{Q})_+\backslash\left(G(\mathbb{A}_f)\times G(\mathbb{R})_+\right) & \to & \mathbb{C} \\ & [x,g] & \mapsto & j\left([x, g\delta^{-1}]\right). \end{array}$$

The function $j_{\delta^{-1}}$ is right-invariant under $K_{\delta^{-1}} := \delta^{-1}K_0\delta$. Define K' as the compact open subgroup $K_0 \cap K_{\delta^{-1}}$ of $G(\mathbb{A}_f)$. The diagram (2.14) can be rewritten as

$$\begin{array}{ccc} \mathrm{Sh}_{K'}(\mathrm{GL}_2, \mathbb{H}_1)(\mathbb{C}) & \xrightarrow{[x,g]\mapsto[x,g\delta^{-1}]} & \mathrm{Sh}_{\delta K'\delta^{-1}}(\mathrm{GL}_2, \mathbb{H}_1)(\mathbb{C}) \\ \downarrow & & \downarrow \\ \mathcal{S} & & \mathcal{S}. \end{array} \tag{2.15}$$

Let K'' be a normal subgroup of finite index in K_0 contained in K'. We let K_0 act (on the left) on the set of modular functions for K'' as follows: if $k \in K_0$ and f is such a function, we define

$$k \cdot f \colon [x, g] \mapsto f([x, gk]).$$

The subgroup K' of K_0 is contained in the stabilizer of $j_{\delta^{-1}}$. Therefore the coefficients of the polynomial

$$Q_\ell := \prod_{\gamma \in K_0/K'} \left(Y - \gamma \cdot j_{\delta^{-1}}\right)$$

are modular functions of level K_0. In the world of classical modular forms, Q_ℓ corresponds exactly to P_ℓ, as inversion induces a bijection between right cosets of $\Gamma^0(\ell)$ in $\mathrm{SL}_2(\mathbb{Z})$ and left cosets of K' in K_0. The general definition of modular equations uses analogues of diagram (2.15), which make sense for a general Shimura variety and any $\delta \in G(\mathbb{A}_f)$; these diagrams are called *Hecke correspondences*.

2.4.2 Hecke correspondences

Let us fix a PEL datum $(B, *, V, \psi, G, X_+)$ as in §2.2.2, as well as a compact open subgroup $K \subset G(\mathbb{A}_f)$. Following §2.4.1, we look at the diagram

$$\begin{array}{ccc} \mathrm{Sh}_{K'}(G, X_+)(\mathbb{C}) & \xrightarrow{[x,g]\mapsto[x,g\delta]} & \mathrm{Sh}_{\delta^{-1}K'\delta}(G, X_+)(\mathbb{C}) \\ \downarrow & & \downarrow \\ \mathrm{Sh}_K(G, X_+)(\mathbb{C}) & & \mathrm{Sh}_K(G, X_+)(\mathbb{C}) \end{array} \tag{2.16}$$

where $\delta \in G(\mathbb{A}_f)$, and $K' := K \cap \delta K \delta^{-1}$. The induced map

$$H_\delta \colon \mathrm{Sh}_{K'}(\mathbb{C}) \to \mathrm{Sh}_K(\mathbb{C}) \times \mathrm{Sh}_K(\mathbb{C})$$

is called the *Hecke correspondence* of level δ on $\mathrm{Sh}_K(\mathbb{C})$. Alternatively, H_δ can be seen as a correspondence on $\mathrm{Sh}_K(\mathbb{C}) \times \mathrm{Sh}_K(\mathbb{C})$ consisting of all pairs of the form $\big([x,g],[x,g\delta]\big)$ for $[x,g] \in \mathrm{Sh}_{K'}(\mathbb{C})$. Hecke correspondences are algebraic: the diagram (2.16) is the analytification of a diagram existing at the level of algebraic varieties. Moreover, Hecke correspondences are defined over the reflex field $E(G,X_+)$ [Mil05, Thm. 13.6].

We define the *degree* of H_δ to be the index

$$d(\delta) = [K : K'] = [K : K \cap \delta K \delta^{-1}].$$

This index is finite as both K and K' are compact open subgroups of $G(\mathbb{A}_f)$, and is the degree of the projection map $\mathrm{Sh}_{K'} \to \mathrm{Sh}_K$. One can also consider H_δ as a map from Sh_K to its $d(\delta)$-th symmetric power, sending $z \in \mathrm{Sh}_K$ to the set $\{z' \in \mathrm{Sh}_K : (z,z') \in H_\delta\}$.

It is easy to see how H_δ behaves with respect to connected components of $\mathrm{Sh}_K(\mathbb{C})$: if z lies in the connected component indexed by $t \in T(\mathbb{A}_f)$ (in the notation of §2.2.2), then its images under H_δ all lie in the connected component indexed by $t\,\nu(\delta)$.

We call the Hecke correspondence H_δ *absolutely irreducible* if for every connected component $\mathcal{S}$ of $\mathrm{Sh}_K(\mathbb{C})$ with field of definition L, the preimage of $\mathcal{S}$ in $\mathrm{Sh}_{K'}$ is absolutely irreducible as a variety defined over L (or equivalently, connected as a variety over $\mathbb{C}$). A sufficient condition for H_δ to be absolutely irreducible is that $\nu(K') = \nu(K)$.

Modular interpretation of Hecke correspondences. In the modular interpretation, Hecke correspondences describe isogenies of a certain type between polarized abelian varieties. Let $\Lambda_0, \mathcal{C}$, and $\mathcal{O}$ be as in §2.2.3, and write

$$K = \bigsqcup_{i=1}^{d(\delta)} \kappa_i K',$$

where $\kappa_i \in G(\mathbb{A}_f)$ for $1 \le i \le d(\delta)$. Let $c \in \mathcal{C}$, denote by $\mathcal{S}_c$ the connected component of $\mathrm{Sh}_K(\mathbb{C})$ indexed by c, and consider the lattice with PEL structure $(\Lambda_c, x, \iota, \psi_c, cK)$ associated with a point $[x,c] \in \mathcal{S}_c$ by proposition 2.2.8.

In order to construct the lattices associated with $[x,c]$ via the Hecke correspondence H_δ, we partition the orbit cK into the K'-orbits $c\kappa_i K'$ for $1 \le i \le d(\delta)$. Each $c\kappa_i\delta \in G(\mathbb{A}_f)$ is then a $\widehat{\mathbb{Z}}$-linear embedding of $\mathcal{O}$-modules $\widehat{\Lambda}_0 \hookrightarrow V(\mathbb{A}_f)$; it is well defined up to right multiplication by $\delta^{-1}K'\delta$, hence by K. Let $\Lambda_i \subset V(\mathbb{Q})$ be the lattice such that $\Lambda_i \otimes \widehat{\mathbb{Z}}$ is the image of this embedding. There is still a natural action of $\mathcal{O}$ on Λ_i.

The decomposition $c\kappa_i\delta K = q_i c' K$, with $q_i \in G(\mathbb{Q})_+$ and $c' \in \mathcal{C}$, is well defined, and the element c' does not depend on i.

Proposition 2.4.1. *Let $\delta \in G(\mathbb{A}_f)$, let $z = [x, c] \in \mathcal{S}_c$, and construct Λ_i, q_i, c' as above. Then the image of z under the Hecke correspondence H_δ in the modular interpretation of proposition 2.2.9 is given by the $d(\delta)$ isomorphism classes of tuples with representatives*

$$\Big(\Lambda_i, x, \frac{\lambda_{c'}}{\lambda_c}\psi_c\big(\mu(q_i^{-1})\cdot,\cdot\big), c\kappa_i\delta K\Big) \qquad \text{for} \quad 1 \le i \le d(\delta).$$

Proof. By construction, the images of $[x, c]$ under the Hecke correspondence are the points $[q_i^{-1}x, c']$ of $\mathrm{Sh}_K(\mathbb{C})$. The relation $c\kappa_i\delta K = q_i c' K$ shows that the map q_i^{-1} sends the lattice Λ_i to $\Lambda_{c'}$. This map also respects the action of $\mathcal{O}$, and sends the complex structure x to $q_i^{-1}x$. Finally, it sends the polarization $(u, v) \mapsto \psi_c(u, v)$ on Λ_i to $(u, v) \mapsto \psi_c\big(\mu(q_i)u, v\big)$ on $\Lambda_{c'}$. □

After multiplying δ by a unique suitable element in $\mathbb{Q}_+^\times$, which does not change H_δ, we can assume that $\delta(\widehat{\Lambda}_0) \subset \widehat{\Lambda}_0$ and $\delta(\widehat{\Lambda}_0) \not\subset p\widehat{\Lambda}_0$ for every prime p; we say that δ is *normalized* with respect to Λ_0. In this case, we define the *isogeny degree* of H_δ as the unique integer $l(\delta) \ge 1$ such that $l(\delta)^{-1}\det(\delta)$ is a unit in $\widehat{\mathbb{Z}}$. In other words,

$$l(\delta) = \#\big(\widehat{\Lambda}_0/\delta(\widehat{\Lambda}_0)\big).$$

For a general $\delta \in G(\mathbb{A}_f)$, we set $l(\delta) = l(\lambda\delta)$ where $\lambda \in \mathbb{Q}_+^\times$ is chosen such that $\lambda\delta$ is normalized with respect to Λ_0.

Corollary 2.4.2. *Let $\delta \in G(\mathbb{A}_f)$. Then, in the modular interpretation of proposition 2.2.9, the Hecke correspondence H_δ sends an abelian variety A with PEL structure to $d(\delta)$ abelian varieties $A_1, \dots, A_{d(\delta)}$ such that for every $1 \le i \le d(\delta)$, there exists an isogeny $A_i \to A$ of degree $l(\delta)$.*

Proof. We can assume that δ is normalized with respect to Λ_0. Then, in proposition 2.4.1, each Λ_i for $1 \le i \le d(\delta)$ is a sublattice of Λ_c endowed with the same complex structure x. Moreover, for every $1 \le i \le d(\delta)$, we have $\Lambda_c/\Lambda_i \simeq \widehat{\Lambda}_0/\delta(\widehat{\Lambda}_0)$, so the index of each Λ_i in Λ_c is $l(\delta)$. □

Example 2.4.3. In the case of the classical modular curve of level one, we have $G = \mathrm{GL}_2$ and $K = K_0 = \mathrm{GL}_2(\widehat{\mathbb{Z}})$. Consider the Hecke correspondence of level $\delta = \left(\begin{smallmatrix}\ell & 0\\ 0 & 1\end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{A}_f)$, where ℓ is a prime. In this case

$$K' = K \cap \delta K \delta^{-1} = \left\{\left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right) \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) \colon b = 0 \bmod \ell\right\};$$

this is the Hecke correspondence considered in §2.4.1. The Hecke correspondence H_δ is absolutely irreducible since $\nu(K_0) = \nu(K') = \widehat{\mathbb{Z}}^\times$. We have $d(\delta) = \ell + 1$ and $l(\delta) = \ell$. In the classical modular interpretation, H_δ sends an elliptic curve E to the quotients of E by the $\ell + 1$ cyclic subgroups of $E[\ell]$; these quotient isogenies are dual to the isogenies given in corollary 2.4.2.

A relation between degrees. For later purposes, we state an inequality relating $d(\delta)$ with a power of $l(\delta)$. Since $K \subset G(\mathbb{A}_f)$ is open, there exists a smallest integer $N \geq 1$ such that

$$\big\{g \in G(\mathbb{A}_f) \cap \mathrm{GL}(\widehat{\Lambda}_0) \colon g = 1 \bmod N\widehat{\Lambda}_0\big\} \subset K,$$

which we call the *level* of K with respect to $\widehat{\Lambda}_0$.

Proposition 2.4.4. *There exists a constant C depending on K and Λ_0 such that for every $\delta \in G(\mathbb{A}_f)$, we have $d(\delta) \leq C\, l(\delta)^{(\dim V)^2}$. We can take $C = N^{(\dim V)^2}$, where N is the level of K with respect to $\widehat{\Lambda}_0$.*

Proof. We can assume that δ is normalized with respect to $\widehat{\Lambda}_0$. Then the subgroup $K \cap \delta K \delta^{-1}$ contains all the elements $g \in G(\mathbb{A}_f) \cap \mathrm{GL}(\widehat{\Lambda}_0)$ that are the identity modulo $\widehat{\Lambda} = l(\delta) N \widehat{\Lambda}_0$. In other words we have a morphism of groups $K \to \mathrm{GL}(\Lambda_0 / N\, l(\delta)\Lambda_0)$ whose kernel is contained in $K \cap \delta K \delta^{-1}$. This yields the result since $\#\,\mathrm{GL}(\Lambda_0 / N\, l(\delta)\Lambda_0) \leq (N\, l(\delta))^{(\dim V)^2}$. □

Remark 2.4.5. The upper bound on $d(\delta)$ given in proposition 2.4.4 is far from optimal in many cases: for instance, if δ is normalized with respect to $\widehat{\Lambda}_0$, if $l(\delta)$ is prime to N, and if moreover δ normalizes the image of K in $\mathrm{GL}(\Lambda_0 / N\Lambda_0)$, then $d(\delta) \leq l(\delta)^{(\dim V)^2}$. But in general, the level of K does enter into account. As an example, take $G = \mathrm{GL}_2$, $\delta = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, and

$$K = \Big\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) \colon a = d = 1 \bmod N \text{ and } c = 0 \bmod N \Big\}.$$

Then $d(\delta) = N$ even though $l(\delta) = 1$. In the modular interpretation, the Hecke correspondence H_δ has the effect of forgetting the initial K-level structure entirely.

2.4.3 Modular equations on PEL Shimura varieties

When defining the elliptic modular polynomial Φ_ℓ, besides choosing the Hecke correspondence, one uses a particular generator of the function field of the modular curve, namely the j-invariant. In the general case, we also fix a choice of invariants on Shimura varieties.

Fix a PEL datum $(B, *, V, \psi, G, X_+)$ as above, let $K \subset G(\mathbb{A}_f)$ be a compact open subgroup, and let Σ be a finite group of automorphisms of V as in §2.2.4. Let n be the complex dimension of X_+; we assume that $n \geq 1$. Let $\mathcal{S}$, $\mathcal{T}$ be connected components of $\mathrm{Sh}_K(G, X_+)(\mathbb{C})$, and let L be their field of definition.

Since the field $L(\mathcal{S})$ of modular functions on $\mathcal{S}$ has transcendence degree n over L, the field $L(\mathcal{S})^\Sigma$ of modular functions on $\mathcal{S}$ that are symmetric under Σ also has transcendence degree n over L. Choose a transcendence basis $(j_1, \ldots, j_n)$ of $L(\mathcal{S})^\Sigma$ over L, and another symmetric function j_{n+1} that generates the remaining finite extension, whose degree is denoted by e. On $\mathcal{S}$, the function j_{n+1} satisfies a minimal relation of the form

$$E(j_1, \ldots, j_{n+1}) = 0 \tag{2.17}$$

where

$$E = \sum_{k=0}^{e} E_k(J_1, \ldots, J_n)\, J_{n+1}^k \in L[J_1, \ldots, J_{n+1}]$$

and E is irreducible. If $L(\mathcal{S})^\Sigma$ is purely transcendental over L (if $\Sigma = \{1\}$, this means that $\mathcal{S}$ is birational to $\mathbb{P}^n$), then we can take $j_{n+1} = 1$, ignore eq. (2.17), and work with n invariants only. This simplification will happen for our three main examples of PEL Shimura varieties, namely the modular curve $X(1)$, the Siegel threefold, and the Hilbert surface for $F = \mathbb{Q}(\sqrt{5})$.

We proceed similarly to define coordinates on $\mathcal{T}$: no confusion will arise if we also denote them by $j_1, \ldots, j_{n+1}$. We refer to the data defined up to now as a choice of *PEL setting*.

Given a PEL setting, let $\delta \in G(\mathbb{A}_f)$ define an absolutely irreducible Hecke correspondence H_δ that intersects $\mathcal{S} \times \mathcal{T}$ nontrivially. We want to define explicit polynomials with coefficients in $L(\mathcal{S})$, called the *modular equations of level* δ, describing H_δ in the product $\mathcal{S} \times \mathcal{T}$. To do this, we mimic the definition of elliptic modular polynomials in the language of PEL Shimura varieties given in §2.4.1. As in §2.4.2, we write

$$K' = K \cap \delta K \delta^{-1}.$$

Let K'' be a normal subgroup of finite index in K, contained in K', and stabilized by Σ. Let $\mathcal{S}''$ be the preimage of $\mathcal{S}$ in $\mathrm{Sh}_{K''}(\mathbb{C})$ (which is possibly not connected). There is a left action of $K \rtimes \Sigma$ on the space of modular functions on $\mathcal{S}''$, given by

$$(k, \sigma) \cdot f \colon [x, g] \mapsto \sigma \cdot f([x, gk]).$$

The modular functions that are invariant under $K' \rtimes \{1\}$ (resp. $K \rtimes \Sigma$) are exactly the functions on $H_\delta \cap (\mathcal{S} \times \mathcal{T})$ defined over $\mathbb{C}$ (resp. the functions on $\mathcal{S}$ defined over $\mathbb{C}$ and invariant under Σ). The modular functions

$$j_{i,\delta} \colon [x, g] \mapsto j_i([x, g\delta])$$

for $1 \leq i \leq n+1$ are defined over L and generate the function field of $H_\delta \cap (\mathcal{S} \times \mathcal{T})$. We define the decreasing chain of subgroups

$$K \rtimes \Sigma = K_0 \supset K_1 \supset \cdots \supset K_{n+1} \supset K'$$

as follows: for each $1 \leq i \leq n+1$, the subgroup K_i is the stabilizer of the modular functions $j_{1,\delta}, \ldots, j_{i,\delta}$.

Galois theory applied to the Galois covering $\mathcal{S}'' \to \mathcal{S}$ tells us that for every $1 \leq i \leq n+1$, the field $L(j_1, \ldots, j_{n+1}, j_{1,\delta}, \ldots, j_{i,\delta})$ is the function field of the preimage of $\mathcal{S}$ in the Shimura variety Sh_{K_i}, and consists of all modular functions on $\mathcal{S}''$ defined over L that are invariant under K_i. In other words, we have a tower of function fields:

$$\begin{array}{c} L(j_1, \ldots, j_{n+1}, j_{1,\delta}, \ldots, j_{n+1,\delta}) = L(H_\delta \cap (\mathcal{S} \times \mathcal{T})) \\ \text{degree } d_{n+1} \Big| \\ \vdots \\ \text{degree } d_2 \Big| \\ L(j_1, \ldots, j_{n+1}, j_{1,\delta}) \\ \text{degree } d_1 \Big| \\ L(\mathcal{S})^\Sigma, \end{array}$$

where $d_i = [K_{i-1} : K_i]$ for $1 \leq i \leq n+1$. The modular equations of level δ are defining equations for the successive extensions in the tower.

Definition 2.4.6. The *modular equations* of level δ on $\mathcal{S} \times \mathcal{T}$ are the tuple $(\Psi_{\delta,1}, \ldots, \Psi_{\delta,n+1})$ defined as follows: for each $1 \leq m \leq n+1$, $\Psi_{\delta,m}$ is the multivariate polynomial in the m variables $Y_1, \ldots, Y_m$ defined by

$$\Psi_{\delta,m} = \sum_{\gamma \in K_0/K_{m-1}} \left(\left(\prod_{i=1}^{m-1} \prod_{\gamma_i} (Y_i - \gamma_i \cdot j_{i,\delta}) \right) \prod_{\gamma_m \in K_{m-1}/K_m} (Y_m - \gamma\gamma_m \cdot j_{m,\delta}) \right),$$

where the middle product is over all $\gamma_i \in K_0/K_i$ such that $\gamma_i = \gamma$ modulo K_{i-1}, but $\gamma_i \neq \gamma$ modulo K_i. The expression for $\Psi_{\delta,m}$ makes sense, because multiplying γ on the right by an element in K_{m-1} only permutes the factors in the last product.

Definition 2.4.6 generalizes the definition of elliptic modular polynomials; the precise formula is inspired from preexisting definitions of modular equations for abelian surfaces [BL09, Mil15, Mar20, MR20]. We will return to these examples in §2.4.4.

Let us give elementary properties of modular equations. We need a technical lemma.

Lemma 2.4.7. *Let $\gamma, \gamma' \in K_0$ and $1 \leq i \leq n+1$. Assume that the equality $\gamma \cdot j_{i,\delta} = \gamma' \cdot j_{i,\delta}$ holds on one connected component of $\mathcal{S}''$. Then it holds on all connected components of $\mathcal{S}''$.*

Proof. Write $\gamma = (k, \sigma)$ and $\gamma' = (k', \sigma')$ where $k, k' \in K$ and $\sigma, \sigma' \in \Sigma$. Let $c \in \mathcal{C} \subset G(\mathbb{A}_f)$ be an element defining the connected component $\mathcal{S}$ in $\mathrm{Sh}_K(\mathbb{C})$, so that $\mathcal{S} = \Gamma_c \backslash X_+$ with $\Gamma_c = G(\mathbb{Q})_+ \cap cKc^{-1}$. By assumption, there exists an element $g \in G(\mathbb{A}_f)$ such that $g = c$ in the double quotient space $G(\mathbb{Q})_+ \backslash G(\mathbb{A}_f)/K$, and

$$\forall x \in X_+,\ j_{i,\delta}\big([\sigma^{-1}(x), \sigma^{-1}(gk)]\big) = j_{i,\delta}\big([\sigma'^{-1}(x), \sigma'^{-1}(gk')]\big). \tag{2.18}$$

Since H_δ is absolutely irreducible, we have

$$G(\mathbb{Q})_+ \backslash G(\mathbb{A}_f)/K = G(\mathbb{Q})_+ \backslash G(\mathbb{A}_f)/K'.$$

Using the description of connected components of a PEL Shimura variety in §2.2.2, and the fact that the action of Σ leaves ν invariant, we find that there exist $\gamma_1, \gamma_2 \in G(\mathbb{Q})_+$ such that $gk = \gamma_1 \sigma(c) \bmod \sigma(K')$ and $gk' = \gamma_2 \sigma'(c) \bmod \sigma'(K')$. Then eq. (2.18) is equivalent to the following:

$$\forall x \in X_+,\ j_{i,\delta}\big([x, c]\big) = j_{i,\delta}\big([\sigma'^{-1}(\gamma_2^{-1}\gamma_1 \sigma(x)), c]\big). \tag{2.19}$$

Note that $\gamma_2^{-1}\gamma_1$ is well-defined and independent of g, up to multiplication on the left by an element of $G(\mathbb{Q})_+ \cap \sigma'(cK'c^{-1})$, and multiplication on the right by an element of $G(\mathbb{Q})_+ \cap \sigma(cK'c^{-1})$. Therefore eq. (2.19) holds for each $g \in G(\mathbb{A}_f)$ such that $g = c$ in $G(\mathbb{Q})_+ \backslash G(\mathbb{A}_f)/K$. In other words, the equality $\gamma \cdot j_{i,\delta} = \gamma' \cdot j_{i,\delta}$ holds on every connected component of $\mathcal{S}''$. □

Proposition 2.4.8. *Let $1 \leq m \leq n+1$, and let $\gamma \in K_0/K_{m-1}$. Then, up to multiplication by an element in $L(j_1, \ldots, j_{n+1}, \gamma \cdot j_{1,\delta}, \ldots, \gamma \cdot j_{m-1,\delta})^\times$, we have*

$$\Psi_{\delta,m}(\gamma \cdot j_{1,\delta}, \ldots, \gamma \cdot j_{m-1,\delta}, Y_m) = \prod_{\gamma_m \in K_{m-1}/K_m} \Big(Y_m - \gamma\gamma_m \cdot j_{m,\delta}\Big).$$

Proof. By definition 2.4.6, the above equality holds true after multiplying the right hand side by

$$f = \prod_{i=1}^{m-1} \prod_{\substack{\gamma_i \in K_0/K_i \\ \gamma_i \neq \gamma \\ \gamma_i = \gamma \bmod K_{i-1}}} \left(\gamma \cdot j_{i,\delta} - \gamma_i \cdot j_{i,\delta}\right).$$

The function f a product of nonzero modular functions on $\mathcal{S}''$ defined over L. In order to show that $f \in L(j_1, \ldots, j_{n+1}, \gamma \cdot j_{1,\delta}, \ldots, \gamma \cdot j_{m-1,\delta})$, we check that f is invariant under the action of $\gamma K_{m-1} \gamma^{-1}$. By definition of the subgroups K_i, no factor of f is identically zero on $\mathcal{S}''$. Therefore f is invertible by lemma 2.4.7. □

Let $1 \leq m \leq n+1$. Proposition 2.4.8 implies that up to scaling, the univariate polynomial $\Psi_{\delta,m}(j_{1,\delta}, \ldots, j_{m-1,\delta}, Y_m)$ is the minimal polynomial of $j_{m,\delta}$ over the field $L(j_1, \ldots, j_{n+1}, j_{1,\delta}, \ldots, j_{m-1,\delta})$. In other words, when the multiplicative coefficient does not vanish, which is generically the case, $\Psi_{\delta,m}$ provides all the possible values for $j_{m,\delta}$ once $j_1, \ldots, j_{n+1}$ and $j_{1,\delta}, \ldots, j_{m-1,\delta}$ are known. In particular, modular equations vanish on H_δ as promised.

We could also define other modular equations $\Phi_{\delta,m}$ for which there is true equality in proposition 2.4.8, as in the case of elliptic modular polynomials, but they have a more complicated expression. In practice, using the polynomials $\Psi_{\delta,m}$ is more convenient as they are typically smaller.

Proposition 2.4.9. *Let $1 \leq m \leq n+1$. The coefficients of $\Psi_{\delta,m}$ lie in $L(j_1, \ldots, j_{n+1})$. The degree of $\Psi_{\delta,m}$ in Y_m is $[K_m : K_{m-1}]$, and for each $1 \leq i < m$, the degree of $\Psi_{\delta,m}$ in Y_i is at most $[K_i : K_{i-1}] - 1$.*

Proof. It is clear from definition 2.4.6 that the action of K_0 leaves $\Psi_{\delta,m}$ invariant. Hence the coefficients of $\Psi_{\delta,m}$ are functions on $\mathcal{S}$ invariant under Σ and defined over L, so the first statement holds. The second part is obvious. □

In general, using a nontrivial Σ increases the degree of modular equations. This has a geometric interpretation: modular equations describe the Hecke correspondence H_δ and its conjugates under Σ simultaneously.

Let $J_1, \ldots, J_{n+1}$ be indeterminates, and let $1 \leq m \leq n+1$. By the equation (2.17) satisfied by j_{n+1} on $\mathcal{S}$, there exists a unique element of the ring $L(J_1, \ldots, J_n)[J_{n+1}, Y_1, \ldots, Y_m]$ of degree at most $e-1$ in J_{n+1} which, when evaluated at $J_i = j_i$ for $1 \leq i \leq n+1$, yields $\Psi_{\delta,m}$. In the sequel, we also denote it by $\Psi_{\delta,m}$. Therefore the coefficients of $\Psi_{\delta,m}$ will be either

functions on $\mathcal{S}$, i.e. elements of $L(j_1, \dots, j_{n+1})$, or multivariate rational fractions in the indeterminates $J_1, \dots, J_{n+1}$ that are polynomial in J_{n+1} of degree at most $e-1$, depending on the context.

Remark 2.4.10. In several cases, the function $j_{1,\delta}$ already generates the whole extension of function fields, so that $K_1 = \cdots = K_{n+1} = K'$,

$$\Psi_{\delta,1} = \prod_{\gamma_1 \in K_0/K'} \big(Y_1 - \gamma_1 \cdot j_{1,\delta}\big),$$

and for every $2 \leq m \leq n+1$,

$$\Psi_{\delta,m} = \sum_{\gamma \in K_0/K'} \left(\left(\prod_{\gamma_1 \neq \gamma} \big(Y_1 - \gamma_1 \cdot j_{1,\delta}\big) \right) \big(Y_m - \gamma \cdot j_{m,\delta}\big) \right). \tag{2.20}$$

In this case, for each $2 \leq m \leq n+1$, we have

$$\Psi_{\delta,m}(j_{1,\delta}) = \partial_{Y_1} \Psi_{\delta,1}(j_{1,\delta}) \cdot (Y_m - j_{m,\delta}),$$

where ∂_{Y_1} denotes derivative with respect to Y_1. Therefore $\Psi_{\delta,m}$ is just the expression of $j_{m,\delta}$ as an element of $L(\mathcal{S})^{\Sigma}[\, j_{1,\delta}]$ in a compact representation inspired from [GHK+06].

In this case, we will often keep only the (negative of the) constant term in (2.20), and consider the modular equations $\Psi_{\delta,m}$ for $2 \leq m \leq n+1$ as "univariate" polynomials, i.e. elements of the ring $L(J_1, \dots, J_n)[J_{n+1}, Y]$ of degree at most $e-1$ in J_{n+1}, defined by the relations

$$\Psi_{\delta,m}(j_1, \dots, j_{n+1}) = \sum_{\gamma \in K_0/K'} \big(\gamma \cdot j_{m,\delta}\big) \prod_{\gamma_1 \neq \gamma} \big(Y - \gamma_1 \cdot j_{1,\delta}\big). \tag{2.21}$$

Then, we simply have $j_{m,\delta} = \Psi_{\delta,m}(j_{1,\delta}) / \partial_{Y_1} \Psi_{\delta,1}(j_{1,\delta})$.

2.4.4 Modular equations for abelian surfaces

To conclude chapter 2, we show that modular equations of Siegel and Hilbert type in dimension 2 [BL09, Mil15, Mar20, MR20] are special cases of modular equations as defined above.

Modular equations of Siegel type. We saw in §2.3.1 that the Siegel threefold $\mathrm{Sp}_4(\mathbb{Z})\backslash\mathbb{H}_2$ is the PEL Shimura variety associated with the algebraic group $G = \mathrm{GSp}_4(\mathbb{Q})$, with compact open subgroup $K = \mathrm{GSp}_4(\widehat{\mathbb{Z}})$. The Shimura variety Sh_K is connected, defined over $\mathbb{Q}$, and its function field is generated by the three algebraically independent Igusa invariants j_1, j_2, j_3.

Let ℓ be a prime, and consider the Hecke correspondence of level

$$\delta = \begin{pmatrix} \ell\, I_2 & 0 \\ 0 & I_2 \end{pmatrix} \qquad \text{as a } 4\times 4 \text{ matrix in } 2\times 2 \text{ blocks.}$$

The subgroup $K \cap \delta K \delta^{-1} \cap G(\mathbb{Q})_+$ of $\mathrm{Sp}_4(\mathbb{Z})$ is usually denoted by $\Gamma^0(\ell)$, and the degree of H_δ is

$$d(\delta) = \ell^3 + \ell^2 + \ell + 1.$$

The Hecke correspondence H_δ is absolutely irreducible, and describes all principally polarized abelian surfaces ℓ-isogenous to a given one; the degree of these isogenies is $l(\delta) = \ell^2$. In this case, the function $j_{1,\delta}$ generates the function field on the Hecke correspondence [BL09, Lem. 4.2], so that $d_1 = d(\delta)$ and $d_2 = d_3 = 1$, in the notation of §2.4.3. The modular equations from definition 2.4.6 (or rather remark 2.4.10) are the usual modular equations of Siegel type and level ℓ, denoted by $\Psi_{\ell,m} \in \mathbb{Q}(J_1, J_2, J_3)[Y]$ for $1 \le m \le 3$. They have been computed for $\ell = 2$ and $\ell = 3$ [Mil15].

By the dictionary between modular forms in the sense of PEL Shimura varieties and classical modular forms, we have the following formulæ between meromorphic functions on $\mathrm{Sp}_4(\mathbb{Z})\backslash\mathbb{H}_1$ [BL09, Mil15]:

$$\Psi_{\ell,1}\big(j_1(\tau), j_2(\tau), j_3(\tau)\big) = \prod_{\gamma \in \Gamma^0(\ell)\backslash \mathrm{Sp}_4(\mathbb{Z})} \Big(Y - j_1\big(\tfrac{1}{\ell}\gamma\tau\big)\Big),$$

and for $m \in \{2, 3\}$,

$$\Psi_{\ell,m}\big(j_1(\tau), j_2(\tau), j_3(\tau)\big) = \sum_{\gamma \in \Gamma^0(\ell)\backslash \mathrm{Sp}_4(\mathbb{Z})} j_m\big(\tfrac{1}{\ell}\gamma\tau\big) \prod_{\gamma' \neq \gamma} \Big(Y - j_1\big(\tfrac{1}{\ell}\gamma'\tau\big)\Big).$$

Modular equations of Hilbert type. As we saw in §2.3.3, the Hilbert surface $\Gamma_F(1)\backslash\mathbb{H}_1^2$ for the quadratic field $F = \mathbb{Q}(\sqrt{5})$ is identified with a connected component $\mathcal{S}$ of the PEL Shimura variety associated with the algebraic group $G = \mathrm{GL}_2(F)$ and a well-chosen congruence subgroup K of $\mathrm{GL}(\Lambda_0 \otimes \widehat{\mathbb{Z}})$, where $\Lambda_0 = \mathbb{Z}_F \oplus \partial_F^{-1}$.

Let Σ be the group of order two generated by the real conjugation σ in F. The algebraic variety $\mathcal{S}$ is defined over $\mathbb{Q}$, and the field of symmetric modular functions on $\mathcal{S}$ is generated by the two algebraically independent Gundlach invariants g_1, g_2.

Let $\beta \in \mathbb{Z}_F$ be totally positive and prime, and let

$$\delta = \begin{pmatrix} \beta & 0 \\ 0 & 1 \end{pmatrix} \in G(\mathbb{A}_f).$$

The Hecke correspondence H_δ is absolutely irreducible, and has degree $d(\delta) = N_{F/\mathbb{Q}}(\beta) + 1$. In the modular interpretation, H_δ parametrizes β-isogenies between abelian surfaces with real multiplication by $\mathbb{Z}_F$; the degree of these isogenies is $l(\delta) = N_{F/\mathbb{Q}}(\beta)$. One can check that H_δ intersects $\mathcal{S} \times \mathcal{S}$ nontrivially. Being able to consider this Hecke correspondence is the reason for our choice of G in §2.2.2.

The associated modular equations are called the *Hilbert modular polynomials* of level β in Gundlach invariants, denoted by $\Psi_{\beta,m} \in \mathbb{Q}(J_1, J_2)[Y]$ for $m \in \{1, 2\}$ [MR20]. They have been computed up to $N_{F/\mathbb{Q}}(\beta) = 59$ [Mil]; since the Gundlach invariants are symmetric, they describe both β- and $\sigma(\beta)$-isogenies. In order to write out Hilbert modular equations of level β in terms of classical Hilbert modular forms, we define

$$\Gamma^0(\beta) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_F(1) \colon \beta | b \right\},$$

and for all $\lambda \in F$ and $\tau = (\tau_1, \tau_2) \in \mathbb{H}_1^2$, we write

$$\lambda\tau := (\lambda\tau_1, \overline{\lambda}\tau_2).$$

Then we have the following equalities of meromorphic functions on $\mathbb{H}_1^2$:

$$\Psi_{\beta,1}\big(g_1(\tau), g_2(\tau)\big) = \prod_{\gamma \in \Gamma^0(\beta)\backslash(\Gamma_F(1) \rtimes \Sigma)} \Big(Y - g_1\big(\tfrac{1}{\beta}\gamma\tau\big)\Big), \qquad \text{and}$$

$$\Psi_{\beta,2}\big(g_1(\tau), g_2(\tau)\big) = \prod_{\gamma \in \Gamma^0(\beta)\backslash(\Gamma_F(1) \rtimes \Sigma)} g_2\big(\tfrac{1}{\beta}\gamma\tau\big) \prod_{\gamma' \neq \gamma} \Big(Y - g_1\big(\tfrac{1}{\beta}\gamma'\tau\big)\Big).$$

Other types of Hilbert modular equations for $\mathbb{Q}(\sqrt{5})$ using nonsymmetric invariants have also been defined and computed [Mar20].

Finally, we can also consider modular equations of Hilbert type using pullbacks of the three Igusa invariants under the Hilbert embedding as invariants on $\mathcal{S}$. The equation (2.17) satisfied by j_1, j_2, j_3 on $\mathcal{S}$ is the Humbert equation attached to F. When $F = \mathbb{Q}(\sqrt{5})$, there is no reason to use Hilbert modular equations in Igusa invariants in practice, as they have a larger number of variables and larger degrees compared to the modular equations in Gundlach invariants; their only advantage is that they can be defined for every real quadratic field F.

Chapter 3

Computing isogenies between abelian surfaces

This chapter corresponds to the preprint [KPR19]. We present an algorithm solving the following problem: given two principally polarized abelian surfaces A and A' over a field k which are isogenous, compute such an isogeny explicitly. More precisely, we assume that either

1. A and A' are ℓ-isogenous, where ℓ is a prime (the *Siegel case*); or
2. A and A' have real multiplication by $\mathbb{Z}_F$, where F is a fixed real quadratic field, and are β-isogenous where $\beta \in \mathbb{Z}_F$ is a totally positive prime (the *Hilbert case*).

We refer to §2.1.4 for the definition of ℓ- and β-isogenies. During the algorithm, we make several genericity assumptions on A and A'; in particular we assume that A and A' are Jacobians of genus 2 hyperelliptic curves over $\bar{k}$. We also assume that the characteristic of k is sufficiently large with respect to ℓ or β. We give precise statements in theorems 3.5.2 and 3.5.3.

While the Siegel and Hilbert cases do not cover all possible isogenies between principally polarized abelian surfaces, they "almost" do so, and the isogeny algorithm can be adapted to the cases when $\mathrm{End}(A)^\dagger$ is not maximal or when $\beta \in \mathbb{Z}_F$ is totally positive but not prime; see [DJRV17, Thm. 1.1] for the full classification.

Our algorithm is the generalization of the isogeny algorithm for elliptic curves described in §1.2.2 to the case of modular equations of Siegel and Hilbert type for abelian surfaces: the crucial step is to compute the action of the isogeny on differential forms using derivatives of modular equations. This computation has a geometric interpretation in terms of the Kodaira–Spencer isomorphism, and this formulation yields a clear roadmap to generalize the isogeny algorithm to higher dimensions.

Using the results of this chapter, one can already implement toy examples of Elkies's point counting method on abelian surfaces using precomputed modular equations of small levels. However, the key issue of the *complexity* of manipulating modular equations is left aside for the moment. We return to this question in chapter 5, where we present general size bounds for modular equations in terms of their level, and in chapter 6, where we investigate the complexity of *evaluating* modular equations of Siegel and Hilbert type, i.e. providing the input to the isogeny algorithm.

3.1 Principle of the algorithm

Let us describe the outline of our algorithm in the case of ℓ-isogenies from a geometric point of view, in any dimension g. The central object is the Hecke correspondence

$$H_\ell = (H_{\ell,1}, H_{\ell,2})\colon \mathcal{A}_{g,\ell} \to \mathcal{A}_g \times \mathcal{A}_g$$

where $\mathcal{A}_{g,\ell}$ denotes the moduli space of principally polarized abelian varieties of dimension g endowed with the kernel of an ℓ-isogeny, and $\mathcal{A}_g$ denotes the moduli space of principally polarized abelian varieties of dimension g. In the modular interpretation, the map H_ℓ is given by $(A, K) \mapsto (A, A/K)$. Over $\mathbb{C}$, the map H_ℓ takes the form

$$\begin{array}{ccc} \Gamma^0(\ell)\backslash\mathbb{H}_g & \to & \mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g \times \mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g \\ \tau & \mapsto & (\tau, \tau/\ell), \end{array}$$

where the subgroup $\Gamma^0(\ell) \subset \mathrm{Sp}_{2g}(\mathbb{Z})$ is defined by

$$\Gamma^0(\ell) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sp}_{2g}(\mathbb{Z})\colon b = 0 \bmod \ell \right\}. \tag{3.1}$$

If $g = 2$, then H_ℓ is exactly the Hecke correspondence used in §2.4.4 to construct the Siegel modular equations of level ℓ.

Let $\varphi\colon A \to A'$ be an ℓ-isogeny, so that (A, A') lies in the image of H_ℓ. Assume further that (A, A') are sufficiently generic, so that $H_{\ell,1}$ and $H_{\ell,2}$ are étale at (A, A'). Denote by $T_A(\mathcal{A}_g)$ the tangent space of $\mathcal{A}_g$ at A, and denote by $T_0(A)$ the tangent space of A at its neutral point. Then there is a close relation between two maps:

- the *deformation map* $\mathcal{D}(\varphi)\colon T_A(\mathcal{A}_g) \to T_{A'}(\mathcal{A}_g)$ defined as

$$\mathcal{D}(\varphi) := dH_{\ell,2} \circ (dH_{\ell,1})^{-1}, \quad \text{and}$$

- the *tangent map* $d\varphi\colon T_0(A) \to T_0(A')$.

This relation stems from a canonical isomorphism, called the *Kodaira–Spencer map*, between $T_A(\mathcal{A}_g)$ and the vector space $\mathrm{Sym}^2 T_0(A)$ [And17]. Therefore, in any dimension g, an isogeny algorithm could run as follows.

1. Compute the deformation map at (A, A') by differentiating certain modular equations giving a local model of $\mathcal{A}_{g,\ell}$ and $\mathcal{A}_g$.

2. Compute $d\varphi$ from the deformation map using an explicit version of the Kodaira–Spencer isomorphism, i.e. an explicit way to map a pair (A, w) where $w \in \mathrm{Sym}^2 T_0(A)$ to the corresponding point of $T_A(\mathcal{A}_g)$ in the local model of $\mathcal{A}_g$.

3. Reconstruct φ as a rational map by solving a differential system in power series and performing a multivariate rational reconstruction.

In step 3, the characteristic p of k should be large with respect to ℓ. If p is too small, then a standard solution [JL06] is to lift the isogeny to the ring of Witt vectors of k, which has characteristic zero, and to control the p-adic precision losses in step 3. This outline follows the steps of the isogeny algorithm for elliptic curves (§1.2.2), albeit in a very abstract formulation.

We mention that the genericity assumption on (A, A') can be removed if one is willing to work with $\mathcal{A}_g$ as a *Deligne–Mumford stack*, rather than a scheme over k; then the maps $H_{\ell,1}$ and $H_{\ell,2}$ are étale everywhere. We refer to [KPR19, §4] for an account on this formalism. In practice, working with stacks would involve adding a level structure of level $n \geq 3$ prime to ℓ and keeping track of automorphisms. For simplicity, we choose to keep the genericity hypothesis.

In the case $g = 2$, we consider the local model of $\mathcal{A}_g$ given by the Igusa invariants, defined in §2.3.2; the local model of $\mathcal{A}_{g,\ell}$ in step 1 is given by the Siegel modular equations of level ℓ (recall §1.3.1 and §2.4.4). This choice introduces further singularities, which can be avoided by choosing other models of $\mathcal{A}_g$; this would be necessary for instance if A or A' is a product of elliptic curves, so that its Igusa invariants are not defined. In order to compute the deformation map, it is enough to evaluate the Siegel modular equations of level ℓ and their derivatives at (A, A'). In step 2, we choose to encode a basis of $T_0(A)$ as the choice of a hyperelliptic curve equation $\mathcal{C}$ such that $A = \mathrm{Jac}(\mathcal{C})$. Then, the explicit Kodaira–Spencer isomorphism is simply an expression for certain Siegel modular forms, namely derivatives of Igusa invariants, in terms of the coefficients of the curve (see theorem 3.2.15). Since the curve $\mathcal{C}$ embeds in its Jacobian, we are able to compute with power series in only one variable in step 3.

3.2 An explicit Kodaira–Spencer map

Let $\mathcal{C}$ be a genus 2 hyperelliptic curve over $\mathbb{C}$. Then a choice of hyperelliptic equation for $\mathcal{C}$ defines a basis ω of the space of global differential forms on $\mathcal{C}$, hence on $\mathrm{Jac}(\mathcal{C})$. As we saw in remark 2.3.3, any Siegel modular form f with values in a vector space W can be evaluated on the pair $(\mathrm{Jac}(\mathcal{C}), \omega)$. This gives rise to a map, denoted by $\mathrm{Cov}(f)$, from an open subset of $\mathbb{C}_6[x]$ to W. We show that $\mathrm{Cov}(f)$ is a *covariant* of the hyperelliptic equation; the main result of this section is an algorithm to obtain this covariant explicitly given the q-expansion of f. These calculations have also been done in [CFv17].

The existence of the Kodaira–Spencer isomorphism translates into the fact that derivatives of Igusa invariants are modular forms of weight Sym^2 (recall proposition 2.3.9). The explicit Kodaira–Spencer isomorphism that we are looking for is the expression of their associated covariants in terms of the coefficients of the curve.

We use the following notation. If m is a matrix, we denote the transpose of m by m^t, and we denote the inverse of m^t by m^{-t}. We denote the diagonal $n \times n$ matrix with diagonal entries $x_1, \ldots, x_n$ by $\mathrm{Diag}(x_1, \ldots, x_n)$.

3.2.1 Hyperelliptic equations

Let $\mathcal{C}$ be a smooth hyperelliptic *equation* of genus 2 over $\mathbb{C}$:

$$\mathcal{C} \colon v^2 = E_{\mathcal{C}}(u),$$

with $\deg E_{\mathcal{C}} \in \{5, 6\}$. Then $\mathcal{C}$ is naturally endowed with the basis of differential forms

$$\omega(\mathcal{C}) = \Big(\frac{u\,du}{v}, \frac{du}{v}\Big). \tag{3.2}$$

Recall from §2.1.2 that the Jacobian $\mathrm{Jac}(\mathcal{C})$ is a principally polarized abelian surface over $\mathbb{C}$. Choosing a base point P on $\mathcal{C}$ gives an embedding

$$\eta_P \colon \mathcal{C} \hookrightarrow \mathrm{Jac}(\mathcal{C}), \quad Q \mapsto [Q - P].$$

The pullback map η_P^* on the spaces of differential forms is an isomorphism which is independent of P. Therefore we can see $\omega(\mathcal{C})$ as a basis of differential forms on $\mathrm{Jac}(\mathcal{C})$. This basis depends on the particular hyperelliptic equation chosen.

Lemma 3.2.1. *Let $\mathcal{C}$ be a genus 2 hyperelliptic equation over $\mathbb{C}$, and let $r = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{C})$. Let $E_{\mathcal{C}'}$ be the image of $E_{\mathcal{C}}$ by $\det^{-2}\mathrm{Sym}^6(r)$, and*

let $\mathcal{C}'$ be the genus 2 hyperelliptic curve defined by the equation $y'^2 = E_{\mathcal{C}'}(x')$. Let $\eta\colon \mathcal{C}' \to \mathcal{C}$ be the isomorphism defined by

$$\eta(x', y') = \left(\frac{ax' + c}{bx' + d}, \frac{(\det r)\, y'}{(bx' + d)^3} \right).$$

Then the matrix of $\eta^*\colon \Omega^1(\mathcal{C}) \to \Omega^1(\mathcal{C}')$ in the bases $\omega(\mathcal{C})$ and $\omega(\mathcal{C}')$ is r.

Proof. Write $(x, y) = \eta(x', y')$. Then direct calculation shows that

$$\frac{dx}{y} = (bx' + d)\frac{dx'}{y'} \quad \text{and} \quad \frac{x\, dx}{y} = (ax' + c)\frac{dx'}{y'}. \qquad \square$$

Proposition 3.2.2. *Let A be a principally polarized abelian surface over $\mathbb{C}$ that is not a product of two elliptic curves, and let ω be a basis of $\Omega^1(A)$. Then there exists a unique hyperelliptic curve equation $\mathcal{C}$ of genus 2 over $\mathbb{C}$ such that $\big(\mathrm{Jac}(\mathcal{C}), \omega(\mathcal{C})\big)$ is isomorphic to (A, ω).*

Proof. Since A is not a product of elliptic curves, there exists a genus 2 curve equation $\mathcal{C}_0$ over $\mathbb{C}$ such that A is isomorphic to $\mathrm{Jac}(\mathcal{C}_0)$. Then ω differs from $\omega(\mathcal{C}_0)$ by a linear transformation in $\mathrm{GL}_2(\mathbb{C})$. By lemma 3.2.1, we can make a suitable change of variables on $\mathcal{C}_0$ to find the correct $\mathcal{C}$. The curve $\mathcal{C}$ is unique because every isomorphism between hyperelliptic curves comes from a matrix r as in lemma 3.2.1. $\square$

Let $g \geq 1$. Recall that the complex abelian variety of dimension g attached to $\tau \in \mathbb{H}_g$ is $A(\tau) = \mathbb{C}^g/\Lambda(\tau)$, where $\Lambda(\tau) = \mathbb{Z}^g \oplus \tau\mathbb{Z}^g$. Recall also that for every $\tau \in \mathbb{H}_g$ and every $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sp}_{2g}(\mathbb{Z})$ in $g \times g$ blocks, we write

$$\gamma\tau = (a\tau + b)(c\tau + d)^{-1} \quad \text{and} \quad \gamma^*\tau = c\tau + d.$$

There is a natural basis of differential forms on $A(\tau)$ given by

$$\omega(\tau) = (2\pi i\, dz_1, \ldots, 2\pi i\, dz_g), \tag{3.3}$$

where $z_1, \ldots, z_g$ are the coordinates of $\mathbb{C}^g$. This basis of differential forms satisfies a simple transformation rule under the action of $\mathrm{Sp}_{2g}(\mathbb{Z})$.

Proposition 3.2.3 ([BL04, Rem. 8.1.4])**.** *Let $\tau \in \mathbb{H}_g$, and let $\gamma \in \mathrm{Sp}_{2g}(\mathbb{Z})$. Then there exists an isomorphism*

$$\eta_{\gamma,\tau}\colon A(\tau) \to A(\gamma\tau), \quad z \mapsto (\gamma^*\tau)^{-t} z.$$

Similarly, let F be a totally real number field of degree g, seen as a subset of $\mathbb{R}^g$ via its g real embeddings. By §2.3.3, for every $\tau = (\tau_1, \dots, \tau_g) \in \mathbb{H}_1^g$, the abelian variety of dimension g with real multiplication by $\mathbb{Z}_F$ attached to τ is $A_F(\tau) = \mathbb{C}^g/\Lambda_F(\tau)$, where $\Lambda_F(\tau) = \mathbb{Z}_F \oplus \mathrm{Diag}(\tau_1, \dots, \tau_g)\partial_F^{-1}$. The tuple of differential forms $\omega(\tau)$ defined as in eq. (3.3) also defines a basis of differential forms on $A_F(\tau)$.

Definition 3.2.4. Let us return to the case $g = 2$. The bases of differential forms defined above allow us to define particular curve equations attached to a point of $\mathbb{H}_2$ or $\mathbb{H}_1^2$.

1. Let $\tau \in \mathbb{H}_2$, and assume that $\chi_{10}(\tau) \neq 0$. Then, by proposition 3.2.2, there exists a unique hyperelliptic equation $\mathcal{C}(\tau)$ over $\mathbb{C}$ such that

$$\big(\mathrm{Jac}(\mathcal{C}(\tau)), \omega(\mathcal{C}(\tau))\big) \simeq \big(A(\tau), \omega(\tau)\big).$$

   We call $\mathcal{C}(\tau)$ the *standard curve* attached to τ. We define the functions $a_i(\tau)$ on $\mathbb{H}_2$ for $0 \leq i \leq 6$ as the coefficients of $\mathcal{C}(\tau)$:

$$\mathcal{C}(\tau)\colon y^2 = \sum_{i=0}^{6} a_i(\tau) x^i.$$

2. Let F be a real quadratic field, and let $R \in \mathrm{GL}_2(\mathbb{R})$ be a matrix defining a Hilbert embedding (§2.3.3). Let $\tau \in \mathbb{H}_1^2$, and assume that $\chi_{10}(H_R(\tau)) \neq 0$. Then, by proposition 3.2.2, there exists a unique hyperelliptic equation $\mathcal{C}_F(\tau)$ over $\mathbb{C}$ such that

$$\big(\mathrm{Jac}(\mathcal{C}_F(\tau)), \omega(\mathcal{C}_F(\tau))\big) \simeq \big(A_F(\tau), \omega(\tau)\big).$$

   We call $\mathcal{C}_F(\tau)$ the *standard curve with real multiplication by* $\mathbb{Z}_F$ attached to τ; this is an abuse of language, since the Jacobian of $\mathcal{C}_F(\tau)$ has real multiplication by $\mathbb{Z}_F$, not the curve itself.

Proposition 3.2.5. *The function $\tau \mapsto \mathcal{C}(\tau)$ is a Siegel modular function of weight $\det^{-2}\mathrm{Sym}^6$ which has no poles on the open set $\mathcal{V} = \{\chi_{10} \neq 0\} \subset \mathbb{H}_2$.*

Proof. Let $\mathcal{Z}$ be the set of isomorphism classes of pairs $(\mathrm{Jac}(\mathcal{C}), \omega)$ where $\mathcal{C}$ is a genus 2 hyperelliptic curve over $\mathbb{C}$ and ω is a basis of $\Omega^1(\mathcal{C})$. By proposition 3.2.3, $\mathcal{Z}$ is the quotient of $\mathcal{V} \times \mathrm{GL}_2(\mathbb{C})$ by the action of $\mathrm{Sp}_4(\mathbb{Z})$ given by

$$\gamma \cdot (\tau, r) = \big(\gamma\tau, (\gamma^*\tau)^t\, r\big),$$

and thus inherits a natural complex structure. Let $\mathcal{U} \subset \mathbb{C}_6[x]$ be the open set consisting of polynomials with simple roots. Then the Torelli map

from $\mathcal{U}$ to $\mathcal{Z}$ sending $\mathcal{C}$ to $(\mathrm{Jac}(\mathcal{C}), \omega(\mathcal{C}))$, which is bijective by proposition 3.2.2, is actually biholomorphic. The map $\mathcal{V} \to \mathcal{Z}$ sending τ to (τ, I_2) is holomorphic, therefore $\tau \mapsto \mathcal{C}(\tau)$ is holomorphic on $\mathcal{V}$. Combining proposition 3.2.3 and lemma 3.2.1 shows the transformation rule. □

Finally, one can relate the curves $\mathcal{C}_F(\tau)$ and $\mathcal{C}(H_R(\tau))$ for every $\tau \in \mathbb{H}_1^2$.

Proposition 3.2.6. *Let F, R be as above. Then for every $\tau \in \mathbb{H}_1^2$, left multiplication by R^t on $\mathbb{C}^2$ induces an isomorphism $A_F(\tau) \to A(H_R(\tau))$.*

Proof. As above, we view F as a subset of $\mathbb{R}^2$ via its two real embeddings. By definition of R, we have $\mathbb{Z}_F = R\,\mathbb{Z}^2$ and $\partial_F^{-1} = R^{-t}\,\mathbb{Z}^2$. Then a direct computation shows that

$$\forall t \in \mathbb{H}_1^2,\ \Lambda(H_R(\tau)) = R^t\,\Lambda_F(\tau).$$ □

Proposition 3.2.7. *Let F, R be as above. Then for every $\tau \in \mathbb{H}_1^2$, we have*

$$\mathcal{C}_F(\tau) = \det{}^{-2}\,\mathrm{Sym}^6(R)\,\mathcal{C}(H_R(\tau)).$$

Proof. Combine proposition 3.2.6 and lemma 3.2.1. □

3.2.2 Covariants

As indicated in remark 2.3.3, if f is a Siegel modular form of weight ρ in dimension g, if A is a principally polarized abelian variety of dimension g over $\mathbb{C}$, and if ω is a basis of $\Omega^1(A)$, then $f(A, \omega)$ is well defined. To compute this quantity, choose $\tau \in \mathbb{H}_g$ and an isomorphism $\eta\colon A \to A(\tau)$. Let $r \in \mathrm{GL}_g(\mathbb{C})$ be the matrix of the pullback $\eta^*\colon \Omega^1(A(\tau)) \to \Omega^1(A)$ in the bases $\omega(\tau)$ and ω. Then

$$f(A, \omega) = \rho(r) f(\tau).$$

We can check using proposition 3.2.3 that $f(A, \omega)$ does not depend on the choice of τ and η. As a consequence, for every Siegel modular form f in dimension 2, we have a map

$$\mathrm{Cov}(f)\colon \mathcal{C} \mapsto f\big(\mathrm{Jac}(\mathcal{C}), \omega(\mathcal{C})\big).$$

We apply the results of §3.2.1 to show that $\mathrm{Cov}(f)$ is a covariant of the curve equation. A recent reference on covariants is [Mes91].

Definition 3.2.8. Let $\rho\colon \mathrm{GL}_2(\mathbb{C}) \to \mathrm{GL}(V)$ be a finite-dimensional holomorphic representation of $\mathrm{GL}_2(\mathbb{C})$. A *covariant*, or *polynomial covariant*, of weight ρ is a map

$$C\colon \mathbb{C}_6[x] \to V$$

which is polynomial in the coefficients, and such that the following transformation rule holds: for every $r \in \mathrm{GL}_2(\mathbb{C})$ and $P \in \mathbb{C}_6[x]$,

$$C\big(\det{}^{-2}\,\mathrm{Sym}^6(r)\,P\big) = \rho(r)\,C(P). \tag{3.4}$$

If $\dim V \geq 2$, then C is said to be *vector-valued*, and otherwise *scalar-valued*. A *fractional covariant* is a map satisfying (3.4) which is only required to have a fractional expression in terms of the coefficients.

As for Siegel modular forms, it is enough to consider covariants of weight $\det^k \mathrm{Sym}^n$ for $k \in \mathbb{Z}$ and $n \in \mathbb{N}$. What we call a vector-valued covariant of weight $\det^k \mathrm{Sym}^n$ is called in [Mes91] a covariant of order n and degree $k + n/2$; what we call a scalar-valued covariant of weight $\det^k$ is called in [Mes91] an invariant of degree k. The reason for this change of terminology is the following.

Proposition 3.2.9. *Let ρ be a representation of $\mathrm{GL}_2(\mathbb{C})$ as above. If f is a Siegel modular function of weight ρ, then $\mathrm{Cov}(f)$ is a fractional covariant of weight ρ. Conversely, if F is a fractional covariant of weight ρ, then the meromorphic function $\tau \mapsto F(\mathcal{C}(\tau))$ is a Siegel modular function of weight ρ. These operations are inverse of each other.*

Proof. If f is a Siegel modular function, then $\mathrm{Cov}(f)$ is well defined on a Zariski open set of $\mathbb{C}_6[x]$ and is algebraic. Therefore $\mathrm{Cov}(f)$ has a fractional expression in terms of the coefficients. It is easy to check the transformation rules using lemma 3.2.1 and proposition 3.2.5. □

Proposition 3.2.9 gives a bijection between Siegel modular *functions* and *fractional* covariants, but we need more. The following theorem establishes a relation between Siegel modular *forms* and *polynomial* covariants, and was first proved in [CFv17, §4].

Theorem 3.2.10. *Let f be a Siegel modular form. Then $\mathrm{Cov}(f)$ is a polynomial covariant. Moreover, if f is a cusp form, then $\mathrm{Cov}(f/\chi_{10})$ is also a polynomial covariant.*

Sketch of proof. The main difficulty is that singular hyperelliptic equations form a codimension 1 subset of all degree 6 polynomials: therefore, if f is a Siegel modular form, then the proof of proposition 3.2.9 only

shows that $\mathrm{Cov}(f)$ is a polynomial divided by some power of the discriminant. However, one can show that f extends to a certain compactification of $\mathcal{A}_2(\mathbb{C})$ called the *toroidal compactification*, and this implies that $\mathrm{Cov}(f)$ is well defined on all curve equations with at most one node. Since the complementary of this set has codimension 2, the result follows. □

Unlike the graded algebra of vector-valued Siegel modular forms, the graded $\mathbb{C}$-algebra generated by polynomial covariants is finitely generated.

Theorem 3.2.11 ([Cle72, p. 296])**.** *The graded $\mathbb{C}$-algebra of covariants is generated by 26 elements defined over $\mathbb{Q}$. The number of generators of weight $\det^k \mathrm{Sym}^n$ is indicated in the following table:*

$n \setminus k$	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	15
0						1		1		1				1		1
2						1		1		1	1		1		1	
4				1		1	1		1		1					
6		1		1	1		2									
8		1	1		1											
10			1													
12	1															

We only need to manipulate a small subset of these generators. Take our scalar generators of even weight to be the Igusa–Clebsch covariants I_2, I_4, I_6, I_{10}, in Mestre's notation A', B', C', D' [Mes91], and set

$$I_6' := (I_2 I_4 - 3I_6)/2.$$

Denote the generator of weight $\det^{15}$ by R, and denote by y_1, y_2, y_3 the generators of weights $\det^2 \mathrm{Sym}^2$, $\det^4 \mathrm{Sym}^2$, and $\det^6 \mathrm{Sym}^2$ respectively; they are constructed explicitly in [Mes91, §1]. Note that the integers m and n on page 315 of [Mes91] should be the *orders* of f and g, and not their degrees. As a sanity check, we mention that the coefficient of $a_1^5 a_4^{10}$ in R is $2^{-2}3^{-6}5^{-10}$. Finally, the generator of weight $\det^{-2} \mathrm{Sym}^6$, denoted by X, is the degree 6 polynomial itself.

3.2.3 From q-expansions to covariants

We now explain how to compute the polynomial covariant associated with a Siegel modular form whose q-expansion is known up to a certain precision. The works of Igusa already provide the answer in the case of scalar-valued covariants. Let $\psi_4, \psi_6, \chi_{10}$ and χ_{12} be the Siegel modular

forms defined in §2.3.2, and let χ_{35} be the unique Siegel modular form of weight 35 [Igu62] whose q-expansion starts as follows:

$$\chi_{35}(\tau) = q_1^2 q_2^2 (q_1 - q_2)(q_3 - q_3^{-1}) + O(q_1^4, q_2^4).$$

The modular form χ_{35} also has an explicit expression in terms of genus 2 theta constants.

Theorem 3.2.12. *We have*

$$\begin{aligned} 4\ \mathrm{Cov}(\psi_4) &= I_4, \\ 4\ \mathrm{Cov}(\psi_6) &= I_6', \\ 2^{12}\ \mathrm{Cov}(\chi_{10}) &= I_{10}, \\ 2^{15}\ \mathrm{Cov}(\chi_{12}) &= I_2 I_{10}, \quad \textit{and} \\ 2^{37} 3^{-9} 5^{-10}\ \mathrm{Cov}(\chi_{35}) &= I_{10}^2 R. \end{aligned}$$

Proof. By [Igu62, p. 848], there exists a constant $\lambda \in \mathbb{C}^\times$ such that these relations hold up to a factor of λ^k, for $k \in \{4, 6, 10, 12, 35\}$ respectively. Note that Igusa's covariant E is $-2^5 3^9 5^{10} R$. In order to find λ, we apply Thomae's formulæ [Mum84, Thm. IIIa.8.1], [Tho70, pp. 216–217] on a hyperelliptic curve $\mathcal{C}_0$ with real roots; for instance

$$\mathcal{C}_0 \colon y^2 = x(x-1)(x-2)(x-3)(x-4)(x-5).$$

This yields the values of both the left and right hand sides of theorem 3.2.12 for the curve $\mathcal{C}_0$. We obtain that $\lambda = 1$. □

In particular, the Igusa invariants satisfy

$$\mathrm{Cov}(j_1) = \frac{I_4 I_6'}{I_{10}}, \quad \mathrm{Cov}(j_2) = \frac{I_2 I_4^2}{I_{10}}, \quad \mathrm{Cov}(j_3) = \frac{I_4^5}{I_{10}^2}. \tag{3.5}$$

In order to give an analogue of theorem 3.2.12 for vector-valued modular forms, we compute the q-expansion of the standard curve $\mathcal{C}(\tau)$. Recall the Siegel modular form $f_{8,6}$ of weight $\det^8 \mathrm{Sym}^6$ introduced in example 2.3.10.

Proposition 3.2.13. *In the notation of §3.2.2, we have* $\mathrm{Cov}(f_{8,6}/\chi_{10}) = X$. *In other words, for every* $\tau \in \mathbb{H}_2$ *such that* $\chi_{10}(\tau) \neq 0$, *we have*

$$\mathcal{C}(\tau) = \frac{f_{8,6}(\tau)}{\chi_{10}(\tau)}.$$

Proof. Since $f_{8,6}$ is a cusp form, by theorem 3.2.10, $\mathrm{Cov}(f_{8,6}/\chi_{10})$ is a nonzero polynomial covariant of weight $\det^{-2}\mathrm{Sym}^6$. By theorem 3.2.11, this space of covariants has dimension 1 and is generated by X, so the relation holds up to a factor $\lambda \in \mathbb{C}^\times$. This yields q-expansions for the coefficients $a_i(\tau)$ of $\mathcal{C}(\tau)$ up to a factor λ. Then, the relations from theorem 3.2.12 imply $\lambda^4 = \lambda^6 = \lambda^{35} = 1$, hence $\lambda = 1$. □

Given a Siegel modular form f of weight ρ defined over $\mathbb{Q}$ whose q-expansion can be computed, the following algorithm recovers the expression of $\mathrm{Cov}(f)$ as a polynomial.

Algorithm 3.2.14.
1. Compute a basis $\mathcal{B}$ of the vector space of polynomial covariants of weight ρ using theorem 3.2.11.
2. Choose a precision $n \geq 1$ and compute the q-expansion of f modulo the ideal (q_1^n, q_2^n) in $\mathbb{Q}(q_3)[[q_1, q_2]]$.
3. For every $B \in \mathcal{B}$, compute the q-expansion of the Siegel modular function $\tau \mapsto B(\mathcal{C}(\tau))$ using proposition 3.2.13.
4. Do linear algebra; if the matrix does not have full rank, go back to step 2 with a larger n.

Sturm-type bounds [BGP17] provide a theoretical limit for the precision n that we need to consider; for the examples given in this thesis, taking $n = 3$ is sufficient.

We now apply algorithm 3.2.14 to derivatives of Igusa invariants. Recall from proposition 2.3.9 that for each $1 \leq k \leq 3$, the derivative

$$\frac{dj_k}{d\tau} := \frac{1}{2\pi i}\Big(\frac{\partial j_k}{\partial z_1}x^2 + \frac{\partial j_k}{\partial z_3}x + \frac{\partial j_k}{\partial z_2}\Big),$$

where we write $\tau = \begin{pmatrix} z_1(\tau) & z_3(\tau) \\ z_3(\tau) & z_2(\tau) \end{pmatrix}$, is a modular function of weight Sym^2.

Theorem 3.2.15. *In the notation of §3.2.2, we have*

$$\begin{aligned}
\mathrm{Cov}\Big(\frac{dj_1}{d\tau}\Big) &= \frac{1}{I_{10}}\Big(\frac{153}{8}I_2^2I_4y_1 - \frac{135}{2}I_2I_6y_1 + \frac{135}{2}I_4^2y_1 + \frac{46575}{4}I_2I_4y_2 \\
&\qquad - 30375\,I_6y_2 + 1366875\,I_4y_3\Big), \\
\mathrm{Cov}\Big(\frac{dj_2}{d\tau}\Big) &= \frac{1}{I_{10}}\big(90\,I_2^2I_4y_1 + 900\,I_2^2y_1 + 40500\,I_2I_4y_2\big), \quad \textit{and} \\
\mathrm{Cov}\Big(\frac{dj_3}{d\tau}\Big) &= \frac{1}{I_{10}^2}\big(225\,I_2I_4^4y_1 + 101250\,I_4^4y_2\big).
\end{aligned}$$

Proof. Let $1 \leq k \leq 3$. The function $\chi_{10}^2 j_k$ has no poles on $\mathcal{A}_2(\mathbb{C})$. Therefore, the Siegel modular function

$$f_k = \chi_{10}^3 \frac{dj_k}{d\tau}$$

is holomorphic on $\mathcal{A}_2(\mathbb{C})$. Its q-expansion can be computed from the q-expansion of $\chi_{10}^2 j_k$ by formal differentiation. Since

$$\frac{1}{2\pi i}\frac{\partial}{\partial z_j} = q_j \frac{\partial}{\partial q_j}$$

for each $1 \leq j \leq 3$, we check that f_k is a cusp form defined over $\mathbb{Q}$. By theorem 3.2.10, $\mathrm{Cov}(f_k/\chi_{10})$ is a polynomial covariant of weight $\det^{20} \mathrm{Sym}^2$. By theorem 3.2.11, a basis of this space of covariants is given by covariants of the form Iy where $y \in \{y_1, y_2, y_3\}$ and I is a scalar-valued covariant of the appropriate even weight. Algorithm 3.2.14 succeeds with $n = 3$; the computations were done using Pari/GP [The19]. □

Remark 3.2.16. Theorems 3.2.12 and 3.2.15 can be checked numerically. Computing big period matrices of hyperelliptic curves using the algorithm of [MN19] provides pairs $(\tau, \mathcal{C}(\tau))$ with $\tau \in \mathbb{H}_2$. We can evaluate Igusa invariants and their derivatives at a given τ to high precision using their expression in terms of theta constants (see chapter 6); from this data, we identify the associated covariants to high precision using numerical linear algebra. The computations were done using the libraries `hcperiods` [Mol18] and `cmh` [ET14].

Remark 3.2.17. From theorem 3.2.15, we can compute the covariants associated with derivatives of other modular functions, or even invariants for abelian surfaces with extra structure such as theta constants. For instance, consider the invariants

$$h_1 = \frac{\psi_6^2}{\psi_4^3}, \quad h_2 = \frac{\chi_{12}}{\psi_4^3}, \quad h_3 = \frac{\chi_{10}\psi_6}{\psi_4^4}$$

which are generically well-defined on products of elliptic curves. Then we

obtain

$$\begin{aligned}
\mathrm{Cov}\Big(\frac{dh_1}{d\tau}\Big) &= \frac{1}{I_4^4}\Big(-\frac{297}{8}y_1I_4^2I_2^3 + -\frac{54675}{4}y_2I_4^2I_2^2 + \frac{1701}{8}y_1I_6I_4I_2^2 + \frac{135}{2}y_1I_4^3I_2 \\
&\quad + 1366875y_3I_4^2I_2 + \frac{346275}{4}y_2I_6I_4I_2 - \frac{1215}{4}y_1I_6^2I_2 + -\frac{405}{2}y_1I_6I_4^2 \\
&\quad - 4100625y_3I_6I_4 - \frac{273375}{2}y_2I_6^2\Big), \\
\mathrm{Cov}\Big(\frac{dh_2}{d\tau}\Big) &= \frac{1}{I_4^4}\big(-135y_1I_{10}I_2^2 - 60750y_2I_{10}I_2 + 900y_1I_{10}I_4\big), \\
\mathrm{Cov}\Big(\frac{dh_3}{d\tau}\Big) &= \frac{1}{I_4^5}\Big(-\frac{747}{8}y_1I_{10}I_4I_2^2 - \frac{155925}{4}y_2I_{10}I_4I_2 + 270y_1I_{10}I_6I_2 \\
&\quad + \frac{135}{2}y_1I_{10}I_4^2 + 1366875y_3I_{10}I_4 + 121500y_2I_{10}I_6\Big).
\end{aligned}$$

3.3 Computing the action on tangent spaces

Let $\varphi\colon \mathrm{Jac}(\mathcal{C}) \to \mathrm{Jac}(\mathcal{C}')$ be an ℓ-isogeny over $\mathbb{C}$, where ℓ is a prime. Using the explicit formula for the Kodaira–Spencer isomorphism provided by theorem 3.2.15, we relate the deformation map of φ, computed from derivatives of Siegel modular equations of level ℓ, to the tangent map of φ. Then we adapt the computations to the Hilbert case, and explain why the results remains valid over any field.

We use an explicit and well-known description of ℓ- and β-isogenies over $\mathbb{C}$: namely, they are all described by the Hecke correspondences introduced in §2.4.4.

Proposition 3.3.1. *Let ℓ be a prime number, let F be a real quadratic field, and let $\beta \in \mathbb{Z}_F$ be a totally positive prime.*

1. *For every $\tau \in \mathbb{H}_2$, the identity map on $\mathbb{C}^2$ induces an ℓ-isogeny*

$$A(\tau) \to A(\tau/\ell).$$

Let A, A' be principally polarized abelian surfaces over $\mathbb{C}$, and let $\varphi\colon A \to A'$ be an ℓ-isogeny. Then there exists $\tau \in \mathbb{H}_2$ such that there is a commutative diagram

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & A' \\ \downarrow\wr & & \downarrow\wr \\ A(\tau) & \xrightarrow{z\mapsto z} & A(\tau/\ell). \end{array}$$

where the vertical arrows are isomorphisms.

2. *For every $\tau = (\tau_1, \tau_2) \in \mathbb{H}_1^2$, write $\tau/\beta := (\tau_1/\beta, \tau_2/\overline{\beta})$. The identity map on $\mathbb{C}^2$ induces a β-isogeny*

$$A_F(\tau) \to A_F(\tau/\beta).$$

Let A, A' be principally polarized abelian surfaces over $\mathbb{C}$ with real multiplication by $\mathbb{Z}_F$, and let $\varphi\colon A \to A'$ be a β-isogeny. Then there exists $\tau \in \mathbb{H}_1^2$ such that there is a commutative diagram

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & A' \\ \downarrow\wr & & \downarrow\wr \\ A_F(\tau) & \xrightarrow{z\mapsto z} & A_F(\tau/\beta) \end{array}$$

where the vertical arrows are isomorphisms of principally polarized abelian surfaces with real multiplication.

We used the analogous statement for ℓ-isogenies between elliptic curves in chapter 1.

Proof. Let $\varphi\colon A \to A'$ be an ℓ-isogeny over $\mathbb{C}$. Then there exists a $\tau \in \mathbb{H}_2$ such that $A(\tau)$ is isomorphic to A. The kernel of φ, seen as a subgroup of $A(\tau)[\ell]$, is maximal isotropic for the symplectic form induced by the polarization; this property characterizes ℓ-isogenies [BL09, §3]. If $\ker\varphi$ is equal to $\Lambda(\tau/\ell)$, then we obtain the required diagram. Otherwise, we use the facts that $\mathrm{Sp}_4(\mathbb{Z}/\ell\mathbb{Z})$ acts transitively on maximal isotropic subgroups of $A[\ell]$, and that the reduction map $\mathrm{Sp}_4(\mathbb{Z}) \to \mathrm{Sp}_4(\mathbb{Z}/\ell\mathbb{Z})$ is surjective, to replace τ by a suitable $\mathrm{Sp}_4(\mathbb{Z})$-conjugate.

The proof in the Hilbert case is similar: if A has real multiplication by $\mathbb{Z}_F$, then β-isogenies φ with domain A are characterized by the property that $\ker\varphi$ is a maximal isotropic subgroup of $A[\beta] \simeq (\mathbb{Z}/\ell\mathbb{Z})^2$. □

3.3.1 The Siegel case

Let $\mathcal{C}, \mathcal{C}'$ be equations of genus 2 hyperelliptic curves over $\mathbb{C}$, let A, A' be their Jacobians, and let $\varphi\colon A \to A'$ be an ℓ-isogeny. The choice of curve equations encodes a choice of bases of $\Omega^1(A)$ and $\Omega^1(A')$, or equivalently, by taking dual bases, a choice of bases of the tangent spaces $T_0(A)$ and $T_0(A')$. By an abuse of notation, we identify the tangent map $d\varphi\colon T_0(A) \to T_0(A')$ with its matrix written in these bases.

Definition 3.3.2. It is convenient to introduce matrix notations.

- For $\tau \in \mathbb{H}_2$, we define

$$\partial j(\tau) := \left(\frac{1}{2\pi i}\frac{\partial j_k}{\partial \tau_l}(\tau)\right)_{1\le k,l\le 3} \cdot \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

where we write $\tau = \begin{pmatrix} \tau_1 & \tau_2 \\ \tau_2 & \tau_3 \end{pmatrix}$. In other words, if we set

$$v_1 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix},$$

then the l-th column of $\partial j(\tau)$ contains (up to $2\pi i$) the derivatives of the three Igusa invariants at τ in the direction v_l. More generally, for each $r \in \mathrm{GL}_2(\mathbb{C})$, the l-th column of $\partial j(\tau)\,\mathrm{Sym}^2(r)$ contains the derivatives of Igusa invariants at τ in the direction $r\,v_l\,r^t$.

Let (A, ω) be a principally polarized abelian surface over $\mathbb{C}$ with a basis of differential forms, let $\eta\colon A \to A(\tau)$ be an isomorphism for some $\tau \in \mathbb{H}_2$, and let r be the matrix of η^* in the bases $\omega(\tau)$ and ω. Since derivatives of Igusa invariants have weight Sym^2, we have

$$\partial j(A, \omega) = \partial j(\tau)\,\mathrm{Sym}^2(r^t).$$

We denote by

$$\mathcal{C} \mapsto \partial j(\mathcal{C})$$

the associated fractional covariant; theorem 3.2.15 expresses the entries of this matrix in terms of the coefficients of $\mathcal{C}$.

- Consider the Siegel modular equations $\Psi_{\ell,1}, \Psi_{\ell,2}, \Psi_{\ell,3}$ of level ℓ as elements of the ring $\mathbb{Q}[J_1, J_2, J_3, J_1', J_2', J_3']$, after multiplying by their denominators. We define

$$D\Psi_{\ell,L} = \left(\frac{\partial \Psi_{\ell,n}}{\partial J_k}\right)_{1\le n,k\le 3} \quad \text{and} \quad D\Psi_{\ell,R} = \left(\frac{\partial \Psi_{\ell,n}}{\partial J_k'}\right)_{1\le n,k\le 3}.$$

Definition 3.3.3. Let φ be an ℓ-isogeny as above, and write j (resp. j') as a shorthand for the Igusa invariants (j_1, j_2, j_3) of A (resp. (j_1', j_2', j_3') of A'). We say that the isogeny φ is *generic* if the 3×3 matrices $D\Psi_{\ell,L}(j, j')$, $D\Psi_{\ell,R}(j, j')$, $\partial j(\mathcal{C})$ and $\partial j(\mathcal{C}')$ are invertible. In this case, we define the *deformation matrix* $\mathcal{D}(\varphi)$ of φ as

$$\mathcal{D}(\varphi) = -\partial j(\mathcal{C}')^{-1} \cdot D\Psi_{\ell,R}(j, j')^{-1} \cdot D\Psi_{\ell,L}(j, j') \cdot \partial j(\mathcal{C}).$$

The matrix $\mathcal{D}(\varphi)$ is the matrix of the deformation map of φ in the bases of $T_A(\mathcal{A}_2)$ and $T_{A'}(\mathcal{A}_2)$ associated with $\omega(\mathcal{C})$ and $\omega(\mathcal{C}')$ via the Kodaira–Spencer isomorphism.

Proposition 3.3.4. *If φ is generic in the sense of definition 3.3.3, then we have*

$$\mathrm{Sym}^2(d\varphi) = \ell\,\mathcal{D}(\varphi).$$

Proof. By proposition 3.3.1, there exist $\tau \in \mathbb{H}_2$ and isomorphisms η, η' such that there is a commutative diagram

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & A' \\ \downarrow{\scriptstyle \eta} & & \downarrow{\scriptstyle \eta'} \\ A(\tau) & \xrightarrow{z \mapsto z} & A(\tau/\ell). \end{array}$$

Let r be the matrix of η^* in the bases $\omega(\tau)$ and $\omega(\mathcal{C})$, and define r' similarly. Then we have $d\varphi = r'^t r^{-t}$. By the definition of modular equations, we have

$$\Psi_{\ell,n}\big(j_1(\tau), j_2(\tau), j_3(\tau), j_1(\tau/\ell), j_2(\tau/\ell), j_3(\tau/\ell)\big) = 0 \quad \text{for } 1 \le n \le 3.$$

We differentiate with respect to the entries of τ and obtain

$$D\Psi_{\ell,L}(j, j') \cdot \partial j(\tau) + \frac{1}{\ell} D\Psi_{\ell,R}(j, j') \cdot \partial j(\tau/\ell) = 0.$$

We can rewrite this relation as

$$-\ell\, D\Psi_{\ell,L}(j, j') \cdot \partial j(\mathcal{C}) \cdot \mathrm{Sym}^2(r^t) = D\Psi_{\ell,R}(j, j') \cdot \partial j(\mathcal{C}') \cdot \mathrm{Sym}^2(r'^t). \quad \square$$

Once we compute the deformation matrix $\mathcal{D}(\varphi)$, the matrix $d\varphi$ can be computed up to sign using proposition 3.3.4. This sign indeterminacy is not an issue in the context of the isogeny algorithm, as the input does not distinguish φ from $-\varphi$.

3.3.2 The Hilbert case

The methods of §3.3.1 can be adapted to modular equations of Hilbert type and level β. In this case, the Kodaira–Spencer isomorphism takes a different form. If A is a principally polarized abelian surface over $\mathbb{C}$ with real multiplication by $\mathbb{Z}_F$, then $T_0(A)$ is a $\mathbb{Z}_F \otimes_{\mathbb{Z}} \mathbb{C}$-module, and we have a canonical isomorphism [Rap78, Prop. 1.6 and 1.9]

$$T_A(\mathcal{A}_{2,F}) \simeq \mathrm{Hom}_{\mathbb{Z}_F \otimes_{\mathbb{Z}} \mathbb{C}}\big(T_0(A), T_0(A)\big).$$

Concretely, this means that we can evaluate Hilbert modular forms on a basis ω of $\Omega^1(A)$ only if ω is *Hilbert-normalized*.

Definition 3.3.5. Let A be a principally polarized abelian surface over $\mathbb{C}$ endowed with a real multiplication embedding $\iota\colon \mathbb{Z}_F \hookrightarrow \mathrm{End}(A)^\dagger$, and let ω be a basis of $\Omega^1(A)$. We say that (A, ι, ω) is *Hilbert-normalized* if for every $\alpha \in \mathbb{Z}_F$, the matrix of $\iota(\alpha)^*\colon \Omega^1(A) \to \Omega^1(A)$ in the basis ω is $\left(\begin{smallmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{smallmatrix}\right)$.

In more technical terms, we have an isomorphism $F \otimes_{\mathbb{Q}} \mathbb{C} = F \times F$ via the identity and real conjugation; and ω is Hilbert-normalized if and only if it induces a trivialization of $\Omega^1(A)$ as a $F \otimes_{\mathbb{Q}} \mathbb{C}$-module.

If (A, ι, ω) is Hilbert-normalized and if f is a Hilbert modular form of weight (w_1, w_2), then the quantity $f(A, \iota, \omega)$ is computed as follows. Choose $\tau \in \mathbb{H}_1^2$, and choose an isomorphism $\eta\colon (A, \iota) \to \big(A_F(\tau), \iota_F(\tau)\big)$, where $\iota_F(\tau)$ denotes the canonical real multiplication embedding on $A_F(\tau)$. Let r be matrix of η^* in the bases $\omega(\tau)$ and ω. Then r is diagonal; if we write $r = \left(\begin{smallmatrix} r_1 & 0 \\ 0 & r_2 \end{smallmatrix}\right)$, then

$$f(A, \iota, \omega) = r_1^{w_1} r_2^{w_2} f(\tau).$$

Given a genus 2 curve $\mathcal{C}$ and a real multiplication embedding ι on $\mathrm{Jac}(\mathcal{C})$, we say that the equation of $\mathcal{C}$ is *Hilbert-normalized* if $(\mathrm{Jac}(\mathcal{C}), \iota, \omega(\mathcal{C}))$ is. Not all curve equations are Hilbert-normalized: the degree of freedom in choosing a Hilbert-normalized curve equation is $\mathrm{GL}_1 \times \mathrm{GL}_1$, not GL_2.

Computing the tangent matrix. For the moment, assume that there exists a β-isogeny $\varphi\colon (A, \iota) \to (A', \iota')$ between complex abelian surfaces with real multiplication by $\mathbb{Z}_F$, and that we are given curve equations $\mathcal{C}$ and $\mathcal{C}'$ that are Hilbert-normalized with respect to these real multiplication embeddings. We address the question of constructing $\mathcal{C}$ and $\mathcal{C}'$ at the end of this section.

Definition 3.3.6. As in the Siegel case, we introduce matrix notations.

- Denote by j_1, j_2, j_3 the pullbacks of Igusa invariants to $\mathbb{H}_1^2$ via the Hilbert embedding. For $\tau \in \mathbb{H}_1^2$, we define

$$\partial j^{(F)}(\tau) = \left(\frac{1}{\pi i} \frac{\partial j_k}{\partial \tau_l}(\tau) \right)_{1 \le k \le 3, 1 \le l \le 2}, \qquad \text{where } \tau = (\tau_1, \tau_2).$$

  If $\mathcal{C}$ is a curve equation such that $\omega(\mathcal{C})$ is Hilbert-normalized, then we denote by $\partial j^{(F)}(\mathcal{C})$ the value of this modular form on $\mathcal{C}$.

- We define the 3×3 matrices $D\Psi_{\beta,L}$ and $D\Psi_{\beta,R}$ in the case of Hilbert modular equations of level β in Igusa invariants as in definition 3.3.2.

- Write j as a shorthand for the Igusa invariants (j_1, j_2, j_3) of A, and j' for the invariants (j'_1, j'_2, j'_3) of A'. We say that the isogeny φ is *generic* if the 3×2 matrices

$$D\Psi_{\beta,L}(j, j') \cdot \partial j^{(F)}(\mathcal{C}) \quad \text{and} \quad D\Psi_{\beta,R}(j, j') \cdot \partial j^{(F)}(\mathcal{C}')$$

  have rank 2.

Since derivatives of Igusa invariants with respect to τ_1 and τ_2 on $\mathbb{H}_1^2$ are Hilbert modular functions of weight $(2,0)$ and $(0,2)$ respectively, the weight of the modular function $\partial j^{(F)}$ is the following.

Lemma 3.3.7. *Let (A, ι, ω) be a Hilbert-normalized abelian surface, and let $\tau \in \mathbb{H}_1^2$ such that there is an isomorphism $\eta\colon (A, \iota) \to (A_F(\tau), \iota_F(\tau))$. Let r be the matrix of η^* in the bases $\omega(\tau)$ and ω. Then*

$$\partial j^{(F)}(A, \iota, \omega) = \partial j^{(F)}(\tau) \cdot r^2.$$

In order to compute the value of $\partial j^{(F)}$ on a curve equation, we relate it with its Siegel analogue ∂j.

Proposition 3.3.8. *Let (A, ι, ω) be a Hilbert-normalized abelian surface. Then*

$$\partial j^{(F)}(A, \iota, \omega) = \partial j(A, \omega) \cdot T \qquad \text{where } T = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{smallmatrix}\right).$$

Proof. Let τ, η, r be as in lemma 3.3.7, and let $R \in \mathrm{GL}_2(\mathbb{R})$ be a matrix defining a Hilbert embedding for F. By the expression (2.12) of the Hilbert embedding, the columns of $\partial j^{(F)}(\tau)$ contain the derivatives of Igusa invariants at $H_R(\tau)$ in the directions

$$\frac{1}{\pi i} R^t \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} R \quad \text{and} \quad \frac{1}{\pi i} R^t \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} R.$$

Hence we have

$$\partial j^{(F)}(\tau) = \partial j\big(H_R(\tau)\big) \cdot \mathrm{Sym}^2(R^t) \cdot T.$$

By proposition 3.2.6, we have an isomorphism $\zeta\colon A_F(\tau) \to A\big(H_R(\tau)\big)$ such that the matrix of ζ^* in the bases $\omega(\tau)$ and $\omega\big(H_R(\tau)\big)$ is R. Therefore

$$\partial j^{(F)}(A, \iota, \omega) = \partial j^{(F)}(\tau) \cdot r^2 \quad \text{and} \quad \partial j(A, \omega) = \partial j(\tau)\, \mathrm{Sym}^2\big((rR)^t\big). \quad \square$$

It is natural that R does not appear in the result of proposition 3.3.8: the Kodaira–Spencer isomorphisms are intrinsic and independent of the choice of Hilbert embedding.

Proposition 3.3.9. *Let $\varphi\colon A \to A'$ be a β-isogeny, and let $\mathcal{C}$ and $\mathcal{C}'$ be Hilbert-normalized curve equations as above. Then the tangent matrix $d\varphi$ is diagonal, and*

$$D\Psi_{\beta,L}(j,j')\cdot \partial j^{(F)}(\mathcal{C}) = -D\Psi_{\beta,R}(j,j')\cdot \partial j^{(F)}(\mathcal{C}')\cdot \begin{pmatrix} 1/\beta & 0 \\ 0 & 1/\overline{\beta} \end{pmatrix} \cdot (d\varphi)^2.$$

Proof. By proposition 3.3.1, we can find $\tau \in \mathbb{H}_1^2$ as well as isomorphisms η and η' such that there is a commutative diagram

$$\begin{array}{ccc} (A,\iota) & \xrightarrow{\quad\varphi\quad} & (A',\iota') \\ \downarrow{\scriptstyle \eta} & & \downarrow{\scriptstyle \eta'} \\ (A_F(\tau),\iota_F(\tau)) & \xrightarrow{z\mapsto z} & (A_F(\tau/\beta),\iota_F(\tau/\beta)). \end{array}$$

Let r be the matrix of η^* in the bases $\omega(\tau)$ and $\omega(\mathcal{C})$, and define r' similarly; they are diagonal by definition 3.3.5. We have $d\varphi = r'^t r^{-t} = r' r^{-1}$. Differentiating the modular equations, we obtain

$$D\Psi_{\beta,L}(j,j')\cdot \partial j^{(F)}(\tau) + D\Psi_{\beta,R}(j,j')\cdot \partial j^{(F)}(\tau/\beta)\cdot \mathrm{Diag}(1/\beta, 1/\overline{\beta}) = 0.$$

The result follows by lemma 3.3.7 since

$$\partial j^{(F)}(\tau) = \partial j^{(F)}(\mathcal{C})\cdot r^2 \quad \text{and} \quad \partial j^{(F)}(\tau/\beta) = \partial j^{(F)}(\mathcal{C}')\cdot r'^2. \qquad \square$$

Proposition 3.3.9 allows us to compute $(d\varphi)^2$ from derivatives of modular equations of level β when φ is generic. However, in contrast with the Siegel case, the knowledge of $(d\varphi)^2$ does not allow us to recover the diagonal matrix $d\varphi$ up to sign, as we have to perform two a priori uncorrelated root extractions: we obtain two possible candidates for $\pm d\varphi$.

Constructing Hilbert-normalized curves. Let (A,ι) be a principally polarized abelian surface over $\mathbb{C}$ with real multiplication by $\mathbb{Z}_F$. Given the Igusa invariants (j_1,j_2,j_3) of A, we want to construct a curve equation $\mathcal{C}$ such that $A \simeq \mathrm{Jac}(\mathcal{C})$ and $\big(A,\iota,\omega(\mathcal{C})\big)$ is Hilbert-normalized. First, we compute any curve equation using Mestre's algorithm [Mes91]; then we look for a suitable change of variables.

However, we are missing some information, as the two pairs (A,ι) and $(A,\overline{\iota})$, where $\overline{\iota}$ denotes the real conjugate of ι, have the same Igusa invariants. The best we can hope for is to compute an equation $\mathcal{C}$ such that either $\big(A,\iota,\omega(\mathcal{C})\big)$ or $\big(A,\overline{\iota},\omega(\mathcal{C})\big)$ is Hilbert-normalized. In this case, we say that $\mathcal{C}$ is *potentially Hilbert-normalized*. This ambiguity is a direct consequence of using symmetric invariants on the Hilbert surface.

Proposition 3.3.10. *Let $\mathcal{C}$ be a hyperelliptic curve equation of genus 2 over $\mathbb{C}$ whose Jacobian $\mathrm{Jac}(\mathcal{C})$ has real multiplication by $\mathbb{Z}_F$. Denote the Igusa invariants of $\mathrm{Jac}(\mathcal{C})$ by (j_1, j_2, j_3), and assume that $j_3 \neq 0$. Then $\mathcal{C}$ is potentially Hilbert-normalized if and only if the two columns of the 3×2 matrix*

$$\partial j(\mathcal{C}) \cdot T \qquad \text{where } T = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$$

define tangent vectors to the Humbert surface at (j_1, j_2, j_3).

Proof. Let $R \in \mathrm{GL}_2(\mathbb{R})$ be a matrix defining a Hilbert embedding for F, and choose $\tau \in \mathbb{H}_1^2$ such that there is an isomorphism $\eta\colon \mathrm{Jac}(\mathcal{C}) \to A_F(\tau)$. Let r be the matrix of η^* in the bases $\omega(\tau)$ and $\omega(\mathcal{C})$. Then the columns of $\partial j(\mathcal{C}) \cdot T$ contain, up to πi, the derivatives of Igusa invariants at $H_R(\tau)$ in the directions

$$R^t r \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} r^t R \quad \text{and} \quad R^t r \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} r^t R.$$

These directions are tangent to the Humbert surface if and only if r is is either diagonal or anti-diagonal. Moreover, the condition $j_3 \neq 0$ ensures that Igusa invariants are local coordinates at $H_R(\tau) \in \mathcal{A}_2(\mathbb{C})$. □

Assume that the equation of the Humbert surface for F in terms of Igusa invariants is given: this precomputation depends only on F. Given Igusa invariants (j_1, j_2, j_3) on the Humbert surface such that $j_3 \neq 0$, the algorithm to reconstruct a potentially Hilbert-normalized curve equation runs as follows.

Algorithm 3.3.11.

1. Construct any curve equation $\mathcal{C}_0$ with Igusa invariants (j_1, j_2, j_3) using Mestre's algorithm.
2. Compute $r \in \mathrm{GL}_2(\mathbb{C})$ such that the two columns of the matrix

$$\partial j(\mathcal{C}_0) \cdot \mathrm{Sym}^2(r^t) \cdot T$$

   are tangent to the Humbert surface at (j_1, j_2, j_3).
3. Output $\det^{-2} \mathrm{Sym}^6(r)\, \mathcal{C}_0$.

In step 2, if a, b, c, d denote the entries of r, we only have to solve a quadratic equation in a, c, and a quadratic equation in b, d. Therefore algorithm 3.3.11 involves $O_F(1)$ square roots and elementary operations (i.e. a constant number of operations, when F is fixed).

In practice, when computing a β-isogeny $\varphi\colon A \to A'$ in the Hilbert case, we are only given the Igusa invariants of A and A'. Constructing potentially

Hilbert-normalized curves is then equivalent to making a choice of real multiplication embedding for each abelian surface. If these embeddings are incompatible via φ, we obtain antidiagonal matrices when computing the tangent matrix using proposition 3.3.9; in this case, we apply the change of variables $x \mapsto 1/x$ on one of the curve equations to make them compatible. Even if they are compatible, φ will be either a β- or a $\overline{\beta}$-isogeny depending on the choices of real multiplication embeddings. Therefore we really obtain four candidates for the tangent matrix of φ up to sign, among which only one is usually correct.

3.3.3 Extension to arbitrary fields

The statements of propositions 3.3.4 and 3.3.9, which we proved over $\mathbb{C}$, make sense over an arbitrary field k, at least if the degree of the isogeny φ is prime to the characteristic of k. In fact, they continue to hold in this generalized setting, along with the fundamental property of modular equations: if A and A' are abelian surfaces over k with suitable PEL structure satisfying a certain genericity hypothesis, then modular equations of Siegel or Hilbert type vanish at (A, A') if and only if there exists an isogeny $\varphi\colon A \to A'$ of the corresponding type over $\bar{k}$.

These extended statements are easily proved if k has characteristic zero using Lefschetz's principle. We can restrict to the case where k is finitely generated over $\mathbb{Q}$; then k embeds in $\mathbb{C}$, so that the complex theory applies.

If k has positive characteristic, then the proof uses more advanced algebraic geometry, namely the notion of *moduli stacks* that we already mentioned in §3.1. We do not define these objects here, and refer to [LM00] for the theory; instead, we only look at examples and apply some of their properties in the context of the isogeny algorithm.

Moduli stacks of abelian varieties. The moduli space of abelian varieties of dimension $g \geq 1$, denoted by $\mathscr{A}_g$, is an algebraic object endowed with a universal family $\mathscr{X}_g \to \mathscr{A}_g$ such that the following property holds: if S is any scheme, and if $X \to S$ is an abelian scheme of dimension g (if $S = \operatorname{Spec} k$, this just means that X is an abelian variety of dimension g over k), then there exists a unique morphism $f\colon S \to \mathscr{A}_g$ such that X is the pullback of $\mathscr{X}_g$ via f. In particular we have a bijection between k-points of $\mathscr{A}_g$ and isomorphism classes of abelian varieties of dimension g over k.

We stress that $\mathscr{A}_g$ is not a scheme, but its associated *coarse moduli scheme* $\mathcal{A}_g$ is; it is even a quasi-projective scheme by geometric invariant theory [MFK94]. The $\bar{k}$-points of $\mathcal{A}_g$ still classify isomorphism classes of

principally polarized abelian varieties of dimension g over $\bar{k}$, but the correspondence may fail to hold over non algebraically closed fields. As an example, take $g = 1$; the j-invariant realizes an isomorphism from $\mathcal{A}_1$ to the affine space $\mathbb{A}^1$. Two elliptic curves having the same j-invariant over k are not necessarily isomorphic over k, but they are isomorphic over $\bar{k}$.

The set of $\mathbb{C}$-points of $\mathcal{A}_g$ is identified with the quotient $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g$. Informally, we view $\mathscr{A}_g$ as the scheme $\mathcal{A}_g$ endowed with an additional structure of *inertia*, i.e. the data of a finite group of automorphisms of x, also called *stabilizers* of x, for each point x of $\mathcal{A}_g$. For instance the $\mathbb{C}$-points of $\mathscr{A}_g$ are identified with the quotient $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g$ seen as an *orbifold*, the same kind of object in the world of complex varieties.

Other examples of moduli stacks are given by the moduli stacks of principally polarized abelian varieties of dimension g endowed with

- a level ℓ structure for some fixed $\ell \geq 1$,
- the kernel of an ℓ-isogeny,
- a real multiplication structure by $\mathbb{Z}_F$, where F is a fixed totally real number field of degree g over $\mathbb{Q}$,

denoted by $\mathscr{A}_g(\ell)$, $\mathscr{A}_{g,\ell}$, and $\mathscr{A}_{g,F}$ respectively.

Not all these moduli problems are well defined over $\mathbb{Z}$. For instance the stacks $\mathscr{A}_g(\ell)$ and $\mathscr{A}_{g,\ell}$ classify isomorphism classes of abelian schemes with suitable structure over S when S is a scheme over $\mathbb{Z}[1/\ell]$, i.e. when S has "characteristic prime to ℓ". Both $\mathscr{A}_g(\ell)$ and $\mathscr{A}_{g,\ell}$ can be extended to the whole of $\operatorname{Spec}\mathbb{Z}$ as algebraic stacks, but giving a modular interpretation becomes more difficult: this is done in [DR73] in the case of moduli stacks of elliptic curves.

Smoothness and étaleness. In many aspects, moduli stacks of abelian varieties satisfy nicer properties than their coarse moduli schemes. For instance $\mathscr{A}_g$ is *smooth* over $\mathbb{Z}$ [FC90], and similarly $\mathscr{A}_g(\ell)$ and $\mathscr{A}_{g,\ell}$ are smooth over $\mathbb{Z}[1/\ell]$. As smoothness implies formal smoothness [The18, Tag 02GZ], we have the following lifting property.

Proposition 3.3.12. *Let k be a field of characteristic prime to ℓ, and denote by $W(k)$ the Witt ring of k. Let A and A' be principally polarized abelian varieties of dimension $g \geq 1$ over k, and let $\varphi\colon A \to A'$ be an ℓ-isogeny defined over k. Then there exists an ℓ-isogeny $\widetilde{\varphi}\colon \widetilde{A} \to \widetilde{A}'$ defined over $W(k)$ between principally polarized abelian varieties $\widetilde{A}, \widetilde{A}'$ over $W(k)$ such that $\widetilde{A}, \widetilde{A}'$ and $\widetilde{\varphi}$ have good reduction, and which reduces to the initial ℓ-isogeny φ over k.*

As a consequence, the algorithm to compute the tangent matrix $d\varphi$ in the Siegel case, presented in §3.3.1, is universally valid over fields of characteristic prime to ℓ.

Besides proving the existence of lifts to characteristic zero, moduli stacks help to better understand the geometric situation behind the isogeny algorithm [KPR19, §4]. Indeed, the two maps

$$H_{\ell,1} \text{ and } H_{\ell,2} \colon \mathcal{A}_{g,\ell} \to \mathcal{A}_g,$$

that we used to define the deformation map in §3.1, are well-defined and étale everywhere at the level of stacks [KPR19, §4.2.1]. If A is a principally polarized abelian variety of dimension g over k with generic automorphisms, i.e. such that $\mathrm{Aut}_{\bar{k}}(A) = \{\pm 1\}$, then the scheme $\mathcal{A}_g$ and the stack $\mathscr{A}_g$ are essentially isomorphic locally around A; we refer to [KPR19, §4.1] for more details, notably the statement of Luna's fundamental lemma. From this, we obtain a sufficient condition for an ℓ-isogeny to be generic in the sense of definition 3.3.3.

Proposition 3.3.13 ([KPR19, §4.5])**.** *Let k be a field, and let $\mathcal{U}$ be the open set of $\mathcal{A}_2$ consisting of abelian surfaces A such that $\mathrm{Aut}_{\bar{k}}(A) = \{\pm 1\}$ and $j_3(A) \neq 0$. Let $\varphi \colon A \to A'$ be an ℓ-isogeny over k such that A and A' lie in $\mathcal{U}$, and such that the subvariety cut out by Siegel modular equations of level ℓ is normal at $(j(A), j(A'))$, where j denotes the collection of the three Igusa invariants. Then φ is generic in the sense of definition 3.3.3.*

In fact, all the computations of §3.3.1 can be interpreted algebraically, and hence make sense over any scheme of characteristic prime to ℓ [KPR19, §4.5]; this is arguably a more intrinsic way of showing their universal validity than using Lefschetz's principle.

The Hilbert case. The situation in the Hilbert case is analogous: we have a map

$$\mathscr{H}_\beta = (\mathscr{H}_{\beta,1}, \mathscr{H}_{\beta,2}) \colon \mathscr{A}_{g,\beta} \to \mathscr{A}_{g,F} \times \mathscr{A}_{g,F}$$

where $\mathscr{A}_{g,\beta}$ denotes the moduli stack of principally polarized abelian varieties of dimension g with real multiplication by $\mathbb{Z}_F$ endowed with the kernel of a β-isogeny, and both $\mathscr{H}_{\beta,1}$ and $\mathscr{H}_{\beta,2}$ are everywhere étale. All these stacks are smooth over $\mathbb{Z}[1/\ell]$, where $\ell = N_{F/\mathbb{Q}}(\beta)$ [KPR19, §4.2.3], [Rap78]. As in proposition 3.3.12, we obtain a sufficient condition for a β-isogeny φ to be generic, at least if we are using nonsymmetric invariants on the Hilbert surface [KPR19, §4.5].

To interpret the computation of the tangent matrix in proposition 3.3.9 algebraically if $\varphi \colon A \to A'$ is a β-isogeny defined over a field k, we consider

the tangent spaces of A and A' at zero as well as the tangent spaces to the Hilbert surface at A and A' as $\mathbb{Z}_F \otimes_{\mathbb{Z}} k$-modules of rank 1 [KPR19, §4.3.2]. Up to enlarging k, we fix an isomorphism $\mathbb{Z}_F \otimes_{\mathbb{Z}} k \simeq k \times k$; concretely, we choose a particular value of $\sqrt{\Delta}$ in k, where Δ denotes the discriminant of F. Then we define a *Hilbert-normalized* basis of differential forms on A as a basis of $\Omega^1(A)$ in which the action of $\mathbb{Z}_F \otimes_{\mathbb{Z}} k$ by multiplication takes the form of the action of $k \times k$ by diagonal matrices; when $k = \mathbb{C}$, we recover definition 3.3.5.

If A and A' are endowed with Hilbert-normalized bases of differential forms, then $d\varphi$ is a diagonal matrix and can be computed as in proposition 3.3.9. Moreover, if (ω_1, ω_2) is a Hilbert-normalized basis of differential forms on A, then $(\omega_1 \otimes \omega_1, \omega_2 \otimes \omega_2)$ is naturally identified with a basis of $T_A(\mathcal{A}_{2,F})$ [Rap78, Prop. 1.6]: this provides an algebraic interpretation of proposition 3.3.8, and also shows that algorithm 3.3.11 to construct Hilbert-normalized curve equations is universally valid.

When $\sqrt{\Delta} \notin k$, we could also use another presentation of $\mathbb{Z}_F \otimes_{\mathbb{Z}} k$ instead of diagonal matrices; since $\mathbb{Z}_F \otimes_{\mathbb{Z}} k$ has no zero divisors in this case, we should be able to decover $\pm d\varphi$ from $(d\varphi)^2$ without ambiguity.

3.4 Computing isogenies from tangent maps

3.4.1 General strategy

Let A, A' be principally polarized abelian varieties of dimension g defined over a field k, and assume that we are given the tangent map $d\varphi$ of a separable isogeny $\varphi\colon A \to A'$. In general, the task of *computing φ explicitly* is the following: given models of A and A', i.e. given very ample line bundles $\mathcal{L}_A, \mathcal{L}_{A'}$ on A, A' and a choice of global sections (a_i) (resp. (a'_j)) which give a projective embedding of A (resp. A'), express the functions $\varphi^* a'_j$ on A as rational fractions in terms the coordinates (a_i).

One method to determine φ given $d\varphi$ is to use the formal groups of A and A'. Let $x_1, \dots, x_g$ be uniformizers at the neutral point 0_A of A, and let $y_1, \dots, y_g$ be uniformizers at $0_{A'}$. Knowing the map $d\varphi$ is equivalent to expressing the differential forms $\varphi^* dy_j$ in term of the differential forms dx_i on A. This allows us to write a differential system satisfied by the functions $\varphi^* a'_j$. We can try to solve this differential system using a multivariate Newton algorithm, possibly over an extension of the formal group. If the algorithm succeeds, we recover the functions $\varphi^* a'_j$ as power series in $x_1, \dots, x_g$ up to some precision. Finally we obtain φ as a rational map by multivariate rational reconstruction. In order for the rational reconstruc-

tion algorithm to succeed, the power series precision must be large enough when compared to the degrees of the result in the variables (a_i). These degrees can be estimated from the intersection degree of $\varphi^*\mathcal{L}_{A'}$ and $\mathcal{L}_A$, or alternatively from the intersection degree of $\varphi_*\mathcal{L}_A$ and $\mathcal{L}_{A'}$.

This strategy to compute φ is not new: the idea of using a differential equation to compute isogenies in dimension 1 appears in [Elk98], and [BMSS08] uses a Newton algorithm to solve this differential equation, as we explained in §1.2.2. These ideas are extended to dimension 2 in [CE15]. The method is further extended to compute endomorphisms of Jacobians over a number field in [CMSV19]. In [CMSV19, §6], the endomorphism is represented as a divisorial correspondence; the interpolation of this divisor is done via linear algebra on Riemann–Roch spaces.

A necessary condition for the whole method to work is that φ be completely determined by its tangent map. In general, this will be the case when $\operatorname{char} k$ is large with respect to the degree of φ. For instance, we have the following statement in the case of ℓ-isogenies.

Proposition 3.4.1. *Let A and A' be principally polarized abelian varieties over a field k, and let $M\colon T_0(A) \to T_0(A')$ be a linear map. Assume that either $\operatorname{char} k = 0$ or $\operatorname{char} k > 4N$. Then there exists at most one ℓ-isogeny $\varphi\colon A \to A'$ with $\ell \leq N$ such that $d\varphi = M$.*

Proof. Let φ_1 and φ_2 be two such isogenies. Then $\varphi_1 = \varphi_2 + \psi$ where ψ is inseparable. If $\operatorname{char} k = 0$, this implies $\psi = 0$ and hence $\varphi_1 = \varphi_2$. Otherwise, write $p = \operatorname{char} k$ and denote by $\widehat{\varphi}_1$ the contragredient isogeny. Then we have

$$\psi\widehat{\psi} = \varphi_2\widehat{\varphi}_2 + \varphi_1\widehat{\varphi}_1 - \varphi_1\widehat{\varphi}_2 - \varphi_2\widehat{\varphi}_1.$$

If $\psi \neq 0$, then $\psi\widehat{\psi}$ is equal to p^m for some $m \geq 1$, and $\varphi_1\widehat{\varphi}_1 = \ell_1$, $\varphi_2\widehat{\varphi}_2 = \ell_2$ with $\max\{\ell_1, \ell_2\} \leq N$ by hypothesis. Therefore we obtain $p^m \leq 4N$. □

In practice, Newton iterations will fail to reach sufficiently high power series precision if $\operatorname{char} k$ is too small.

In the rest of this section, we carry out this strategy in detail when A and A' are the Jacobians of genus 2 hyperelliptic curves $\mathcal{C}$ and $\mathcal{C}'$. Concretely, we are given the matrix of $d\varphi$ in the bases of $T_0(A)$ and $T_0(A')$ that are dual to $\omega(\mathcal{C})$ and $\omega(\mathcal{C}')$ respectively; this is precisely the input provided by the algorithm of §3.3. In this case, a nice simplification occurs: the isogeny φ is completely determined by the compositum

$$\mathcal{C} \xhookrightarrow{Q\mapsto[Q-P]} \operatorname{Jac}(\mathcal{C}) \xrightarrow{\varphi} \operatorname{Jac}(\mathcal{C}') \dashrightarrow \mathcal{C}'^{<2>} \dashrightarrow \mathbb{A}^4 \tag{3.6}$$

where P is any point on $\mathcal{C}$, and m is the rational map given by

$$\{(x_1, y_1), (x_2, y_2)\} \mapsto \left(x_1 + x_2,\ x_1 x_2,\ y_1 y_2,\ \tfrac{y_2 - y_1}{x_2 - x_1}\right).$$

The compositum (3.6) is a tuple of four rational fractions $s, p, q, r \in k(u, v)$ that we call the *rational representation of φ at the base point P*. We choose a uniformizer z of $\mathcal{C}$ around P and perform the Newton iterations and rational reconstruction using univariate power series in z.

We explain how we choose the base point P and solve the differential system in §3.4.2. One difficulty is that the differential system we obtain is singular; therefore we use the geometry of the curves to compute the first few terms in the series before switching to Newton iterations. In §3.4.3, we estimate the degrees of the rational fractions that we want to compute and present the rational reconstruction step.

3.4.2 Solving the differential equation

We write the curve equations $\mathcal{C}, \mathcal{C}'$ and the tangent matrix as

$$\mathcal{C}\colon v^2 = E_{\mathcal{C}}(u), \quad \mathcal{C}'\colon y^2 = E_{\mathcal{C}'}(x), \quad \text{and} \quad d\varphi = \begin{pmatrix} m_{1,1} & m_{1,2} \\ m_{2,1} & m_{2,2} \end{pmatrix}.$$

We assume that φ is separable, so that $d\varphi$ is invertible. If P is a base point on $\mathcal{C}$, we denote by η_P the associated embedding $\mathcal{C} \to \mathrm{Jac}(\mathcal{C})$, as in §2.1.2. Since $\mathcal{C}$ is smooth and $\mathcal{C}'^{<2>}$ is proper, the compositum

$$\mathcal{C} \xhookrightarrow{\eta_P} \mathrm{Jac}(\mathcal{C}) \xrightarrow{\varphi} \mathrm{Jac}(\mathcal{C}') \dashrightarrow \mathcal{C}'^{<2>}$$

extends to a morphism from $\mathcal{C}$ to $\mathcal{C}'^{<2>}$ that we denote by φ_P.

Step 1: choice of base point and power series. Let P be a point on $\mathcal{C}$ which is not a point at infinity; after enlarging k, we assume that $P \in \mathcal{C}(k)$. Since $\varphi_P(P)$ is zero in $\mathrm{Jac}(\mathcal{C}')$, we have

$$\varphi_P(P) = \{Q, i(Q)\}$$

for some point Q on $\mathcal{C}'$, where i denotes the hyperelliptic involution on $\mathcal{C}'$. We say that φ_P is of *Weierstrass type* if Q is a Weierstrass point of $\mathcal{C}'$, and of *generic type* otherwise. If z is a local uniformizer of $\mathcal{C}$ at P, and if R is an étale extension of $k[[z]]$, then we define a *local lift of φ_P at P with*

coefficients in R to be a tuple $\widetilde{\varphi}_P = (x_1, x_2, y_1, y_2) \in R^4$ such that we have a commutative diagram

$$\begin{array}{ccccc} \operatorname{Spec} R & & \xrightarrow{(x_1,y_1),(x_2,y_2)} & & \mathcal{C}'^2 \\ \downarrow & & & & \downarrow \\ \operatorname{Spec} k[[z]] & \longrightarrow & \mathcal{C} & \xrightarrow{\varphi_P} & \mathcal{C}'^{<2>}. \end{array} \tag{3.7}$$

If the power series x_1, x_2, y_1, y_2 define a local lift of φ_P, then they satisfy the following differential system:

$$\begin{cases} \dfrac{x_1\,dx_1}{y_1} + \dfrac{x_2\,dx_2}{y_2} &= (m_{1,1}u + m_{1,2})\dfrac{du}{v} \\ \dfrac{dx_1}{y_1} + \dfrac{dx_2}{y_2} &= (m_{2,1}u + m_{2,2})\dfrac{du}{v} \\ y_1^2 = E_{\mathcal{C}'}(x_1) \\ y_2^2 = E_{\mathcal{C}'}(x_2), \end{cases} \tag{3.8}$$

where we consider the coordinates u, v on $\mathcal{C}$ as elements of $k[[z]]$, and where the letter d denotes derivation with respect to z.

We will show how to solve (3.8) when φ_P is of generic type. Proposition 3.4.4 below explains how to choose the base point P in order to enforce this condition; in order to prove it, we first study the existence of local lifts for arbitrary base points.

Lemma 3.4.2. *Let z be a uniformizer of $\mathcal{C}$ at P. Then there exists a quadratic extension k'/k and a local lift of φ_P at P with coefficients in $R = k'[[\sqrt{z}]]$. Moreover, if φ_P is of generic type, or if P is a Weierstrass point of $\mathcal{C}$, then the same statement holds with $R = k'[[z]]$.*

Proof. First, we assume that φ_P is of generic type. Since the unordered pair $\{Q, i(Q)\}$ is defined over k, there exists a quadratic extension k'/k such that Q is defined over k'. The map $\mathcal{C}'^2 \to \mathcal{C}'^{<2>}$ is étale at $(Q, i(Q))$, so it induces an isomorphism of completed local rings. Therefore a local lift of φ_P exists over $k'[[z]]$.

Second, we assume that φ_P is of Weierstrass type. In diagram (3.7), the bottom map $\operatorname{Spec} k[[z]] \to \mathcal{C}'^{<2>}$ defines a $k((z))$-point of $\mathcal{C}'^{<2>}$, and there exists a preimage of this point defined over an extension K of $k((z))$ of degree 2. Let R be the integral closure of $k[[z]]$ in K. Then R is contained in $k'[[\sqrt{z}]]$ for some quadratic extension k' of k [The18, Tag 09E8]. By the valuative criterion of properness, our K-point of $\mathcal{C}'^2$ extends to an R-point uniquely, so a local lift exists over R.

Finally, we assume that φ_P is of Weierstrass type and that P is a Weierstrass point of $\mathcal{C}$. Let (x_1, x_2, y_1, y_2) be a local lift of φ_P over $k'[[\sqrt{z}]]$. The completed local ring of the Kummer line of $\mathcal{C}$ at P is $k[[z^2]]$, and the unordered pair $\{x_1, x_2\}$ is defined on the Kummer line; by the same argument as above, x_1 and x_2 are actually defined over $k'[[z]]$. The differential system (3.8) can be written as

$$\begin{pmatrix} 1/y_1 \\ 1/y_2 \end{pmatrix} = \begin{pmatrix} x_1\,dx_1 & x_2\,dx_2 \\ dx_1 & dx_2 \end{pmatrix}^{-1} \begin{pmatrix} R_1(z) \\ R_2(z) \end{pmatrix}$$

for some power series $R_1, R_2 \in k[[z]]$. Therefore y_1 and y_2 are defined over $k'[[z]]$ as well. □

Consider the tangent space $T_{(Q,i(Q))}(\mathcal{C}'^2)$ of $\mathcal{C}'^2$ at $(Q, i(Q))$. It decomposes as

$$T_{(Q,i(Q))}(\mathcal{C}'^2) = T_Q(\mathcal{C}') \oplus T_{i(Q)}(\mathcal{C}') \simeq (T_Q(\mathcal{C}'))^2 \tag{3.9}$$

where the last isomorphism in (3.9) is given by the hyperelliptic involution on the second term.

Lemma 3.4.3. *Assume that a local lift $\widetilde{\varphi}_P$ of φ_P to $k'[[z]]$ exists. Then under the isomorphism (3.9), the tangent vector $d\widetilde{\varphi}_P/dz$ at $z = 0$ is not of the form (v, v) where $v \in T_Q(\mathcal{C}')$.*

Proof. Assume the contrary. Since the direction $(1, 1)$ is contracted to zero in $\mathrm{Jac}(\mathcal{C}')$, every differential form on $\mathrm{Jac}(\mathcal{C}')$ is pulled back to zero via φ_P. This is a contradiction because φ^* is nonzero. □

Proposition 3.4.4. *The unordered pair $\{Q, i(Q)\} = \varphi_P(P)$ is uniquely determined by the property that, up to a scalar factor,*

$$\varphi^* \omega'_Q = \omega_P$$

where ω_P (resp. ω'_Q) is a nonzero differential form on $\mathcal{C}$ (resp. $\mathcal{C}'$) vanishing at P (resp. Q).

Proof. First, assume that a local lift $\widetilde{\varphi}_P$ exists over $k'[[z]]$. By lemma 3.4.3, the tangent vector $d\widetilde{\varphi}_P/dz$ at $z = 0$ is of the form $(v + w, w)$ for some $v, w \in T_Q(\mathcal{C}')$ such that $v \neq 0$. Let ω' be the unique nonzero differential form pulled back to ω_P by φ. Then ω' vanishes on $(v, 0)$, in other words ω' vanishes at Q.

Second, assume that no such lift exists. By lemma 3.4.2, Q is a Weierstrass point on $\mathcal{C}'$, and P is not a Weierstrass point on $\mathcal{C}$. After a change

of variables, we may assume that Q is not at infinity. Write $P = (u_0, v_0)$ with $v_0 \neq 0$, and $Q = (x_0, 0)$. We have to show that

$$x_0 = \frac{m_{1,1}u_0 + m_{1,2}}{m_{2,1}u_0 + m_{2,2}}.$$

Let (x_1, y_1, x_2, y_2) be a lift over $k'[[\sqrt{z}]]$ as in lemma 3.4.2, and look at the differential system (3.8). Write the lift as

$$y_1 = v_1\sqrt{z} + t_1 z + O(z^{3/2}) \quad \text{and} \quad y_2 = v_2\sqrt{z} + t_2 z + O(z^{3/2}).$$

Then the relation $y^2 = E_{\mathcal{C}'}(x)$ forces x_1 and x_2 to have no term in $\sqrt{z}$, so that we can write

$$x_1 = x_0 + w_1 z + O(z^{3/2}) \quad \text{and} \quad x_2 = x_0 + w_2 z + O(z^{3/2}).$$

Using the relation $dx/y = 2\,dy/E'_{\mathcal{C}'}(x)$, we obtain the equalities

$$\begin{cases} 2x_1\dfrac{dy_1}{E'_{\mathcal{C}'}(x_1)} + 2x_2\dfrac{dy_2}{E'_{\mathcal{C}'}(x_2)} = (m_{1,1}u + m_{1,2})\dfrac{du}{v}, \\ 2\dfrac{dy_1}{E'_{\mathcal{C}'}(x_1)} + 2\dfrac{dy_2}{E'_{\mathcal{C}'}(x_2)} = (m_{2,1}u + m_{2,2})\dfrac{du}{v}. \end{cases} \tag{3.10}$$

Inspection of the coefficients in $(\sqrt{z})^{-1}$ gives the relation $v_1 = -v_2$. Write $e = E'_{\mathcal{C}'}(x_0)$. Then the constant term of the series on the left hand side of (3.10) are respectively

$$2x_0\left(\tfrac{t_1}{e} + \tfrac{t_2}{e}\right) \quad \text{and} \quad 2\left(\tfrac{t_1}{e} + \tfrac{t_2}{e}\right).$$

The differential forms on the right hand side do not vanish simultaneously at P, therefore $m_{2,1}u_0 + m_{2,2}$ is nonzero. Taking the constant term in the quotient of the two lines of (3.10) gives the result. □

Using proposition 3.4.4, we choose a base point P on $\mathcal{C}$ such that φ_P is of generic type. By lemma 3.4.2, a local lift $\widetilde{\varphi}_P = (x_1, x_2, y_1, y_2)$ of φ_P exists over $k'[[z]]$, where k' is a quadratic extension of k. By the proof of lemma 3.4.2, k' is the field of definition of Q.

Let U and D be the power series in z with respective constant terms u_0 and d_0 such that $u = U(z)$ and $du/v = D(z)\,dz$. Then we can rewrite (3.8) as follows:

$$\begin{cases} \dfrac{x_1x_1'}{y_1} + \dfrac{x_2x_2'}{y_2} &= (m_{1,1}U + m_{2,1})D \\ \dfrac{x_1'}{y_1} + \dfrac{x_2'}{y_2} &= (m_{2,1}U + m_{2,2})D \\ y_1^2 = E_{\mathcal{C}'}(x_1) \\ y_2^2 = E_{\mathcal{C}'}(x_2). \end{cases} \tag{3.11}$$

where the prime denotes derivation with respect to z.

Step 2: initialization. Now we explain how to compute the power series x_1, x_2, y_1, y_2 up to $O(z^2)$, keeping the notation above. First, we compute the point $Q = (x_0, y_0)$ using proposition 3.4.4. Write

$$x_1 = x_0 + v_1 z + O(z^2), \quad x_2 = x_0 + v_2 z + O(z^2).$$

Then, using the curve equations, we can compute y_1, y_2 up to $O(z^2)$ in terms of v_1, v_2 respectively. Then (3.11) gives

$$v_1 + v_2 = \frac{y_0}{x_0}(m_{1,1}u_0 + m_{2,1})d_0 = y_0(m_{2,1}u_0 + m_{2,2})d_0. \tag{3.12}$$

Combining the two lines of (3.11), we also obtain

$$(x_1 - x_0)\frac{x_1'}{y_1} + (x_2 - x_0)\frac{x_2'}{y_2} = R,$$

where $R = r_1 z + O(z^2)$ has no constant term. At order 1, this yields

$$v_1^2 + v_2^2 = y_0 r_1. \tag{3.13}$$

Combining (3.12) and (3.13) yields a quadratic equation satisfied by v_1 and v_2.

Step 3: Newton iterations. Assume that the series x_1, x_2, y_1, y_2 are known up to $O(z^n)$ for some $n \geq 2$. The system (3.11) is satisfied up to $O(z^{n-1})$ for the first two lines, and $O(z^n)$ for the last two lines. We attempt to double the current precision, and write

$$x_1 = x_1^0(z) + \delta x_1(z) + O(z^{2n}), \text{ etc.}$$

where x_1^0 is the polynomial of degree at most $n-1$ that has already been computed. The unknown series δx_i and δy_i for $i \in \{1, 2\}$ start at the term z^n. As above, we denote by x' the derivative of a power series x with respect to z.

Proposition 3.4.5. *The power series δx_1 and δx_2 satisfy a linear differential equation of the form*

$$M(z)\begin{pmatrix}\delta x_1' \\ \delta x_2'\end{pmatrix} + N(z)\begin{pmatrix}\delta x_1 \\ \delta x_2\end{pmatrix} = R(z) + O(z^{2n-1}) \tag{3.14}$$

where $M, N \in \mathrm{Mat}_{2\times 2}(k'[[z]])$ and $R \in \mathrm{Mat}_{2\times 1}(k'[[z]])$ have explicit expressions in terms of x_1^0, x_2^0, y_1^0, y_2^0, u, v, $E_{\mathcal{C}}$ and $E_{\mathcal{C}'}$. In particular,

$$M(z) = \begin{pmatrix} x_1^0/y_1^0 & x_2^0/y_2^0 \\ 1/y_1^0 & 1/y_2^0 \end{pmatrix}$$

and, writing $e = E'_{\mathcal{C}'}(x_0)$*, the constant term of* N *is*

$$\begin{pmatrix} \frac{v_1}{y_0} - \frac{x_0 v_1}{2y_0^3}e & \frac{v_2}{y_0} - \frac{x_0 v_2}{2y_0^3}e \\ -\frac{v_1}{2y_0^3}e & -\frac{v_2}{2y_0^3}e \end{pmatrix}.$$

Proof. Linearize the system (3.11). We omit the calculations. □

In order to solve the system (3.11) in quasi-linear time in the required precision, it is enough to solve equation (3.14) in quasi-linear time in n. One difficulty here, that does not appear in similar works [CE15, CMSV19], is that the matrix M is not invertible in $k'[[z]]$. Still, we can adapt the generic divide-and-conquer algorithm from [BCG+17, §13.2].

Lemma 3.4.6. *The determinant* $\det M(z) = \dfrac{x_1^0 - x_2^0}{y_1^0 y_2^0}$ *has valuation one in* z*.*

Proof. We know that y_1^0 and y_2^0 have constant term $\pm y_0 \neq 0$. The polynomials x_1^0 and x_2^0 have the same constant term x_0, but they do not coincide at order 1: if they did, then so would y_1 and y_2 because of the curve equation, contradicting lemma 3.4.3. □

By lemma 3.4.6, we can find a matrix $I \in \mathrm{Mat}_{2\times 2}(k'[[z]])$ such that $IM = \left(\begin{smallmatrix} z & 0 \\ 0 & z \end{smallmatrix}\right)$.

Lemma 3.4.7. *Let* $\kappa \geq 1$*, and assume that* $\operatorname{char} k > \kappa + 1$*. Let* $A = IN$*. Then the matrix* $A + \kappa$ *has an invertible constant term.*

Proof. By lemma 3.4.6, the leading term of $\det(M)$ is λz for some nonzero $\lambda \in k'$. By proposition 3.4.5, the constant term of $\det(A+\kappa)$ is $\lambda^2\kappa(\kappa+1)$. We omit the calculations. □

Proposition 3.4.8. *Let* $1 \leq m \leq 2n-1$*, and assume that* $\operatorname{char} k > m$*. Then we can solve* (3.14) *and compute* δx_1 *and* δx_2 *up to precision* $O(z^m)$ *using* $\widetilde{O}(m)$ *operations in* k'*.*

Proof. Write $\theta = \begin{pmatrix} \delta x_1 \\ \delta x_2 \end{pmatrix}$. Multiplying (3.14) by I, we obtain the equation

$$z\theta' + (A+\kappa)\theta = B + O(z^d), \quad \text{where } d = 2n-1 \text{ and } \kappa = 0.$$

We show that θ can be computed from this kind of equation up to $O(z^d)$ using a divide-and-conquer strategy. If $d > 1$, write $\theta = \theta_1 + z^{d_1}\theta_2$ where $d_1 = \lfloor d/2 \rfloor$. Then we have

$$z\theta_1' + (A+\kappa)\theta_1 = B + O(z^{d_1}).$$

By induction, we can recover θ_1 up to $O(z^{d_1})$. Then

$$z\theta_2' + (A + \kappa + d_1)\theta_2 = E + O(z^{d-d_1})$$

where E can be computed from θ_1. By induction, we recover θ_2 up to precision $O(z^{d-d_1})$, hence we can recover θ up to $O(z^d)$. We initialize the induction with the case $d = 1$, where we have to solve for the constant term in an equation of the form

$$(A + \kappa)\theta = B.$$

Since θ starts at z^2, the values of κ that occur are $2, \dots, m-1$ when solving (3.14) up to precision $O(z^m)$. By lemma 3.4.7, the constant term of $A + \kappa$ is invertible. This concludes the induction, and the result follows from standard lemmas in computer algebra [BCG+17, Lem. 1.12]. □

Proposition 3.4.9. *Let $n \geq 1$, and assume that $\operatorname{char} k > n$. Then we can compute the lift $\widetilde{\varphi}_P$ up to precision $O(z^n)$ within $\widetilde{O}(n)$ operations in k'.*

Proof. Combine proposition 3.4.8 and [BCG+17, Lem. 1.12]. □

3.4.3 Rational reconstruction

Finally, we want to recover the rational representation (s, p, q, r) of φ at P from its power series expansion $\widetilde{\varphi}_P$ at some finite precision. First, we estimate the degrees of the rational fractions that we want to compute; then we present the reconstruction algorithm.

Degree estimates. The degrees of s, p, q, r as morphisms from $\mathcal{C}$ to $\mathbb{P}^1$ can be computed as the intersection numbers of certain divisors on $\mathrm{Jac}(\mathcal{C}')$, namely $\varphi_P(\mathcal{C})$ and the polar divisors of s, p, q and r as functions on $\mathrm{Jac}(\mathcal{C}')$. These degrees are already known in the case of an ℓ-isogeny.

Proposition 3.4.10 ([CE15, §6.1])**.** *Let $\varphi\colon \mathrm{Jac}(\mathcal{C}) \to \mathrm{Jac}(\mathcal{C}')$ be an ℓ-isogeny, and let $P \in \mathcal{C}(k)$. Let (s, p, q, r) be the rational representation of φ at the base point P. Then the degrees of s, p, q and r as morphisms from $\mathcal{C}$ to $\mathbb{P}^1$ are 4ℓ, 4ℓ, 12ℓ, and 8ℓ respectively.*

Let F be a real quadratic field, and assume that $\mathrm{Jac}(\mathcal{C})$ and $\mathrm{Jac}(\mathcal{C}')$ have real multiplication by $\mathbb{Z}_F$ given by embeddings $\iota : \mathbb{Z}_F \to \mathrm{End}(\mathrm{Jac}(\mathcal{C}))^\dagger$ and ι'. Let

$$\varphi\colon (\mathrm{Jac}(\mathcal{C}), \iota) \to (\mathrm{Jac}(\mathcal{C}'), \iota')$$

be a β-isogeny. Denote the theta divisors on $\mathrm{Jac}(\mathcal{C})$ and $\mathrm{Jac}(\mathcal{C}')$ by Θ and Θ' respectively, and denote by $\eta_P\colon \mathcal{C} \to \mathrm{Jac}(\mathcal{C})$ the map $Q \mapsto [Q - P]$. Then $\eta_P(\mathcal{C})$ is algebraically equivalent to Θ.

Lemma 3.4.11. *The polar divisors of the rational functions s, p, q and r on $\mathrm{Jac}(\mathcal{C}')$ are algebraically equivalent to $2\Theta'$, $2\Theta'$, $6\Theta'$ and $4\Theta'$ respectively.*

Proof. See [CE15, §6.1]. For instance, $s = x_1 + x_2$ has a pole of order 1 along each of the two divisors $\{(\infty_\pm, Q) \colon Q \in \mathcal{C}'\}$; here $\infty_\pm$ are the two points at infinity on $\mathcal{C}'$, assuming that we chose a degree 6 hyperelliptic model. Each of these two divisors is algebraically equivalent to Θ'. The proof for p, q, and r is similar. □

Recall that linear equivalence classes of divisors on $\mathrm{Jac}(\mathcal{C}')$ are in bijective correspondence with isomorphism classes of line bundles (§2.1.1). Recall from §2.1.4 that if (A, ι) is a principally polarized abelian surface with real multiplication by $\mathbb{Z}_F$, then there is a bijection (2.2) between $\mathbb{Z}_F$ and the Néron–Severi group of A, denoted by $\alpha \mapsto \mathcal{L}_A^{\iota(\alpha)}$.

Lemma 3.4.12. *Let φ be a β-isogeny as above. Then the divisor $\varphi_P(\mathcal{C})$ is algebraically equivalent to the divisor corresponding to the line bundle $\mathcal{L}_{\mathrm{Jac}(\mathcal{C}')}^{\iota'(\overline{\beta})}$ on $\mathrm{Jac}(\mathcal{C}')$.*

Proof. There exists an $\alpha \in \mathbb{Z}_F$ such that the divisor $\varphi_P(\mathcal{C})$ corresponds to the line bundle $\mathcal{L}_{\mathrm{Jac}(\mathcal{C}')}^{\iota'(\alpha)}$ up to algebraic equivalence. Consider the pullback $\varphi^* (\varphi_P(\mathcal{C}))$ as a divisor on $\mathrm{Jac}(\mathcal{C})$. By definition, we have

$$\varphi^* (\varphi_P(\mathcal{C})) = \sum_{x \in \ker \varphi} (x + \eta_P(\mathcal{C})) .$$

Therefore, up to algebraic equivalence, we have

$$\varphi^* (\varphi_P(\mathcal{C})) = (\#\ker \varphi)\Theta = N_{F/\mathbb{Q}}(\beta)\Theta.$$

Since φ is a β-isogeny, the pullback $\varphi^*\Theta'$ corresponds to the line bundle $\mathcal{L}_{\mathrm{Jac}(\mathcal{C})}^{\iota(\beta)}$ up to algebraic equivalence. Therefore, for every $\gamma \in \mathbb{Z}_F$, we have

$$\varphi^* \mathcal{L}_{\mathrm{Jac}(\mathcal{C}')}^{\iota'(\gamma)} = \mathcal{L}_{\mathrm{Jac}(\mathcal{C})}^{\iota(\gamma\beta)}.$$

Therefore $\alpha\beta = N_{F/\mathbb{Q}}(\beta)$, hence $\alpha = \overline{\beta}$. □

The next step is to compute the intersection degree of Θ' and the divisor corresponding to $\mathcal{L}_{\mathrm{Jac}(\mathcal{C}')}^{\iota(\alpha)}$ on $\mathrm{Jac}(\mathcal{C}')$, for every $\alpha \in \mathbb{Z}_F$.

Proposition 3.4.13 ([Kan19, Rem. 16])**.** *Let (A, ι) be a principally polarized abelian surface with real multiplication by $\mathbb{Z}_F$, and let Θ be its theta divisor. Then the quadratic form*

$$D \mapsto (D \cdot \Theta)^2 - 2(D \cdot D)$$

on $\mathrm{NS}(A)$ *corresponds via the bijection* (2.2) *to the quadratic form on* $\mathbb{Z}_F$ *given by*

$$\alpha \mapsto 2\,\mathrm{Tr}_{F/\mathbb{Q}}(\alpha^2) - \frac{1}{2}\,\mathrm{Tr}_{F/\mathbb{Q}}(\alpha)^2.$$

Corollary 3.4.14. *Let* (A, ι) *be a principally polarized abelian surface with real multiplication by* $\mathbb{Z}_F$*, and let* Θ *be its theta divisor. Let* $\alpha \in \mathbb{Z}_F$*. Then we have*

$$\left(\mathcal{L}_A^{\iota(\alpha)} \cdot \Theta\right)^2 = \mathrm{Tr}_{F/\mathbb{Q}}(\alpha)^2.$$

Proof. Write $\alpha = a + b\sqrt{\Delta}$, where Δ denotes the discriminant of F. By proposition 3.4.13, we have

$$\left(\mathcal{L}_A^{\iota(\alpha)} \cdot \Theta\right)^2 - 2\left(\mathcal{L}_A^{\iota(\alpha)} \cdot \mathcal{L}_A^{\iota(\alpha)}\right) = 2\,\mathrm{Tr}(\alpha^2) - \frac{1}{2}\,\mathrm{Tr}(\alpha)^2 = 4b^2\Delta.$$

On the other hand, the Riemann–Roch theorem [Mil86a, Thm. 11.1] gives

$$\left(\mathcal{L}_A^{\iota(\alpha)} \cdot \mathcal{L}_A^{\iota(\alpha)}\right) = 2\,\chi\left(\mathcal{L}_A^{\iota(\alpha)}\right) = 2\sqrt{\deg \iota(\alpha)} = 2(a^2 - b^2\Delta).$$

The result follows by combining the two equations. □

Proposition 3.4.15. *Let* φ *be a* β*-isogeny as above, and let* (s, p, q, r) *be the rational representation of* φ *at* P*. Then the respective degrees of* s*,* p*,* q*, and* r *considered as morphisms from* $\mathcal{C}$ *to* $\mathbb{P}^1$ *are* $2\,\mathrm{Tr}_{F/\mathbb{Q}}(\beta)$*,* $2\,\mathrm{Tr}_{F/\mathbb{Q}}(\beta)$*,* $6\,\mathrm{Tr}_{F/\mathbb{Q}}(\beta)$ *and* $4\,\mathrm{Tr}_{F/\mathbb{Q}}(\beta)$*.*

Proof. The degrees of s, p, q, r can be computed as the intersection degrees of the polar divisors from lemma 3.4.11 and the divisor $\varphi_P(\mathcal{C})$. By lemma 3.4.12, the line bundle associated with $\varphi_P(\mathcal{C})$ is $\mathcal{L}_{\mathrm{Jac}(\mathcal{C}')}^{\iota'(\overline{\beta})}$ up to algebraic equivalence. Its intersection number with Θ' is nonnegative, hence by corollary 3.4.14, we have

$$\left(\varphi_P(\mathcal{C}) \cdot \Theta'\right) = \mathrm{Tr}_{F/\mathbb{Q}}(\overline{\beta}) = \mathrm{Tr}_{F/\mathbb{Q}}(\beta).$$

The result follows by lemma 3.4.11. □

Rational reconstruction. Let us present the rational reconstruction algorithm, and compute the power series precision that is precisely needed to recover rational representations of ℓ- and β-isogenies.

Lemma 3.4.16. *Let* $s\colon \mathcal{C} \to \mathbb{P}^1$ *be a morphism of degree* $d \geq 1$*.*

1. *If* s *is invariant under the hyperelliptic involution, then we can write* $s(u, v) = X(u)$ *where the degree of the rational fraction* X *is bounded above by* $d/2$*.*

2. *In general, let X, Y be the rational fractions such that*

$$s(u, v) = X(u) + v\,Y(u).$$

Then the degrees of X and Y are bounded above by d and $d + 3$ respectively.

Proof. The function $u\colon \mathcal{C} \to \mathbb{P}^1$ has degree 2; item 1 follows. For item 2, write

$$s(u, v) + s(u, -v) = 2X(u), \qquad \frac{s(u, v) - s(u, -v)}{v} = 2Y(u).$$

The degrees of these morphisms are bounded above by $2d$ and $2d + 6$ respectively, and item 1 applies. □

Proposition 3.4.17. *Let $\widetilde{\varphi}_P$ and $\widetilde{\varphi}_{i(P)}$ be local lifts of φ_P at P and $i(P)$ in the uniformizers z and $i(z)$, where i denotes the hyperelliptic involution on $\mathcal{C}$. Let $n = 8\ell + 7$ in the Siegel case, and $n = 4\,\mathrm{Tr}_{F/\mathbb{Q}}(\beta) + 7$ in the Hilbert case. Then, given $\widetilde{\varphi}_P$ and $\widetilde{\varphi}_{i(P)}$ at precision $O(z^n)$, we can compute the rational representation of φ at P within $\widetilde{O}(n)$ field operations.*

Proof. It is sufficient to recover the rational fractions s and p; afterwards, q and r can be computed from the equation of $\mathcal{C}'$.

First, assume that P is a Weierstrass point of $\mathcal{C}$. Then s, p are invariant under the hyperelliptic involution. Therefore we have to recover univariate rational fractions in u of degree $d \leq 2\ell$ (resp. $d \leq \mathrm{Tr}_{F/\mathbb{Q}}(\beta)$), by propositions 3.4.10 and 3.4.15. This can be done in $\widetilde{O}(n)$ field operations from the power series expansions of s and p up to precision $O(u^{2d+1})$ [BCG+17, §7.1]. Since u has valuation 2 in z, we need to compute $\widetilde{\varphi}_P$ at precision $O(z^{4d+1})$.

Second, assume that P is not a Weierstrass point of $\mathcal{C}$. Then the series defining $s(u, -v)$ and $p(u, -v)$ are given by $\widetilde{\varphi}_{i(P)}$. We have to compute rational fractions of degree $d \leq 4\ell + 3$ (resp. $d \leq 2\,\mathrm{Tr}_{F/\mathbb{Q}}(\beta) + 3$) in u, by lemma 3.4.16. Since u has valuation 1 in z, this can be done in $\widetilde{O}(n)$ field operations if $\widetilde{\varphi}_P$ and $\widetilde{\varphi}_{i(P)}$ are known up to precision $O(z^{2d+1})$. □

3.5 Summary of the algorithm

We summarize the isogeny algorithm and state precise complexity results in both the Siegel case (theorem 3.5.2) and the Hilbert case (theorem 3.5.3). Let k be any field, and let F be a fixed real quadratic field.

Algorithm 3.5.1. Let A and A' be principally polarized abelian surfaces over k. Assume either that A and A' are ℓ-isogenous where ℓ is a prime (the Siegel case), or that A and A' have real multiplication by $\mathbb{Z}_F$ and are β-isogenous where $\beta \in \mathbb{Z}_F$ is a totally positive prime (the Hilbert case). Moreover, assume that the isogeny, called φ, is generic in the sense of definition 3.3.3 or definition 3.3.6.

1. Use Mestre's algorithm [Mes91] to construct curve equations $\mathcal{C}$ and $\mathcal{C}'$ whose Jacobians are isomorphic to A and A' respectively. In the Hilbert case, use algorithm 3.3.11 to ensure that $\mathcal{C}$ and $\mathcal{C}'$ are potentially Hilbert-normalized.
2. Compute at most 4 candidates for the tangent matrix of the isogeny φ using proposition 3.3.4 in the Siegel case, or proposition 3.3.9 in the Hilbert case. Run the rest of the algorithm for all these candidates; in general, only one will produce meaningful results.
3. Choose a base point P on $\mathcal{C}$ such that φ_P is of generic type using proposition 3.4.4, and compute the power series $\widetilde{\varphi}_P$ and $\widetilde{\varphi}_{i(P)}$ up to precision $O\left(z^{8\ell+7}\right)$, resp. $O\left(z^{4\operatorname{Tr}(\beta)+7}\right)$ using proposition 3.4.9.
4. Recover the rational representation of φ at P using proposition 3.4.17.

Theorem 3.5.2. *Let ℓ be a prime, and let k be a field such that* $\operatorname{char} k = 0$ *or* $\operatorname{char} k > 8\ell+7$. *Assume that there exists an algorithm to evaluate derivatives of Siegel modular equations of level ℓ at given Igusa invariants (j, j') over k using $C_{\text{eval}}(\ell)$ operations in k.*

Let A, A' be principally polarized abelian surfaces over k whose Igusa invariants $j(A), j(A')$ are well defined, and assume that there exists an ℓ-isogeny $\varphi\colon A \to A'$ defined over k which is generic in the sense of definition 3.3.3. Then, given $j(A)$ and $j(A')$, algorithm 3.5.1 succeeds and returns

1. *a tower k'/k of at most three quadratic extensions,*
2. *hyperelliptic curve equations $\mathcal{C}$ and $\mathcal{C}'$ over k' whose Jacobians are isomorphic to A and A' respectively,*
3. *a point $P \in \mathcal{C}(k')$,*
4. *the rational representation $(s, p, q, r) \in k'(u, v)^4$ of an ℓ-isogeny $\varphi\colon \operatorname{Jac}(\mathcal{C}) \to \operatorname{Jac}(\mathcal{C}')$ at P.*

The cost of algorithm 3.5.1 in the Siegel case is $O\left(C_{\text{eval}}(\ell)\right) + \widetilde{O}(\ell)$ elementary operations and $O(1)$ square roots in k'.

Proof. Mestre's algorithm returns curve equations $\mathcal{C}$ and $\mathcal{C}'$ defined over extensions of k of degree at most 2, and costs $O(1)$ operations in k and $O(1)$

square roots. Under our hypotheses, proposition 3.3.4 applies and allows us to recover the matrix $\mathrm{Sym}^2(d\varphi)$ using $O(C_{\mathrm{eval}}(\ell)) + O(1)$ operations in k. We recover $d\varphi$ up to sign using $O(1)$ square roots and elementary operations; since φ is defined over k, extending the base field is not necessary. We choose the base point P on $\mathcal{C}$ such that φ_P is of generic type using proposition 3.4.4, perhaps taking another extension of degree 2. By proposition 3.4.9, we can compute the local lifts $\widetilde{\varphi}_P$ and $\widetilde{\varphi}_{i(P)}$ up to precision $8\ell+7$ within $\widetilde{O}(\ell)$ field operations; this is where we use the hypothesis on $\operatorname{char} k$. Finally, we recover the rational representation at P using $\widetilde{O}(\ell)$ field operations by proposition 3.4.17. The total number of quadratic extensions taken is at most 3. □

We conclude with the analogue of theorem 3.5.2 in the Hilbert case.

Theorem 3.5.3. *Let F be a real quadratic field, and let $\beta \in \mathbb{Z}_F$ be a totally positive prime. Let k be a field such that either $\operatorname{char} k = 0$ or $\operatorname{char} k > 4\operatorname{Tr}_{F/\mathbb{Q}}(\beta)+7$. Assume that there exists an algorithm to evaluate derivatives of Hilbert modular equations of level β in Igusa invariants at a given point (j, j') over k using $C_{\mathrm{eval}}(\beta)$ operations in k.*

Let A, A' be principally polarized abelian surfaces over k with real multiplication by $\mathbb{Z}_F$ whose Igusa invariants $j(A), j(A')$ are well defined, and assume that there exists a β-isogeny $\varphi\colon A \to A'$ defined over k which is generic in the sense of definition 3.3.6. Then, given $j(A)$ and $j(A')$, algorithm 3.5.1 succeeds and returns

1. *a tower k'/k of at most three quadratic extensions,*
2. *hyperelliptic curve equations $\mathcal{C}$ and $\mathcal{C}'$ over k' whose Jacobians are isomorphic to A and A' respectively,*
3. *a point $P \in \mathcal{C}(k')$,*
4. *at most 4 possible candidates for the rational representation $(s, p, q, r) \in k'(u, v)^4$ of a β- or $\overline{\beta}$-isogeny $\varphi\colon \mathrm{Jac}(\mathcal{C}) \to \mathrm{Jac}(\mathcal{C}')$ at P.*

The cost of algorithm 3.5.1 is $O\left(C_{\mathrm{eval}}(\beta)\right) + \widetilde{O}\left(\operatorname{Tr}_{F/\mathbb{Q}}(\beta)\right) + O_F(1)$ elementary operations and $O(1)$ square roots in k'; the implied constants, $O_F(1)$ excepted, are independent of F.

Note that $C_{\mathrm{eval}}(\beta)$ also depends on F. We expect that the algorithm returns only one answer for the rational representation of φ at P; if the algorithm outputs several answers, we could implement tests of correctness, but they might be more expensive than the isogeny algorithm itself.

Proof. We use algorithm 3.3.11 to construct the curve equations $\mathcal{C}$ and $\mathcal{C}'$. We obtain potentially Hilbert-normalized curves, and each of them is defined over an extension of k of degree dividing 4. This requires $O_F(1)$ elementary operations and $O(1)$ square roots in k. We may assume that $\mathcal{C}$ and $\mathcal{C}'$ are Hilbert-normalized for some choice of real multiplication embeddings that are compatible via φ, which becomes either a β- or a $\overline{\beta}$-isogeny.

Under our hypotheses, proposition 3.3.9 applies and allows us to recover two possible values for $(d\varphi)^2$ within $O(C_{\text{eval}}(\beta)) + O(1)$ operations in k, and hence 4 possible values for $d\varphi$, by taking $O(1)$ square roots. We can now make a change of variables to the (not necessarily Hilbert-normalized) curves output by Mestre's algorithm, so that each curve is defined over an extension of k of degree at most 2. The end of the algorithm is similar to the Siegel case: we take an extension of degree 2 to find the base point, then we try to compute the rational representation for each value of $d\varphi$ using $\widetilde{O}(\mathrm{Tr}_{F/\mathbb{Q}}(\beta))$ operations in k. For the correct value of $d\varphi$, rational reconstruction will succeed and output fractions of the correct degrees. □

3.6 An example in the Hilbert case for $\mathbb{Q}(\sqrt{5})$

To conclude this chapter, we present a variant of our algorithm in the case of principally polarized abelian varieties with real multiplication by $\mathbb{Z}_F$ where $F = \mathbb{Q}(\sqrt{5})$. In this case, the Humbert surface is rational: its function field can be generated by only two elements called *Gundlach invariants*, introduced in §2.3.4. Having only two coordinates reduces the size of modular equations, as we mentioned in §2.4.4.

We work over $\mathbb{C}$, but the methods of §3.3.3 show that the computations remain valid in general. We illustrate the algorithm with an example of cyclic isogeny of degree 11 over a finite field.

3.6.1 Variants in the isogeny algorithm

There are two improvements. First, in the construction of Hilbert-normalized curves, we take advantage of the known structure of the ring of Hilbert modular forms to find an expression for the standard curve $\mathcal{C}_F(\tau)$ when $\tau \in \mathbb{H}_1^2$. Second, reducing the number of variables allows us to compute $(d\varphi)^2$ using only 2×2 matrices.

Constructing potentially Hilbert-normalized curves. An alternative method to algorithm 3.3.11 is given by proposition 2.3.17, which gives

the expression of the "universal" Siegel modular form $f_{8,6}$ as a Hilbert modular form. We keep the notation used in this proposition and in §3.2.1. By propositions 3.2.7 and 3.2.13, the standard curve $\mathcal{C}_F(\tau)$ attached to $\tau \in \mathbb{H}_1^2$ is proportional to the curve $y^2 = \sum b_i(\tau)x^i$. The algorithm to compute a potentially Hilbert-normalized curve $\mathcal{C}$ from its Igusa invariants (j_1, j_2, j_3) runs as follows.

Algorithm 3.6.1.

1. Compute Gundlach invariants (g_1, g_2) mapping to the Igusa invariants (j_1, j_2, j_3) via the Hilbert embedding using proposition 2.3.16, and compute values for the generators G_2, F_6, F_{10} giving these values of g_1, g_2.
2. Compute b_3^2, $b_1 b_5$, etc. using proposition 2.3.17.
3. Recover values for the coefficients: let b_3 be any square root of b_3^2; choose any value for b_1, which gives b_5; finally, solve a quadratic equation to find b_0 and b_6.

We can always choose values of G_2, F_6, and F_{10} in such a way that b_3^2 is a square in k. Then the output is defined over a quadratic extension of k. Even if arbitrary choices are made during algorithm 3.6.1, the output will be potentially Hilbert-normalized.

Computing the tangent matrix. Consider the Hilbert modular equations in Gundlach invariants of level β, denoted by $\Psi_{\beta,1}$ and $\Psi_{\beta,2}$, as elements of the ring $\mathbb{Q}[G_1, G_2, G_1', G_2']$ after multiplication by their denomiators. Define the 2×2 matrices

$$D\Psi_{\beta,L} = \left(\frac{\partial \Psi_n}{\partial G_k}\right)_{1 \le n,k \le 2} \quad \text{and} \quad D\Psi_{\beta,R} = \left(\frac{\partial \Psi_n}{\partial G_k'}\right)_{1 \le n,k \le 2}.$$

Then the analogue of proposition 3.3.9 holds, where we replace derivatives of Igusa invariants by derivatives of Gundlach invariants. The relation between these derivatives is given by proposition 2.3.16.

3.6.2 An example of cyclic isogeny

We illustrate our algorithm in the Hilbert case with $F = \mathbb{Q}(\sqrt{5})$ by computing a β-isogeny between Jacobians with real multiplication by $\mathbb{Z}_F$, where

$$\beta = 3 + \frac{1+\sqrt{5}}{2} \in \mathbb{Z}_F,$$

so that $\mathrm{N}_{F/\mathbb{Q}}(\beta) = 11$ and $\mathrm{Tr}_{F/\mathbb{Q}}(\beta) = 7$. We work over the prime finite field $k = \mathbb{F}_{56311}$, whose characteristic is large enough for our purposes. We

choose a trivialization of $\mathbb{Z}_F \otimes_{\mathbb{Z}} k$, in other words a square root of 5 in k, so that $\beta = 26213$.

Consider the Gundlach invariants

$$(g_1, g_2) = \big(23, 56260\big), \quad (g_1', g_2') = \big(8, 36073\big).$$

The corresponding Igusa invariants are

$$(j_1, j_2, j_3) = \big(14030, 9041, 56122\big), \quad (j_1', j_2', j_3') = \big(13752, 42980, 12538\big);$$

they lie on the Humbert surface as expected.

In order to reconstruct a Hilbert-normalized curve, we apply algorithm 3.6.1. We obtain the curve equations

$$\begin{aligned} \mathcal{C}\colon v^2 &= 13425u^6 + 34724u^5 + 102u^3 + 54150u + 11111 \\ \mathcal{C}'\colon y^2 &= 47601x^6 + 35850x^5 + 40476x^3 + 24699x + 40502. \end{aligned}$$

The derivatives of Gundlach invariants are given by the matrices

$$\partial g^{(F)}(\mathcal{C}) = \begin{pmatrix} 43658 & 17394 \\ 16028 & 26656 \end{pmatrix}, \quad \partial g^{(F)}(\mathcal{C}') = \begin{pmatrix} 15131 & 739 \\ 50692 & 49952 \end{pmatrix}.$$

After computing derivatives of the modular equations as in proposition 3.3.9, we find that the isogeny φ is compatible with the real multiplication embeddings for which $\mathcal{C}$ and $\mathcal{C}'$ are Hilbert-normalized. We do not known whether φ is a β- or a $\overline{\beta}$-isogeny, so we have four candidates for the tangent matrix up to sign:

$$\begin{aligned} d\varphi_{\beta,\pm} &= \begin{pmatrix} 38932\alpha + 19466 & 0 \\ 0 & \pm(53318\alpha + 26659) \end{pmatrix}, \\ d\varphi_{\overline{\beta},\pm} &= \begin{pmatrix} 50651\alpha + 53481 & 0 \\ 0 & \pm(11076\alpha + 5538) \end{pmatrix} \end{aligned}$$

where $\alpha^2 + \alpha + 2 = 0$. We see that φ is only defined over the quadratic extension $k(\alpha)$.

The curve $\mathcal{C}$ has the rational Weierstrass point $\big(36392, 0\big)$. We can bring it to $(0, 0)$, so that $\mathcal{C}$ is of the more standard form

$$\mathcal{C}\colon v^2 = 33461u^6 + 7399u^5 + 16387u^4 + 34825u^3 + 14713u^2 + u.$$

This change of variables multiplies the tangent matrix $d\varphi$ on the right by

$$\begin{pmatrix} 44206 & 18649 \\ 0 & 7615 \end{pmatrix}.$$

Choose $P = (0,0)$ as a base point on $\mathcal{C}$, and choose $z = \sqrt{u}$ as a uniformizer; P is a Weierstrass point, and we check that φ_P is of generic type. We solve the differential system (3.11) up to precision $O(z^{35})$, or any higher precision. It turns out that the correct tangent matrix is $d\varphi_{\overline{\beta},+}$ as the other powers series do not come from rational fractions of the prescribed degrees. We obtain

$$s(u) = \frac{50255u^6 + 40618u^5 + 17196u^4 + 9527u^3 + 22804u^2 + 49419u + 11726}{u^6 + 40883u^5 + 22913u^4 + 41828u^3 + 18069u^2 + 14612u + 7238},$$
$$p(u) = \frac{35444u^6 + 9569u^5 + 52568u^4 + 3347u^3 + 9325u^2 + 32206u + 7231}{u^6 + 40883u^5 + 22913u^4 + 41828u^3 + 18069u^2 + 14612u + 7238}.$$

The degrees of s and p agree with proposition 3.4.15. The isogeny is k-rational at the level of Kummer surfaces, but not on the Jacobians themselves: α appears on the numerator of $r(u,v)$.

Chapter 4

Heights of rational fractions and interpolation

This chapter corresponds to the preprint [Kie20c]. Let F be a univariate polynomial of rational fraction of degree d defined over a number field. We present bounds on the *height* of F, which encodes the size of the coefficients of F, in terms of the heights of its evaluations at small integers: we review well-known bounds obtained from interpolation algorithms given values of F at $d+1$ or $2d+1$ points, and obtain tighter results when considering a larger number of evaluation points. While interesting in their own right, and readable independently, the results of this chapter have a particular purpose in this thesis: they will be essential in proving height bounds for modular equations in chapter 5.

4.1 Presentation of the problem

Let F be a univariate rational fraction of degree d defined over $\mathbb{Q}$. To define the height $h(F)$ of F, write $F = P/Q$ where $P, Q \in \mathbb{Z}[Y]$ are coprime; then $h(F)$ is the maximum value of $\log|c|$, where c runs through the nonzero coefficients of P and Q. In particular, if $x = p/q$ is a rational number in irreducible form, then $h(x) = \log\max\{|p|, |q|\}$.

Heights can be generalized to arbitrary number fields, and are a basic tool in diophantine geometry [HS00, Part B]. They are also meaningful from an algorithmic point of view: the amount of memory needed to store F in a computer is in general $O(d\,h(F))$, and the cost of manipulating F grows with the size of its coefficients.

As announced, this chapter presents relations between the height of F and the heights of evaluations $F(x)$, where x is an integer. One direction

is easy: by [HS00, Prop. B.7.1], we have

$$h(F(x)) \leq d\,h(x) + h(F) + \log(d+1). \tag{4.1}$$

In the other direction, when we want to bound $h(F)$ from the heights of its values, matters are more complicated.

An easy case is when $F \in \mathbb{Z}[Y]$ is a polynomial with integer coefficients of degree at most $d \geq 1$. Then looking at the archimedean absolute value of the coefficients of F is sufficient to bound $h(F)$. Moreover, given height bounds on $d+1$ values of F, the Lagrange interpolation formula allows us to bound $h(F)$ in a satisfactory way. For instance, assuming that

$$h(F(i)) \leq H \quad \text{for every } i \in [\![0, d]\!]\,, \tag{4.2}$$

we easily obtain

$$h(F) \leq H + d\log(2d) + \log(d+1). \tag{4.3}$$

This result can be refined and adapted to other sets of intepolation points [BS10, Lem. 20], [Paz19, Lem. 4.1]; in any case the bound on $h(F)$ is roughly H up to additional terms in $O(d \log d)$. This is consistent with inequality (4.1).

When F is a rational fraction or even a polynomial with rational coefficients, this nice result breaks down, and surprisingly little information appears in the literature despite the simplicity of the question.

Polynomials. First, consider the case where F is a polynomial in $\mathbb{Q}[Y]$ of degree at most $d \geq 1$. Then F is determined by its values at $d+1$ distinct points. Let $x_1, \ldots, x_{d+1}$ be distinct integers, let $H \geq 1$, and assume that $h(F(x_i)) \leq H$ for every i. This time, the Lagrange interpolation formula yields a bound on $h(F)$ which is roughly $O(dH)$ (see proposition 4.4.2). This is intuitive enough: in general, computing F from its values $F(x_i)$ involves reducing the rational numbers $F(x_i)$ to the same denominator, thus multiplying the heights of the input by the number of evaluation points. But then, inequality (4.1) is very pessimistic at each of the evaluation points x_i: massive cancellations occur with the denominator of F, and the "expected" size of $F(x_i)$ is divided by d.

However, if we consider *more* than $d+1$ evaluation points $x_1 \ldots, x_N$ where $h(F(x_i)) \leq H$, we will likely find an evaluation point where inequality (4.1) is accurate, and hence obtain a bound on $h(F)$ of the form $O(H)$ rather than $O(dH)$. We prove the following result in this direction.

Proposition 4.1.1. *Let L be a number field, and let $[\![A, B]\!]$ be an interval in $\mathbb{Z}$. Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $d \geq 1$, let $F \in L[Y]$ be a polynomial of degree at most d, let $N \geq d + 1$, and let $x_1, \ldots, x_N$ be distinct elements of $[\![A, B]\!]$. Assume that $h(F(x_i)) \leq H$ for all $1 \leq i \leq N$. Then we have*

$$h(F) \leq \frac{N}{N-d} H + D \log(D) + d \log(2M) + \log(d+1).$$

For instance, we obtain a bound on $h(F)$ which is linear in H when considering $N = 2d$ evaluation points.

Rational fractions. Second, consider the case where $F \in \mathbb{Q}(Y)$ is a rational fraction of degree at most $d \geq 1$. Then F is determined by its values at $2d + 1$ points. If $x_1, \ldots, x_{2d+1}$ are distinct integers which are not poles of F, and $h(F(x_i)) \leq H$ for every i, then a direct analysis of the interpolation algorithm yields a bound on $h(F)$ which is roughly $O(d^2 H)$ (see proposition 4.5.7). As above, it makes sense to ask for a bound which is linear in H when more evaluation points are given.

In this case we could imagine cases where $F = P/Q$ has a very large height, but massive cancellations happen in all the quotients $P(x_i)/Q(x_i)$. This makes the result more intricate.

Proposition 4.1.2. *Let L be a number field, and let $[\![A, B]\!]$ be an interval in $\mathbb{Z}$. Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L(Y)$ be a rational fraction of degree at most $d \geq 1$. Let S be a subset of $[\![A, B]\!]$ containing no poles of F, let $\eta \geq 1$, and let $H \geq \max\{4, \log(2M)\}$. Assume further that*

1. *$h(F(x)) \leq H$ for every $x \in S$.*
2. *S contains at least D/η elements.*
3. *$D \geq \max\{\eta d^3 H, 4\eta d d_L\}$.*

Then we have

$$h(F) \leq H + C_L \eta d \log(\eta d H) + d \log(2M) + \log(d+1),$$

where C_L is a constant depending only on L. We can take $C_{\mathbb{Q}} = 960$.

The number of evaluation points in this result is quite large, and depends on H. Still, proposition 4.1.2 is the main result that we need in chapter 3, and is strong enough to imply the following.

Proposition 4.1.3. *Let $c \geq 1$, let $d \geq 1$, and let $F \in \mathbb{Q}(Y)$ be a rational fraction of degree at most d. Let $V \subset \mathbb{Z}$ be a finite set such that F has no poles in $\mathbb{Z}\backslash V$. Assume that for every $x \in \mathbb{Z}\backslash V$, we have*

$$h(F(x)) \leq c \max\{1, d \log d + d\, h(x)\}.$$

Then there exists a constant $C = C(c, \#V)$ such that

$$h(F) \leq C d \log(4d).$$

We can take $C = (4c + 1923)\big(12 + \log \max\{1, \#V\} + 2\log(c)\big)$.

It would be interesting to know whether we can obtain an efficient bound on $h(F)$ using only $O(d)$ evaluation points, as in the case of polynomials, instead of $O(d^3 H)$. The constants in propositions 4.1.2 and 4.1.3 are not optimal: tighter bounds can be obtained following the same ideas of proof, at the cost of lengthier expressions. The rest of this chapter is devoted to the definition of heights over number fields and the proofs of propositions 4.1.1 to 4.1.3.

4.2 Definition of heights

Let L be a number field of degree d_L over $\mathbb{Q}$. We denote by $\mathcal{V}_L^0$ (resp. $\mathcal{V}_L^\infty$) the set of all nonarchimedean (resp. archimedean) places of L, and write $\mathcal{V}_L = \mathcal{V}_L^0 \sqcup \mathcal{V}_L^\infty$. Let $\mathcal{P}_\mathbb{Q}$ (resp. $\mathcal{P}_L$) be the set of primes in $\mathbb{Z}$ (resp. prime ideals in the ring of integers $\mathbb{Z}_L$ of L).

For each place v of L, the local degree of $L/\mathbb{Q}$ at v is $d_v = [L_v : \mathbb{Q}_v]$, where subscripts denote completion. Denote by $|\cdot|_v$ the normalized absolute value associated with v: when $v \in \mathcal{V}_L^0$, and $p \in \mathcal{P}_\mathbb{Q}$ is the prime below v, we have $|p|_v = 1/p$. If $\mathfrak{p} \in \mathcal{P}_L$, we denote the $\mathfrak{p}$-adic valuation by $v_\mathfrak{p}$. By convention, the $\mathfrak{p}$-adic valuation of zero is $+\infty$.

The (absolute logarithmic Weil) *height* of projective tuples, affine tuples, polynomials and rational fractions over L is defined as follows [HS00, §B.2 and §B.7].

Definition 4.2.1. Let $n \geq 1$, and let $a_0, \ldots, a_n \in L$.

1. The projective height of $(a_0 : \cdots : a_n) \in \mathbb{P}_L^n$ is

$$h(a_0 : \ldots : a_n) = \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log\big(\max_{0 \leq i \leq n} |a_i|_v\big).$$

2. The affine height of $(a_1, \dots, a_n) \in L^n$ is defined as the projective height of the tuple $(1 : a_1 : \cdots : a_n)$:

$$h(a_1, \dots, a_n) \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log\bigl(\max\{1, \max_{1 \le i \le n} |a_i|_v\}\bigr).$$

In particular, for $a \in L$, we have

$$h(a) = h(1 : a) = \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log\bigl(\max\{1, |a|_v\}\bigr).$$

3. The height of a polynomial $P \in L[Y_1, \dots, Y_n]$ is the height of the affine tuple formed by all its coefficients: if

$$P = \sum_{k=(k_1,\dots,k_n) \in \mathbb{N}^n} c_k Y_1^{k_1} \cdots Y_n^{k_n},$$

then for $v \in \mathcal{V}_L$, we write

$$|P|_v = \max_{k \in \mathbb{N}^n} |c_k|_v$$

and we define $h(P)$ as

$$h(P) = \sum_{v \in \mathcal{V}_L} \frac{d_v}{d_L} \log\bigl(\max\{1, |P|_v\}\bigr).$$

If $\mathfrak{p} \in \mathcal{P}_L$ is a prime ideal, we also define the $\mathfrak{p}$-adic valuation of P as

$$v_{\mathfrak{p}}(P) = \min_{k \in \mathbb{N}^n} v_{\mathfrak{p}}(c_k).$$

4. Let $F \in L(Y_1, \dots, Y_n)$ be a multivariate rational fraction over L, and choose coprime polynomials $P, Q \in L[Y_1, \dots, Y_n]$ such that $F = P/Q$. Then we define $h(F)$ as the height of the projective tuple formed by all the coefficients of P and Q.

Elementary properties of heights.

1. The projective height of $(a_0 : \cdots : a_n)$ is indeed invariant under multiplication by a common scalar factor, by the product formula [HS00, Lem. B.2.1(a)]. Therefore the height of a fraction is also well defined.

2. Heights are independent of the ambient field [HS00, Lem. B.2.1(c)]. This is another consequence of the product formula: in particular, we have
$$\sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} = 1.$$
3. If $x, y \in L$, then $h(xy) \leq h(x) + h(y)$; if $x \neq 0$, then $h(1/x) = h(x)$.
4. Heights satisfy the *Northcott property*: for every $H \in \mathbb{R}$, the number of projective tuples $(a_0 : \cdots : a_n) \in \mathbb{P}^n_L$ such that $h(a_0 : \cdots : a_n) \leq H$ is finite [HS00, Thm. B.2.3].
5. If $L = \mathbb{Q}$, then definition 4.2.1 coincides with the naive definition of heights given in §4.1.

As definition 4.2.1 suggests, in order to obtain height bounds for polynomials and rational fractions, we will try to bound their coefficients from above in all the absolute values associated with places of L.

4.3 Heights of values and roots of polynomials

In this section, we relate the height of a univariate polynomial over L with the height of its roots. We also give bounds from above on the height of evaluations of polynomials, generalizing eq. (4.1). The statements are all easy consequences of the formulæ in definition 4.2.1, and will be used on several occasions.

Let us start with the evaluation of polynomials; the following proposition is a slight generalization of [HS00, Prop. B.7.1].

Proposition 4.3.1. *Let $P \in L[Y_1, \ldots, Y_n]$ be a multivariate polynomial of total degree at most $d \geq 1$, let $1 \leq m \leq n$, and let $y_1, \ldots, y_m \in L$. Let $Q = P(y_1, \ldots, y_m, Y_{m+1}, \ldots, Y_n)$. Then we have*
$$h(Q) \leq h(P) + m \log(d+1) + d\, h(y_1, \ldots, y_n).$$
More generally, if $\mathcal{I}_1 \sqcup \cdots \sqcup \mathcal{I}_r$ is a partition of $[\![1, m]\!]$, and if d_k denotes the total degree of P in the variables Y_i for $i \in \mathcal{I}_k$, then we have
$$h(Q) \leq h(P) + \sum_{k=1}^{r} (\#\mathcal{I}_k) \log(d_k + 1) + \sum_{k=1}^{r} d_k\, h\big((y_i)_{i \in \mathcal{I}_k}\big).$$

Proof. It is enough to prove the second statement. Each term of Q is a sum of terms of the form
$$c \cdot y_1^{a_1} \cdots y_m^{a_m} \cdot Y_{m+1}^{a_{m+1}} \cdots Y_n^{a_n}$$

where c is a coefficient of P, and $a_i \leq d_k$ for each $1 \leq i \leq m$ and $1 \leq k \leq r$ such that $i \in \mathcal{I}_k$. For each $v \in \mathcal{V}_L$, we have

$$|c \cdot y_1^{a_1} \cdots y_m^{a_m}|_v \leq |P|_v \prod_{k=1}^{r} \Big(\max\big\{1, \max_{i \in \mathcal{I}_k} |y_i|_v\big\} \Big)^{d_k}.$$

If $v \in \mathcal{V}_L^0$, then the ultrametric inequality implies that

$$|P(y_1, \ldots, y_m, Y_{m+1}, \ldots, Y_n)|_v \leq |P|_v \prod_{k=1}^{r} \Big(\max\big\{1, \max_{i \in \mathcal{I}_k} |y_i|_v\big\} \Big)^{d_k}. \tag{4.4}$$

If $v \in \mathcal{V}_L^\infty$, then inequality (4.4) holds after multiplying the right hand side by the number of possible monomials in $Y_1, \ldots, Y_m$, which is

$$\prod_{k=1}^{r} (d_k + 1)^{\#\mathcal{I}_k}.$$

Taking logarithms and summing gives the result. □

As a consequence, we can bound the height of a monic polynomial given the height of its roots.

Proposition 4.3.2. *Let $Q \in L[Y]$ be monic polynomial of degree $d \geq 1$, and let $\alpha_1, \ldots, \alpha_d$ be its roots in the algebraic closure of L. Then*

$$h(Q) \leq \sum_{i=1}^{d} h(\alpha_k) + d \log 2.$$

Proof. Apply proposition 4.3.1 on the multivariate polynomial

$$P = \prod_{k=1}^{d} (Y_{d+1} - Y_k)$$

with $m = d$, $y_k = \alpha_k$, and $\mathcal{I}_k = \{k\}$. Since the coefficients of P all belong to $\{-1, 0, 1\}$, we have $h(P) = 0$. □

Conversely, the height of a univariate polynomial over L controls the height of its roots.

Proposition 4.3.3. *Let $P \in L[Y] \backslash \{0\}$, and let α be a root of P. Then*

$$h(\alpha) \leq h(P) + \log(2).$$

Proof. We reproduce the proof given in a lecture by F. Pazuki. We can assume that P is monic. Let $v \in \mathcal{V}_L$. We want to show that $|\alpha|_v \leq |P|_v$ if $v \in \mathcal{V}_L^0$, and $|\alpha|_v \leq 2\,|P|_v$ if $v \in \mathcal{V}_L^\infty$. Since P is monic, we always have $|P|_v \geq 1$. Write $P = Y^n + \sum_{k=0}^{n-1} c_k Y^k$, for some $n \geq 1$.

If $v \in \mathcal{V}_L^0$, we can assume that $|\alpha|_v \geq 1$. Then

$$|\alpha|_v = \Big|\sum_{i=0}^{n-1} c_k \alpha^k\Big|_v \leq |P|_v\,|\alpha|_v^{n-1}\,,$$

so $|\alpha|_v \leq |P|_v$.

If $v \in \mathcal{V}_L^\infty$, we can assume that $|\alpha|_v \geq 2$. Then, by the triangle inequality, we obtain

$$|\alpha|_v \leq |P|_v\,|\alpha_v|^{n-1}\left(1 + \frac{1}{|\alpha|_v} + \cdots + \frac{1}{|\alpha|_v^{n-1}}\right) \leq 2\,|\alpha|_v^{n-1}\,|P|_v\,,$$

so $|\alpha|_v \leq 2\,|P|_v$. Taking logarithms and summing over all $v \in \mathcal{V}_L$ yields the result. □

4.4 Heights of polynomials from their values

This section presents bounds on the height of a polynomial $F \in L[Y]$ of degree at most $d \geq 1$ in terms of the heights of evaluations of F. We take our evaluation points to be integers in an interval $[\![A, B]\!] \subset \mathbb{Z}$, and we write $D = B - A$ and $M = \max\{|A|\,, |B|\}$. Our main tool is the Lagrange interpolation formula: if $x_1, \ldots, x_{d+1} \in [\![A, B]\!]$ are distinct, then

$$F = \frac{1}{D!}\sum_{i=0}^{d+1} F(x_i)Q_i \qquad \text{where} \quad Q_i = D!\frac{\prod_{j\neq i}(Y - x_j)}{\prod_{j\neq i}(x_i - x_j)} \in \mathbb{Z}[Y]. \tag{4.5}$$

Lemma 4.4.1. *In the notation of eq. (4.5), we have* $|Q_i| \leq D!\,(2M)^d$.

Proof. Since the denominator $\prod_{k\neq i}(x_i - x_k)$ divides $D!$, we have

$$Q_i = N_i \prod_{j\neq i}(X - x_j)$$

for some $N_i \in \mathbb{Z}$ dividing $D!$. Therefore, for every $0 \leq k \leq d$, if c_k denotes the coefficient of X^{d-k} in Q_i, we have

$$|c_k| \leq |N_i| \binom{d}{k} M^k \leq D!\,2^d M^d.$$ □

A straightforward application of the Lagrange interpolation formula on $d+1$ evaluation points yields the following result.

Proposition 4.4.2. *Let $d \geq 1$, let $F \in L[Y]$ be a polynomial of degree at most d, and let $x_1, \dots, x_{d+1}$ be distinct integers in $[\![A, B]\!]$. Write $D = B - A$ and $M = \max\{A, B\}$. Assume that $h(F(x_i)) \leq H$ for every $1 \leq i \leq d+1$. Then we have*

$$h(F) \leq (d+1)H + D\log(D) + d\log(2M) + \log(d+1).$$

Proof. Let v be a place of L. If $v \in \mathcal{V}_L^0$, then (4.5) yields

$$\begin{aligned}\max\{1, |F|_v\} &\leq \left|\frac{1}{D!}\right|_v \max\{1, |F(x_1)|_v, \dots, |F(x_{d+1})|_v\} \\ &\leq \left|\frac{1}{D!}\right|_v \prod_{i=1}^{n+1} \max\{1, |F(x_i)|_v\}.\end{aligned}$$

If $v \in \mathcal{V}_L^\infty$, then lemma 4.4.1 yields

$$\max\{1, |F|_v\} \leq \sum_{i=1}^{d+1} |F(x_i)|_v \, 2^d M^d \leq (d+1)2^d M^d \prod_{i=1}^{d+1} \max\{1, |F(x_i)|_v\}.$$

Since $h(1/D!) = h(D!) \leq D\log(D)$, taking logarithms and summing gives the result. □

It is interesting to compare proposition 4.4.2 with [HS00, Cor. B.2.6], using the evaluation maps at x_i as linear maps from $L[Y]$ to L: under the hypotheses of the proposition, the height of the *tuple* $(F(x_1), \dots, F(x_{d+1}))$ can be as large as $(d+1)H$.

Better bounds on $h(F)$ can be obtained given height bounds on more than $d+1$ values of F: this is the content of proposition 4.1.1, which we recall here for convenience.

Proposition 4.4.3. *Let L be a number field, and let $[\![A, B]\!]$ be an interval in $\mathbb{Z}$. Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $d \geq 1$, let $F \in L[Y]$ be a polynomial of degree at most d, let $N \geq d+1$, and let $x_1, \dots, x_N$ be distinct elements of $[\![A, B]\!]$. Assume that $h(F(x_i)) \leq H$ for all $1 \leq i \leq N$. Then we have*

$$h(F) \leq \frac{N}{N-d}H + D\log(D) + d\log(2M) + \log(d+1).$$

We need a lemma.

Lemma 4.4.4. *Keep the notation from proposition 4.4.3, and let $v \in \mathcal{V}_L^0$ (resp. $v \in \mathcal{V}_L^\infty$). Then the number of elements $x \in [\![A, B]\!]$ satisfying the inequality*

$$|F(x)|_v < |D!\, F|_v \quad \left(\text{resp. } |F(x)|_v < \frac{|F|_v}{(2M)^d(d+1)}\right)$$

is at most d.

Proof of lemma 4.4.4. We argue by contradiction. Let $(x_i)_{1\leq i\leq d+1}$ be distinct elements of $[\![A, B]\!]$ satisfying the given inequality. If $v \in \mathcal{V}_L^0$, then the Lagrange interpolation formula (4.5) gives

$$|D!\, F|_v \leq \max_i |F(x_i)|_v < |D!\, F|_v$$

which is a contradiction. If $v \in \mathcal{V}_L^\infty$, then the contradiction is

$$|F|_v \leq (2M)^d \sum_{i=1}^{d+1} |F(x_i)|_v < |F|_v$$

by lemma 4.4.1. □

Proof of proposition 4.4.3. Let v be a place of L. If $v \in \mathcal{V}_L^0$, then by lemma 4.4.4, we have $|F(x_i)|_v \geq |D!\, F|_v$ for at least $N - d$ values of i. Therefore

$$\prod_{i=1}^{N} \max\{1, |F(x_i)|_v\} \geq |D!\, F|_v^{N-d},$$

hence

$$\log \max\{1, |F|_v\} \leq \log \left|\frac{1}{D!}\right|_v + \frac{1}{N-d} \sum_{i=1}^{N} \log \max\{1, |F(x_i)|_v\}.$$

If $v \in \mathcal{V}_L^\infty$, then by lemma 4.4.4, at least $N - d$ of the values $F(x_i)$ satisfy the inequality $|F(x_i)|_v \geq |F|_v / (2M)^d(d+1)$, hence

$$\log \max\{1, |F|_v\} \leq d \log(2M) + \log(d+1) + \frac{1}{N-d} \sum_{i=1}^{N} \log \max\{1, |F(x_i)|_v\}.$$

Since $h(1/D!) \leq D \log D$, summing over $\mathcal{V}_L$ yields the result. □

4.5 Heights of fractions from their values

In this section, we study the more difficult question of bounding the height of rational fractions from above in terms of heights of their values. Let $F \in L(Y)$ be a fraction of degree at most $d \geq 1$, and write $F = P/Q$ where $P, Q \in \mathbb{Z}_L[Y]$ are coprime in $L[Y]$. Let S be a collection of evaluation points (in $\mathbb{Z}$, for instance) and assume that $h(F(x)) \leq H$ for each $x \in S$. If we can prove that both $P(x)$ and $Q(x)$ are "small" for each $x \in S$, then we can apply proposition 4.1.1 to obtain an upper bound on $h(F)$. However, it can very well happen that $P(x)$ and $Q(x)$ have a large common factor even if $F(x)$ has small height.

In order to prove proposition 4.1.2, we decompose $P(x)$ and $Q(x)$ in ideals. Define ideals $\mathfrak{s}_x$, $\mathfrak{n}_x$ and $\mathfrak{d}_x$ of $\mathbb{Z}_L$ as follows:

$$\mathfrak{s}_x = \gcd(P(x), Q(x)), \quad (P(x)) = \mathfrak{n}_x \mathfrak{s}_x, \quad (Q(x)) = \mathfrak{d}_x \mathfrak{s}_x.$$

Then $(F(x)) = \mathfrak{n}_x \mathfrak{d}_x^{-1}$. The ideal $\mathfrak{s}_x$ encodes the simplifications that occur when evaluating P/Q at x. The main point of the proof is to show that the ideal $\mathfrak{s}_x$ is "small" for at least some values of $x \in S$. The height of an ideal in $\mathbb{Z}_L$ is not well defined, but its norm is. Therefore we start by studying the relation between heights and norms of elements in $\mathbb{Z}_L$.

4.5.1 Heights and norms of integers

We denote the norm of elements and ideals in L by $N_{L/\mathbb{Q}}$, and we make the following definition.

Definition 4.5.1. Let $x \in L \backslash \{0\}$. Then we define

$$\widetilde{h}(x) = \frac{1}{d_L} \log \left| N_{L/\mathbb{Q}}(x) \right| = \sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} \log |x|_v .$$

If $\mathfrak{a}$ is a fractional ideal in L, we also write

$$\widetilde{h}(\mathfrak{a}) = \frac{1}{d_L} \log N_{L/\mathbb{Q}}(\mathfrak{a}).$$

Note that $\widetilde{h}$ and h are equal on $\mathbb{Z}$. They are not equal on $\mathbb{Z}_L$ in general. For instance, $\widetilde{h}$ is invariant under multiplication by units; this is not the case for h as soon as L admits a fundamental unit, by the Northcott property.

Lemma 4.5.2. *Let $x \in \mathbb{Z}_L \backslash \{0\}$. Then we have*

$$0 \leq \widetilde{h}(x) \leq h(x).$$

Equality holds on the right if and only if $|x|_v \geq 1$ for every $v \in \mathcal{V}_L^\infty$.

Proof. We have $N_{L/\mathbb{Q}}(x) \in \mathbb{Z}\backslash\{0\}$, so $\left|N_{L/\mathbb{Q}}(x)\right| \geq 1$ and $\widetilde{h}(x) \geq 0$. The rest is obvious. □

Proposition 4.5.3. *There exists a constant C depending only on the number field L such that for every $x \in \mathbb{Z}_L\backslash\{0\}$, there exists a unit $\varepsilon \in \mathbb{Z}_L^\times$ such that*

$$h(\varepsilon x) \leq \max\{C, \widetilde{h}(x)\}.$$

We can take $C = 2d_L \max_{i\in I} h(\varepsilon_i)$, where $(\varepsilon_i)_{i\in I}$ is any basis of units in $\mathbb{Z}_L$.

Proof. Let $m = \#\mathcal{V}_L^\infty$. In $\mathbb{R}^m$, we define the hyperplane H_s for $s \in \mathbb{R}$ by

$$H_s = \{(t_1, \dots, t_m) \in \mathbb{R}^m \colon t_1 + \cdots + t_m = s\},$$

and we define the convex cone Δ_s by

$$\Delta_s = \big\{(t_1, \dots, t_m) \in \mathbb{R}^m \colon \forall i,\ t_i \geq -s\big\}.$$

The image of $\mathbb{Z}_L^\times$ by the logarithmic embedding

$$\mathrm{Log} = \Big(\frac{d_v}{d_L} \log|\cdot|_v\Big)_{v\in\mathcal{V}_L^\infty}$$

is a full rank lattice Λ in H_0. There exists a real number $s_{\min}$ such that for every $s \geq s_{\min}$, the convex set $H_0 \cap \Delta_s$ contains a fundamental cell V of Λ; we can take $s_{\min} = \max_{i\in I} h(\varepsilon_i)$ in the notation above. Translating in the direction $(1, \dots, 1)$, we also have:

1. For every $s \geq ms_{\min}$, the set $H_s \cap \Delta_0$ contains a translate of V.
2. For every $s \geq 0$, the set $H_s \cap \Delta_{s_{\min}}$ contains a translate of V.

Let $x \in \mathbb{Z}_L\backslash\{0\}$, and consider the point

$$\mathrm{Log}(x) = \Big(\frac{d_v}{d_L} \log|x|_v\Big)_{v\in\mathcal{V}_L^\infty} \in \mathbb{R}^m.$$

The sum of its coordinates is $\widetilde{h}(x)$. If $\widetilde{h}(x) \geq ms_{\min}$, then by (1) there exists a unit $\varepsilon \in \mathbb{Z}_L^\times$ such that $\mathrm{Log}(x) + \mathrm{Log}(\varepsilon)$ belongs to Δ_0. Then $|\varepsilon x|_v \geq 1$ for every $v \in \mathcal{V}_L^\infty$, so

$$h(\varepsilon x) = \widetilde{h}(\varepsilon x) = \widetilde{h}(x)$$

by lemma 4.5.2.

On the other hand, if $0 \leq \widetilde{h}(x) < ms_{\min}$, then by (2) we can still find a unit ε such that $\mathrm{Log}(x) + \mathrm{Log}(\varepsilon) \in \Delta_{s_{\min}}$, in other words

$$\frac{d_v}{d_L} \log|\varepsilon x|_v \geq -s_{\min}$$

for all $v \in \mathcal{V}_L^\infty$. Then

$$h(\varepsilon x) = \sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} \log \max\{1, |\varepsilon x|_v\} \leq \widetilde{h}(\varepsilon x) + \sum_{v \in \mathcal{V}_\infty} s_{\min} \leq 2ms_{\min}.$$

This proves the proposition with $C = 2ms_{\min} \leq 2d_L s_{\min}$. □

Corollary 4.5.4. *Let C be as in proposition 4.5.3. Then every principal ideal $\mathfrak{a}$ of $\mathbb{Z}_L$ admits a generator $a \in \mathbb{Z}_L$ such that*

$$h(a) \leq \max\{C, \widetilde{h}(\mathfrak{a})\}.$$

Proof. Apply proposition 4.5.3 with x an arbitrary generator of $\mathfrak{a}$. □

This allows us to bound the height of a common denominator for a given polynomial $P \in \mathbb{Z}_L[Y]$.

Proposition 4.5.5. *There exists a constant C depending only on L such that for every polynomial $P \in L[Y]$, there exists an element $a \in \mathbb{Z}_L$ such that $aP \in \mathbb{Z}_L[Y]$ and $\max\{h(a), h(aP)\} \leq h(P) + C$.*

Proof. Let $\mathfrak{C}$ be a set of ideals of $\mathbb{Z}_L$ that are representatives for the class group of L, and let $P \in L[Y]$, which we may assume to be nonzero. Let

$$\mathfrak{a} = \prod_{\mathfrak{p} \in \mathcal{P}_L} \mathfrak{p}^{\max\{0, -v_\mathfrak{p}(P)\}}$$

be the denominator ideal of P. Then

$$\widetilde{h}(\mathfrak{a}) = \sum_{\mathfrak{p} \in \mathcal{P}_L} \frac{d_\mathfrak{p}}{d_L} \log \max\{1, |P|_\mathfrak{p}\} \leq h(P).$$

Let $\mathfrak{c} \in \mathfrak{C}$ such that $\mathfrak{ca}$ is principal. By corollary 4.5.4, if C denotes the constant from proposition 4.5.3, we can find a generator a of $\mathfrak{ca}$ such that

$$h(a) \leq \max\{C, \widetilde{h}(\mathfrak{ca})\} \leq h(P) + C'$$

with

$$C' = \max\{C, \max_{\mathfrak{c} \in \mathfrak{C}} \widetilde{h}(\mathfrak{c})\}.$$

Then aP has integer coefficients, and we have

$$\begin{aligned} h(aP) &\leq \sum_{v \in \mathcal{V}_L^\infty} \frac{d_v}{d_L} \big(\log \max\{1, |P|_v\} + \log \max\{1, |a|_v\}\big) \\ &= h(P) + h(a) - \sum_{v \in \mathcal{V}_L^0} \frac{d_v}{d_L} \log \max\{1, |P|_v\} \\ &= h(P) + h(a) - \widetilde{h}(\mathfrak{a}) \leq h(P) + C'. \end{aligned}$$

□

4.5.2 A naive height bound for fractions

Let L be a number field, and let $F \in L(Y)\backslash\{0\}$ be a rational fraction of degree at most $d \geq 1$. Write $F = P/Q$ where P, Q are coprime polynomials in $L[Y]$, and let d_P and d_Q be the degrees of P and Q respectively. Let x_i for $1 \leq i \leq d_P + d_Q + 1$ be distinct elements in an interval $[\![A, B]\!] \subset \mathbb{Z}$ that are not poles of F.

Let us recall the interpolation algorithm to reconstruct F given the pairs $(x_i, F(x_i))$ [BCG+17, §7.1]. Let $S \in L[Y]$ be the polynomial of degree at most $d_P + d_Q$ interpolating the points $(x_i, F(x_i))$. Let a be a common denominator for the coefficients of S, so that $T = aS$ has coefficients in $\mathbb{Z}_L$. We compute the d_P-th subresultant of T and the polynomial

$$Z = \prod_{i=1}^{s+1} (Y - x_i) \in \mathbb{Z}[Y].$$

We obtain a Bézout relation of the form

$$UT + VZ = R$$

where $U, V, R \in \mathbb{Z}_L[Y]$, and moreover $\deg(U) \leq d_Q$ and $\deg(R) \leq d_P$. Then $F = R/aU$.

In order to obtain an upper bound on $h(F)$, we first bound $h(S)$ from above using proposition 4.4.2. Then, we use the following well-known fact about the sizes of subresultants in $\mathbb{Z}_L[Y]$.

Lemma 4.5.6. *Let $P, Q \in \mathbb{Z}_L[Y]\backslash\{0\}$ be polynomials of degrees d_P and d_Q respectively, and let $0 \leq k \leq \min\{d_P, d_Q\} - 1$. Let R be the k-th subresultant of P and Q, and let U and V be the associated Bézout coefficients. Write $s = d_P + d_Q$. Then we have*

$$h(R) \leq (d_Q - k)\, h(P) + (d_P - k)\, h(Q) + \frac{s - 2k}{2} \log(s - 2k),$$
$$h(U) \leq (d_Q - k - 1)\, h(P) + (d_P - k)\, h(Q) + \frac{s - 2k - 1}{2} \log(s - 2k - 1),$$
$$h(V) \leq (d_Q - k)\, h(P) + (d_P - k - 1)\, h(Q) + \frac{s - 2k - 1}{2} \log(s - 2k - 1).$$

For instance, lemma 4.5.6 bounds the sizes of the coefficients appearing in the subresultant version of the Euclidean algorithm in $\mathbb{Q}(Y)$.

Proof. Let $v \in \mathcal{V}_L^\infty$. Every coefficient r of R has an expression as the determinant of a square matrix of size $d_P + d_Q - 2k$; its entries in the

first $d_Q - k$ columns are coefficients of P, and its entries in the last $d_P - k$ columns are coefficients of Q. By Hadamard's lemma, we can bound $|r|_v$ by the product of L^2-norms of the columns in this determinant. Hence

$$|r|_v \leq \left(\sqrt{d_P + d_Q - 2k}\ |P|_v\right)^{d_Q - k} \left(\sqrt{d_P + d_Q - 2k}\ |Q|_v\right)^{d_P - k}.$$

Taking logarithms and summing over v, we obtain the desired height bound on R. Similarly, the coefficients of U (resp. V) are determinants of size $d_P + d_Q - 2k - 1$, with one column less coming from P (resp. Q). □

Proposition 4.5.7. *Let $[\![A, B]\!]$ be an interval in $\mathbb{Z}$, and let $d \geq 1$. Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L(Y)\backslash\{0\}$ be a rational fraction of degree d. Let d_P and d_Q be the degrees of the numerator and denominator of F respectively. Let x_i for $1 \leq i \leq d_P + d_Q + 1$ be distinct elements of $[\![A, B]\!]$ that are not poles of F, and assume that $h(F(x_i)) \leq H$ for every i. Then we can write $F = P/Q$ where $P, Q \in \mathbb{Z}_L[Y]$ satisfy* $\deg P = d_P$, $\deg Q = d_Q$, *and*

$$\begin{aligned}\max\{h(P),\, h(Q)\} \leq\ & (d+1)(2d+1)H + (d+1)D\log(D) \\ & + (4d^2 + 3d)\log(2M) + (2d+2)\log(2d+1) \\ & + (d+1)C,\end{aligned}$$

where C denotes the constant from proposition 4.5.5.

Proof. Let S, a, T, R, U, V be as in the interpolation algorithm above; to choose a, we use proposition 4.5.5, so that

$$\max\{h(a), h(T)\} \leq h(S) + C.$$

By proposition 4.4.2, we have

$$h(S) \leq (2d+1)H + D\log(D) + 2d\log(2M) + \log(2d+1). \tag{4.6}$$

The absolute values of the coefficients of Z are bounded above by $(2M)^{2d+1}$, hence

$$h(Z) \leq (2d+1)\log(2M).$$

By lemma 4.5.6, we have

$$\begin{aligned}h(R) &\leq (d+1)\,h(T) + d(2d+1)\log(2M) + \frac{2d+1}{2}\log(2d+1), \\ h(U) &\leq d\,h(T) + d(2d+1)\log(2M) + d\log(2d+1).\end{aligned}$$

Then $F = R/aU$. Therefore

$$\begin{aligned} h(F) &\leq \max\{h(R), h(a) + h(U)\} \\ &\leq (d+1)(h(S)+C) + d(2d+1)\log(2M) + \frac{2d+1}{2}\log(2d+1). \end{aligned}$$

We conclude using the upper bound (4.6) on $h(S)$. □

As mentioned in §4.1, the upper bound on $h(F)$ in proposition 4.5.7 is roughly $O(d^2H)$. This motivates a result like proposition 4.1.2, where the dependency on H is only linear.

4.5.3 Some preparatory lemmas

In this section, we state preparatory lemmas for the proof of proposition 4.1.2 in §4.5.4; the reader might wish to skip them until their use in the proof becomes apparent.

Lemma 4.5.8. *Let $[\![A, B]\!] \subset \mathbb{Z}$, let $D = B - A$, and let $\eta \geq 1$. Let S be a subset of $[\![A, B]\!]$ containing at least D/η elements, and let $1 \leq k \leq \frac{D}{2\eta}$ be an integer. Then there exists a subinterval of $[\![A, B]\!]$ of amplitude at most $\lceil 2\eta k \rceil$ containing at least $k+1$ elements of S.*

Recall that the *amplitude* of $[\![A, B]\!]$ is $B - A$.

Proof. Assume the contrary. We can partition $[\![A, B]\!]$ in at most $\left\lceil \frac{D}{2\eta k} \right\rceil$ intervals of amplitude at most $\lceil 2\eta k \rceil$, hence

$$\frac{D}{\eta} \leq \#S \leq k \left\lceil \frac{D}{2\eta k} \right\rceil < \frac{D}{2\eta} + k.$$

This is absurd because $k \leq \frac{D}{2\eta}$. □

Lemma 4.5.9. *Let $R \in \mathbb{Z}_L \backslash \{0\}$ be a non-unit. Then*

$$\sum_{\substack{\mathfrak{p} \in \mathcal{P}_L,\, \mathfrak{p} | R \\ \mathfrak{p} | p \in \mathcal{P}_\mathbb{Q}}} \frac{\log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} \leq d_L(2\log\log\left|N_{L/\mathbb{Q}}(R)\right| + 3.5). \tag{4.7}$$

The sum is over all prime ideals $\mathfrak{p}$ of L diving R, and $p \in \mathcal{P}_\mathbb{Q}$ denotes the prime lying under $\mathfrak{p}$.

Proof. First, we assume that $L = \mathbb{Q}$, so that $R \in \mathbb{Z}$ and $|R| \geq 2$. Let m be the number of prime factors of R, and let $(p_i)_{i\geq 1}$ be the sequence of primes in increasing order. It is enough to prove (4.7) for $R' = \prod_{i=1}^{m} p_i$, which has both a greater left hand side, since $\log(p)/(p-1)$ is a decreasing function of p, and a smaller right hand side, since $R' \leq R$. We can assume that $m \geq 2$. Then

$$\sum_{i=1}^{m} \frac{\log(p_i)}{p_i - 1} = \sum_{i=1}^{m} \frac{\log(p_i)}{p_i} + \sum_{i=1}^{m} \frac{\log(p_i)}{p_i(p_i-1)} \leq \log(p_m) + 3$$

by Mertens's first theorem [Mer74], and because the sum of the second series is less than 0.76. By [Ros41], we have $p_m < m \log m + m \log\log m$ if $m \geq 6$; so the rough bound $p_m \leq m^2$ holds. Since $m \leq \log(R')/\log(2)$, the result in the case $L = \mathbb{Q}$ follows.

In the general case, if $\mathfrak{p}|R$ lies above $p \in \mathcal{P}_{\mathbb{Q}}$, then p divides $N_{L/\mathbb{Q}}(R)$, and $\left|N_{L/\mathbb{Q}}(R)\right| \geq 2$. We apply lemma 4.5.9 to the integer $N_{L/\mathbb{Q}}(R) \in \mathbb{Z}$: hence

$$\begin{aligned}\sum_{\mathfrak{p}|R} \frac{\log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} &\leq \sum_{p|N_{L/\mathbb{Q}}(R)} \frac{\sum_{\mathfrak{p}|p} \log(N_{L/\mathbb{Q}}(\mathfrak{p}))}{p-1} \\ &= d_L \sum_{p|N_{L/\mathbb{Q}}(R)} \frac{\log(p)}{p-1} \leq d_L(2\log\log\left|N_{L/\mathbb{Q}}(R)\right| + 3.5). \square\end{aligned}$$

Lemma 4.5.10. *Let $\mathfrak{p} \in \mathcal{P}_L$ be a prime ideal lying over $p \in \mathcal{P}_{\mathbb{Q}}$, and let $L_{\mathfrak{p}}$ be the $\mathfrak{p}$-adic completion of L. Let $d \geq 0$, let $Q \in L_{\mathfrak{p}}[Y]\backslash\{0\}$ be a polynomial of degree d, and assume that $v_{\mathfrak{p}}(Q) = 0$. Let $x_1, \ldots, x_n$ be distinct elements of $[\![A, B]\!]$, and write $D = B - A$; assume that $D \geq 1$. Let $\beta \in \mathbb{N}$. Then*

$$\sum_{i=1}^{n} \min\{\beta, v_{\mathfrak{p}}(Q(y_i))\} \leq d\left(\beta + d_L \frac{\log(D)}{\log N_{L/\mathbb{Q}}(\mathfrak{p})} + \frac{D}{p-1}\right). \tag{4.8}$$

Proof. We can assume that $d \geq 1$. Let λ be the leading coefficient of Q, and let $\alpha_1, \ldots, \alpha_d$ be the roots of Q in an algebraic closure of $L_{\mathfrak{p}}$, where we extend $|\cdot|_{\mathfrak{p}}$ and $v_{\mathfrak{p}}$. Up to reindexation, we may assume that $|\alpha_j|_{\mathfrak{p}} \leq 1$ for $1 \leq j \leq t$, and $|\alpha_j|_{\mathfrak{p}} > 1$ for $t+1 \leq j \leq d$. For every i, we have

$$|Q(x_i)|_{\mathfrak{p}} = |\lambda|_{\mathfrak{p}} \prod_{i=1}^{d} |x_i - \alpha_j|_{\mathfrak{p}} = \left(|\lambda|_{\mathfrak{p}} \prod_{j=t+1}^{d} |\alpha_j|_{\mathfrak{p}}\right) \prod_{j=1}^{t} |x_i - \alpha_j|_{\mathfrak{p}}.$$

We must have

$$|\lambda|_{\mathfrak{p}} \prod_{j=t+1}^{d} |\alpha_j|_{\mathfrak{p}} \geq 1,$$

for otherwise all the coefficients of Q would belong to $\mathfrak{p}$. Therefore

$$v_{\mathfrak{p}}(Q(x_i)) \leq \sum_{j=1}^{t} v_{\mathfrak{p}}(x_i - \alpha_j).$$

Let $k \in \mathbb{N}$ such that $p^k \leq D < p^{k+1}$. Since the x_i are all distinct modulo p^{k+1}, there are at most d values of i such that $v_{\mathfrak{p}}(x_i - \alpha_j) > k$ for some j. For these indices i, we bound $\min\{\beta, v_{\mathfrak{p}}(Q(x_i))\}$ by β. This accounts for the term $d\beta$ in inequality (4.8).

For all other values of i (say $i \in I$), we have $v_{\mathfrak{p}}(x_i - \alpha_j) \leq k$. Thus, for every $1 \leq j \leq t$, we have

$$v_{\mathfrak{p}}(x_i - \alpha_j) = \int_0^k \mathbf{1}_{u \leq v_{\mathfrak{p}}(x_i - \alpha_j)} du.$$

Any two x_i that fall in the same disk $\{u \leq v_{\mathfrak{p}}(x - \alpha_j)\}$ coincide modulo $p^{\lceil u \rceil}$. Therefore, for a given α_j, and a given $u \in]l, l+1]$, there exist at most $\lceil D/p^{l+1} \rceil$ values of i such that x_i belongs to this disk. Therefore

$$\begin{aligned}
\sum_{i \in I} v_{\mathfrak{p}}(Q(x_i)) &\leq \sum_{i \in I} \sum_{j=1}^{t} v_{\mathfrak{p}}(x_i - \alpha_j) \\
&= \sum_{i \in I} \sum_{j=1}^{t} \sum_{l=0}^{k-1} \int_l^{l+1} \mathbf{1}_{u \leq v_{\mathfrak{p}}(x_i - \alpha_j)} du \\
&= \sum_{j=1}^{t} \sum_{l=0}^{k-1} \int_l^{l+1} \left(\sum_{i \in I} \mathbf{1}_{u \leq v_{\mathfrak{p}}(x_i - \alpha_j)} \right) du \\
&\leq t \sum_{l=0}^{k-1} \left\lceil \frac{D}{p^{l+1}} \right\rceil \leq t \sum_{l=0}^{k-1} \left(\frac{D}{p^{l+1}} + 1 \right) \leq tk + \frac{tD}{p-1}.
\end{aligned}$$

We have $t \leq d$, and

$$k \leq \frac{\log(D)}{\log(p)} \leq d_L \frac{\log(D)}{\log N_{L/\mathbb{Q}}(\mathfrak{p})}.$$

This accounts for the two remaining terms in inequality (4.8). □

4.5.4 Main result

Finally, we prove propositions 4.1.2 and 4.1.3. We recall the statements for convenience.

Proposition 4.5.11. *Let L be a number field, and let $[\![A, B]\!]$ be an interval in $\mathbb{Z}$. Write $D = B - A$ and $M = \max\{|A|, |B|\}$. Let $F \in L(Y)$ be a rational fraction of degree at most $d \geq 1$. Let S be a subset of $[\![A, B]\!]$ containing no poles of F, let $\eta \geq 1$, and let $H \geq \max\{4, \log(2M)\}$. Assume further that*

1. *$h(F(x)) \leq H$ for every $x \in S$.*
2. *S contains at least D/η elements.*
3. *$D \geq \max\{\eta d^3 H, 4\eta d d_L\}$.*

Then we have

$$h(F) \leq H + C_L \eta d \log(\eta d H) + d \log(2M) + \log(d+1),$$

where C_L is a constant depending only on L. We can take $C_{\mathbb{Q}} = 960$.

Proof. We can assume that $F \neq 0$. We have $D \geq 4\eta d$, so by lemma 4.5.8 with $k = 2d$, we can find a subinterval $[\![A', B']\!]$ of $[\![A, B]\!]$ with amplitude at most $\lceil 4\eta d \rceil$ containing $2d+1$ elements of S, denoted by $x_1, \ldots, x_{2d+1}$. We use these x_i as evaluation points in proposition 4.5.7: we can write $F = P/Q$ where $P, Q \in \mathbb{Z}_L[X]$ have the correct degrees and

$$\begin{aligned}\max\{h(P), h(Q)\} &\leq (d+1)(2d+1)H + 2d\lceil 4\eta d \rceil \log(\lceil 4\eta d \rceil) \\ &\quad + (4d^2 + 3d)\log(2M) + (2d+2)\log(2d+1) + (d+1)C_1 \\ &\leq (27 + C_1)\eta d^2 H\end{aligned}$$

where C_1 is the constant from proposition 4.5.3, which depends only on L. To simplify the right hand side, we use the inequalities $1 \leq d$, $1 \leq \eta$, $\lceil 4\eta d \rceil \leq D \leq 2M$, $\lceil 4\eta d \rceil \leq 5\eta d$, and $\log(2M) \leq H$.

Let $x \in S$, and define the ideals $\mathfrak{s}_x$, $\mathfrak{n}_x$ and $\mathfrak{d}_x$ of $\mathbb{Z}_L$ as above:

$$\mathfrak{s}_x = \gcd(P(x), Q(x)), \quad (P(x)) = \mathfrak{n}_x \mathfrak{s}_x, \quad (Q(x)) = \mathfrak{d}_x \mathfrak{s}_x.$$

Let $\mathfrak{r}$ be the greatest common divisor of all the coefficients of P and Q.

Claim 4.5.12. *There exists a constant C_L depending only on L such that the following property holds: there exist at least $2dd_L + 1$ elements x of S such that*

$$\widetilde{h}(\mathfrak{s}_x) \leq \widetilde{h}(\mathfrak{r}) + C_L \eta d \log(\eta d H).$$

Let us explain how to finish the proof of proposition 4.5.11 if claim 4.5.12 holds. By lemma 4.4.4, we can find an $x \in S$ among these $2dd_L + 1$ values such that for every $v \in \mathcal{V}_L^\infty$, we have

$$|P(x)|_v \geq \frac{|P|_v}{(2M)^d(d+1)} \quad \text{and} \quad |Q(x)|_v \geq \frac{|Q|_v}{(2M)^d(d+1)}.$$

Then, we can calculate $h(F)$ as

$$\begin{aligned}
&\sum_{v\in\mathcal{V}_L^\infty}\frac{d_v}{d_L}\log\max\{|P|_v,|Q|_v\}-\widetilde{h}(\mathfrak{r})\\
&\leq\sum_{v\in\mathcal{V}_L^\infty}\frac{d_v}{d_L}\log\max\{|P(x)|_v,|Q(x)|_v\}-\widetilde{h}(\mathfrak{r})+d\log(2M)+\log(d+1)\\
&\leq\sum_{v\in\mathcal{V}_L}\frac{d_v}{d_L}\log\max\{|P(x)|_v,|Q(x)|_v\}+\widetilde{h}(\mathfrak{s}_x)-\widetilde{h}(\mathfrak{r})+d\log(2M)+\log(d+1)\\
&\leq H+C_L\eta d\log(\eta dH)+d\log(2M)+\log(d+1),
\end{aligned}$$

as claimed.

In order to prove claim 4.5.12, a crucial remark is that $\mathfrak{s}_x$ divides the resultant R of P and Q. By lemma 4.5.6, we have

$$h(R)\leq d\,h(P)+d\,h(Q)+d\log(2d)\leq(55+2C_1)\eta d^3H.$$

Let $\mathfrak{p}\in\mathcal{P}_L$ be a prime factor of R with valuation $\beta_\mathfrak{p}$, and let I be a subset of S with n elements. We claim that

$$\sum_{x\in I}v_\mathfrak{p}(\mathfrak{s}_x)\leq n\,v_\mathfrak{p}(\mathfrak{r})+d\left(\beta_\mathfrak{p}+d_L\frac{\log(D)}{\log N_{L/\mathbb{Q}}(\mathfrak{p})}+\frac{D}{p-1}\right). \tag{4.9}$$

To prove (4.9), we can work in the $\mathfrak{p}$-adic completion $L_\mathfrak{p}$ of L. Let π be a uniformizer of $L_\mathfrak{p}$, and let $r=\min\{v_\mathfrak{p}(P),v_\mathfrak{p}(Q)\}$ be the $\mathfrak{p}$-adic valuation of $\mathfrak{r}$. Write $P_1=P/\pi^r$, $Q_1=Q/\pi^r$. Then one of P_1 and Q_1 is not divisible by π; for instance, assume that π does not divide Q_1. For every $x\in S$, we have

$$v_\mathfrak{p}(\mathfrak{s}_x)\leq\min\big\{\beta_\mathfrak{p},v_\mathfrak{p}(Q(x))\big\}\leq v_\mathfrak{p}(\mathfrak{r})+\min\big\{\beta_\mathfrak{p},v_\mathfrak{p}(Q_1(x))\big\}.$$

Thus inequality (4.9) follows from lemma 4.5.10.

Summing over the prime factors $\mathfrak{p}$ of R, we obtain an upper bound on the product of the norms of the ideals $\mathfrak{s}_x$, for $x\in I$. We can assume that R is not a unit, otherwise claim 4.5.12 holds trivially. Then

$$\begin{aligned}
\left|\prod_{x\in I}N_{L/\mathbb{Q}}(\mathfrak{s}_x)\right|&\leq N_{L/\mathbb{Q}}(\mathfrak{r})^n\left|N_{L/\mathbb{Q}}(R)\right|^d\\
&\quad\cdot\exp\Bigg(\sum_{\substack{\mathfrak{p}\in\mathcal{P}_L,\,\mathfrak{p}|R\\ \mathfrak{p}|p\in\mathcal{P}_\mathbb{Q}}}\left(dd_L\log(D)+dD\frac{\log N_{L/\mathbb{Q}}(\mathfrak{p})}{p-1}\right)\Bigg)\\
&\leq N_{L/\mathbb{Q}}(\mathfrak{r})^n\left|N_{L/\mathbb{Q}}(R)\right|^d\\
&\quad\cdot\exp\big(dd_L\log(D)\log\left|N_{L/\mathbb{Q}}(R)\right|/\log(2)\\
&\qquad+dd_LD(\log\log\left|N_{L/\mathbb{Q}}(R)\right|+3.5)\big).
\end{aligned}$$

Indeed, R has at most $\log\left|N_{L/\mathbb{Q}}(R)\right| / \log(2)$ prime factors, and we can apply lemma 4.5.9. Since $\widetilde{h}(R) \leq (55 + 2C_1)\eta d^3 H$, we obtain

$$\begin{aligned}\sum_{x\in I} \widetilde{h}(\mathfrak{s}_x) &\leq n\,\widetilde{h}(\mathfrak{r}) + d\,\widetilde{h}(R) + dd_L \frac{\log(D)}{\log(2)}\,\widetilde{h}(R) \\ &\quad + dD(\log\log\left|N_{L/\mathbb{Q}}(R)\right| + 3.5) \\ &\leq n\,\widetilde{h}(\mathfrak{r}) + C_2(\eta d^4 H\log(D) + dD\log(\eta dH))\end{aligned}$$

with

$$C_2 = \max\left\{\frac{3d_L(55+2C_1)}{2\log(2)}, 6.5 + \log(d_L) + \log(55+2C_1)\right\}.$$

Here we use that $\log(\eta dH) \geq 1$, and $\log(D) \geq 2\log 2$.

Now we use the assumption that D and S are sufficiently large. Since $D \geq \eta d^3 H \geq 4 > \exp(1)$, and since the function $t/\log(t)$ is increasing for $t > \exp(1)$, we have

$$\frac{D}{\log(D)} \geq \frac{\eta d^3 H}{3\log(\eta dH)}.$$

Moreover,

$$\#S - 2dd_L \geq \frac{D}{\eta} - \frac{D}{2\eta} = \frac{D}{2\eta}.$$

Therefore,

$$\begin{aligned}\sum_{x\in I} \widetilde{h}(\mathfrak{s}_x) &\leq n\,\widetilde{h}(\mathfrak{r}) + 4C_2 dD\log(\eta dH) \\ &\leq n\,\widetilde{h}(\mathfrak{r}) + 8C_2\eta d\log(\eta dH)(\#S - 2dd_L).\end{aligned}$$

This shows that in every subset of $\#S - 2dd_L$ elements of S, at least one must satisfy $\widetilde{h}(\mathfrak{s}_x) \leq \widetilde{h}(\mathfrak{r}) + 8C_2\eta d\log(\eta dH)$. Therefore claim 4.5.12 holds with $C = 8C_2$.

If $L = \mathbb{Q}$, we have $C_1 = 0$, hence we can take $C_2 = 120$. □

We conclude this chapter with the proof of proposition 4.1.3.

Proposition 4.5.13. *Let $c \geq 1$, let $d \geq 1$, and let $F \in \mathbb{Q}(Y)$ be a rational fraction of degree at most d. Let $V \subset \mathbb{Z}$ be a finite set such that F has no poles in $\mathbb{Z}\backslash V$. Assume that for every $x \in \mathbb{Z}\backslash V$, we have*

$$h(F(x)) \leq c\max\{1, d\log d + d\,h(x)\}.$$

Then there exists a constant $C = C(c, \#V)$ such that

$$h(F) \leq Cd\log(4d).$$

We can take $C = (4c + 1923)\big(12 + \log\max\{1, \#V\} + 2\log(c)\big)$.

Proof. We want to use proposition 4.1.2 on an interval of the form $[\![0, D]\!]$ for some integer $D \geq 4d$, with $\eta = 2$ and $S = [\![0, D]\!] \setminus V$. The set S will contain at least D/η elements as soon as $D \geq 2\#V$.

For every $x \in S$, we have $\widetilde{h}(x) \leq \log(D)$, hence

$$h(F(x)) \leq c \max\{1, d \log d + d \log D\}.$$

Hence, if we define $H(D)$ as

$$H(D) = \max\{4, \log(2D), c(d \log d + d \log D)\},$$

then we can apply proposition 4.1.2 with $H = H(D)$ as soon as

$$D \geq 2d^3 H(D).$$

We check that we can take

$$D = \max\{2\#V, \lceil 4cd^4 \log(4cd^4) \rceil\}.$$

Then, proposition 4.1.2 yields

$$h(F) \leq H(D) + 1920d \log(2dH(D)) + d \log(2D) + \log(d + 1).$$

We have $H(D) \leq 4cd \log(dD)$ and $2dH(D) \leq D$, hence

$$\begin{aligned} h(F) &\leq 4cd \log(dD) + 1920d \log(D) + d \log(2D) + \log(d + 1) \\ &\leq (4c + 1923)d \log(dD) \\ &\leq (4c + 1923)d(\log(2d \max\{1, \#V\}) + \log(5cd^5 \log(4cd^4))) \end{aligned}$$

To simplify this expression further, we write

$$\log(5cd^5 \log(4cd^4)) \leq \log(20c^2 d^9) \leq 3 + 2 \log(c) + 9 \log(d),$$

hence, after other simplifications,

$$h(F) \leq Cd \log(4d)$$

with

$$C = (4c + 1923)(12 + \log \max\{1, \#V\} + 2 \log(c)),$$

as claimed. □

Chapter 5

Degree and height bounds for modular equations

This chapter, which corresponds to the preprint [Kie20a], presents upper bounds on the size of modular equations on PEL Shimura varieties in terms of their level. Let us recall the notation introduced in chapter 2.

We fix a PEL datum $(B, *, V, \psi, G, X_+)$. Let K be a compact open subgroup of $G(\mathbb{A}_f)$, and let Σ be a finite group of automorphisms of V as in §2.2.4. Let n be the complex dimension of X_+; we assume that $n \geq 1$. Let $\mathcal{S}$ and $\mathcal{T}$ be connected components of $\mathrm{Sh}_K(G, X_+)(\mathbb{C})$, and let L be their field of definition. We fix a choice of invariants $j_1, \ldots, j_{n+1}$ as in §2.4. These modular functions satisfy eq. (2.17):

$$E(j_1, \ldots, j_{n+1}) = 0, \tag{5.1}$$

where

$$E = \sum_{k=0}^{e} E_k(J_1, \ldots, J_n)\, J_{n+1}^k \in L[J_1, \ldots, J_{n+1}]$$

and E is irreducible. Our estimates on the size of modular equations involve constants depending only on this data.

Let $\delta \in G(\mathbb{A}_f)$ be an element of G defining an absolutely irreducible Hecke correspondence H_δ that intersects $\mathcal{S} \times \mathcal{T}$ nontrivially. In §2.4.3, we defined the modular equations $\Psi_{\delta,m}$ for $1 \leq m \leq n+1$; they are elements of the ring $L(J_1, \ldots, J_n)[J_{n+1}, Y_1, \ldots, Y_m]$ of degree at most $e-1$ in J_{n+1}, defining the Hecke correspondence H_δ as a subvariety of $\mathcal{S} \times \mathcal{T}$. The degree of H_δ is denoted by $d(\delta)$, and the degree of the isogenies described by H_δ in the modular interpretation (recall corollary 2.4.2) is denoted by $l(\delta)$. Our main theorem, stated as theorem 1.4.3 in chapter 1, is the following.

Theorem 5.0.1. *Using the notation above, there exist constants C_1 and C_2 (independent of δ) with the following property. Let $\delta \in G(\mathbb{A}_f)$ be as above, and let $F \in L(J_1, \dots, J_n)$ be a rational fraction obtained as a coefficient of one of the modular equations $\Psi_{\delta,m}$ for $1 \leq m \leq n+1$. Then the total degree of F is bounded above by $C_1\, d(\delta)$, and the height of F is bounded above by $C_2\, d(\delta) \log(l(\delta))$.*

We refer to §4.2 for the definition of $h(F)$. This result allows us to bound the complexity of algorithms manipulating modular equations, and was previously known only in the case of elliptic modular polynomials [Coh84]. In the case of Siegel modular equations for abelian surfaces, and Hilbert modular equations in Gundlach invariants for $F = \mathbb{Q}(\sqrt{5})$ (recall §2.4.4), the constants C_1 and C_2 are explicit: see propositions 5.1.15 and 5.1.17, and theorem 5.2.13. In particular, the degree bounds we obtain match experimental data exactly; on the other hand the constant C_2 is far too pessimistic for algorithmic use.

The strategy to prove the degree bounds is to exhibit a particular modular form that behaves as the denominator of $\Psi_{\delta,m}$, and to control its weight. Then, we show that rewriting quotients of modular forms in terms of the chosen coordinates $j_1, \dots, j_{n+1}$ translates bounded weights into bounded degrees. Giving an explicit expression of the denominator will also be useful in chapter 6. The proof of height bounds is inspired by previous works on elliptic modular polynomials [Paz19]. We prove height bounds on *evaluations* of modular equations at certain points using well-known results on the Faltings height of isogenous abelian varieties [Fal83]. Then we use the tight relation between the height of a rational fraction over a number field and the height of its evaluations at sufficiently many points provided by proposition 4.1.2, the main result of chapter 4.

5.1 Degree bounds for modular equations

5.1.1 Denominators of modular equations

We keep the notation used in §2.4.3: in particular

$$K' = K \cap \delta K \delta^{-1}, \quad K_0 = K \rtimes \Sigma,$$

and K'' is a normal subgroup of finite index in K, contained in K' and stabilized by Σ. The natural action of K_0 on modular functions of level K'' extends to modular forms.

For each $1 \leq i \leq n+1$, fix a nonzero modular form χ_i invariant under Σ and defined over L such that $\chi_i j_i$ is again a modular form (i.e. has no poles);

we say that χ_i is a *denominator* of j_i. This is possible by proposition 2.2.13. For each i, the function

$$\chi_{i,\delta} \colon [x, g] \mapsto \chi_i([x, g\delta])$$

is a modular form of weight $\mathrm{wt}(\chi_i)$ on the preimage of $\mathcal{S}$ in $\mathrm{Sh}_{K'}(\mathbb{C})$. We define the functions $g_{\delta,m}$ for $1 \leq m \leq n+1$ as

$$g_{\delta,m} = \prod_{i=1}^{m} \prod_{\gamma \in K_0/K'} \gamma \cdot \chi_{i,\delta}.$$

Lemma 5.1.1. *For every $1 \leq m \leq n+1$, the function $g_{\delta,m}$ is a nonzero symmetric modular form on $\mathcal{S}$, and*

$$\mathrm{wt}(g_{\delta,m}) = (\#\Sigma)\, d(\delta) \sum_{i=1}^{m} \mathrm{wt}(\chi_i).$$

Proof. We have $\#(K_0/K') = (\#\Sigma)\, d(\delta)$. Therefore the function $g_{\delta,m}$ is a modular form of level K'' and weight $\sum_{i=1}^{m} \#(K_0/K')\ \mathrm{wt}(\chi_i)$ by construction. Each modular form $\gamma \cdot \chi_{i,\delta}$ is nonzero on every connected component of $\mathrm{Sh}_{K''}(\mathbb{C})$ above $\mathcal{S}$, so $g_{\delta,m}$ is nonzero as well.

Acting by an element of K_0 permutes the factors in the product defining $g_{\delta,m}$, so $g_{\delta,m}$ is in fact a symmetric modular form on $\mathcal{S}$. □

Proposition 5.1.2. *For every $1 \leq m \leq n+1$, the coefficients of the multivariate polynomial $g_{\delta,m}\Psi_{\delta,m}$ are symmetric modular forms on $\mathcal{S}$.*

Proof. By definition 2.4.6, the polynomial $\Psi_{\delta,m}$ is a sum of terms of the form

$$\left(\prod_{i=1}^{m-1} \prod_{\gamma_i} \left(Y_i - \gamma_i \cdot j_{i,\delta} \right) \right) \prod_{\gamma_m \in K_{m-1}/K_m} \left(Y_m - \gamma\gamma_m \cdot j_{m,\delta} \right)$$

where $\gamma \in K_0$ is fixed, and the middle product is over all $\gamma_i \in K_0/K_i$ such that $\gamma_i = \gamma$ modulo K_{i-1}, but $\gamma_i \neq \gamma$ modulo K_i. In this expression, all the cosets γ_i and $\gamma\gamma_m$ are simultaneously disjoint as subsets of K_0/K'. Each denominator is accounted for by some factor in the product defining $g_{\delta,m}$, so the coefficients of $g_{\delta,m}\Psi_{\delta,m}$ are modular forms. □

When the invariants $j_1, \ldots, j_{n+1}$ have similar denominators, it is possible to make a better choice for $g_{\delta,m}$.

Proposition 5.1.3. *Assume that there exists a modular form χ on $\mathcal{S}$ such that $\chi_i = \chi^{\alpha_i}$ for some integer $\alpha_i \geq 0$, for every $1 \leq i \leq n+1$. Let $1 \leq m \leq n+1$, and define*

$$g_{\delta,m} = \Big(\prod_{\gamma \in K_0} \gamma \cdot \chi_\delta \Big)^{\alpha}, \qquad \text{where } \alpha = \max_{1 \leq i \leq m} \alpha_i.$$

Then $g_{\delta,m}$ is a nonzero symmetric modular form on $\mathcal{S}$, and

$$\mathrm{wt}(g_{\delta,m}) = (\#\Sigma)\, d(\delta)\, \alpha\, \mathrm{wt}(\chi).$$

Moreover, the coefficients of $g_{\delta,m}\Psi_{\delta,m}$ are symmetric modular forms on $\mathcal{S}$.

The proof is similar to that of proposition 5.1.2, and omitted.

5.1.2 Writing modular functions in terms of invariants

Let f/g be a quotient of symmetric modular forms of weight w on $\mathcal{S}$. We show that when we rewrite such a quotient in terms of the invariants $j_1, \ldots, j_{n+1}$, the degree of the rational fractions we obtain is bounded above linearly in w, with a proportionality constant depending on our choice of invariants. In order to make this dependency explicit, we define the *symmetric geometric complexity* of our invariants as follows.

Definition 5.1.4. Let f_k for $1 \leq k \leq r$ be nonzero generators over L for the graded ring of symmetric modular forms on $\mathcal{S}$, with respective weights w_k. For each $1 \leq k \leq r-1$, let $\beta_k \geq 1$ be the minimal integer such that

$$\beta_k w_k \in \mathbb{Z} w_{k+1} + \cdots + \mathbb{Z} w_r.$$

We can find nonzero modular forms $\lambda_k, \xi_k \in L[f_{k+1}, \ldots, f_r]$ such that $\mathrm{wt}(\lambda_k) - \mathrm{wt}(\xi_k) = \beta_k w_k$. For every $1 \leq k \leq r-1$, the function $\xi_k f_k^{\beta_k}/\lambda_k$ is a quotient of two symmetric modular forms of the same weight on $\mathcal{S}$; hence there exist polynomials $P_k, Q_k \in L[J_1, \ldots, J_{n+1}]$ such that

$$\frac{\xi_k f_k^{\beta_k}}{\lambda_k} = \frac{P_k(j_1, \ldots, j_{n+1})}{Q_k(j_1, \ldots, j_{n+1})}.$$

We define the *symmetric geometric complexity* of $j_1, \ldots, j_{n+1}$ relative to the choice of $f_k, \lambda_k, \psi_k, P_k, Q_k$ to be the positive rational number given by, either

1.
$$\left(1+\max_{1\le k\le r-1}\frac{\mathrm{wt}(\xi_k)}{\beta_k w_k}\right)\max_{1\le k\le r-1}\frac{\deg(P_k)}{\beta_k w_k+\mathrm{wt}(\xi_k)},$$
if the following conditions are satisfied: for every $1\le k\le r-1$, the modular forms λ_k and ξ_k are powers of f_r and f_{r-1} respectively (in particular $\xi_{r-1}=1$), and $Q_k=1$; or

2.
$$\sum_{k=1}^{r-1}\left(\frac{1}{\beta_k w_k}\max\{\deg(P_k),\deg(Q_k)\}\prod_{l=1}^{k-1}\left(1+\frac{\mathrm{wt}(\xi_l)}{\beta_l w_l}\right)\right),$$
otherwise.

Note that formula 1, when it applies, yields a smaller result than formula 2.

We define the *symmetric geometric complexity* of the chosen coordinates $j_1,\dots,j_{n+1}$, denoted by $\mathrm{SGC}(j_1,\dots,j_{n+1})$, to be the infimum of this quantity over all possible choices of modular forms f_k,λ_k,ξ_k and polynomials P_k,Q_k.

Given definition 5.1.4, explicit upper bounds on the geometric complexity are easy to obtain if a generating set of modular forms is known. Note that the symmetric geometric complexity is invariant under permutations of the invariants $j_1,\dots,j_{n+1}$, in contrast with their *geometric complexity* to be defined in the next section, which takes into account the fact that j_{n+1} is considered differently in eq. (5.1).

Proposition 5.1.5. *Let f, g be symmetric modular forms on $\mathcal{S}$ of weight w such that $g\neq 0$. Then there exist polynomials $P,Q\in L[J_1,\dots,J_{n+1}]$ of degree at most $\mathrm{SGC}(j_1,\dots,j_{n+1})w$ such that*
$$\frac{f}{g}=\frac{P(j_1,\dots,j_{n+1})}{Q(j_1,\dots,j_{n+1})}.$$

Proof. We keep the notation used in definition 5.1.4, and choose generators f_k for $1\le k\le r$, modular forms λ_k,ξ_k for $1\le k\le r-1$, and polynomials $P_k,Q_k\in L[J_1,\dots,J_{n+1}]$ for $1\le k\le r-1$. Let C be the symmetric geometric complexity of $j_1,\dots,j_{n+1}$ relative to this choice.

Let f, g be as in the proposition. Then f and g can be expressed as a sum of monomial terms of the form
$$cf_1^{\alpha_1}\cdots f_r^{\alpha_r}\qquad\text{with } c\in L \text{ and } \sum_{k=1}^{r}\alpha_k w_k=w.$$

We give algorithms to rewrite the fraction $P/Q = f/g$ (currently a rational fraction in terms of the modular forms f_k) as a fraction in the invariants $j_1, \dots, j_{n+1}$, and bound the total degree of the output.

Case 1 of definition 5.1.4. We assume that λ_k and ξ_k are powers of f_r and f_{r-1} respectively for every $1 \leq k \leq r-1$. In this case, for each $k \leq r-2$, the integer β_k can be seen as the order of w_k in the group $\mathbb{Z}/(\mathbb{Z}w_{r-1}+\mathbb{Z}w_r)$. We can write

$$w = \sum_{k=1}^{r-2} s_k w_k \pmod{\mathbb{Z}w_{r-1} + \mathbb{Z}w_r}$$

for some integers $0 \leq s_k < \beta_k$, and this determines the integers s_k uniquely (if such a linear combination vanishes, considering the smallest nonzero s_k yields a contradiction). Then each monomial appearing in P and Q is divisible by $f_1^{s_1} \cdots f_{r-2}^{s_{r-2}}$. After simplifying by this common factor, we can assume that the common weight w of P and Q satisfies $w \in \mathbb{Z}w_{r-1} + \mathbb{Z}w_r$. Then, for each $1 \leq k \leq r-2$, the exponent of f_k in each monomial of P and Q is divisible by β_k. For convenience, write

$$a = \max_{1 \leq k \leq r-1} \frac{\mathrm{wt}(\xi_k)}{\beta_k w_k}.$$

In order to rewrite P/Q in terms of invariants, we proceed as follows.

1. Multiply P and Q by $f_{r-1}^{\lfloor aw/\,\mathrm{wt}(f_{r-1})\rfloor}$.
2. For each $1 \leq k \leq r-2$, replace each occurence of $f_k^{\beta_k}$ by $\lambda_k P_k/\xi_k$ in P and Q.
3. Let $0 \leq s_{r-1} < \beta_{r-1}$ be such that $w = s_{r-1}w_{r-1} \bmod w_r$, and divide P and Q by $f_{r-1}^{s_{r-1}}$.
4. Replace each occurence of $f_{r-1}^{\beta_{r-1}}$ by $\lambda_{r-1}P_{r-1}$ in P and Q.
5. Finally, divide P and Q by $f_r^{(w-s_{r-1}w_{r-1})/w_r}$.

This algorithm runs independently on each monomial of P and Q. Let $M = c\prod_{k=1}^{r} f_k^{\alpha_k}$, with $c \in L$, be such a monomial after step 1. Let us show that the exponent of f_{r-1} in M remains nonnegative after step 2. In this step, we introduce a denominator given by

$$\prod_{k=1}^{r-2} \xi_k^{\alpha_k/\beta_k} = \prod_{k=1}^{r-2} f_{r-1}^{\frac{\mathrm{wt}(\xi_k)\alpha_k}{\mathrm{wt}(f_{r-1})\beta_k}}.$$

We have

$$\sum_{k=1}^{r-2} \frac{\mathrm{wt}(\xi_k)\alpha_k}{\mathrm{wt}(f_{r-1})\beta_k} \leq a \sum_{k=1}^{r-2} \frac{\alpha_k w_k}{\mathrm{wt}(f_{r-1})} \leq \frac{aw}{\mathrm{wt}(f_{r-1})},$$

hence

$$\sum_{k=1}^{r-2} \frac{\mathrm{wt}(\xi_k)\alpha_k}{\mathrm{wt}(f_{r-1})\beta_k} \leq \left\lfloor \frac{aw}{\mathrm{wt}(f_{r-1})} \right\rfloor \leq \alpha_{r-1} \qquad \text{by step 1}$$

because the left hand side is an integer. Therefore, at the end of step 2, M belongs to the polynomial ring $L[J_1, \ldots, J_{n+1}][f_{r-1}, f_r]$. Therefore we have $M \in L[J_1, \ldots, J_{n+1}][f_{r-1}^{\beta_{r-1}}, f_r]$ after step 3, and finally $M \in L[J_1, \ldots, J_{n+1}]$ after step 5.

It remains to bound the total degree of M after step 5. To do this, we consider the total weight of M in $f_1, \ldots, f_{r-1}$. For each $1 \leq k \leq r-1$, the modular form λ_k is a power of f_r; hence replacing $f_k^{\beta_k}$ by $\lambda_k P_k/\xi_k$ in steps 2 or 4 reduces this weight by $\beta_k w_k + \mathrm{wt}(\xi_k)$, and increases the total degree of M in $J_1, \ldots, J_{n+1}$ by at most $\deg(P_k)$. At the beginning of step 2, the total weight of M in $f_1, \ldots, f_{r-1}$ is at most $(1+a)w$. Therefore the total degree of M in $J_1, \ldots, J_{n+1}$ at the end of the algorithm is bounded above by

$$(1+a)w \max_{1 \leq k \leq r-1} \frac{\deg(P_k)}{\beta_k w_k + \deg(\xi_k)} = Cw$$

which is the desired result.

Case 2 of definition 5.1.4. In the general case, we perform replacements and simplifications in a sequential way. We start by defining integers z_k, d_k for $0 \leq k \leq r-1$ and s_k, a_k for $1 \leq k \leq r-1$ by induction as follows:

- $z_0 = w$ and $d_0 = 0$;
- For each $1 \leq k \leq r$, the integer $0 \leq s_k < \beta_k$ is defined by the relation

$$z_{k-1} = s_k w_k \pmod{\mathbb{Z}w_{k+1} + \cdots + \mathbb{Z}w_r};$$

- $a_k = \left\lfloor \dfrac{z_{k-1}}{\beta_k w_k} \right\rfloor$ for each $1 \leq k \leq r-1$;
- $z_k = z_{k-1} - s_k w_k + a_k \,\mathrm{wt}(\xi_k)$ for each $1 \leq k \leq r-1$;
- $d_k = d_{k-1} + a_k \max\{\deg(P_k), \deg(Q_k)\}$ for each $1 \leq k \leq r-1$.

In order to rewrite P/Q in terms of invariants, we proceed as follows. For $k = 1$ up to $r-1$, do:

1. Divide P and Q by $f_k^{s_k}$;
2. Replace each occurence of $f_k^{\beta_k}$ by $\dfrac{\lambda_k P_k}{\xi_k Q_k}$ in P and Q;
3. Multiply P and Q by $\xi_k^{a_k} Q_k^{a_k}$.

Finally, simplify the remaining occurences of f_r. We prove the following statement $(\star)_k$ by induction for every $1 \leq k \leq r$:

At the beginning of the k-th loop, the polynomials P and Q are elements of weight z_{k-1} in the ring $L[J_1, \ldots, J_{n+1}][f_k, \ldots, f_r]$, with total degree at most d_{k-1} in $J_1, \ldots, J_{n+1}$, such that

$$\frac{f}{g} = \frac{P(j_1, \ldots, j_{n+1})}{Q(j_1, \ldots, j_{n+1})}.$$

The statement $(\star)_1$ is true by definition of z_0 and d_0; assume that $(\star)_k$ is true. Then we see, in order, that during the k-th loop:

- $z_{k-1} \in \sum_{i=k}^{r} \mathbb{Z} w_i$, hence s_k is well defined.
- In each monomial of P and Q, the exponent of f_k is of the form $a\beta_k + s_k$ for some integer $a \leq a_k$. Therefore step 1 is an exact division, and after step 2 there are no more occurences of f_k in P or Q.
- After step 3, P and Q are elements of $L[J_1, \ldots, J_{n+1}][f_{k+1}, \ldots, f_r]$ of weight
$$z_{k-1} - s_k w_k + a_k \operatorname{wt}(\xi_k) = z_k.$$

It remains to show that the degree of P, Q in $J_1, \ldots, J_{n+1}$ is bounded by d_k after step 3. This comes from the following observation: during the k-th loop, we only multiply the polynomials in $J_1, \ldots, J_{n+1}$ already present by $P_k^b Q_k^{a_k - b}$ for some $0 \leq b \leq a_k$. This proves $(\star)_k$ for all $1 \leq k \leq r$.

At the end of the algorithm, all the occurences of f_r disappear. Therefore we obtain polynomials P, Q with total degree at most d_{r-1} such that

$$\frac{f}{g} = \frac{P(j_1, \ldots, j_{n+1})}{Q(j_1, \ldots, j_{n+1})}.$$

By induction, we obtain

$$z_k \leq w \prod_{l=1}^{k} \left(1 + \frac{\operatorname{wt}(\xi_l)}{\beta_l w_l}\right)$$

and

$$d_{r-1} \leq \sum_{k=1}^{r-1} \left(\frac{w}{\beta_k w_k} \max\{\deg(P_k), \deg(Q_k)\} \prod_{l=1}^{k-1} \left(1 + \frac{\operatorname{wt}(\xi_l)}{\beta_l w_l}\right)\right) = Cw. \quad \square$$

5.1.3 Degree bounds in canonical form

Recall that the modular function j_{n+1} satisfies eq. (5.1) on $\mathcal{S}$: we have $E(j_1, \ldots, j_{n+1}) = 0$ where

$$E = \sum_{k=0}^{e} E_k(J_1, \ldots, J_n)\, J_{n+1}^k \in L[J_1, \ldots, J_{n+1}]$$

has degree e in J_{n+1} and is irreducible. Let d_E denote the total degree of E in the variables $J_1, \ldots, J_n$.

In this section, we work in the ring $L(J_1, \ldots, J_n)[J_{n+1}]$ modulo E. We say that a fraction $R \in L(J_1, \ldots, J_{n+1})$ is in *canonical form* if R is a polynomial in J_{n+1} of degree at most $e-1$.

Proposition 5.1.6. *Let $P, Q \in L[J_1, \ldots, J_{n+1}]$ be polynomials with total degree at most d, and assume that $Q(j_1, \ldots, j_{n+1})$ is not identically zero. Let $R \in L(J_1, \ldots, J_n)[J_{n+1}]$ be the fraction in canonical form such that $P/Q = R \bmod E$. Then the total degree of R in $J_1, \ldots, J_n$ is bounded above by $(e + 2d_E)d$.*

Proof. In this proof, degrees and coefficients are taken with respect to the variable J_{n+1} unless otherwise specified. First, we invert the denominator Q. Consider the resultant

$$Z = \mathrm{Res}_{J_{n+1}}(Q, E) \in L[J_1, \ldots, J_n],$$

which is nonzero by hypothesis. Let $U, V \in L[j_1, \ldots, j_{n+1}]$ be the associated Bézout coefficients, so that

$$Z = UQ + VE.$$

The inverse of Q modulo E is U/Z, hence $P/Q = UP/Z \bmod E$.

Recall that Z (resp. Q) has a polynomial expression of degree e (resp. degree $e-1$) in the coefficients of Q, and degree $\deg(Q)$ in the coefficients of E. Since the total degree of Q is at most d, the total degrees of Z and UP in $J_1, \ldots, J_n$ are bounded above by $d(e + d_E)$. The degree of UP in J_{n+1} is at most $d + e - 1$.

Now, we reduce UP/Z modulo E to obtain a numerator of degree at most $e-1$ in J_{n+1}. We can decrease this degree by 1 by multiplying above and below by $E_e(J_1, \ldots, J_n)$ and using the relation

$$E_e J_{n+1}^e = -\sum_{k=0}^{e-1} E_k\, J_{n+1}^k \qquad \bmod E.$$

When doing so, the total degree in $J_1, \ldots, J_n$ increases by at most d_E. This operation is done at most d times; the result has total degree at most $(e + 2d_E)d$ in $J_1, \ldots, J_n$ and $e - 1$ in J_{n+1}. □

Definition 5.1.7. The *geometric complexity* of the invariants $j_1, \ldots, j_{n+1}$ is defined as

$$\mathrm{GC}(j_1, \ldots, j_{n+1}) = (e + 2d_E)\,\mathrm{SGC}(j_1, \ldots, j_{n+1}) + e - 1.$$

This quantity depends on the choice of j_{n+1} as a distinguished invariant.

Proposition 5.1.8. *Let f and g be symmetric modular forms on $\mathcal{S}$ of weight w, and assume that $g \neq 0$. Let $R \in L(J_1, \ldots, J_n)[J_{n+1}]$ be the rational fraction in canonical form such that*

$$\frac{f}{g} = R(j_1, \ldots, j_{n+1}).$$

Then the degree of R in $J_1, \ldots, J_{n+1}$ is bounded above by $\mathrm{GC}(j_1, \ldots, j_{n+1})w$.

Proof. Combine propositions 5.1.5 and 5.1.6. □

We are ready to prove the first part of theorem 5.0.1 on degree bounds for modular equations, with an explicit expression for the constant C_1.

Theorem 5.1.9. *Let H_δ be an absolutely irreducible Hecke correspondence on $\mathcal{S} \times \mathcal{T}$ defined by $\delta \in G(\mathbb{A}_f)$, and let $d(\delta)$ be the degree of H_δ. For each $1 \leq i \leq n+1$, let χ_i be a denominator of the modular function j_i as in §5.1.1. Let $F \in L(J_1, \ldots, J_n)[J_{n+1}]$ be a coefficient of one of the modular equations $\Psi_{\delta,m}$ for $1 \leq m \leq n+1$. Then the total degree of F is bounded above by $C\,d(\delta)$, where C is a constant independent of δ; more precisely we can take $C = \mathrm{GC}(j_1, \ldots, j_{n+1})\,(\#\Sigma) \sum_{i=1}^{m} \mathrm{wt}(\chi_i)$.*

Proof. Let $g_{\delta,m}$ be the modular form on $\mathcal{S}$ defined in §5.1.1. By proposition 5.1.2, the modular function $F(j_1, \ldots, j_{n+1})$ is of the form $f/g_{\delta,m}$, where f is a modular form on $\mathcal{S}$ of weight $\mathrm{wt}(g_{\delta,m})$. By lemma 5.1.1, we have

$$\mathrm{wt}(g_{\delta,m}) = (\#\Sigma)\,d(\delta) \sum_{i=1}^{m} \mathrm{wt}(\chi_i),$$

so the result follows from proposition 5.1.8. □

5.1.4 Complete denominators in dimension 2

In the case of Siegel modular equations for abelian surfaces, and Hilbert modular equations for the real quadratic field $F = \mathbb{Q}(\sqrt{5})$, it is possible to refine the choice of denominator $g_{\delta,m}$ in such a way that the coefficients of $g_{\delta,m}\Psi_{\delta,m}$ are defined over $\mathbb{Z}$, i.e. have integral Fourier coefficients. This refinement is not necessary to obtain explicit degree bounds for these modular equations (see §5.1.5), but will be useful in the evaluation algorithm of chapter 6.

Recall from §2.4.4 that for each prime ℓ, the Siegel modular equations of level ℓ are denoted by $\Psi_{\ell,k} \in \mathbb{Q}(J_1, J_2, J_3)[Y]$ for $1 \leq k \leq 3$. Similarly, for each totally positive prime $\beta \in \mathbb{Z}_F$, the Hilbert modular equations of level β in Gundlach invariants are denoted by $\Psi_{\beta,1}, \Psi_{\beta,2} \in \mathbb{Q}(J_1, J_2)[Y]$. In both cases, the underlying moduli space $\mathcal{S}$ is rational.

The Siegel case. We call a polynomial $D_\ell \in \mathbb{Z}[J_1, J_2, J_3]$ a *complete denominator* of the Siegel modular equations $\Psi_{\ell,k}$ if for each $1 \leq k \leq 3$, we have

$$D_\ell \Psi_{\ell,k} \in \mathbb{Z}[J_1, J_2, J_3, Y].$$

Our goal is to construct a complete denominator for the Siegel modular equations, given by an analytic formula. Let C_ℓ be a set of representatives for the quotient $\Gamma^0(\ell)\backslash \mathrm{Sp}_4(\mathbb{Z})$, the congruence subgroup $\Gamma^0(\ell)$ being defined by eq. (3.1). For every $\tau \in \mathbb{H}_2$, we define

$$g_\ell(\tau) = \prod_{\eta \in C_\ell} \det(\eta^*\tau)^{-20} h_{10}^2(\tfrac{1}{\ell}\eta\tau).$$

One can check that the function g_ℓ is independent of the choice of C_ℓ, and is a Siegel modular form of weight

$$w_\ell = 20(\ell^3 + \ell^2 + \ell + 1).$$

Under the correspondence of §2.3.2, the Siegel modular form g_ℓ is the classical analogue of the modular form $g_{\delta,m}$ for $m = 2$ or $m = 3$ defined in proposition 5.1.3, if we choose the modular forms $h_{10}, h_{10}^2, h_{10}^2$ as denominators for the Igusa invariants j_1, j_2, j_3 respectively.

For every $\tau \in \mathbb{H}_2$ and $0 \leq i \leq \ell^3 + \ell^2 + \ell + 1$, we define $f_{\ell,k}^{(i)}(\tau)$ as the coefficient of Y^i in the polynomial $g_\ell(\tau)\Psi_{\ell,k}(j_1(\tau), j_2(\tau), j_3(\tau)) \in \mathbb{C}[Y]$. The functions $f_{\ell,k}^{(i)}$ are holomorphic on $\mathbb{H}_2$, and also are Siegel modular forms of weight w_ℓ by proposition 5.1.3.

Proposition 5.1.10. *The modular forms g_ℓ and $f_{\ell,k}^{(i)}$ are defined over $\mathbb{Z}$.*

Proof. Both the Hecke correspondence and the modular form h_{10} are defined over $\mathbb{Q}$, therefore the modular form g_ℓ is defined over $\mathbb{Q}$ as well. This is also the case of the modular forms $f_{\ell,k}^{(i)}$, because the coefficients of Siegel modular equations are defined over $\mathbb{Q}$ as modular functions. Therefore we only have to show that their Fourier coefficients are algebraic integers.

Let f be a Siegel modular form on $\mathbb{H}_2$ of weight k defined over $\mathbb{Z}$, and let $\eta \in C_\ell$. We claim that the function

$$h(\tau) = \det(\eta^*\tau)^{-k} f(\tfrac{1}{\ell}\eta\tau)$$

has a Fourier expansion in terms of $\exp(2\pi i z_j(\tau)/\ell^2)$ for $1 \leq j \leq 3$, where we write

$$\tau = \begin{pmatrix} z_1(\tau) & z_3(\tau) \\ z_3(\tau) & z_2(\tau) \end{pmatrix},$$

with coefficients in the ring $\mathbb{Z}[\exp(2\pi i/\ell^2)]$. This claim implies proposition 5.1.10 because g_ℓ and $f_{\ell,k}^{(i)}$ are polynomials in such functions.

In order to compute the Fourier expansion of $h(\tau)$, we compute a matrix $\eta^R \in \mathrm{Sp}_4(\mathbb{Z})$ such that the transformation $\tau \mapsto \eta^R(\frac{1}{\ell}\eta\tau)$ leaves the cusp at infinity "invariant". More precisely, writing 4×4 matrices in 2×2 blocks, we require that

$$\eta^R \begin{pmatrix} a & b \\ \ell c & \ell d \end{pmatrix} = \begin{pmatrix} A_\eta & B_\eta \\ 0 & D_\eta \end{pmatrix}, \quad \text{where } \eta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

To compute η^R, we proceed as follows. Denote by $\langle \cdot, \cdot \rangle$ the standard alternating form $\left(\begin{smallmatrix} 0 & I_2 \\ -I_2 & 0 \end{smallmatrix}\right)$ on $\mathbb{Q}^4$. Let $u_1, u_2 \in \mathbb{Z}^4$ be the two columns of the 4×2 matrix $\left(\begin{smallmatrix} -\ell c \\ a \end{smallmatrix}\right)$. Then $\langle u_1, u_2 \rangle = 0$, hence u_1 and u_2 are contained in an isotropic subspace $V \subset \mathbb{Q}^4$ of dimension 2. The two last lines of η^R are given by a basis of $\mathbb{Z}^4 \cap V$, and we complete them into a symplectic basis of $\mathbb{Z}^4$ to obtain η^R.

A possible choice of C_ℓ consists of the $\ell^3 + \ell^2 + \ell + 1$ following matrices [Dup06, Prop. 10.1]:

- $T_1(a,b,c) = \left(\begin{smallmatrix} -I_2 & \begin{smallmatrix} a & b \\ b & c \end{smallmatrix} \\ 0 & -I_2 \end{smallmatrix}\right)$ for $a, b, c \in [\![0, \ell - 1]\!]$,
- $T_2(a,b,c) = \left(\begin{smallmatrix} 0 & -I_2 \\ I_2 & \begin{smallmatrix} -a & -b \\ -b & -c \end{smallmatrix} \end{smallmatrix}\right)$ for $a, b, c \in [\![0, \ell - 1]\!]$ such that $ac = b^2 \bmod \ell$,
- $T_3(a) = \left(\begin{smallmatrix} -1 & -a & 0 & 0 \\ 0 & 0 & -a & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{smallmatrix}\right)$ for $a \in [\![0, \ell - 1]\!]$, and
- $T_4 = \left(\begin{smallmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 0 & 0 \end{smallmatrix}\right)$.

Let us detail a possible choice of η^R for all these matrices.

- If $\eta = T_1(a,b,c)$, we take
$$\eta^R = I_4, \quad D_\eta = -\ell I_2, \quad \det D_\eta = \ell^2.$$
- If $\eta = T_2(a,b,c)$, we take
$$\eta^R = \begin{pmatrix} 0 & I_2 \\ -I_2 & 0 \end{pmatrix}, \quad D_\eta = I_2, \quad \det D_\eta = 1.$$
- If $\eta = T_3(a)$, we take
$$\eta^R = \begin{pmatrix} 0&0&0&-1 \\ 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \end{pmatrix}, \quad D_\eta = \begin{pmatrix} -a & 1 \\ -\ell & 0 \end{pmatrix}, \quad \det D_\eta = \ell.$$
- If $\eta = T_4$, we take
$$\eta^R = \begin{pmatrix} 1&0&0&0 \\ 0&\ell&0&1 \\ -\ell&\ell&1&1 \\ 1&-1&0&0 \end{pmatrix}, \quad D_\eta = \begin{pmatrix} 2\ell & \ell \\ -1 & 0 \end{pmatrix}, \quad \det D_\eta = \ell.$$

Then we have
$$\begin{aligned} h(\tau) &= \det(\eta^*\tau)^{-k} \det\bigl(\eta^{R*}(\tfrac{1}{\ell}\eta\tau)\bigr)^{-k} f\bigl((A_\eta\tau + B_\eta)D_\eta^{-1}\bigr) \\ &= \ell^{2k} \det(D_\eta)^{-k} f\bigl((A_\eta\tau + B_\eta)D_\eta^{-1}\bigr). \end{aligned}$$
We develop $f\bigl((A_\eta\tau + B_\eta)D_\eta^{-1}\bigr)$ using the Fourier expansion of f, and obtain an expansion with coefficients in $\mathbb{Z}[\exp(2\pi i/\ell^2)]$ in terms of the quantities $\exp(2\pi i z_j(\tau)/\ell^2)$ for $1 \le j \le 3$. Moreover $\det(D_\eta)^{-k}\ell^{2k} \in \mathbb{Z}$. □

The computations in the proof of proposition 5.1.10 show that g_ℓ is divisible by $\ell^{20(2\ell^2+\ell+1)}$, but we do not need this fact. We finally define
$$D_\ell(\tau) = 12^{w_\ell} \frac{h_4(\tau)^{\lfloor w_\ell/6 \rfloor}}{h_{10}(\tau)^a h_4(\tau)^b} \prod_{\eta \in C_\ell} \det(\eta^*\tau)^{-20} h_{10}^2(\tfrac{1}{\ell}\eta\tau), \tag{5.2}$$
where a, b are such that $4\lfloor w_\ell/6 \rfloor + w_\ell = 10a + 4b$ with $0 \le b \le 4$.

Proposition 5.1.11. *$D_\ell \in \mathbb{Z}[J_1, J_2, J_3]$, and $D_\ell \Psi_{\ell,k} \in \mathbb{Z}[J_1, J_2, J_3, Y]$ for each $1 \le k \le 3$.*

Proof. By propositions 2.3.7 and 5.1.10, we know that

$$12^{w_\ell} g_\ell \in \mathbb{Z}[h_4, h_6, h_{10}, h_{12}], \quad \text{and} \quad 12^{w_\ell} f_{\ell,k}^{(i)} \in \mathbb{Z}[h_4, h_6, h_{10}, h_{12}] \text{ for all } k, i.$$

Moreover, using the equalities

$$h_4 h_6 = j_1 h_{10}, \quad h_4^2 h_{12} = j_2 h_{10}^2, \quad h_4^5 = j_3 h_{10}^2,$$

one can show (following the proof of proposition 5.1.5) that for every modular form $f \in \mathbb{Z}[h_4, h_6, h_{10}, h_{12}]$ of weight k, we have

$$\frac{h_4^{\lfloor k/6 \rfloor} f}{h_{10}^a h_4^b} \in \mathbb{Z}[j_1, j_2, j_3]$$

where a, b are such that $4\lfloor k/6 \rfloor + k = 10a + 4b$ and $0 \leq b \leq 4$. □

The Hilbert case. Fix $F = \mathbb{Q}(\sqrt{5})$, and let $\beta \in \mathbb{Z}_F$ be a totally positive prime. We call $D_\beta \in \mathbb{Z}[J_1, J_2]$ a *complete denominator* of the Hilbert modular polynomials $\Psi_{\beta,k}$ if for each $1 \leq k \leq 2$, we have

$$D_\beta \Psi_{\beta,k} \in \mathbb{Z}[J_1, J_2, Y].$$

We construct a complete denominator given by an analytic formula, as in the Siegel case.

Recall that the subgroup $\Gamma^0(\beta)$ of $\Gamma(1)_F = \mathrm{SL}(\mathbb{Z}_F \oplus \partial_F^{-1})$ is defined by

$$\Gamma_F^0(\beta) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(1)_F \colon b = 0 \bmod \beta \right\}.$$

Let $\Delta = 5$ be the discriminant of F, and write $\ell = N_{F/\mathbb{Q}}(\beta)$. A set C_β of representatives for the quotient $\Gamma^0(\beta)\backslash\Gamma(1)_F$ consists of the $\ell + 1$ following matrices:

$$\begin{pmatrix} 0 & \sqrt{\Delta} \\ -1/\sqrt{\Delta} & 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & a\sqrt{\Delta} \\ 0 & 1 \end{pmatrix} \quad \text{for } a \in [\![0, \ell - 1]\!],$$

Let σ denote the real conjugation automorphism in F; it induces an automorphism of $\mathbb{H}_1^2$ given by $\sigma((\tau_1, \tau_2)) = (\tau_2, \tau_1)$. A set of representatives for the quotient $\Gamma^0(\beta)\backslash\big(\Gamma(1)_F \rtimes \langle\sigma\rangle\big)$ is given by $C_\beta^\sigma := C_\beta \sqcup C_\beta\sigma$.

For every $\tau \in \mathbb{H}_1^2$, we define

$$g_\beta(\tau) = \prod_{\eta \in C_\beta^\sigma} (\eta^*\tau)^{-10} F_{10}(\tfrac{1}{\beta}\eta\tau).$$

We can show that g_β is independent of the choice of C_β, and that g_β is a symmetric Hilbert modular form of weight

$$w_\beta = 10(2\ell + 2).$$

As in the Siegel case, the Hilbert modular form g_β is the classical analogue of the modular form $g_{\delta,m}$ for $m = 1$ or $m = 2$ defined in proposition 5.1.3, if we choose F_{10} as the denominator of Gundlach invariants; this choice is valid by proposition 2.3.16. For every $0 \leq i \leq 2\ell+2$ and $1 \leq k \leq 2$, we define $f^{(i)}_{\beta,k}(\tau)$ as the coefficient of Y^i in the polynomial $g_\beta(\tau)\Psi_{\beta,k}(g_1(\tau), g_2(\tau)) \in \mathbb{C}[Y]$. It is also a symmetric Hilbert modular form of weight w_β.

Proposition 5.1.12. *The modular forms g_β and $f^{(i)}_{\beta,k}$ are defined over $\mathbb{Z}$.*

Proof. As above, both the Hecke correspondence and F_{10} are defined over $\mathbb{Q}$, hence g_β is defined over $\mathbb{Q}$. This is also the case of the modular forms $f^{(i)}_{\beta,k}$ because the coefficients of Hilbert modular equations are defined over $\mathbb{Q}$ as modular functions.

As in proposition 5.1.10, it is enough to show the following: if f is a Hilbert modular form of weight k and $\eta \in C^\sigma_\beta$, then the function

$$h(\tau) = (\eta^*\tau)^{-k} f(\tfrac{1}{\beta}\eta\tau)$$

on $\mathbb{H}_1^2$ has an expansion in terms of the quantities $\exp(2\pi i(n\tau_1 + \overline{n}\tau_2)/\ell)$, where n runs through totally positive elements of $\mathbb{Z}_F$, with coefficients in the ring $\mathbb{Z}[\exp(2\pi i/\ell)]$. The computations are easier than in the Siegel case due to the simpler form of coset representatives:

- If $\eta = \left(\begin{smallmatrix} 0 & \sqrt{\Delta} \\ -1/\sqrt{\Delta} & 0 \end{smallmatrix}\right)$, then we make η act again, and we have

$$(\eta^*\tau)^{-k} f(\tfrac{1}{\beta}\eta\tau) = \ell f(\beta\tau).$$

- If $\eta = \left(\begin{smallmatrix} 1 & a\sqrt{\Delta} \\ 0 & 1 \end{smallmatrix}\right)$, then we directly have

$$(\eta^*\tau)^{-k} f(\tfrac{1}{\beta}\eta\tau) = f\big(\tfrac{1}{\beta}(t + (a\sqrt{\Delta}, -a\sqrt{\Delta}))\big).$$

Therefore the Fourier expansion of h has integer coefficients. □

For every $\tau \in \mathbb{H}_1^2$, we define

$$D_\beta(\tau) = 2^{w_\beta} \frac{G_2(\tau)^{2\lfloor w_\beta/6 \rfloor}}{F_{10}(\tau)^a G_2(\tau)^b} \prod_{\eta \in C^\sigma_\beta} (\eta^*\tau)^{-10} F_{10}(\tfrac{1}{\beta}\eta\tau), \tag{5.3}$$

where a, b are such that $2\lfloor w_\beta/6 \rfloor + w_\beta = 10a + 2b$ with $0 \leq b \leq 4$.

Proposition 5.1.13. *Let $F = \mathbb{Q}(\sqrt{5})$, let $\beta \in \mathbb{Z}_F$ be a totally positive prime, and define D_β as in eq. (5.3). Then D_β is a complete denominator of the Hilbert modular equations of level β in Gundlach invariants.*

Proof. By propositions 2.3.15 and 5.1.12, we know that

$$2^{w_\beta} g_\beta \in \mathbb{Z}[G_2, F_6, F_{10}], \quad \text{and} \quad 2^{w_\eta} f^{(i)}_{\beta,k} \in \mathbb{Z}[G_2, F_6, F_{10}] \text{ for all } k, i.$$

Moreover, using the equalities

$$G_2^2 F_6 = g_2 F_{10}, \quad G_2^5 = g_1 F_{10},$$

one can show (following the proof of proposition 5.1.5) that for every modular form $f \in \mathbb{Z}[G_2, F_6, F_{10}]$ of weight k, we have

$$\frac{G_2^{2\lfloor k/6 \rfloor} f}{F_{10}^a G_2^b} \in \mathbb{Z}[g_1, g_2]$$

where a, b are such that $4\lfloor k/6 \rfloor + k = 10a + 2b$ and $0 \leq b \leq 4$. □

In the case of Hilbert modular equations in Igusa invariants for a general real quadratic field F, we are not able to determine a complete denominator, because the structure of the ring of Hilbert modular forms defined over $\mathbb{Z}$ is not known a priori.

5.1.5 Explicit degree bounds in dimension 2

Our methods provide new results about the degrees of the coefficients of modular equations of Siegel and Hilbert type for abelian surfaces. In the Hilbert case, we restrict to the quadratic field $F = \mathbb{Q}(\sqrt{5})$, and we consider modular equations in Gundlach invariants.

In both cases, the variety $\mathcal{S}$ is a rational. We can choose $j_{n+1} = 1$, and take $E = J_{n+1} - 1$ as the equation satisfied by j_{n+1} on $\mathcal{S}$. Then the notions of geometric complexity and symmetric geometric complexity coincide.

Lemma 5.1.14. *Let j_1, j_2, j_3 denote the Igusa invariants on the Siegel threefold $\mathrm{Sp}_4(\mathbb{Z})\backslash\mathbb{H}_2$, as defined in §2.3.2. Then we have*

$$\mathrm{GC}(j_1, j_2, j_3, 1) \leq \frac{1}{6}.$$

Proof. By theorem 2.3.5, the graded ring of Siegel modular forms is generated by

$$f_1 = I'_6, \quad f_2 = I_{12}, \quad f_3 = I_4, \quad \text{and} \quad f_4 = I_{10}.$$

We are in case 1 of definition 5.1.7, since

$$\frac{I_6' I_4}{I_{10}} = j_1, \quad \frac{I_{12} I_4^2}{I_{10}^2} = j_2, \quad \text{and} \quad \frac{I_4^5}{I_{10}^2} = j_3.$$

The definition gives

$$\mathrm{SGC}(j_1, j_2, j_3, 1) \leq \left(1 + \frac{2}{3}\right) \cdot \frac{1}{10} = \frac{1}{6}. \qquad \square$$

Proposition 5.1.15. *Let ℓ be a prime number, and let $\Psi_{\ell,m}$ for $1 \leq m \leq 3$ denote the Siegel modular equations of level ℓ. Let $F \in \mathbb{Q}(J_1, J_2, J_3)$ be a coefficient of $\Psi_{\ell,1}$ (resp. $\Psi_{\ell,2}$ or $\Psi_{\ell,3}$). Then the total degree of F is bounded above by $5\, d(\ell)/3$ (resp. $10\, d(\ell)/3$), where $d(\ell) = \ell^3 + \ell^2 + \ell + 1$.*

Proof. The integer $d(\ell)$ is the degree of the Hecke correspondence. The denominators of j_1, j_2, j_3 can be taken to be the Siegel modular forms h_{10}, h_{10}^2, and h_{10}^2. Let $g_{\ell,m}$ for $1 \leq m \leq 3$ be the common denominators of the modular equations $\Psi_{\ell,m}$ as in proposition 5.1.3, so that $g_{\ell,2} = g_{\ell,3} = g_{\ell,1}^2$ and $\mathrm{wt}(g_{\ell,1}) = 10\, d(\ell)$. The modular form $g_{\ell,2}$ corresponds to the classical modular form g_ℓ studied in §5.1.4.

Let F be a coefficient of $\Psi_{\ell,1}$ (resp. $\Psi_{\ell,2}$ or $\Psi_{\ell,3}$). Then $F(j_1, j_2, j_3)$ is the quotient of two modular forms of degree $10\, d(\ell)$ (resp. $20\, d(\ell)$) on $\mathcal{S}$, by proposition 5.1.3. Therefore the result follows from lemma 5.1.14 and proposition 5.1.8. $\square$

Lemma 5.1.16. *Let $F = \mathbb{Q}(\sqrt{5})$, and let g_1, g_2 denote the Gundlach invariants on the Hilbert surface $\mathrm{SL}(\mathbb{Z}_F \oplus \partial_F^{-1}) \backslash \mathbb{H}_1^2$, as defined in §2.3.4. Then we have*

$$\mathrm{GC}(g_1, g_2, 1) \leq \frac{1}{6}.$$

Proof. Choose F_6, F_2, F_{10} as generators for the graded ring of Hilbert modular forms for F. We have

$$\frac{F_6 F_2^2}{F_{10}} = g_2 \quad \text{and} \quad \frac{F_2^5}{F_{10}} = g_1.$$

We are in case 1 of definition 5.1.7, and

$$\mathrm{GC}(g_1, g_2, 1) \leq \left(1 + \frac{2}{3}\right) \cdot \frac{1}{10} = \frac{1}{6}. \qquad \square$$

Proposition 5.1.17. *Let $F = \mathbb{Q}(\sqrt{5})$, let $\beta \in \mathbb{Z}_F$ be a totally positive prime, and let $\Psi_{\beta,m}$ for $m \in \{1, 2\}$ denote the Hilbert modular equations of level β in Gundlach invariants. Let $F \in \mathbb{Q}(J_1, J_2)$ be a coefficient of $\Psi_{\beta,1}$ or $\Psi_{\beta,2}$. Then the total degree of F is bounded above by $10\, d(\beta)/3$, where $d(\beta) = N_{F/\mathbb{Q}}(\beta) + 1$.*

Proof. The integer $d(\beta)$ is the degree of the Hecke correspondence, and the automorphism group Σ used to define the Hilbert modular equations has order 2. We can take the modular F_{10} as denominator of both g_1 and g_2. The common denominator $g_{\beta,1} = g_{\beta,2}$ from proposition 5.1.3 has weight $20\, d(\beta)$; it corresponds to the classical Hilbert modular form g_β studied in §5.1.4. The result follows from lemma 5.1.14 and proposition 5.1.8. □

The degree bounds in propositions 5.1.15 and 5.1.17 are both reached experimentally. In the Siegel case with $\ell = 2$, the maximum degree is 25; in the Hilbert case with $N_{F/\mathbb{Q}}(\beta) = 41$, the maximum degree is 140 [Mil]. In particular, the inequalities in lemmas 5.1.14 and 5.1.16 are in fact equalities.

5.2 Height bounds for modular equations

We keep the notation introduced at the beginning of chapter 5. We also write $\mathcal{S} = \Gamma\backslash X_+$, where Γ is a subgroup of $G(\mathbb{Q})_+$. As we mentioned above, there are two main steps in the proof of height bounds in theorem 5.0.1: first, we study the heights of evaluations of modular equations; then we apply the results of chapter 4 to obtain height bounds on their coefficients.

5.2.1 Heights of abelian varieties

Different types of heights can be defined for an abelian variety A over a number field L. The *Faltings height* $h_F(A)$ is defined in [Fal83, §3] in terms of Arakelov degrees of metrized line bundles on A. If A is given a principal polarization $\mathcal{L}$, and $r \geq 2$ is an even integer, we can also define the *theta height of level* r of $(A, \mathcal{L})$, denoted by $h_{\Theta,r}(A, \mathcal{L})$, as the projective height of level r theta constants of $(A, \mathcal{L})$ [Paz12, Def. 2.6]. Finally, if A is an abelian variety with PEL structure over L given by a point $z \in \mathcal{S}$ where $j_1, \dots, j_{n+1}$ are well defined, we can define the *j-height* of A as

$$h_j(A) = h\big(j_1(A), \dots, j_{n+1}(A)\big).$$

We also write $\overline{h}_F(A) = \max\{1, h_F(A)\}$ and define $\overline{h}$, $\overline{h}_{\Theta,r}$, $\overline{h}_j$ similarly.

The goal of this section is to relate the j-heights of isogenous abelian varieties, under mild conditions related to the geometry of the moduli space. Such a relation is known for instance in the case of elliptic curves, taking the usual j-invariant as coordinate [Paz19, Thm. 1.1]. To this end, we relate the j-height with the Faltings height, since the latter behaves well with respect to isogenies. Theta heights are an intermediate step between concrete values of invariants and the Faltings height. More precisely, we use the two following results.

Proposition 5.2.1. *Let A, A' be abelian varieties over $\overline{\mathbb{Q}}$, and assume that an isogeny $\varphi\colon A \to A'$ exists. Then*

$$\big|h_F(A) - h_F(A')\big| \leq \tfrac{1}{2}\log(\deg\varphi).$$

Proof. This is a consequence of [Fal83, Lem. 5]. □

Theorem 5.2.2 ([Paz12, Cor. 1.3])**.** *For every $g \geq 1$, and every even $r \geq 2$, there exists a constant $C(g,r)$ such that the following holds. Let $(A, \mathcal{L})$ be a principally polarized abelian variety of dimension g defined over $\overline{\mathbb{Q}}$. Then*

$$\big|\overline{h}_{\Theta,r}(A,\mathcal{L}) - \tfrac{1}{2}\,\overline{h}_F(A)\big| \leq C(g,r)\log\big(\min\{\overline{h}_F(A), \overline{h}_{\Theta,r}(A,\mathcal{L})\} + 2\big).$$

We can take

$$C(g,r) = 1000 r^{2g}\log^5(r^{2g}).$$

5.2.2 Relating the j-height and the Faltings height

Using theorem 5.2.2, we can prove that the j-height and the Faltings height of a generic abelian variety with PEL structure are related.

Proposition 5.2.3. *There exists a nonzero polynomial $P \in L[Y_1, \dots Y_{n+1}]$ and a positive constant C such that the following holds: if A is the abelian variety with PEL structure associated with a point $z \in \mathcal{S}$ where $j_1, \dots, j_{n+1}$ are well defined and $P(j_1, \dots, j_{n+1}) \neq 0$, and if A is defined over $\overline{\mathbb{Q}}$, then*

$$\frac{1}{C}\,\overline{h}_F(A) \leq \overline{h}_j(A) \leq C\,\overline{h}_F(A).$$

Proof. By [Mil05, Thm. 5.17], we can write $\mathcal{S} = \Gamma'\backslash X_+$ where Γ' is a congruence subgroup of G^{der}. Since $G^{\mathrm{der}} \subset \ker(\det)$, it embeds into $\mathrm{GSp}_{2g}(\mathbb{Q})$, where $2g = \dim_{\mathbb{Q}} V$. Therefore, by [Mil05, Thm. 5.16], we can find a congruence subgroup Γ'' of G^{der} and an even integer $r \geq 4$ such that $\Gamma''\backslash X_+$ embeds in the moduli space $\mathcal{A}_{\Theta,r}$ of principally polarized abelian varieties of dimension g with level r theta structure. We have a diagram

$$\begin{array}{ccccc} & & \widetilde{\mathcal{S}} = \widetilde{\Gamma}\backslash X_+ & & \\ & \overset{p'}{\swarrow} & & \overset{p''}{\searrow} & \\ \mathcal{S} = \Gamma'\backslash X_+ & & & & \mathcal{S}'' = \Gamma''\backslash X_+ \overset{\iota}{\hookrightarrow} \mathcal{A}_{\Theta,r} \end{array} \tag{5.4}$$

where $\widetilde{\Gamma} = \Gamma' \cap \Gamma''$. The maps p' and p'' are finite coverings, and all the varieties and maps in this diagram are defined over $\overline{\mathbb{Q}}$.

The modular interpretation of diagram (5.4) is the following. Let (Λ, ψ) be the standard polarized lattice associated with the connected component $\mathcal{S}$, as in proposition 2.2.8. We can find a sublattice $\Lambda'' \subset \Lambda$, and $\lambda \in \mathbb{Q}^\times$ such that $(\Lambda'', \lambda\psi)$ is principally polarized. A point $z \in \mathcal{S}$ defines a complex structure x on $\Lambda \otimes \mathbb{R} = V(\mathbb{R})$, up to action of Γ. Lifting z to $\widetilde{z} \in \widetilde{\mathcal{S}}$ corresponds to considering x up to action of $\widetilde{\Gamma}$ only, and this group leaves Λ'' and its level r theta structure stable. Then the image of $\widetilde{z}$ in $\mathcal{A}_{\Theta,r}$ is given by $(\Lambda'', x, \lambda\psi)$.

In particular, if $\widetilde{z} \in \widetilde{\mathcal{S}}$, and if A and A'' are the abelian varieties corresponding to the points $p'(\widetilde{z}) \in \mathcal{S}$ and $\iota \circ p''(\widetilde{z}) \in \mathcal{A}_{\Theta,r}$ respectively, then A and A'' are linked by an isogeny of degree $d = \#(\Lambda/\Lambda'')$. Hence, by proposition 5.2.1 and theorem 5.2.2, we have

$$\begin{aligned}
&\left|\overline{h}_F(A) - 2\,\overline{h}_{\Theta,r}(A'')\right| \\
&\quad \leq \frac{\log(d)}{2} + C(g,r)\log\Big(\min\{\overline{h}_F(A), \overline{h}_{\Theta,r}(A'')\} + 2 + \tfrac{1}{2}\log(d)\Big) \\
&\quad \leq C_F \min\{\overline{h}_F(A), \overline{h}_{\Theta,r}(A'')\}
\end{aligned}$$

with $C_F = (2 + \log(d))C(g,r)$. Therefore

$$\overline{h}_F(A) \leq (2 + C_F)\,\overline{h}_{\Theta,r}(A'') \quad \text{and} \quad \overline{h}_{\Theta,r}(A'') \leq \frac{1 + C_F}{2}\,\overline{h}_F(A). \tag{5.5}$$

Now we relate the theta height and the j-height using relations between modular functions; the genericity hypothesis encoded in the polynomial P appears in this step. Denote by $\theta_0, \ldots, \theta_k$ the theta constants of level r. They define a projective embedding of $\mathcal{A}_{\Theta,r}$, therefore the pullbacks of the modular functions $\theta_1/\theta_0, \ldots, \theta_k/\theta_0$ generate the function field of $\mathcal{S}''$. By definition, $j_1, \ldots, j_{n+1}$ are coordinates on $\mathcal{S}$. To ease notation, we identify all these functions with their pullbacks to $\widetilde{\mathcal{S}}$.

By the primitive element theorem, there exists a function f on $\widetilde{\mathcal{S}}$ such that both of the tuples $(j_1, \ldots, j_{n+1}, f)$ and $(\theta_1/\theta_0, \ldots, \theta_k/\theta_0, f)$ are generating families for the function field of $\widetilde{\mathcal{S}}$ over $\overline{\mathbb{Q}}$. We choose polynomials

$$P_J \in \overline{\mathbb{Q}}[Y_1, \ldots, Y_{n+1}, X] \quad \text{and} \quad P_\Theta \in \overline{\mathbb{Q}}[Y_1, \ldots, Y_k, X]$$

such that $P_J(j_1, \ldots, j_{n+1}, X)$ and $P_\Theta(\theta_1/\theta_0, \ldots, \theta_k/\theta_0, X)$ are (non necessarily monic) minimal polynomials of f over the function fields of $\mathcal{S}$ and $\mathcal{S}''$ respectively. We also choose polynomials $N_{J,i}, D_{J,i} \in \overline{\mathbb{Q}}[Y_1, \ldots, Y_k, X]$ for $1 \leq i \leq n+1$, and $N_{\Theta,i}, D_{\Theta,i} \in \overline{\mathbb{Q}}[Y_1, \ldots, Y_{n+1}, X]$ for $1 \leq i \leq k$

such that the following equalities hold on $\widetilde{\mathcal{S}}$:

$$j_i = \frac{N_{J,i}}{D_{J,i}}(\theta_1/\theta_0, \dots, \theta_k/\theta_0, f) \qquad \text{for each } 1 \le i \le n+1, \text{ and}$$
$$\theta_i/\theta_0 = \frac{N_{\Theta,i}}{D_{\Theta,i}}(j_1, \dots, j_{n+1}, f) \qquad \text{for each } 1 \le i \le k.$$

Let $\widetilde{F}$ be the smallest Zariski closed subset of $\widetilde{\mathcal{S}}$ such that outside $\widetilde{F}$, the following properties are all statisfied:

- all the functions f, j_i for $1 \le i \le n+1$ and θ_i/θ_0 for $1 \le i \le k$ are well defined;
- the polynomials $P_J(j_1, \dots, j_{n+1}, X)$ and $P_\Theta(\theta_1/\theta_0, \dots, \theta_k/\theta_0, X)$ do not vanish;
- All of the quantities $D_{J,i}(\theta_1/\theta_0, \dots, \theta_k/\theta_0, f)$ for $1 \le i \le n+1$ and $D_{\Theta,i}(j_1, \dots, j_{n+1}, f)$ for $1 \le i \le k$ do not vanish.

Then $\widetilde{F}$ has codimension 1 in $\widetilde{\mathcal{S}}$, hence $\mathcal{U} = \mathcal{S} \backslash p'(\widetilde{F})$ is open dense in $\mathcal{S}$. Let $P \in L[Y_1, \dots, Y_{n+1}]$ be a polynomial such that $\{P(j_1, \dots, j_{n+1}) \neq 0\} \subset \mathcal{U}$.

Let $z \in \mathcal{S}$ be a point where $j_1, \dots, j_{n+1}$ are well defined, take values in $\overline{\mathbb{Q}}$, and satisfy $P(j_1, \dots, j_{n+1}) \neq 0$. We look at the diagram (5.4), from left to right. Lift z to a point $\widetilde{z} \in \widetilde{\mathcal{S}}$; by construction, $\widetilde{z} \notin \widetilde{F}$. By propositions 4.3.1 and 4.3.3, we have

$$\overline{h}\big(j_1(\widetilde{z}), \dots, j_{n+1}(\widetilde{z}), f(\widetilde{z})\big) \le C\, \overline{h}\big(j_1(z), \dots, j_{n+1}(z)\big) \tag{5.6}$$

with $C = h(P_J) + (n+1)\log(d_J + 1) + d_J + 1$, where d_J denotes the total degree of P_J in $Y_1 \dots, Y_{n+1}$. Writing $z'' = p''(\widetilde{z})$, we also have for every $1 \le i \le k$,

$$\overline{h}(\theta_i/\theta_0(\widetilde{z})) \le C\, \overline{h}\big(j_1(\widetilde{z}), \dots, j_{n+1}(\widetilde{z}), f(\widetilde{z})\big) \tag{5.7}$$

with $C = h(N_{\Theta,i}) + h(D_{\Theta,i}) + (n+2)\big(\log(\deg(N_{\Theta,i}) + 1) + \log(\deg(D_{\Theta,i}) + 1)\big) + \deg(N_{\Theta,i}) + \deg(D_{\Theta,i})$. Combining (5.6) and (5.7), we obtain

$$\overline{h}\Big(\frac{\theta_1}{\theta_0}(z''), \dots, \frac{\theta_k}{\theta_0}(z'')\Big) \le C_\Theta\, \overline{h}\big(j_1(z), \dots, j_{n+1}(z)\big)$$

where the constant C_Θ has an explicit expression in terms of the heigts and degrees of the polynomials P_J and $N_{\Theta,i}, D_{\Theta,i}$ for $1 \le i \le k$. Therefore, if A and A'' denote the abelian varieties associated with z and z'' respectively, we have

$$\overline{h}_{\Theta,r}(A'') \le C_\Theta\, \overline{h}_j(A),$$

hence by (5.5)

$$\overline{h}_F(A) \leq (2 + C_F)C_\Theta\, \overline{h}_j(A).$$

Going through the diagram (5.4) from right to left gives the reverse inequality

$$\overline{h}_j(A) \leq \frac{(1 + C_F)C_J}{2}\, \overline{h}_F(A)$$

where C_J is defined in a similar way to C_Θ in terms of the polynomials P_Θ and $N_{J,i}, D_{J,i}$ for $1 \leq i \leq n + 1$. □

Assume that the integers r and d, the modular function f, and the polynomials $P_J, P_\Theta, N_{J,i}, D_{J,i}, N_{\Theta,i}, D_{\Theta,i}$ can be explicitly determined. Then both the polynomial P and the constant C in proposition 5.2.3 can be determined explicitly as well. We will do this computation in a slightly different way in §5.2.5 in the case of Igusa invariants on the Siegel threefold.

In the sequel, we define $\mathcal{U}$ to be the Zariski open set in $\mathcal{S}$ where $j_1, \ldots, j_{n+1}$ are well defined and $P(j_1, \ldots, j_{n+1}) \neq 0$.

Corollary 5.2.4. *Let C be the constant from proposition 5.2.3, let z and z' be points of $\mathcal{U}$, and let A and A' be the abelian varieties with PEL structure associated with z and z'. Assume that A and A' are defined over $\overline{\mathbb{Q}}$, and are linked by an isogeny of degree ℓ. Then*

$$\overline{h}_j(A') \leq C^2(\overline{h}_j(A) + \log \ell).$$

Proof. Combine propositions 5.2.1 and 5.2.3. □

Remark 5.2.5. We can presumably do better than corollary 5.2.4. For instance, when studying j-invariants of isogenous elliptic curves, one can prove that $|\, h(j(E)) - h(j(E'))|$ is bounded by logarithmic terms [Paz19, Thm. 1.1]. This is also the kind of bound provided by theorem 5.2.2. The rough estimate in corollary 5.2.4 is sufficient for our purposes, but has the drawback that the constants we derive from it are very pessimistic.

5.2.3 Heights of evaluated modular equations

Let $\mathcal{U}$ (resp. $\mathcal{U}'$) be an open subset of $\mathcal{S}$ (resp. $\mathcal{T}$) where a relation bewteen the j-height and the Faltings height holds, as in proposition 5.2.3. Define $\mathcal{U}_\delta \subset \mathcal{S}$ to be the Zariski open set of all points $[x, g] \in \mathcal{S}$ such that $[x, g] \in \mathcal{U}$, and moreover the images of $[x, g]$ under the (symmetrized) Hecke correspondence H_δ all lie in $\mathcal{U}'$: in other words $[\sigma(x), \sigma(gk\delta)] \in \mathcal{U}'$ for every $(k, \sigma) \in K_0/K_{n+1}$, in the notation of §2.4.3. Finally, we define $\mathcal{V}_\delta \subset L^n$ to be the Zariski open set of all points $(j_1, \ldots, j_n)$ where the equation (5.1)

given by $E(j_1, \dots, j_n, J_{n+1}) = 0$ has e distinct roots and the following property holds: if j_{n+1} is a root of (5.1), then $(j_1, \dots, j_{n+1})$ are the invariants of some point $z \in \mathcal{U}_\delta$. In particular, the modular equations $\Psi_{\delta,m}$ do not have poles on $\mathcal{V}_\delta$.

Lemma 5.2.6. *There exist a constant C independent of δ, and a nonzero polynomial $P_\delta \in L[J_1, \dots, J_n]$ of total degree at most $Cd(\delta)$ such that $\{P_\delta(j_1, \dots, j_n) \neq 0\} \subset \mathcal{V}_\delta$.*

Proof. Let R be the the resultant of E and its derivative with respect to J_{n+1}. If R does not vanish at $(j_1, \dots, j_n) \in L^n$, then the polynomial $E(j_1, \dots, j_n, J_{n+1}) \in L[J_{n+1}]$ has e distinct roots.

Similarly, there exists a nonzero polynomial $Q \in L[J_1, \dots, J_{n+1}]$ such that every tuple $(j_1, \dots, j_{n+1}) \in L^{n+1}$ that satisfies eq. (5.1) and such that $Q(j_1, \dots, j_{n+1}) \neq 0$ lies in the image of $\mathcal{S}$. Let R' be the resultant of Q and E with respect to J_{n+1}. If R' does not vanish at $(j_1, \dots, j_n)$, then for every root j_{n+1} of $E(j_1, \dots, j_n, J_{n+1})$, the tuple $(j_1, \dots, j_{n+1})$ lies in the image of $\mathcal{S}$.

Let λ, λ' be symmetric modular forms on $\mathcal{S}$ and $\mathcal{T}$ respectively, defined over L, such that $\{\lambda \neq 0\} \subset \mathcal{U}$ and $\{\lambda' \neq 0\} \subset \mathcal{U}'$. These modular forms can be chosen independently of δ. As in §5.1.1, we construct the modular form

$$\lambda^\delta = \lambda \prod_{\gamma \in K_0/K'} \gamma \cdot \lambda'_\delta$$

where λ'_δ is the modular form $[x, g] \mapsto \lambda'([x, g\delta])$ of level K'. The modular form λ^δ is defined over L and has weight

$$\operatorname{wt}(\lambda^\delta) = \operatorname{wt}(\lambda) + (\#\Sigma)\, d(\delta) \operatorname{wt}(\lambda').$$

Modular forms realize a projective embedding of $\mathcal{S}$ by theorem 2.2.12; therefore, possibly after increasing the weight by a constant independent of δ, we can find a symmetric modular form ξ defined over L such that $\operatorname{wt}(\lambda^\delta) = \operatorname{wt}(\xi)$ and such that the divisors of λ^δ and ξ have no common codimension 1 components. By proposition 5.1.8, if we write

$$\frac{\lambda^\delta}{\xi} = \sum_{k=0}^{e-1} R_k(j_1, \dots, j_n) j_{n+1}^k \qquad \text{where } R_k \in L(J_1, \dots, J_n),$$

then $\deg R_k \leq \operatorname{GC}(j_1, \dots, j_{n+1}) \operatorname{wt}(\lambda^\delta)$ for every $0 \leq k \leq e-1$. Taking the resultant of $\sum R_k J_{n+1}^k$ and E with respect to J_{n+1} yields a rational fraction $R'' \in L(J_1, \dots, J_n)$ of degree at most

$$(e-1)d_E + e \max_{0 \leq k \leq e-1} \deg(R_k),$$

where d_E denotes the total degree of E in $j_1, \dots, j_n$. If R' and R'' are well defined and do not vanish at $(j_1, \dots, j_n)$, then for every root j_{n+1} of (5.1), the tuple $(j_1, \dots, j_{n+1})$ comes from a point $z \in \mathcal{U}_\delta$.

We take P_δ to be the product of R, R', and the numerator of R''. The polynomials R and R' are independent of δ, and the degree of R'' is bounded above linearly in $d(\delta)$. □

If degree bounds on equations defining $\mathcal{U}$ and $\mathcal{U}'$ are explicitly known, together with the polynomials E and Q, then the proof of lemma 5.2.6 allows us to determine a valid constant C explicitly.

Proposition 5.2.7. *There exists a constant C, independent of δ, such that the following holds. Let $(j_1, \dots, j_n) \in \mathcal{V}_\delta$, and let $1 \leq m \leq n+1$. Then*

$$h\big(\Psi_{\delta,m}(j_1, \dots, j_n)\big) \leq C\, d(\delta)\big(\overline{h}(j_1, \dots, j_n) + \log l(\delta)\big).$$

Proof. Let $\mathcal{J}$ be the set of roots of (5.1) at $(j_1, \dots, j_n)$, and let $j_{n+1} \in \mathcal{J}$. Let $[x, g]$ be a point of $\mathcal{S}$ describing an abelian variety A with PEL structure whose invariants are $(j_1, \dots, j_{n+1})$. For every $\sigma \in \Sigma$, denote by A_σ the abelian variety with PEL structure associated with the point $[\sigma(x), \sigma(g)]$.

By corollary 2.4.2, for each $\gamma = (\sigma, k) \in K_0/K_m$, the point $[\sigma(x), \sigma(gk\delta)]$ describes an abelian variety A_γ which is linked to A_σ by an isogeny of degree $l(\sigma(\delta)) = l(\delta)$. Therefore, by corollary 5.2.4, we have

$$\overline{h}\big(\gamma \cdot j_{1,\delta}([x, g]), \dots, \gamma \cdot j_{n+1,\delta}([x, g])\big) \leq C(\overline{h}\big(j_1, \dots, j_{n+1}) + \log l(\delta)\big),$$

where the constant $C > 0$ is independent of δ. By definition 2.4.6, the multivariate polynomial $\Psi_{\delta,m}(j_1, \dots, j_n, j_{n+1}) \in L[Y_1, \dots, Y_m]$ is the evaluation of a certain multivariate polynomial at the values $\gamma \cdot j_{i,\delta}([x, g])$, for $1 \leq i \leq m$ and $\gamma \in K_0/K_i$, each appearing with degree 1. The number of such values is

$$d_1 + d_1 d_2 + \cdots + d_1 \cdots d_m \leq m\,(\#\Sigma)\, d(\delta).$$

Therefore, by proposition 4.3.1, we have

$$\begin{aligned} h\big(\Psi_{\delta,m}(j_1, \dots, j_{n+1})\big) &\leq m\,(\#\Sigma)\, d(\delta) \log(2) \\ &\quad + m\,(\#\Sigma)\, d(\delta)\, C\big(\overline{h}(j_1, \dots, j_{n+1}) + \log l(\delta)\big) \\ &\leq C'\, d(\delta)\big(\overline{h}(j_1, \dots, j_{n+1}) + \log l(\delta)\big), \end{aligned}$$

where C and C' denote explicit constants independent of δ. In order to obtain $\Psi_{\delta,m}(j_1, \dots, j_n)$, we interpolate a polynomial of degree $e-1$ in j_{n+1}

where $\mathcal{J}$ is the set of interpolation points. By propositions 4.3.1 and 4.3.3, we have

$$h(j_{n+1}) \leq C\,\overline{h}(j_1, \ldots, j_n) \qquad \text{for every } j_{n+1} \in \mathcal{J},$$

where C is another constant independent on δ. The result follows by applying proposition 4.1.1 with $N = d + 1$; recall that $\log d(\delta) = O(\log l(\delta))$ by proposition 2.4.4. □

The proof of proposition 5.2.7 provides an explicit value of C if the constant from corollary 5.2.4 is known.

5.2.4 Heights of coefficients of modular equations

We are ready to prove height bounds on modular equations (the second part of theorem 5.0.1) using proposition 5.2.7 and the results on heights of fractions given in chapter 4. From now on, we add subscripts to constants: for instance $C_{5.2.3}$ denotes a constant *larger than* 1 such that proposition 5.2.3 holds with this value of C. Moreover, we denote by $C_{\log}$ a constant independent of δ such that $\log d(\delta) \leq C_{\log} \max\{1, \log l(\delta)\}$. We can take $C_{\log} = (\dim V)^2 + \log(C_{2.4.4})$, where V denotes the $\mathbb{Q}$-vector space defining the PEL datum.

Definition 5.2.8. We call an (n, N_1, N_2)-*evaluation tree* a rooted tree with depth n, arity N_1 at depths $0, \ldots, n-2$, and arity N_2 at depth $n-1$, such that every vertex but the root is labeled by an element of $\mathbb{Z}$ and the sons of every vertex are distinct.

Let T be an (n, N_1, N_2)-evaluation tree, and let $1 \leq k \leq n$. The k-th *evaluation set* $\mathcal{I}_k(T)$ of T is the set of points $(y_1, \ldots, y_k) \in \mathbb{Z}^k$ such that y_1 is a son of the root, and y_{i+1} is a son of y_i for every $1 \leq i \leq k-1$. We say that T is *bounded by* M if the absolute value of every vertex in T is bounded above by M. We say that T has *amplitude* (D_1, D_2) if for every vertex y of depth $0 \leq r \leq n-2$ (resp. depth $n-1$) in T, the sons of y lie in an integer interval of amplitude at most D_1 (resp. D_2); by definition, the amplitude of $[\![A, B]\!]$ is $B - A$.

Let T be an (n, N_1, N_2)-evaluation tree, let $a = (a_1, \ldots, a_n) \in \mathbb{Z}^n$, and let $M \geq 1$ be an integer. Let F be a coefficient of $\Psi_{\delta,m}$ for some $1 \leq m \leq n+1$, seen as a polynomial in the variables $J_{n+1}, Y_1, \ldots, Y_m$ of degree at most $e-1$ in J_{n+1}; hence $F \in L(J_1, \ldots, J_n)$. Write $F = P/Q$ in irreducible form where $P, Q \in L(J_1, \ldots, J_n)$, and let $d = \deg(F)$; assume that $d \geq 1$. We say that T, a and M are *valid evaluation data for* F if the following conditions are satisfied:

1. T and a are bounded by M.

2. We have $M \geq 2B\log^2(B+1)$, where

$$B = 4C_{5.1.9}^3 C_{5.2.7}\, d(\delta)^4 \max\{1, \log l(\delta)\}.$$

3. $N_1 = 2d$ and $N_2 \geq M$.
4. T has amplitude $(4d, 2M)$.
5. For every $(y_1, \ldots, y_n) \in \mathcal{I}_n(T)$, the point

$$(j_1, \ldots, j_n) = (y_1 y_n + a_1, \ldots, y_{n-1} y_n + a_{n-1}, y_n + a_n)$$

belongs to $\mathcal{V}_\delta$.

6. For every $(y_1, \ldots, y_{n-1}) \in \mathcal{I}_{n-1}(T)$, the polynomials P and Q evaluated at $(y_1 Y + a_1, \ldots, y_{n-1} Y + a_{n-1}, Y + a_n)$ are coprime in $L[Y]$.
7. $Q(a_1, \ldots, a_n) \neq 0$.

Lemma 5.2.9. *There exists a constant C, independent of δ, such that the following holds. Let F be a coefficient of $\Psi_{\delta,m}$ of degree $d \geq 1$. Then there exist valid evaluation data (T, a, M) for F such that*

$$C\, d(\delta)^4 \max\left\{1, \log^3(l(\delta))\right\} \leq M < C\, d(\delta)^4 \max\left\{1, \log^3(l(\delta))\right\} + 1$$

and $M \geq 4d[L : \mathbb{Q}]$. We can take

$$C = \max\{C_1, C_2, C_3\}$$

where

$$C_1 = 24C_{5.1.9}^3 C_{5.2.7}\big(4C_{\log} + \log(24C_{5.1.9}^3 C_{5.2.7}) + 1\big),$$
$$C_2 = 14C_{5.1.9}^2 + 5C_{5.2.6}, \quad \textit{and} \quad C_3 = 4C_{5.1.9}[L : \mathbb{Q}].$$

Proof. Let M be as above. We check that condition 2 in the definition of a valid evaluation tree is satisfied because $C \geq C_1$.

We start by constructing the vector a. Note that $M \geq 2d+1$. Since Q is nonzero, and has degree at most d in Y_1, we can find $a_1 \in \mathbb{Z}$ such that $|a_1| \leq M$ and the polynomial $Q(a_1, Y_2, \ldots, Y_n)$ is nonzero. Iterating, we find $a = (a_1, \ldots, a_n)$ bounded by M such that $Q(a_1, \ldots, a_n) \neq 0$.

We now build the evaluation tree T down from the root. Let P_δ be an equation for the complement of $\mathcal{V}_\delta$ as in lemma 5.2.6, and define

$$R_\delta = P_\delta(Y_1 Y_n + a_1, \ldots, Y_{n-1} Y_n + a_{n-1}, Y_n + a_n)$$

which is a nonzero polynomial of degree at most $2C_{5.2.6}\, d(\delta)$. Let R be the resultant with respect to Y_n of the two polynomials

$$P(Y_1Y_n + a_1, \dots, Y_{n-1}Y_n + a_{n-1}, Y_n + a_n)$$

and

$$Q(Y_1Y_n + a_1, \dots, Y_{n-1}Y_n + a_{n-1}, Y_n + a_n).$$

The polynomial R is nonzero and has degree at most $4d^2$.

We want to choose $2d$ values of y_1, lying in an interval with amplitude at most $4d$, such that neither R_δ nor R vanishes when evaluated at $Y_1 = y_1$; this nonvanishing condition excludes at most $4d^2 + 2C_{5.2.6}\, d(\delta)$ possible values of y_1. At least one of the intervals of the form $[\![5kd, (5k+4)d]\!]$ for $0 \leq k \leq 2d + C_{5.2.6}\, d(\delta)/d$ contains at least $2d$ valid choices of y_1. Then $|y_1|$ is always bounded above by $5(2d^2 + C_{5.2.6}\, d(\delta)) + 4d \leq M$, because $C \geq C_2$.

We iterate this procedure to construct T up to depth $n-1$ with the right arity, bound and amplitude, in such a way that the evaluations of R_δ and R are nonzero at every $(y_1, \dots, y_{n-1}) \in \mathcal{I}_{n-1}(T)$.

We conclude by constructing n-th level of T. Let $(y_1, \dots, y_{n-1}) \in \mathcal{I}_{n-1}(T)$. Then, as before, at most $4d^2 + 2C_{5.2.6}\, d(\delta) \leq M$ values for y_n are forbidden as they make either R_δ or R vanish. This leaves at least M available values for y_n in $[\![-M, M]\!]$.

For every tuple $(y_1, \dots, y_n) \in \mathcal{I}_n(T)$, the fact that the polynomials R_δ and R do not vanish at the point $(y_1, \dots, y_n)$ guarantees conditions 5 and 6 of definition 5.2.8 respectively. Finally, the inequality $C \geq C_3$ ensures that $M \geq 4d[L : \mathbb{Q}]$. □

Theorem 5.2.10. *Let H_δ be an absolutely irreducible Hecke correspondence on $\mathcal{S} \times \mathcal{T}$ defined by an element $\delta \in G(\mathbb{A}_f)$, and let $d(\delta)$ be the degree of H_δ. Let $F \in L(J_1, \dots, J_n)$ be a coefficient of one of the modular equations $\Psi_{\delta,m}$ for $1 \leq m \leq n+1$. Then the height of F is bounded above by $C\, d(\delta)$, where C is a constant independent of δ; more precisely we can take*

$$\begin{aligned} C = 2^{n-1}\big(&2C_{5.2.7}(1 + C'') \\ &+ 2C_{4.5.11}C_{5.1.9}\big(\log(4C_{5.1.9}C_{5.2.7}) + 2C_{\log} + 1 + C''\big) \\ &+ 4C_{5.1.9}(\log(C_{5.1.9}) + C_{\log}) \\ &+ 2C_{5.1.9}(\log(2) + C'') + 2\log(2C_{5.1.9}) + 2\big), \end{aligned}$$

where $C'' = 3 + \log(2C_{5.2.9}) + 4C_{\log}$.

Proof. By lemma 5.2.9, we can find valid evaluation data (T, a, M) for F such that the inequality $M \leq C_{5.2.9}\, d(\delta)^4 \max\{1, \log^3 l(\delta)\} + 1$ holds. After scaling P and Q by an element of $L^\times$, we can assume that $Q(a_1, \dots, a_n) = 1$.

Let $(y_1, \dots, y_{n-1}) \in \mathcal{I}_{n-1}(T)$, and write

$$\widetilde{F}(Y) = F(y_1 Y + a_1, \dots y_{n-1} Y + a_{n-1}, Y + a_n).$$

For every son y_n of y_{n-1} in T, we have

$$h\big(y_1 y_n + a_1, \dots, y_{n-1} y_n + a_n\big) \leq \log\big((M+1)M\big) \leq 2\log(M+1).$$

Therefore, by proposition 5.2.7,

$$\begin{aligned} h(\widetilde{F}(y_n)) &\leq C_{5.2.7}\, d(\delta)\big(2\log(M+1) + \log l(\delta)\big) \\ &\leq 2C_{5.2.7}\, d(\delta)\big(\log(M+1) + \max\{1, \log l(\delta)\}\big). \end{aligned}$$

Denote this last quantity by H. We have $H \geq 4$ and $H \geq \log(2M)$. Moreover, in the notation of definition 5.2.8, the inequality $M \geq 2B\log^2(B+1)$ ensures that

$$\frac{M}{\log(M+1)} \geq B \geq d^3\big(4C_{5.2.7}\, d(\delta)\max\{1, \log l(\delta)\}\big).$$

Therefore $M \geq d^3 H$.

We are in position to apply proposition 4.5.11 for the univariate rational fraction $\widetilde{F}$ on the interval $[\![-M, M]\!]$, with $\eta = 2$, using the sons of $(y_1, \dots, y_{n-1})$ in T as evaluation points. We obtain

$$\begin{aligned} h(\widetilde{F}) &\leq H + 2C_{4.5.11} d \log(2dH) + d\log(2M) + \log(d+1) \\ &\leq C'\, d(\delta) \max\left\{1, \log l(\delta)\right\}, \end{aligned}$$

where C' is a constant independent of δ. In order to obtain an explicit expression for C', we note that

$$\log(M+1) \leq C'' \max\left\{1, \log l(\delta)\right\}.$$

We check that we can take

$$\begin{aligned} C' = {} & 2C_{5.2.7}(1 + C'') \\ & + 2C_{4.5.11}C_{5.1.9}\big(\log(4C_{5.1.9}C_{5.2.7}) + 2C_{\log} + 1 + C''\big) \\ & + C_{5.1.9}(\log(2) + C'') + \log(2C_{5.1.9}) + 1. \end{aligned}$$

In the second part of the proof, we relate the height of $\widetilde{F}$ with the height of F. The quotient

$$\frac{P(y_1 Y + a_1, \dots, y_{n-1} Y + a_{n-1}, Y + a_n)}{Q(y_1 Y + a_1, \dots, y_{n-1} Y + a_{n-1}, Y + a_n)}$$

is a way to write $\widetilde{F}$ in irreducible form in $L(Y)$, and has a coefficient equal to 1. Therefore $h(\widetilde{F})$ is the affine height of the coefficients appearing in the quotient. Hence

$$h\big(P(y_1Y_n + a_1, \dots, y_{n-1}Y_n + a_{n-1}, Y_n + a_n)\big) \leq C'\, d(\delta) \max\{1, \log l(\delta)\}$$

for every $(y_1, \dots, y_{n-1}) \in \mathcal{I}_{n-1}(P)$, and the same inequality holds for Q. Since $N_1 = 2d$, we can interpolate successively the variables $y_{n-1}, \dots, y_1$, using proposition 4.4.3 with $2d$ evaluation points at each vertex of the tree T. Finally we obtain

$$\begin{aligned} h(F) &\leq 2^{n-1}\big(C'\, d(\delta) \max\{1, \log l(\delta)\} + 4d \log(4d) + d \log(2M) + \log(d+1)\big) \\ &\leq 2^{n-1}\big(C' + 4C_{5.1.9}(\log(C_{5.1.9}) + C_{\log}) + C_{5.1.9}(\log(2) + C'') \\ &\qquad + \log(2C_{5.1.9}) + 1\big)\, d(\delta) \max\{1, \log l(\delta)\}, \end{aligned}$$

as claimed. □

5.2.5 Explicit height bounds in dimension 2

To conclude this chapter, we derive explicit height bounds for modular equations of Siegel type and level ℓ for abelian surfaces in the Igusa invariants j_1, j_2, j_3. Our first aim is to provide an explicit value for the constant in corollary 5.2.4, using theta constants of level 4 as an intermediate step. To relate theta heights and j-heights in this setting, we use Mestre's algorithm and Thomae's formulæ instead of writing out polynomials $N_{J,i}$, $D_{J,i}$, $N_{\Theta,i}$, $D_{\Theta,i}$ as in the proof of proposition 5.2.3.

Proposition 5.2.11. *Let A be a principally polarized abelian surface defined over $\overline{\mathbb{Q}}$ whose Igusa invariants j_1, j_2, j_3 are well defined, and assume that $j_3(A) \neq 0$. Then we have*

$$h_j(A) \leq 40\, h_{\Theta,4}(A) + 12 \quad \text{and} \quad h_{\Theta,4}(A) \leq 200\, h_j(A) + 1000.$$

Proof. Recall the expression of the Igusa invariants in terms of the Siegel modular forms h_4, h_6, h_{10}, and h_{12} (§2.3.2):

$$j_1 = \frac{h_4 h_6}{h_{10}}, \quad j_2 = \frac{h_4^2 h_{12}}{h_{10}^2}, \quad j_3 = \frac{h_4^5}{h_{10}^2}.$$

These modular forms have a polynomial expression in terms of the ten even theta constants of level 4, given in definition 2.3.4. The total degrees of the polynomials giving h_4, h_6, h_{10} and h_{12} are $8, 12, 20$ and 24 respectively; they contain respectively $10, 60, 1$ and 15 monomials, and their height is

zero. Up to scaling, we may assume that the first theta constant θ_0 takes the value 1. Then, by proposition 4.3.1, we have

$$h(h_4^5,\, h_4h_6h_{10},\, h_4^2h_{12},\, h_{10}^2) \leq 5\log(10) + 40\, h_{\Theta,4}(A),$$

hence the first inequality

$$h_j(A) \leq 40\, h_{\Theta,4}(A) + 12.$$

For the second inequality, we follow Mestre's algorithm [Mes91]. Starting from $j_1(A), j_2(A)$ and $j_3(A)$, Mestre's algorithm constructs a hyperelliptic curve $y^2 = f(x)$ whose Jacobian is isomorphic to A over $\overline{\mathbb{Q}}$. Choosing $I_{10} = 1$ in eq. (3.5), we see that $j_1(A), j_2(A)$ and $j_3(A)$ are realized by covariants I_2, I_4, I_6, and I_{10} in $\overline{\mathbb{Q}}$ such that

$$h(I_2, I_4, I_6, I_{10}) \leq h_j(A).$$

The roots of f are the intersection points of a conic and a cubic in $\mathbb{P}^2$ whose equations are given explicitly in terms of I_2, I_4, I_6, and I_{10}. In order to obtain the equation $\sum_{i,j=1}^{3} c_{ij} z_i z_j = 0$ of the conic, we start from Mestre's equation $\sum_{i,j=1}^{3} A_{ij} x_i x_j = 0$ and substitute the expressions of A, B, C, and D in terms of I_2, I_4, I_6, and I_{10}. Then we multiply by $2^{11}3^{13}5^{14}$ and make the substitutions

$$z_1 = 202500x_1, \quad z_2 = 225x_2, \quad z_3 = x_3.$$

Then, each coefficient c_{ij} has an expression as a multivariate polynomial in I_2, I_4, and I_6 (recall that $I_{10} = 1$) of total degree at most 7; its coefficients are integers whose absolute values are bounded by $324 \cdot 10^6$. By proposition 4.3.1, we have

$$h\big((c_{ij})_{1\leq i,j\leq 3}\big) \leq 7(h_j(A) + \log(3)) + 19.6 + 3\log(8) \leq 7h_j(A) + 33.6.$$

If we restrict to c_{11}, c_{12}, and c_{22}, then we obtain a smaller upper bound, since the total degree and the height of coefficients are at most 5 and 18.3 respectively. Similarly, the cubic equation, denoted by $\sum_{1\leq i\leq j\leq k\leq 3} c_{ijk} z_i z_j z_k = 0$, has total degree at most 11 in I_2, I_4, and I_6, and has integer coefficients whose heights are at most 33.5.

In order to find the hyperelliptic curve equation f, we parametrize the conic. Let us show that it contains a point P_0 defined over $\overline{\mathbb{Q}}$ such that $h(P_0) \leq 5h_j(A) + 29.9$. We can assume that $c_{11} \neq 0$; otherwise we take $P_0 = (1:0:0)$. Let α be a root of the monic polynomial

$$\alpha^2 + \frac{c_{12}}{c_{11}}\alpha + \frac{c_{22}}{c_{11}} = 0.$$

The point $P_0 = (\alpha : 1 : 0)$ belongs to the conic, and by proposition 4.3.3,

$$\begin{aligned} h(P_0) = h(\alpha) &\leq h(c_{11}, c_{12}, c_{22}) + \log(2) \\ &\leq 5(h_j(A) + \log(3)) + 18.3 + 3\log(6) + \log(2) \\ &\leq 5h_j(A) + 29.9. \end{aligned}$$

We parametrize the conic using P_0 as a base point; for simplicity, we continue to assume that $c_{11} \neq 0$. For $(u : v) \in \mathbb{P}^1(\overline{\mathbb{Q}})$, the point $(z_1 : z_2 : z_3)$ defined by

$$\begin{aligned} z_1 &= \alpha(c_{11}u^2 + c_{13}uv + c_{33}v^2) - u((2c_{11}\alpha + c_{12})u + (c_{13}\alpha + c_{23})v), \\ z_2 &= c_{11}u^2 + c_{13}uv + c_{33}v^2, \\ z_3 &= -v((2c_{11}\alpha + c_{12})u + (c_{13}\alpha + c_{23})v) \end{aligned}$$

runs through the conic. Substituting these expressions in the cubic equation gives the curve equation f. The polynomials we obtain have total degrees at most 29 in I_2, I_4, and I_6; they have degree at most 3 in α; and their coefficients are integers whose heights are bounded above by 86.9. Therefore, by proposition 4.3.1 (separating I_2, I_4, I_6 from α), we have

$$\begin{aligned} h(f) &\leq 29(h_j(A) + \log(3)) + 86.9 + 3(5h_j(A) + 29.9) + 3\log(30) + \log(4) \\ &\leq 44h_j(A) + 220.1. \end{aligned}$$

Making f monic does not change its height.

Thomae's formulæ [Mum84, IIIa.8.1] give an expression of the theta constants of level 4 of A in terms of roots of f: if θ is one of these theta constants, then θ^4 is a product of 18 differences of roots of f (up to a common multiplicative factor). Therefore, by proposition 4.3.3, we obtain

$$h_{\Theta,4}(A, L) \leq \tfrac{1}{4} \cdot 18(h(f) + \log(4)) \leq 198h_j(A) + 997. \qquad \square$$

As a consequence, we obtain an explicit analogue of corollary 5.2.4 in the case of isogenies between principally polarized abelian surfaces.

Proposition 5.2.12. *Let A and A' be principally polarized abelian surfaces over $\overline{\mathbb{Q}}$ where the Igusa invariants j_1, j_2, j_3 are well defined, and assume that $j_3(A)j_3(A') \neq 0$. Let $\ell \geq 1$ be an integer. If A and A' are linked by an isogeny of degree ℓ, then we have*

$$\overline{h}_j(A') \leq 8000\,\overline{h}_j(A) + 1.08 \cdot 10^{11}\log(\overline{h}_j(A)) + 1.67 \cdot 10^{12} + 20\log\ell.$$

Proof. By propositions 5.2.1 and 5.2.11 and theorem 5.2.2 (noting that $C(2,4) \leq 1.35 \cdot 10^9$), we have

$$\begin{aligned}
\overline{h}_{\Theta,4}(A) &\leq 200\,\overline{h}_j(A) + 1000,\\
\tfrac{1}{2}\,\overline{h}_F(A) &\leq \overline{h}_{\Theta,4}(A) + C(2,4)\log(\overline{h}_{\Theta,4}(A)+2)\\
&\leq 200\,\overline{h}_j(A) + C(2,4)\log(1202) + C(2,4)\log(\overline{h}_j(A)),\\
\tfrac{1}{2}\,\overline{h}_F(A') &\leq \tfrac{1}{2}\,\overline{h}_F(A) + \tfrac{1}{4}\log\ell,\\
\overline{h}_{\Theta,4}(A') &\leq \tfrac{1}{2}\,\overline{h}_F(A') + C(2,4)\log(\overline{h}_F(A')+2)\\
&\leq 200\,\overline{h}_j(A) + C(2,4)\log(1202) + 2C(2,4)\log(\overline{h}_j(A)) + \tfrac{1}{4}\log\ell\\
&\qquad + C(2,4)\log\big(402 + 2C(2,4)\log(1202) + C(2,4) + \tfrac{1}{2}\log\ell\big),\\
&\leq 200\,\overline{h}_j(A) + 2C(2,4)\log(\overline{h}_j(A)) + 4.17\cdot 10^{10} + \tfrac{1}{2}\log\ell, \quad \text{and}\\
\overline{h}_j(A') &\leq 40\,\overline{h}_{\Theta,4}(A) + 12\\
&\leq 8000\,\overline{h}_j(A) + 80C(2,4)\log\overline{h}_j(A) + 1.67\cdot 10^{12} + 20\log\ell. \qquad \square
\end{aligned}$$

In lemma 5.2.6, we take $\lambda = I_4$ and $\lambda' = I_4 I_{10}$. We have

$$\mathrm{wt}(\lambda^\delta) = 14\,d(\delta) + 4,$$

which is greater than 16, the minimum weight for which Siegel modular forms define a projective embedding of $\mathcal{S}$. Hence ξ can be chosen to be a modular form of weight $\mathrm{wt}(\lambda^\delta)$. The fraction R'' has degree at most $\frac{7}{3}(d(\delta)+1)$ by lemma 5.1.14; this is also an upper bound on $\deg(P_\delta)$.

We also mimic the proof of proposition 5.2.7 in the Siegel case. Let $[x,g]$ be a point of $\mathcal{S}$ with Igusa invariants $(j_1,j_2,j_3) \in \mathcal{V}_\delta$. For each $1 \leq m \leq 3$, by remark 2.4.10, the polynomial $\Psi_{\delta,m}(j_1,j_2,j_3)$ is the evaluation of a multivariate polynomial in $2\,d(\delta)$ variables. Moreover, the Hecke correspondence describes isogenies of degree ℓ^2. By proposition 5.2.12, $h\big(\Psi_{\delta,m}(j_1,j_2,j_3)\big)$ is bounded above by

$$2\,d(\delta)\big(8000\,\overline{h}(j_1,j_2,j_3) + 1.08\cdot 10^{11}\log(\overline{h}_j(A)) + 1.67\cdot 10^{12} + 40\log\ell\big). \quad (5.8)$$

Therefore, we can take

$$C_{5.2.7} = 3.35\cdot 10^{12}.$$

Moreover, we have $d(\delta) = \ell^3+\ell^2+\ell+1$ and $l(\delta) = \ell^2$. Hence we can take

$$C_{\log} = \frac{3}{2} + \log(2) \leq 2.2.$$

We also take

$$\begin{aligned} C_{4.5.11} &= 960 && \text{because } L = \mathbb{Q}, \\ C_{5.1.9} &= \frac{10}{3} && \text{by proposition 5.1.15, and} \\ C_{5.2.6} &= 15 && \text{since } d(\delta) \geq 15. \end{aligned}$$

In lemma 5.2.9, we can take

$$C_{5.2.9} = 1.36 \cdot 10^{17}$$

and in theorem 5.2.10, we can take

$$C_{5.2.10} = 1.42 \cdot 10^{15}.$$

Since $d(\delta) \leq 2\ell^3$ and $\max\{1, \log \ell(\delta)\} \leq 2\log(\ell)$, we obtain the following result.

Theorem 5.2.13. *Let $\ell \geq 1$ be a prime number, and let $F \in \mathbb{Q}(J_1, J_2, J_3)$ be a coefficient of one of the Siegel modular equations of level ℓ in Igusa invariants. Then we have*

$$h(F) \leq 5.68 \cdot 10^{15} \ell^3 \log(\ell).$$

In order to obtain tighter height bounds on Siegel modular equations, we could repeat the computations of §5.2.4 using an expression of the form (5.8) for the height of evaluated modular equations, instead of the simpler formula used in proposition 5.2.7. However we cannot hope to obtain a constant in theorem 5.2.13 that is much lower than the value of $C(2,4) \simeq 1.35 \cdot 10^9$ using our methods. Experimentally, we observe that the inequalities $h(F) \leq 48.7\,\ell^3 \log(\ell)$ and $h(F) \leq 43.6\,\ell^3 \log(\ell)$ hold for $\ell = 2$ and $\ell = 3$ respectively.

We could also give an analogue of theorem 5.2.13 in the case of modular equations of Hilbert type for $\mathbb{Q}(\sqrt{5})$ in Gundlach invariants. To replace proposition 5.2.11, we would use the relations between Gundlach and Igusa invariants (proposition 2.3.16) and the explicit curve equation given by proposition 2.3.17. We leave the precise calculations for future work.

Chapter 6

Evaluating modular equations for abelian surfaces

In this chapter, we are interested in the complexity of manipulating modular equations of Hilbert or Siegel type for abelian surfaces. The corresponding preprint is [Kie20b]; part of this chapter also covers the article [Kie21], to appear in *Publications Mathématiques de Besançon*.

By the main result of chapter 5, theorem 5.0.1, we know that for every prime ℓ, the Siegel modular equations of level ℓ, denoted by

$$\Psi_{\ell,1}, \Psi_{\ell,2}, \Psi_{\ell,3} \in \mathbb{Q}(J_1, J_2, J_3)[Y],$$

have degree $O(\ell^3)$ in each variable J_1, J_2, J_3, Y, and that each of their coefficients has height $O(\ell^3 \log \ell)$. Therefore the cost of storing these modular equations in full is $O(\ell^{15} \log \ell)$. Similarly, if F is a fixed real quadratic field and if $\beta \in \mathbb{Z}_F$ is a totally positive prime of norm ℓ, then the cost of storing Hilbert modular equations of level β in Igusa invariants, or Gundlach invariants if $F = \mathbb{Q}(\sqrt{5})$, is $O_F(\ell^4 \log \ell)$.

In analogy with elliptic modular polynomials, we expect that both upper bounds are asymptotically accurate. Experimentally, Siegel modular equations are very difficult to compute: even using optimized invariants, the computations have never been carried out for $\ell > 7$ to our knowledge [Mil15]. Indeed, we expect that the Siegel modular equations of level 11 are roughly 10000 times as large as their level 7 analogues. In the Hilbert case for the quadratic field $\mathbb{Q}(\sqrt{5})$, the size of a compressed `tar` file containing the modular equations is close to $30\ell^4 \log \ell$ bytes for $\ell \in \{11, 19, 29, 31, 41\}$, after which the database [Mil] ends.

Given these bounds, in the context of counting points on principally polarized abelian surfaces over finite fields, using Elkies's method with precomputed modular equations does not bring any asymptotic improvement

over Schoof's method described in [GKS11, GS12]. This unfortunate state of affairs has led to the belief that modular equations for abelian surfaces should be avoided in practice. In [CE15], for instance, one can read: *"it is unlikely that modular equations will be of any use to accelerate the computation of torsion points for higher genera, since they are all far too big"*.

However, in most contexts, writing out the modular equations in full is not truly needed. In order to detect isogenies between abelian surfaces, and also to compute these isogenies explicitly using the algorithms of chapter 3, we only need *evaluations* of modular equations and their derivatives: in the Siegel case, this would be the polynomials

$$\Psi_{\ell,m}(j_1, j_2, j_3, Y) \in \mathbb{K}[Y] \quad \text{and} \quad \partial_{J_k}\Psi_{\ell,m}(j_1, j_2, j_3, Y) \in \mathbb{K}[Y]$$

for $1 \leq k, m \leq 3$, where j_1, j_2, j_3 are fixed elements of a certain field $\mathbb{K}$.

These evaluations can be a lot smaller, because they erase the increase of the number of variables coming from the dimension of the moduli space. For instance, if $j_1, j_2, j_3 \in \mathbb{Q}$ have height $O(1)$, then the size of these polynomials is $O(\ell^6 \log \ell)$ by theorem 5.0.1; if $\beta \in \mathbb{Z}_F$ is a totally positive prime of norm ℓ, then the size of evaluations of Hilbert modular equations of level β in Igusa invariants at (j_1, j_2, j_3) is only $O_F(\ell^2 \log \ell)$, the same aymptotic estimate (up to constants depending on F) as in the case of elliptic modular polynomials.

The goal of this chapter is to present an algorithm that directly computes evaluations of modular equations of Siegel or Hilbert type for abelian surfaces, given values of Igusa or Gundlach invariants in a number field. This evaluation algorithm extends to the case of finite fields via lifts. As a special case, we obtain the following result.

Theorem 6.0.1 (Under hypothesis 6.2.15 below)**.** *There exists an algorithm which, given prime numbers p and ℓ, and given $(j_1, j_2, j_3) \in \mathbb{F}_p^3$ where the denominators of the Siegel modular equations of level ℓ do not vanish and such that $j_3 \neq 0$, computes $\Psi_{\ell,m}(j_1, j_2, j_3, Y)$ and $\partial_{J_k}\Psi_{\ell,m}(j_1, j_2, j_3, Y)$ in $\mathbb{F}_p[Y]$ for $1 \leq k, m \leq 3$ within $\widetilde{O}(\ell^3 \log^2 p + \ell^6 \log p)$ binary operations.*

A similar results holds for Hilbert modular equations for $F = \mathbb{Q}(\sqrt{5})$ in Gundlach invariants, with a complexity of only $\widetilde{O}_F(\ell \log^2 p + \ell^2 \log p)$ binary operations. In both cases, we save a factor of $\log p$ when j_1, j_2 and j_3 are quotients of small integers. In the case of Hilbert modular equations for a general F, our algorithm is unfortunately only heuristic.

Let us give a high-level view of the algorithm to evaluate Siegel modular equations, and describe the heuristic assumption that it uses. Let L be a number field, and choose Igusa invariants $(j_1, j_2, j_3) \in L^3$ where the

denominator of Siegel modular equations of level ℓ, denoted by D_ℓ in §5.1.4, does not vanish. Assume for simplicity that $j_1, j_2, j_3 \in \mathbb{Z}_L$. We use complex approximations: for each complex embedding μ of L, we compute a period matrix $\tau \in \mathbb{H}_2$ whose Igusa invariants are $(\mu(j_1), \mu(j_2), \mu(j_3))$. Then, we compute approximations of the numerator and denominator of Siegel modular equations at τ using analytic formulæ (definition 2.4.6 and eq. (5.2)). Finally, we recognize their coefficients as algebraic numbers, more precisely algebraic *integers* by proposition 5.1.11. During the algorithm, we keep track of precision losses in order to obtain a provably correct result.

An essential subroutine in the evaluation algorithm is to convert back and forth between matrices $\tau \in \mathbb{H}_2$ and the values of theta constants at τ, defined in eq. (2.10). This can be done in quasi-linear time in the required precision, at least for certain ranges of inputs, using the arithmetic-geometric mean (AGM) to compute τ from its theta constants [Bor88, BM88, Jar08], and an algorithm of Dupont [Dup06, Chap. 10] in the other direction. Dupont's algorithm combines the AGM with Newton iterations, and works very well in practice, but its proper convergence remains heuristic. Hypothesis 6.2.15, on which our main theorem relies, states that Dupont's algorithm converges uniformly in a certain compact subset of $\mathbb{H}_2$.

This chapter is organized as follows. First, we describe our computational model and analyze precision losses in "basic" operations on complex numbers and polynomials, such as reconstructing polynomials from their roots, Lagrange interpolation, and the reconstruction of integers in number fields from their values in complex embeddings (§6.1). Then we review the AGM method for theta constants in genus 2 and Dupont's algorithm, and make some new contributions (§6.2):

- In order to bound precision losses uniformly in τ, we prove the technical result that the relevant AGM sequences are given by good sign choices [Kie21], a fact that was left as a conjecture in [Dup06, Conj. 9.1] and [Eng09a, Conj. 9];
- Using Dupont's algorithm, we explain how to compute theta constants on a fundamental domain $\mathcal{F}_2 \subset \mathbb{H}_2$ with uniform quasi-linear cost, in a similar way to [Dup11, §6.3] in the genus 1 case;
- We also bound the precision losses when reducing a general $\tau \in \mathbb{H}_2$ to $\mathcal{F}_2$, using techniques inspired from [Str14, §6].

Finally, we present the evaluation algorithm in §6.3, focusing on the case of Hilbert modular equations, which is more intricate due to the presence of the Hilbert embedding.

6.1 Precision losses in polynomial operations

In all algorithms manipulating complex numbers, we use interval arithmetic. Given $z \in \mathbb{C}$ and $N \geq 0$, we define an *approximation of z to precision N* to be a complex ball centered at some $z' \in \mathbb{C}$ of radius 2^{-N} containing z. An *approximation* of a polynomial to precision N is by definition an approximation to precision N coefficient per coefficient.

Approximations of complex numbers centered at dyadic points can be stored in a computer. This model makes more mathematical sense than the usual floating-point or fixed-point approximations; using it, we design algorithms with meaningful input and provably correct output.

Let $M \geq 1$. We say that *the precision loss in an algorithm $\mathcal{A}$ is M bits* if the following property holds: for all $N \geq M$, if the input of $\mathcal{A}$ is given as approximations to precision N, then the output of $\mathcal{A}$ is an approximation of its "theoretical" output to precision $N - M$. For instance, precision losses in elementary operations (additions, multiplications, etc.) can be bounded above in terms of the size of the operands (see §6.1.1). Besides these theoretical upper bounds, precision losses can also be computed on the fly in a precise way. This is done for instance in the Arb library [Joh17], which is therefore the library of choice to implement our algorithms. If we run out of precision during the computation, we can simply double the precision and restart. Therefore, in the theoretical analysis, it is enough to bound the precision losses in the O notation.

We let $\mathcal{M}(N)$ be a quasi-linear, superlinear function of N such that two N-bit integers can be multiplied in $\mathcal{M}(N)$ binary operations. We write log (resp. $\log_2$) for the natural logarithm (resp. logarithm in base 2), and for $x \in \mathbb{R}$, we define

$$\log^+ x = \log \max\{1, x\} \quad \text{and} \quad \log_2^+ x = \log_2 \max\{1, x\}.$$

We denote the absolute value of the largest coefficient in a polynomial P by $|P|$; we also use this notation for vectors and matrices.

6.1.1 Elementary operations

To summarize, additions can be done in linear time with a precision loss of $O(1)$, and multiplications, inversions, and square roots can be done in quasi-linear time with a precision loss given by the size of the input. We state these standard facts without proof.

Proposition 6.1.1. *Let $z \in \mathbb{C}^\times$ and $N \geq -\log_2 |z| + 1$.*

1. *Given an approximation of z to precision N, the inverse $1/z$ can be computed within $O\big(\mathcal{M}(N + \log^+|z|)\big)$ binary operations, with a precision loss of $-2\log_2|z| + O(1)$ bits.*
2. *Given an approximation of z to precision N, an approximation of a square root of z can be computed within $O\big(\mathcal{M}(N + \log^+|z|)\big)$ binary operations, with a precision loss of $-\frac{1}{2}\log_2|z| + O(1)$ bits.*

In proposition 6.1.1, the assumption on N ensures that the ball approximating z does not contain 0.

Precision losses in additions and multiplications of complex numbers are a special case of the following result for univariate polynomials.

Proposition 6.1.2. *Let $P_1, P_2 \in \mathbb{C}[X]$, and $N, N_1, N_2 \geq 1$. Assume that P_1, P_2 and their approximations have degree at most d.*

1. *Given approximations of P_1, P_2 to precision N, the sum $P_1 + P_2$ can be computed within*

$$O\big((d+1)\,(N + \log\max\{1, |P_1|, |P_2|\})\big)$$

binary operations, with a precision loss of $O(1)$ bits.

2. *Given approximations of P_i to precision N_i for $i = 1, 2$, the product P_1P_2 can be computed within*

$$O\big(\mathcal{M}\big((d+1)\max\{N_1 + \log|P_1|, N_2 + \log|P_2|\}\big)\big)$$

binary operations, to precision

$$\min\{N_1 - \log_2^+|P_2|, N_2 - \log_2^+|P_1|\} - \log_2(1+d) - O(1).$$

6.1.2 Reconstruction from roots and interpolation

We start with a technical lemma that we will use several times, when we construct polynomials as products of linear factors.

Lemma 6.1.3. *There exists an algorithm such that the following holds. Let $d \geq 1$, $B \geq 1$, $C \geq 1$, and let x_i, y_i, z_i for $1 \leq i \leq d$ be complex numbers such that*

$$\log^+|x_i| \leq B, \quad \log^+|y_i| \leq B, \quad \log^+|z_i| \leq C, \quad \text{for all } i.$$

Let $N \geq 1$. Then, given approximations of these complex numbers to precision N, the algorithm computes the polynomials

$$P = \prod_{i=1}^{d}(x_iX + y_i) \qquad Q = \sum_{i=1}^{d} z_i \prod_{j\neq i}(x_jX + y_j)$$

within $O\big(\mathcal{M}\big(d(N+C+dB)\big)\log d\big)$ *binary operations, with a precision loss of* $O(C+dB)$ *bits.*

Proof. We use product trees [BCG+17, §I.5.4]. For each $0 \leq m \leq \lceil \log_2(d) \rceil$, the m-th level of the product tree to compute P consists of $2^{\lceil \log_2(d) \rceil - m}$ products of (at most) 2^m factors of the form $x_i X + y_i$. Hence, for every polynomial R appearing at the m-th level, we have

$$\deg(R) \leq 2^m \quad \text{and} \quad \log^+ |R| = O(2^m B).$$

Level 0 is given as input. In order to compute level $m+1$ from level m, we compute one product per vertex, for a total cost of $O\big(\mathcal{M}\big(d(N+dB)\big)\big)$ binary operations; the precision loss is $O(2^m B)$ bits by proposition 6.1.2. Therefore the total precision loss when computing P is $O(dB)$ bits. The number of levels is $O(\log d)$, so the total cost is $O\big(\mathcal{M}\big(d(N+dB)\big)\log d\big)$ binary operations.

The computations are similar for the polynomial Q, with a different product tree. Each vertex at level $m+1$ is a polynomial of the form $N_1 P_2 + N_2 P_1$ where P_i is a vertex of the product tree for P satisfying

$$\deg(P_i) \leq 2^m \quad \text{and} \quad \log^+ |P_i| = O(2^m B),$$

and the polynomials N_i come from the m-th level, and satisfy

$$\deg(N_i) \leq 2^m - 1 \quad \text{and} \quad \log^+ |N_i| = O(C + 2^m B).$$

By induction, the m-th level can be computed to precision $N - O(C + 2^m B)$ using a total of $O\big(\mathcal{M}\big(d(N+C+dB)\big)\big)$ binary operations. □

We apply lemma 6.1.3 to Lagrange interpolation.

Proposition 6.1.4. *There exists an algorithm such that the following holds. Let* $P \in \mathbb{Z}[X]$ *be an irreducible polynomial of degree* $d \geq 1$*, let* $(\alpha_i)_{1 \leq i \leq d}$ *be the roots of* P*, and let* $(t_i)_{1 \leq i \leq d}$ *be complex numbers. Let* $M, C \geq 1$ *such that*

$$\log^+ |P| \leq M, \quad \textit{and} \quad \log^+ |t_i| \leq C \quad \textit{for every } i.$$

Let $N \geq 1$*. Then, given* P *and approximations of the* α_i, t_i*, and* $1/P'(\alpha_i)$ *to precision* N*, the algorithm computes the polynomial* Q *of degree at most* $d-1$ *interpolating the points* (α_i, t_i) *within*

$$O\big(\mathcal{M}\big(d(N+C+dM+d\log d)\big)\log d\big)$$

binary operations. The precision loss is $O(C + dM + d\log d)$ *bits.*

Proof. We write

$$Q = \sum_{i=1}^{d} \frac{t_i}{P'(\alpha_i)} \prod_{j \neq i} (X - \alpha_j).$$

We have $\log^+ |P'| \leq M + \log d$. The discriminant $\mathrm{Disc}(P)$ of P is the resultant of P and P'. Hence we can write

$$UP + VP' = \mathrm{Disc}(P)$$

with $U, V \in \mathbb{Z}[X]$; the coefficients of U, V have expressions as determinants of size $O(d)$ involving the coefficients of P and P', so by Hadamard's lemma, we have in particular

$$\log^+ |V| = O(dM + d \log d).$$

By the proof of proposition 4.3.3, we have $\log^+ |\alpha_i| \leq M + \log(2)$ for every i, hence

$$\log^+ \left| \frac{1}{P'(\alpha_i)} \right| = \log^+ \left| \frac{V(\alpha_i)}{\mathrm{Disc}(P)} \right| = O(dM + d \log d).$$

Therefore the precision loss taken when computing the d complex numbers $z_i = t_i / P'(\alpha_i)$ is $O(C + dM + d \log d)$ bits; the total cost to compute the z_i is

$$O\big(d\mathcal{M}(N + C + dM + d \log d)\big)$$

binary operations. We conclude using lemma 6.1.3. □

6.1.3 Recognizing integers in number fields

We conclude this section with estimates on the necessary precision to recognize integers in a number field L.

We give two results according to the description of the number field. In the first description, the number field is $\mathbb{Q}(\alpha)$ where α is a root of some polynomial $P \in \mathbb{Z}[X]$ with bounded coefficients, and we want to recognize an element $x \in \mathbb{Z}[\alpha]$. This situation arises for instance when lifting from a finite field; not much is known about the number field itself. In the second description, we assume that an LLL-reduced basis of $\mathbb{Z}_L$ is known, and we want to recognize an element $x \in \mathbb{Z}_L$. The necessary precision is given in terms of the discriminant Δ_L of L and the height $h(x)$ of x, defined in §4.2.

Proposition 6.1.5. *There exist an algorithm and an absolute constant C such that the following holds. Let L be a number field of degree d over $\mathbb{Q}$*

defined by a monic irreducible polynomial $P \in \mathbb{Z}[X]$, and let $M \geq 1$ such that $\log^+ |P| \leq M$. Let α be a root of P in L. Let

$$x = \sum_{j=0}^{d-1} \lambda_j \alpha^j \in \mathbb{Z}[\alpha]$$

with $\lambda_j \in \mathbb{Z}$ and $\log^+ |\lambda_j| \leq H$ for every j. Let $N \geq C(H + dM + d \log d)$. Then, given P and approximations of x, α and $1/P'(\alpha)$ to precision N in every complex embedding of L, the algorithm computes x within

$$O\big(\mathcal{M}\big(d(H + dM + d \log d)\big) \log d\big)$$

binary operations.

Proof. Denote the complex embeddings of L by $\mu_1, \ldots, \mu_d$. The polynomial $Q = \sum_{j=0}^{d-1} \lambda_j X^j$ interpolates the points $(\mu_i(\alpha), \mu_i(x))$ for every $1 \leq i \leq d$. By assumption, we have for each i

$$\log^+ |\mu_i(x)| \leq H + O(dM).$$

We are in the situation of proposition 6.1.4: we can compute an approximation of Q with a precision loss of $O(H + dM + d \log d)$ bits. Therefore, for an appropriate choice of the constant C that we do not make explicit, the resulting precision is sufficient to obtain Q exactly by rounding the result to the nearest integers. □

Let L be a number field of degree d over $\mathbb{Q}$. We endow $\mathbb{Z}_L$ with the euclidean metric induced by the map $\mathbb{Z}_L \to \mathbb{C}^d$ given by the d complex embeddings $\mu_1, \ldots, \mu_d$ of L. Then $\mathbb{Z}_L$ becomes a lattice of volume Δ_L in the Euclidean space $\mathbb{Z}_L \otimes_{\mathbb{Z}} \mathbb{R}$. Denote by $1 \leq \lambda_1 \leq \cdots \leq \lambda_d$ the successive minima of $\mathbb{Z}_L$. They satisfy the following inequality [NV09, Chap. 2, Thm. 5]:

$$\prod_{k=1}^{d} \lambda_k \leq \alpha_d^{d/2} \Delta_L,$$

where $\alpha_d \leq 1 + \frac{d}{4}$ denotes Hermite's constant [NV09, Chap. 2, Cor. 3].

There exist several definitions of a *reduced* $\mathbb{Z}$-basis $(a_1, \ldots, a_d)$ of $\mathbb{Z}_L$ in the literature, which are usually formulated in terms of the coefficients of the base-change matrix from $(a_1, \ldots, a_d)$ to its Gram–Schmidt orthogonalization. We do not need the precise definitions here, and we only state the following properties:

- If $(a_1, \dots, a_d)$ is *HKZ-reduced* [NV09, Chap. 2, Thm. 6], then for each $1 \leq k \leq d$, we have

$$\frac{4}{k+3} \leq \left(\frac{\|a_k\|}{\lambda_k}\right)^2 \leq \frac{k+3}{4}.$$

- If $(a_1, \dots, a_d)$ is *LLL-reduced* (with parameter $\delta = \frac{3}{4}$) [NV09, Chap. 2, Thm. 9], then for each $1 \leq k \leq d$, we have

$$\|a_k\| \leq 2^{(d-1)/2}\lambda_k.$$

  Moreover,

$$\prod_{k=1}^{d} \|a_k\| \leq 2^{d(d-1)/4}\Delta_L.$$

HKZ-reduced bases approximate the successive minima closely, but are difficult to compute as the dimension d grows. On the other hand, LLL-reduced bases can be computed in polynomial time in d by the LLL algorithm [LLL82].

Proposition 6.1.6. *There exist an algorithm and an absolute constant C such that the following holds. Let L be a number field of degree d and discriminant Δ_L. Let $(a_1, \dots, a_d)$ be an LLL-reduced basis of $\mathbb{Z}_L$, let $\mu_1, \dots, \mu_d$ be the complex embeddings of L, and let m_L be the matrix $(\mu_i(a_j))_{1 \leq i,j \leq d}$. Let $x \in \mathbb{Z}_L$, and let $H \geq 1$ such that $h(x) \leq H$. Let*

$$N \geq C(\log \Delta_L + dH + d^2).$$

Then, given approximations of $(\mu_i(x))_{1 \leq i \leq d}$ and m_L^{-1} to precision N, the algorithm computes x within $O\big(d^2\mathcal{M}(H + \log \Delta_L + d^2)\big)$ binary operations.

Proof. Let $\lambda_j \in \mathbb{Z}$ such that $x = \sum \lambda_j a_j$. By definition of m_L, we have

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_d \end{pmatrix} = m_L^{-1} \begin{pmatrix} \mu_1(x) \\ \vdots \\ \mu_d(x) \end{pmatrix}.$$

The determinant of m_L is Δ_L, so $|\det m_L| \geq 1$. In order to bound the absolute values of the coefficients of m_L^{-1} from above, we use Hadamard's lemma. Each coefficient of $(\det m_L) \cdot m_L^{-1}$ is the determinant of a submatrix of m_L, and the L^2-norms of the columns of m_L are precisely the $\|a_k\|$ for $1 \leq k \leq d$. Moreover $\|a_k\| \geq 1$ for every k. Therefore,

$$\left|m_L^{-1}\right| \leq \prod_{k=1}^{d} \|a_k\| \leq 2^{d(d-1)/2}\Delta_L,$$

and hence

$$\log^+ \left|m_L^{-1}\right| \leq \log \Delta_L + O(d^2).$$

Since $h(x) \leq H$, we have $\sum_{i=1}^{d} \log^+ |\mu_i(x)| \leq dH$. Therefore, for some choice of the constant C that we do not make explicit, we can recover the coefficients $\lambda_j \in \mathbb{Z}$ exactly. On average, we have $\log^+ |\mu_i(x)| \leq H$, so the cost of each multiplication is on average $O\big(\mathcal{M}(H + \log \Delta_L + d^2)\big)$ binary operations. Therefore the total cost of the matrix-vector product is only $O\big(d^2\mathcal{M}(H + \log \Delta_L + d^2)\big)$ binary operations. □

If $(a_1, \ldots, a_d)$ is instead assumed to be HKZ-reduced in proposition 6.1.6, then a similar proof shows that one can take

$$N \geq C(\log \Delta_L + dH + d \log d)$$

with a cost of $O\big(d^2\mathcal{M}(H + \log \Delta_L + d \log d)\big)$ binary operations. Indeed, in this case we have

$$\prod_{k=1}^{d} \|a_k\| \leq d^d \prod_{k=1}^{d} \lambda_k \leq d^d (1 + \tfrac{d}{4})^{d/2} \Delta_L,$$

hence $\log^+ \left|m_L^{-1}\right| \leq \log(\Delta_L) + O(d \log d)$.

6.2 Computing theta functions in genus 2

In this section, we present the AGM method to compute $\tau \in \mathbb{H}_2$ from its theta constants, and Dupont's algorithm to go in the reverse direction. For convenience, we recall the expression (2.10) of theta constants in any genus: for $a, b \in \{0,1\}^g$ and $\tau \in \mathbb{H}_g$, we write

$$\theta_{a,b}(\tau) = \sum_{m \in \mathbb{Z}^g} \exp\left(i\pi \left(\left(m + \tfrac{a}{2}\right)^t \tau \left(m + \tfrac{a}{2}\right) + \left(m + \tfrac{a}{2}\right)^t \tfrac{b}{2}\right)\right). \tag{6.1}$$

6.2.1 The AGM method

We start with the easier case of the AGM method in genus 1. Let $\mathcal{F}_1$ be the classical fundamental domain for the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}_1$, and assume that the theta quotients $\theta_{0,1}/\theta_{0,0}$ and $\theta_{1,0}/\theta_{0,0}$ at $\tau \in \mathcal{F}_1$ are given. Then the sequence

$$B(\tau) = \left(\frac{\theta_{0,0}^2(2^n\tau)}{\theta_{0,0}^2(\tau)}, \frac{\theta_{0,1}^2(2^n\tau)}{\theta_{0,0}^2(\tau)}\right)_{n \geq 0}$$

is an *AGM sequence*, meaning that each term of $B(\tau)$ is obtained from the previous one by the transformation

$$(x, y) \mapsto \Big(\frac{x+y}{2}, \sqrt{x}\sqrt{y}\Big),$$

for some choice of the square roots. This is a consequence of the duplication formula [Mum83, p. 221], the correct square roots being the theta quotients themselves. In the algorithm, the sign ambiguity is easily removed using the fact that $\sqrt{x}$ and $\sqrt{y}$ lie in a common open quarter plane seen from the origin [Dup11, Thm. 2], i.e. a set of the form

$$\{r \exp(i(\alpha_0 + \alpha)) \mid r > 0 \text{ and } 0 < \alpha < \pi/2\}$$

for some $\alpha_0 \in \mathbb{R}$. We say that the AGM sequence $B(\tau)$ is given by *good sign choices*. The sequence $B(\tau)$ converges quadratically to $1/\theta_{0,0}^2(\tau)$, as the series expansion (6.1) shows.

It turns out that the sequence $B(-1/\tau)$ is also an AGM sequence with good sign choices [Dup11, Prop. 7]. Its first term can be computed from theta quotients at τ using the transformation formulæ for theta constants under $\mathrm{SL}_2(\mathbb{Z})$. The limit of $B(-1/\tau)$ is $1/\theta_{0,0}^2(-1/\tau)$. Finally, we can recover τ using the formula

$$\theta_{0,0}^2\Big(\frac{-1}{\tau}\Big) = -i\tau\theta_{0,0}^2(\tau). \tag{6.2}$$

Since the convergence of both AGM sequences is quadratic, we obtain an algorithm to invert theta functions on $\mathcal{F}_1$ with quasi-linear complexity in the output precision, at least for fixed τ. This method was already known to Gauss [Gau68, X.1, pp. 184–206], and we recommend [Cox84, §3C] for a historical exposition of Gauss's works on the AGM and elliptic functions.

The analogue of this algorithm in genus 2 is based on a generalization of AGM sequences to more than two elements, called Borchardt sequences [Bor88].

Definition 6.2.1. A *Borchardt sequence* is a sequence of complex numbers

$$(s_b^{(n)})_{b \in (\mathbb{Z}/2\mathbb{Z})^2,\, n \geq 0}$$

with the following property: for every $n \geq 0$, there exist complex numbers $t_b^{(n)}$ for $b \in (\mathbb{Z}/2\mathbb{Z})^2$ such that $t_b^{(n)}$ is a square root of $s_b^{(n)}$, and

$$s_b^{(n+1)} = \frac{1}{4} \sum_{b_1 + b_2 = b} t_{b_1}^{(n)} t_{b_2}^{(n)} \qquad \text{for each } b \in (\mathbb{Z}/2\mathbb{Z})^2.$$

The duplication formula [Mum83, p. 221] states that for every $\tau \in \mathbb{H}_2$, the sequence

$$B(\tau) = \left(\theta_{0,b}^2(2^n\tau)\right)_{b\in\{0,1\}^2, n\geq 0}$$

is a Borchardt sequence; the correct choice of square roots at each step is given by the theta constants $\theta_{0,b}(2^n\tau)$ themselves. By the series expansion (6.1), we have

$$\theta_{0,b}(2^n\tau) = \sum_{m\in\mathbb{Z}^2} \exp\left(-2^n\pi m^t \operatorname{Im}(\tau) m\right) \exp\left(i\pi\left(2^n m^t \operatorname{Re}(\tau) m + m^t \tfrac{b}{2}\right)\right).$$

When n tends to infinity, all the terms except $m = 0$ converge rapidly to zero, because $\operatorname{Im}(\tau)$ is positive definite. Therefore the Borchardt sequence $B(\tau)$ converges to $(1, 1, 1, 1)$.

We say that a set of complex numbers is *in good position* when it is included in an open quarter plane. The property of being in good position is invariant by nonzero complex scaling. A Borchardt sequence is given by *good sign choices* if for every $n \geq 0$, the complex numbers $t_b^{(n)}$ for $b \in (\mathbb{Z}/2\mathbb{Z})^2$ are in good position.

In order to present the algorithm to recover $\tau \in \mathbb{H}_2$ from its theta quotients, we introduce four matrices $\gamma_k \in \mathrm{Sp}_4(\mathbb{Z})$ for $0 \leq k \leq 3$. Let

$$S_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; S_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \; S_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and define the matrix $\gamma_k \in \mathrm{Sp}_4(\mathbb{Z})$ for $0 \leq k \leq 3$ by

$$\gamma_0 = I_4, \text{ and } \gamma_k = \begin{pmatrix} -I_2 & -S_k \\ S_k & -I + S_k^2 \end{pmatrix} \text{ for } 1 \leq k \leq 3.$$

Also recall from §2.3.2 that the genus 2 theta constants are numbered as follows [Dup06, §6.2]:

$$\theta_{(a_0,a_1),(b_0,b_1)} =: \theta_j \quad \text{where } j = b_0 + 2b_1 + 4a_0 + 8a_1 \in [\![0, 15]\!].$$

Assuming that the correct choices of square roots in the sequences $B(\gamma_k\tau)$ can be determined (for instance, if the sign choices in these Borchardt sequences are always good), we can compute $\tau \in \mathbb{H}_2$ from its theta quotients as follows.

Algorithm 6.2.2 ([Dup06, §9.2.3])**.**
Input: The vector of squares of theta constants $\theta_j^2(\tau)/\theta_0^2(\tau)$ for $j \in [\![1, 15]\!]$, for some $\tau \in \mathbb{H}_2$.
Output: The matrix τ.

1. For each $0 \leq k \leq 3$, compute the first term of the Borchardt sequence $B(\gamma_k\tau)/\theta_0^2(\gamma_k\tau)$ using the transformation formulæ for theta constants under $\mathrm{Sp}_4(\mathbb{Z})$ (see [Igu72, Thm. 2 p. 175 and Cor. p. 176], or proposition 6.2.6 below);
2. For each $0 \leq k \leq 3$, compute $1/\theta_0^2(\gamma_k\tau)$ as the limit of the Borchardt sequence $B(\gamma_k\tau)/\theta_0^2(\gamma_k\tau)$;
3. Use the input and the newly computed $\theta_0^2(\gamma_0\tau) = \theta_0^2(\tau)$ to compute all squares of theta constants at τ;
4. Recover $\tau = \begin{pmatrix} z_1 & z_3 \\ z_3 & z_2 \end{pmatrix}$ using the relations given in [Dup06, §6.3.1]:

$$\theta_0^2(\gamma_1\tau) = -iz_1\theta_4^2(\tau),\ \theta_0^2(\gamma_2\tau) = -iz_2\theta_8^2(\tau),\ \theta_0^2(\gamma_3\tau) = -\det(\tau)\theta_0^2(\tau).$$

We stress that algorithm 6.2.2 is theoretical: in practice, we need a way of determining the correct sign choices, we handle approximations of all the complex numbers involved, and we have to bound the precision losses incurred in the computations.

The correct sign choices in the AGM algorithm are difficult to describe in general, but as in the genus 1 case, things become easier if we restrict to the fundamental $\mathcal{F}_2$ for the action of $\mathrm{Sp}_4(\mathbb{Z})$ on $\mathbb{H}_2$, defined as follows.

Definition 6.2.3. Let

$$Y = \begin{pmatrix} y_1 & y_3 \\ y_3 & y_2 \end{pmatrix}$$

be a symmetric 2×2 real matrix, and assume that Y is positive definite. We say that Y is *Minkowski-reduced* if

$$y_1 \leq y_2 \quad \text{and} \quad -0 \leq 2y_3 \leq y_1.$$

The domain $\mathcal{F}_2$ is defined as the set of all matrices $\tau \in \mathbb{H}_2$ such that

1. $\mathrm{Im}(\tau)$ is Minkowski-reduced.
2. $|\mathrm{Re}(\tau)| \leq 1/2$.
3. $|\det(\gamma^*\tau)| \geq 1$ for every $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$.

It is enough to check the last condition for an explicit finite set $\mathfrak{S} \subset \mathrm{Sp}_4(\mathbb{Z})$ consisting of 19 elements [Got59], [Dup06, Thm. 6.1]. Recall that for every $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sp}_4(\mathbb{Z})$ and $\tau \in \mathbb{H}_2$, we write $\gamma^*\tau := c\tau + d$.

The original reference [Dup06] makes the observation that the sign choices in the sequences $B(\gamma_k\tau)$ for $0 \leq k \leq 3$ are always good when τ belongs to $\mathcal{F}_2$. This fact is proved in the case of $\gamma_0 = I_4$ in [Dup06, Prop. 9.1],

but is otherwise left as a conjecture [Dup06, Conj. 9.1], [Eng09a, Conj. 9]. We managed to prove that the other three sequences $B(\gamma_k\tau)$ for $1 \leq k \leq 3$ are also given by good sign choices when $\tau \in \mathcal{F}_2$ [Kie21]; we give the idea of the proof in §6.2.2. Then we bound the precision losses in algorithm 6.2.2 in §6.2.3.

6.2.2 Sign choices in the AGM method

In the sequel, we use the following notation. For $\tau \in \mathbb{H}_2$, we write

$$\tau = \begin{pmatrix} z_1(\tau) & z_3(\tau) \\ z_3(\tau) & z_2(\tau) \end{pmatrix} \quad \text{and} \quad \begin{cases} x_j(\tau) = \operatorname{Re} z_j(\tau) \\ y_j(\tau) = \operatorname{Im} z_j(\tau) \end{cases} \quad \text{for } 1 \leq j \leq 3.$$

For $1 \leq j \leq 3$, we also write

$$q_j(\tau) = \exp(-\pi y_j(\tau)).$$

We denote by $\lambda_1(\tau)$ the smallest eigenvalue of $\operatorname{Im}(\tau)$, and define

$$r(\tau) = \min\left\{\lambda_1(\tau), \frac{y_1(\tau)}{2}, \frac{y_2(\tau)}{2}\right\}.$$

We often omit the argument τ to ease notation. We define $\mathcal{F}_2'$ to be the set of all $\tau \in \mathbb{H}_2$ such that the following conditions are satisfied:

$$\begin{aligned} |x_j(\tau)| &\leq \frac{1}{2} \quad \text{for each } 1 \leq j \leq 3, \\ 2\,|y_3(\tau)| &\leq y_1(\tau) \leq y_2(\tau), \\ y_1(\tau) &\geq \frac{\sqrt{3}}{2}, \\ |z_j(\tau)| &\geq 1 \quad \text{for } j \in \{1, 2\}. \end{aligned} \tag{6.3}$$

The domain $\mathcal{F}_2'$ contains the classical fundamental domain $\mathcal{F}_2$ introduced in §6.2.1. Assumptions similar to (6.3) are usual when giving analytic estimates on theta constants: for instance, the domain $\mathcal{B}$ in [Str14] is defined by the first three inequalities of (6.3).

Theorem 6.2.4 ([Kie21])**.** *For every $\tau \in \mathcal{F}_2'$ and every $0 \leq k \leq 3$, the Borchardt sequence $B(\gamma_k\tau)$ contains only good sign choices.*

The proof works by giving estimates on the value of theta constants appearing as the terms of these Borchardt sequences, using the series expansion (6.1). However, the first terms of this series approximate $\theta_{a,b}(\tau)$

efficiently only when τ is close to the cusp at infinity, more precisely when $\operatorname{Im} z_1(\tau)$, $\operatorname{Im} z_2(\tau)$, and $\det \operatorname{Im}(\tau)$ are large. These conditions are satisfied in the case of the sequence $B(\gamma_0\tau)$, whose n-th term consists of theta constants at $2^n\tau$ for every $n \geq 0$, but not in the other cases.

To solve this problem, we construct specific symplectic matrices that bring the matrices $2^n\gamma_k\tau \in \mathbb{H}_2$ for $n \geq 0$ and $1 \leq k \leq 3$ closer to the cusp at infinity, exactly in the same way as we constructed the matrices η^R in the proof of proposition 5.1.10. For every $n \geq 0$, we define

$$\eta_1^{(n)} = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 2^n & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \eta_2^{(n)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2^n \end{pmatrix},$$
$$\eta_3^{(n)} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 2^n & 0 \\ 1 & 0 & 0 & 2^n \end{pmatrix}, \quad \text{and} \quad \eta_4^{(n)} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \eta_3^{(n)}.$$

Proposition 6.2.5. *Let $n \geq 0$.*

1. *For every $1 \leq k \leq 4$, the matrix $\eta_k^{(n)}$ belongs to $\mathrm{Sp}_4(\mathbb{Z})$.*
2. *For every $\tau = \begin{pmatrix} z_1 & z_3 \\ z_3 & z_2 \end{pmatrix} \in \mathbb{H}_2$, we have*

$$\begin{aligned} \tau_1^{(n)} &:= \eta_1^{(n)}(2^n\gamma_1\tau) = \begin{pmatrix} 2^{-n}z_1 & z_3 \\ z_3 & 2^n z_2 \end{pmatrix}, \\ \tau_2^{(n)} &:= \eta_2^{(n)}(2^n\gamma_2\tau) = \begin{pmatrix} 2^n z_1 & z_3 \\ z_3 & 2^{-n}z_2 \end{pmatrix}, \\ \tau_3^{(n)} &:= \eta_3^{(n)}(2^n\gamma_3\tau) = 2^{-n}\tau, \quad \textit{and} \\ \tau_4^{(n)} &:= \eta_4^{(n)}(2^n\gamma_3\tau) = \begin{pmatrix} -2^n/z_1 & -z_3/z_1 \\ -z_3/z_1 & 2^{-n}(z_2 - z_3^2/z_1) \end{pmatrix}. \end{aligned} \tag{6.4}$$

Proof. 1. The lines of each $\eta_k^{(n)}$ define a symplectic basis of $\mathbb{Z}^4$.

2. Recall that the action of $\mathrm{Sp}_4(\mathbb{Z})$ on $\mathbb{H}_2$ extends to an action of the group $\mathrm{GSp}_4(\mathbb{Q})_+$. The matrix $2^n\gamma_k\tau$ is the image of τ under

$$\begin{pmatrix} -2^n I_2 & -2^n S_k \\ S_k & -I + S_i^2 \end{pmatrix} \in \mathrm{GSp}_4(\mathbb{Q})_+.$$

When we multiply this matrix by $\eta_k^{(n)}$ on the left, we obtain

$$\begin{array}{ll} \mathrm{Diag}(-1,-2^n,-2^n,-1) & \text{for } k=1, \\ \mathrm{Diag}(-2^n,-1,-1,-2^n) & \text{for } k=2, \text{ and} \\ \mathrm{Diag}(-1,-1,-2^n,-2^n) & \text{for } k=3. \end{array} \qquad \square$$

In order to relate theta constants at $2^n\gamma_k\tau$ with theta constants at $\tau_k^{(n)}$, we use the following transformation formula. For a square matrix m, we denote by m_0 the column vector containing the diagonal of m.

Proposition 6.2.6 ([Igu72, Thm. 2 p. 175 and Cor. p. 176])**.** *Let $a, b \in \{0,1\}^2$, and let*

$$\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z}).$$

Define

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \gamma^t \begin{pmatrix} a - (CD^t)_0 \\ b - (AB^t)_0 \end{pmatrix}.$$

Then, for every $\tau \in \mathbb{H}_2$, we have

$$\theta_{a,b}(\gamma\tau) = \kappa(\gamma)\, \zeta_8^{\varepsilon(\gamma,a,b)} \det(C\tau + D)^{1/2}\, \theta_{a',b'}(\tau)$$

where

$$\zeta_8 = e^{i\pi/4}, \qquad \begin{pmatrix} a' \\ b' \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \bmod 2,$$

$$\varepsilon(\gamma,a,b) = 2(B\alpha)^t(C\beta) - (D\alpha)^t(B\alpha) - (C\beta)^t(A\beta) + 2\big((AB^t)_0\big)^t(D\alpha - C\beta),$$

and $\kappa(\gamma)$ is an eighth root of unity depending only on γ, with a sign ambiguity coming from the choice of a holomorphic square root of $\det(C\tau + D)$.

Corollary 6.2.7. *For every $\tau \in \mathbb{H}_2$, we have the following equalities of projective tuples:*

$$(\theta_j(2^n\gamma_1\tau))_{0\le j\le 3} = \begin{cases} (\theta_4(\tau) : \theta_0(\tau) : \theta_6(\tau) : \theta_2(\tau)) & \text{if } n = 0, \\ (\theta_0(\tau_1^{(n)}) : \theta_4(\tau_1^{(n)}) : \theta_2(\tau_1^{(n)}) : \theta_6(\tau_1^{(n)})) & \text{if } n \ge 1, \end{cases}$$

$$(\theta_j(2^n\gamma_2\tau))_{0\le j\le 3} = \begin{cases} (\theta_8(\tau) : \theta_9(\tau) : \theta_0(\tau) : \theta_1(\tau)) & \text{if } n = 0, \\ (\theta_0(\tau_2^{(n)}) : \theta_1(\tau_2^{(n)}) : \theta_8(\tau_2^{(n)}) : \theta_9(\tau_2^{(n)})) & \text{if } n \ge 1, \end{cases}$$

$$(\theta_j(2^n\gamma_3\tau))_{0\le j\le 3} = (\theta_0(\tau_3^{(n)}) : \theta_8(\tau_3^{(n)}) : \theta_4(\tau_3^{(n)}) : \theta_{12}(\tau_3^{(n)})) \quad \text{for all } n \ge 0,$$

$$(\theta_j(2^n\gamma_3\tau))_{0\le j\le 3} = (\theta_0(\tau_4^{(n)}) : \theta_8(\tau_4^{(n)}) : \theta_1(\tau_4^{(n)}) : \theta_9(\tau_4^{(n)})) \quad \text{for all } n \ge 0,$$

where the $\tau_j^{(n)}$ are defined as in (6.4).

Proof. Apply proposition 6.2.6 to the matrices $\eta_k^{(n)}$. □

The rest of the proof of theorem 6.2.4 involves technical estimates on the tails of series defining theta constants. The computations are similar to those found in [Kli90, pp. 116–117], [Dup06, §6.2], [HP17, §5.1]. For instance, in the case of the Borchardt sequence $B(\gamma_0\tau)$ for $\tau \in \mathcal{F}_2'$, we can apply the following result.

Proposition 6.2.8. *Let $\tau \in \mathbb{H}_2$.*

1. *If $r(\tau) \geq 0.4$, then the $\theta_j(\tau)$ for $0 \leq j \leq 3$ are in good position.*
2. *If $\lambda_1(\tau) \geq 0.6$, then the $\theta_j(\tau)$ for $0 \leq j \leq 3$ are in good position.*

Proof. 1. Write

$$q = \exp(-\pi r(\tau)).$$

For $0 \leq j \leq 3$, by (6.1), we have

$$\begin{aligned} |\theta_j(\tau) - 1| &\leq 4q^2 + \sum_{n \in \mathbb{Z}^2,\, \|n\|^2 \geq 2} \exp(-\pi \lambda_1(\tau)\, \|n\|^2) \\ &\leq 8q^2 + 4q^4 + 8q^5 + 4q^8 + 4\frac{1+q}{(1-q)^2} q^9. \end{aligned} \tag{6.5}$$

In (6.5), the first term $4q^2$ comes from the four vectors $n \in \mathbb{Z}^2$ with $\|n\| = 1$. Then we separate the terms $n = (n_1, n_2)$ such that $|n_1| \geq 3$ and $|n_2| \geq 3$; this accounts for the term $4q^9(1+q)/(1-q)^2$, as in the proof of [Dup06, Prop. 6.1]. We leave the remaining terms as they are.

If $q \leq 0.287$, then the quantity on the right hand side of (6.5) is less than $\sqrt{2}/2$, and the $\theta_j(\tau)$ are contained in a disk centered at 1 which is itself contained in a quarter plane. We have $q \leq 0.287$ when $r(\tau) \geq 0.4$.

2. Write

$$q = \exp(-\pi \lambda_1(\tau)).$$

Then for $0 \leq j \leq 3$, we have

$$|\theta_j(\tau) - 1| \leq 4q + 4q^2 + 4q^4 + 8q^5 + 4q^8 + 4\frac{1+q}{(1-q)^2} q^9.$$

This quantity is less than $\sqrt{2}/2$ when $\lambda_1(\tau) \geq 0.6$. □

Proposition 6.2.8 implies that $B(\gamma_0\tau)$ for $\tau \in \mathcal{F}_2'$ contains only good sign choices, since

$$r(2^n\tau) = 2^n r(\tau) \geq \sqrt{3}/4 \geq 0.4$$

for every $n \geq 0$.

In the study of $B(\gamma_k\tau)$ for $1 \leq k \leq 3$, we separate different cases according to the value of n. If n is large enough, then $2^n\lambda_1(\gamma_k\tau) \geq 0.6$, and hence proposition 6.2.8 shows that the $n+1$st term of $B(\gamma_k\tau)$ is obtained by a Borchardt iteration with good sign choices. The precise rank where this happens depends on τ.

Proposition 6.2.9. *Let $\tau \in \mathbb{H}_2$ and $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$. Then*

$$\lambda_1(\gamma\tau) \geq \frac{\det \mathrm{Im}(\tau)}{8\,|\gamma|^2\,|\tau|\,(2\,|\tau|+1)^2}.$$

Proof. We have

$$\lambda_1(\gamma\tau) \geq \frac{\det \mathrm{Im}(\gamma\tau)}{\mathrm{Tr}\,\mathrm{Im}(\gamma\tau)}.$$

By [Str10, (5.11) p. 57], we have

$$\mathrm{Im}(\gamma\tau) = (\gamma^*\tau)^{-t}\,\mathrm{Im}(\tau)(\gamma^*\bar{\tau})^{-1},$$

hence

$$\det \mathrm{Im}(\gamma\tau) = \frac{\det \mathrm{Im}(\tau)}{|\det(\gamma^*\tau)|^2}, \qquad \text{and}$$

$$\mathrm{Tr}\,\mathrm{Im}(\gamma\tau) \leq 8\left|(\gamma^*\tau)^{-1}\right|^2 |\mathrm{Im}(\tau)| \leq 8\frac{|\gamma^*\tau|^2\,|\tau|}{|\det(\gamma^*\tau)|^2} \leq 8\frac{|\gamma|^2\,(2\,|\tau|+1)^2\,|\tau|}{|\det(\gamma^*\tau)|^2}.\square$$

One can give more precise versions of proposition 6.2.9 when $\tau \in \mathcal{F}_2'$ and $\gamma \in \{\gamma_1, \gamma_2, \gamma_3\}$ [Kie21]. For instance, we have

$$r(\gamma_1\tau) \geq \frac{9\,y_1(\tau)}{34\,|z_1(\tau)|^2}$$

for all $\tau \in \mathcal{F}_2'$. Hence proposition 6.2.8 applies to $2^n\gamma_1\tau$ if $2^n \geq 1.96 y_1(\tau)$.

On the other hand, when n is smaller, then the matrix

$$\tau_1^{(n)} = \begin{pmatrix} 2^{-n}z_1 & z_3 \\ z_3 & 2^n z_2 \end{pmatrix}$$

is close to the cusp at infinity. By corollary 6.2.7, in order to prove that $B(\gamma_1\tau)$ contains only good choices of signs, we are led to prove that the theta constants $\theta_j\left(\tau_1^{(n)}\right)$ for $j \in \{0,2,4,6\}$ are in good position for $\tau \in \mathcal{F}_2'$ and $2^n \leq 1.96 y_1(\tau)$. This is done by careful estimates on the tails of the series (6.1) and on the complex arguments of the sum of the first few terms. We refer to [Kie21] for a full proof of theorem 6.2.4.

6.2.3 Precision losses in the AGM method

Once we know that the Borchardt sequences in algorithm 6.2.2 are given by good sign choices when $\tau \in \mathcal{F}_2'$, we can estimate the precision losses in this algorithm when giving approximations of theta constants at τ as input. We start with four lemmas: the first two give information on the first terms of the Borchardt sequences obtained from the series expansion (6.1), and the other ones are general results on Borchardt sequences.

Lemma 6.2.10 ([Str14, Prop. 7.6 and Cor. 7.7])**.** *Let $\tau \in \mathcal{F}_2'$, and abbreviate $y_1(\tau)$ as y_1, $z_1(\tau)$ as z_1, etc. Then we have:*

$$\begin{aligned} |\theta_j(\tau) - 1| &< 0.405 \quad \text{for } j \in \{0,1,2,3\}, \\ \left|\frac{\theta_j(\tau)}{2\exp(i\pi z_1/4)} - 1\right| &< 0.348 \quad \text{for } j \in \{4,6\}, \\ \left|\frac{\theta_j(\tau)}{2\exp(i\pi z_2/4)} - 1\right| &< 0.348 \quad \text{for } j \in \{8,9\}, \end{aligned}$$

and

$$\left|\frac{\theta_{12}(\tau)}{2(1+\exp(i\pi z_3))\exp(i\pi(z_1+z_2-2z_3)/4)} - 1\right| < 0.438.$$

In particular,

$$\begin{aligned} 0.59 &< |\theta_j(\tau)| < 1.41 \quad \text{for } j \in \{0,1,2,3\}, \\ 1.3\exp\left(-\tfrac{\pi}{4}y_1\right) &< |\theta_j(\tau)| < 1.37 \quad \text{for } j \in \{4,6\}, \\ 1.3\exp\left(-\tfrac{\pi}{4}y_2\right) &< |\theta_j(\tau)| < 1.37 \quad \text{for } j \in \{8,9\}\text{, and} \\ 1.05\exp\left(-\tfrac{\pi}{4}(y_1+y_2-2y_3)\right) &< |\theta_{12}(\tau)| < 1.56. \end{aligned}$$

Lemma 6.2.11. *There exists an angle $\alpha < \pi/2$ such that the following holds. Let S be one of the sets $\{0,1,2,3\}$, $\{0,2,4,6\}$, $\{0,1,8,9\}$, or $\{0,4,8,12\}$. Then for every $\tau \in \mathcal{F}_2'$, the theta constants $\theta_j(\tau)$ for $j \in S$ are contained in an open angular sector of angle α seen from the origin.*

Proof. Using the estimates of lemma 6.2.10 and the definition of $\mathcal{F}_2'$ in (6.3) is sufficient to conclude, except in the case $S = \{0,4,8,12\}$. Consider for instance the case of $S = \{0,1,8,9\}$. The absolute value of the argument of $\exp(i\pi z_2/4)$ is at most $\pi/8$, hence the angle between the theta constants $\theta_j(\tau)$ for $j \in S$ is bounded above by

$$\frac{\pi}{8} + \arcsin(0.348) + \arcsin(0.405) < \frac{\pi}{2}.$$

In the case of $S = \{0,4,8,12\}$, more delicate estimates are needed. We refer to [Kie21] for the proof. □

Lemma 6.2.12 ([Dup06, §7.2]). *Let $B = \left(s_b^{(n)}\right)_{b\in(\mathbb{Z}/2\mathbb{Z})^2,n\geq 0}$ be a Borchardt sequence with good sign choices only, and assume that $\operatorname{Re}\left(s_b^{(n)}\right) > 0$ for every $b \in (\mathbb{Z}/2\mathbb{Z})^2$ and $n \geq 0$. For every $n \geq 0$, define*

$$m_n = \min_{b\in(\mathbb{Z}/2\mathbb{Z})^2} \operatorname{Re}\left(s_b^{(n)}\right) \quad \text{and} \quad M_n = \max_{b\in(\mathbb{Z}/2\mathbb{Z})^2} \left|s_b^{(n)}\right|.$$

Then the sequences (m_n) and (M_n) are nondecreasing and nonincreasing, respectively.

Note that every Borchardt sequence with good sign choices can be rescaled into a Borchardt sequence to which lemma 6.2.12 applies. Finally, the next lemma shows that once a Borchardt sequence with good sign choices starts to converge, it does so quadratically.

Lemma 6.2.13 ([Dup06, Prop. 7.1]). *Let $B = \left(s_b^{(n)}\right)_{b\in(\mathbb{Z}/2\mathbb{Z})^2,n\geq 0}$ be a Borchardt sequence with good sign choices only, and let s^∞ be its limit. Let $0 < \varepsilon < \sqrt{3/2} - 1$, and let $n_0 \in \mathbb{N}$ such that for every $b \in (\mathbb{Z}/2\mathbb{Z})^2$, we have*

$$\left|s_b^{(n_0)} - s_0^{(n_0)}\right| \leq \varepsilon \left|s_0^{(n_0)}\right|.$$

Then, for every $k \geq 0$, we have

$$\left|s^\infty - s_0^{(n_0+k)}\right| \leq 5M_{n_0}\left(\frac{7\varepsilon}{2}\right)^{2^k},$$

where M_{n_0} is defined as in lemma 6.2.12.

We are now ready to bound the precision losses in the AGM method in genus 2. The contribution we make compared to [Dup06, §9.2.3], besides showing that the sign choices are good, is that we make the dependency on $\tau \in \mathcal{F}_2'$ explicit.

Proposition 6.2.14. *Let $\tau \in \mathcal{F}_2'$ and $N \geq 1$. Then, given approximations of squares of theta quotients at τ to precision N, algorithm 6.2.2 computes an approximation of τ within*

$$O\left(\mathcal{M}(N + |\tau|)\log|\tau| + \mathcal{M}(N)\log N\right)$$

binary operations. The precision loss is $O(\log N + |\tau|\log|\tau|)$ bits.

Proof. By theorem 6.2.4, we obtain the quantities

$$(\theta_j^2(2^n\gamma_i\tau)/\theta_0^2(\gamma_i\tau))_{0\leq j\leq 3}$$

after n Borchardt steps with good sign choices. By proposition 6.2.9, we know that

$$\left|\log \lambda_1(\gamma_i\tau)\right| = O(\log|\tau|).$$

Therefore, there exists $n_0 = O(\log|\tau|)$ such that $\lambda_1(2^{n_0}\gamma_i\tau) \geq 10$, for instance. Even if $|\tau|$ is not known explicitly, this moment can be detected during the algorithm: the four values in the Borchardt sequence become close to each other, so that lemma 6.2.13 applies.

Let us estimate the precision losses in these n_0 steps. Let $B = \left(s_b^{(n)}\right)$ be one of the four Borchardt sequences that we consider. By lemma 6.2.11, after scaling by a complex number of modulus one, we may assume that the absolute values of the arguments of the $s_b^{(0)}$ for $b \in (\mathbb{Z}/2\mathbb{Z})^2$ are bounded above by some $\alpha < \pi/2$ which is independent of τ. Then, by lemma 6.2.10, we have

$$|\log m_0| = O(|\tau|) \quad \text{and} \quad |M_0| = O(1).$$

By lemma 6.2.12, the same estimates hold for m_n and M_n for every $n \geq 0$. Therefore, in the first n_0 Borchardt steps, we perform $O(\log|\tau|)$ elementary operations on complex numbers z that satisfy $|\log(|z|)| = O(|\tau|)$, for a total cost of

$$O(\mathcal{M}(N + |\tau|)\log|\tau|)$$

binary operations, and a precision loss of $O(|\tau|\log|\tau|)$ bits.

Then, we scale the values in such a way that one of the four values of the n_0-th term in the Borchardt sequence equals 1, and we add $O(\log N)$ Borchardt steps: this O-constant and the accuracy of the result can be made explicit by lemma 6.2.13. This costs $O(\mathcal{M}(N)\log N)$ binary operations, and the precision loss is $O(\log N)$ bits. This allows us to compute the quantities $\theta_0^2(\gamma_i\tau)$ for $1 \leq i \leq 3$; the precision loss up to now is $O(\log N + |\tau|\log|\tau|)$ bits.

Finally, we recover the entries of τ in step 4 of algorithm 6.2.2, for a cost of $O(N+|\tau|)$ binary operations and a precision loss of $O(|\tau|)$ bits. □

6.2.4 Dupont's algorithm for theta constants

Dupont's thesis [Dup06] builds a Newton scheme around the AGM method in order to compute theta constants in quasi-linear time. In genus 1, this method yields a provably correct algorithm [Dup11]. In genus 2, the convergence of Dupont's algorithm remains heuristic, even using the provably correct AGM algorithm presented in §6.2.3. In this section, we briefly present Dupont's method, we explain the remaining challenges to obtain a provably correct algorithm, and we state hypothesis 6.2.15 according to

which Dupont's algorithm converges uniformly on a compact set. This assumption is fundamental in the complexity estimates for the evaluation of modular equations for abelian surfaces.

In genus 1, the algorithm is formulated in terms of the function

$$k'(\tau) := \left(\frac{\theta_{0,1}(\tau)}{\theta_{0,0}(\tau)}\right)^2, \quad \text{for } \tau \in \mathbb{H}_1.$$

For each $z \in \mathbb{C}$ with positive real part, we define $M_1(z)$ as the limit of the AGM sequence with good sign choices whose first term is $\{1, z\}$. Define $\mathcal{F}'_1 \subset \mathbb{H}_1$ as in [Dup11, Prop. 7]:

$$\mathcal{F}'_1 = \{\tau \in \mathbb{H}_1 \colon |\operatorname{Re}\tau| < 1,\ |2\tau + 1| > 1,\ |2\tau - 1| > 1\}.$$

Note that $\mathcal{F}'_1$ contains the fundamental domain $\mathcal{F}_1$. Then k' is a surjective analytic map from $\mathcal{F}'_1$ to $\{z \in \mathbb{C} \colon \operatorname{Re}(z) > 0\}$ whose derivative does not vanish. Moreover, the equality

$$M_1(k'(\tau)) = \frac{1}{\theta_{0,0}^2(\tau)}$$

holds for all $\tau \in \mathcal{F}'_1$. By the inverse function theorem, the function M_1 is also analytic.

The Newton scheme to compute $k'(\tau)$ from $\tau \in \mathcal{F}_1$ uses the fact that the function

$$f_\tau(z) := iM_1(z) - \tau M_1(\sqrt{1 - z^2})$$

vanishes at $k'(\tau)$; this equality comes from eq. (6.2), the expression of theta constants at $-1/\tau \in \mathcal{F}'_1$, and the Jacobi identity $\theta_{0,0}^4 = \theta_{0,1}^4 + \theta_{1,0}^4$ [Dup11, §5.1]. The linearization step in the Newton scheme is based on the equality

$$\frac{df_\tau}{dz}(k'(\tau)) = \frac{-2M_1(k'(\tau))^3}{\pi\tau k'(\tau)(1 - k'(\tau)^2)}. \tag{6.6}$$

If $k'(\tau)$ is currently computed to precision N, then the derivative of f_τ at $k'(\tau)$ can also be computed to precision roughly N, and this allows in turn to compute a better approximation of $k'(\tau)$ to precision roughly $2N$. The computations are made precise in [Dup11, §5.2, §6.1, and §6.2]. Before starting the Newton iterations, $k'(\tau)$ is computed at low precision using the series expansions of theta constants (the so-called "naive method", which is *not* quasi-linear in the required precision).

The precision losses in each Newton iteration and the initial precision necessary depend on τ, and grow quickly as $\operatorname{Im}\tau \to +\infty$. However, when τ belongs to a compact set of $\mathcal{F}_1$, these precision losses can be bounded

uniformly. When $\mathrm{Im}(\tau)$ is large, one can compute theta constants at $\tau/2^n$ for some appropriate n and perform n AGM steps to recover theta constants at τ. In this way, we obtain an algorithm to evaluate $k'(\tau)$ at precion $N \geq 1$ using $O(\mathcal{M}(N)\log N)$ operations, uniformly in $\tau \in \mathcal{F}_1$ [Dup11, Thm. 5].

The only remaining issue is that in order to *implement* a provably correct version of Dupont's algorithm in genus 1, one needs *explicit* upper bounds on the precision losses in a given compact subset of $\mathcal{F}_1$. No such bound is given in [Dup11], or in subsequent works to our knowledge [Lab18, LT16]. In practice, a common *heuristic* strategy is to take an arbitrary error margin; if the Newton iterations do converge in an experimental sense, then one accepts the end value as an approximation of the true result.

In genus 2, Dupont's algorithm to compute theta constants follows an analogous strategy. Given $\tau \in \mathcal{F}_2'$, we want to compute the theta quotients $b_j(\tau) := \theta_j(\tau)^2/\theta_0(\tau)^2$ for $j \in \{1,2,3\}$; this is sufficient, because

1. the duplication formulæ express all squares of theta constants at 2τ in terms of these quantities;
2. the theta constants $\theta_j(\tau)^2$ themselves can be computed from this data using one further Borchardt mean.

By theorem 6.2.4, computing an approximation of $\tau \in \mathcal{F}_2'$ from the $b_j(\tau)$ for $1 \leq j \leq 3$ can be done using four Borchardt means with good sign choices. This yields, as in the genus 1 case, a function

$$F_\tau \colon \{z \in \mathbb{C} \colon \mathrm{Re}\, z > 0\}^3 \to \mathbb{C}^3$$

which satisfies

$$F_\tau\big(b_1(\tau), b_2(\tau), b_3(\tau)\big) = 0.$$

In order to build a Newton scheme to compute τ from its theta quotients, one relies on the following heuristic assumption, which holds in practice [Dup06, §10.2]: the Jacobian matrix of F_τ at $(b_1(\tau), b_2(\tau), b_3(\tau))$ is invertible, and can be suitably approximated using finite differences on the function F_τ. This yields a heuristic algorithm to compute theta constants at precision $N \geq 1$ in $O_\tau(\mathcal{M}(N)\log N)$ binary operations, for a *fixed* $\tau \in \mathcal{F}_2'$. As a by-product, the algorithm computes the Jacobian matrix of F_τ at high precision; equivalently, we also recover the values of *derivatives* of theta constants at τ at precision roughly $N/2$ for no additional cost.

For our application to the evaluation of modular equations for abelian surfaces, we need to assume more on the behavior of this algorithm, namely that the Newton iterations converge uniformly on a compact subset of $\mathcal{F}_2'$.

Hypothesis 6.2.15. There exists an algorithm such that the following holds. Let $\tau \in \mathcal{F}_2'$ and $N \geq 1$. Assume that

$$\max\{y_1(\tau), y_2(\tau)\} \leq 10.$$

Then, given an approximation of τ to precision N, the algorithm computes squares of theta constants at τ, and derivatives of theta constants at τ with respect to $z_1(\tau), z_2(\tau), z_3(\tau)$, in $O(\mathcal{M}(N)\log N)$ binary operations with a precision loss of $O(1)$ bits.

Derivatives of theta constants are uniformly bounded on the compact set defined by the conditions in hypothesis 6.2.15, hence the precision loss of $O(1)$ bits. In order to implement an algorithm which hopefully satisfies hypothesis 6.2.15, one simply applies the Newton scheme with a large initial precision, say 1000 bits, using the naive evaluation method for theta constants [Dup06, §10.1], and everything works well.

Using hypothesis 6.2.15, we describe in §6.2.5 an algorithm to compute theta functions in *uniform* quasi-linear time at a given $\tau \in \mathcal{F}_2'$, using duplication formulæ as in the genus 1 case. In general, in order to compute theta constants at a given $\tau \in \mathbb{H}_2$, we first reduce τ to the fundamental domain with controlled precision losses, using an adaptation of the classical reduction algorithm for inexact input (§6.2.6).

6.2.5 Extension to the whole fundamental domain

Here we describe an algorithm to evaluate theta constants and their derivatives at $\tau \in \mathcal{F}_2'$ to precision $N \geq 1$ within $O(\mathcal{M}(N)\log N)$ binary operations, uniformly in τ, assuming hypothesis 6.2.15. There are two easy cases:

1. If τ belongs to the compact set defined in hypothesis 6.2.15, then we use Dupont's algorithm directly.
2. If $y_1(\tau)$ and $y_2(\tau)$ satisfy $\min\{y_1, y_2\} \geq CN$, where C is an absolute constant, then we use the naive algorithm to compute theta constants at τ within $O(\mathcal{M}(N))$ binary operations [Dup06, §10.1].

For other values of $\tau \in \mathcal{F}_2$, we fall back to one of these two cases using duplication formulæ. For every $\tau \in \mathbb{H}_2$, write

$$D_1(\tau) = \frac{\tau}{2} \quad \text{and} \quad D_2(\tau) = \begin{pmatrix} 2z_1(\tau) & z_3(\tau) \\ z_3(\tau) & \frac{1}{2}z_2(\tau) \end{pmatrix}.$$

Lemma 6.2.16. *Let $\tau \in \mathcal{F}_2'$.*

1. *If $D_1(\tau) \in \mathcal{F}_2'$, then the tuple of theta constants $(\theta_j^2(\tau))_{j\in\{0,1,2,3\}}$ is obtained from the tuple of theta constants $\left(\theta_j^2(D_1(\tau))\right)_{j\in\{0,1,2,3\}}$ by a Borchardt iteration with good choice of roots.*
2. *If $D_2(\tau)$ satisfies (6.3), except that the real part of $z_1(D_2(\tau))$ is allowed to be smaller than 1 instead of 1/2, then $(\theta_j^2(\tau))_{j\in\{0,2,4,6\}}$ is obtained from $\left(\theta_j^2(D_2(\tau))\right)_{j\in\{0,2,4,6\}}$ by a Borchardt iteration with good choice of roots.*

Proof. The first item is the classical duplication formula: the choice of roots is given by the theta constants $\theta_j(D_1(\tau))$ for $0 \leq j \leq 3$, and they are in good position by lemma 6.2.10. For the second item, apply the theta transformation formula (proposition 6.2.6) to the symplectic matrix

$$\begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The theta constants $\theta_j(D_2(\tau))$ for $j \in \{0, 2, 4, 6\}$ also are in good position by the proof of [Kie21, Lem. 5.2]. □

Proposition 6.2.17 (Under hypothesis 6.2.15)**.** *There exists an algorithm which, given $\tau \in \mathcal{F}_2'$ to precision $N \geq 1$, computes the squares and derivatives of theta constants at τ in $O(\mathcal{M}(N)\log N)$ binary operations with a precision loss of $O(1)$ bits, uniformly in τ.*

Proof. Since derivatives of theta constants are uniformly bounded on $\mathcal{F}_2'$, the difference between theta constants at τ and theta constants at the center of the complex ball approximating τ is $O(2^{-N})$. Therefore we may assume that τ is given at arbitrarily high precision.

First, we let k_2 be the smallest integer such that

$$2^{k_2} y_1(\tau) \geq \min\{CN,\, 2^{-k_2-2} y_2(\tau)\}$$

where C is an absolute constant. Let τ' be the matrix obtained after applying k_2 times D_2 to τ and reducing the real part at each step. In order to compute theta constants at τ to precision N, we compute theta constants at τ' to some precision $N' \geq N$, then apply k_2 times the duplication formula from lemma 6.2.16. We have $k_2 = O(\log N)$, and the total precision loss when extracting square roots is $O(N)$ bits, hence the total precision loss is $O(N)$ bits. We can take $N' = C'N$ where C' is an absolute constant.

There are now two cases. If $y_1(\tau') \geq CN$, then we have $y_2(\tau') \geq CN$ as well. Hence we can compute theta constants at τ' to precision N' using $O(\mathcal{M}(N))$ operations with the naive algorithm. Otherwise, we have

$$y_1(\tau') \leq y_2(\tau') \leq 4y_1(\tau') \leq 4CN.$$

Therefore we can find an integer $k_1 = O(\log N)$ such that $D_1^{k_1}(\tau')$ belongs to the compact set defined in hypothesis 6.2.15. The precision losses in the duplication formula for D_1 are $O(1)$ bits per step, hence we recover theta constants at τ' to precision N' in $O(\mathcal{M}(N)\log N)$ binary operations. □

6.2.6 An approximate reduction algorithm

In order to evaluate theta constants at a given $\tau \in \mathbb{H}_2$, our strategy is to reduce τ to $\mathcal{F}_2$ and to compute theta constants there using the algorithm of proposition 6.2.17. However, the classical reduction algorithm described in [Str14, §6] is based on inequalities, and this causes problems on the boundary when the input is inexact. Therefore, we rather describe reduction algorithms to *neighborhoods* of $\mathcal{F}_2$; we still write inequalities, but they should be understood as *inclusions of intervals*. Then we show the validity of the reduction algorithm on inexact input provided that the precision remains high enough. We start by defining neighborhoods of $\mathcal{F}_2$.

Definition 6.2.18. Fix $\varepsilon > 0$, and let

$$Y = \begin{pmatrix} y_1 & y_3 \\ y_3 & y_2 \end{pmatrix}$$

be a symmetric 2×2 real matrix. Assume that Y is positive definite. We say that Y is *ε-Minkowski reduced* if

$$y_1 \leq (1+\varepsilon)y_2 \quad \text{and} \quad -\varepsilon y_1 \leq 2y_3 \leq (1+\varepsilon)y_1.$$

Let $\mathfrak{S} \subset \mathrm{Sp}_4(\mathbb{Z})$ be the set of 19 matrices defining the boundary of $\mathcal{F}_2$, as in definition 6.2.3. We define the neighborhood $\mathcal{F}_2^\varepsilon$ of $\mathcal{F}_2$ as the set of all matrices $\tau \in \mathbb{H}_2$ such that

1. $\mathrm{Im}(\tau)$ is ε-Minkowski reduced.
2. $|\mathrm{Re}(\tau)| \leq 1/2 + \varepsilon$.
3. $|\det(\sigma^*\tau)| \geq 1 - \varepsilon$ for every $\sigma \in \mathfrak{S}$.

The fundamental domain $\mathcal{F}_2$ corresponds to the case $\varepsilon = 0$.

The approximate reduction algorithm is copied on [Str14, Alg. 6.8]. The input is a matrix $\tau \in \mathbb{H}_2$ to precision $N \geq 1$, and the output is $\tau' \in \mathcal{F}_2^\varepsilon$ together with $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\tau' = \gamma\tau$. We assume that the precision remains greater than $|\log \varepsilon| + 1$ at any time. If we run out of precision, we stop and output "failure".

Algorithm 6.2.19 (Reduction to $\mathcal{F}_2^\varepsilon$)**.** Start with $\tau' = \tau$ and iterate the following three steps until $\tau' \in \mathcal{F}_2^\varepsilon$, keeping track of a matrix $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\tau' = \gamma\tau$:

1. Reduce $\mathrm{Im}(\tau')$ such that it becomes ε-Minkowski reduced.
2. Reduce $\mathrm{Re}(\tau')$ such that $|\mathrm{Re}(\tau')| \leq 1/2 + \varepsilon$.
3. Apply $\sigma \in \mathfrak{S}$ such that $|\det \sigma^*(\tau')|$ is at most $1 - \varepsilon/2$ and has the smallest upper bound, if such a σ exists.
4. Update $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ and recompute $\tau' = \gamma\tau$.

In order to analyze algorithm 6.2.19, we mimic Streng's analysis of the exact reduction algorithm [Str10, §II.5.3]. For $\tau \in \mathbb{H}_2$, we define

$$\Lambda(\tau) = \log \max\{2, |\tau|, \det(\mathrm{Im}\,\tau)^{-1}\}.$$

Denote by $0 < \lambda_1(\tau) \leq \lambda_2(\tau)$ the two eigenvalues of $\mathrm{Im}(\tau)$, and by $0 < m_1(\tau) \leq m_2(\tau)$ the successive minima of $\mathrm{Im}(\tau)$ on the lattice $\mathbb{Z}^2$. By [Str10, (5.4) p. 54], we always have

$$\frac{3}{4} m_1(\tau) m_2(\tau) \leq \det \mathrm{Im}(\tau) \leq m_1(\tau) m_2(\tau), \tag{6.7}$$

hence

$$\log \max\{\lambda_1(\tau)^{-1}, \lambda_2(\tau), m_1(\tau)^{-1}, m_2(\tau)\} = O(\Lambda(\tau)).$$

First, we detail the Minkowski reduction step.

Lemma 6.2.20. *There exist an algorithm and an absolute constant C such that the following holds. Let $\tau \in \mathbb{H}_2$ and $\varepsilon > 0$. Then, given an approximation of τ to precision $N \geq C(\Lambda(\tau) + |\log \varepsilon|)$, the algorithm computes a matrix $U \in \mathrm{SL}_2(\mathbb{Z})$ such that $U^t \,\mathrm{Im}(\tau) U$ is ε-Minkowski reduced within $O(\mathcal{M}(N) \log N)$ binary operations.*

Proof. Write $\mathrm{Im}(\tau) = R^t R$, and consider the matrix R' obtained by rounding the coefficients of $2^N R$ to the nearest integers. If C is chosen appropriately, then the matrix R' is still invertible. We apply a quasi-linear version

of Gauss's reduction algorithm [Sch91, Yap92], and obtain a Minkowski-reduced basis of the lattice $R'\mathbb{Z}^2$ within $O(\mathcal{M}(N)\log N)$ binary operations. By [Str10, Lem. 5.6], the base change matrix $U \in \mathrm{SL}_2(\mathbb{Z})$ satisfies

$$\log|U| = O(\Lambda(\tau)).$$

Therefore the matrix $U^t \operatorname{Im}(\tau) U$ is ε-Minkowski reduced provided that C is large enough. □

Then, we bound then precision losses during algorithm 6.2.19.

Lemma 6.2.21. *Let $\tau, \tau' \in \mathbb{H}_2$, and assume that there exists a matrix $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\tau' = \gamma\tau$. Then we have*

$$\log^+ \max\{|\gamma^*\tau|, |(\gamma^*\tau)^{-1}|\} = O\big(\max\{\Lambda(\tau), \Lambda(\tau')\}\big),$$
$$\log|\gamma| = O\big(\max\{\Lambda(\tau), \Lambda(\tau')\}\big).$$

Proof. Let R be a real 2×2 matrix such that $R^tR = \operatorname{Im}(\tau)$. Then we have

$$\operatorname{Im}(\tau') = (\gamma^*\tau)^{-t}\operatorname{Im}(\tau)(\gamma^*\overline{\tau})^{-1} = R'^t\overline{R'}$$

with $R' = R(\gamma^*\tau)^{-1}$. Since $|R| \le |\operatorname{Im}(\tau)|^{1/2}$ and $|R'| \le |\operatorname{Im}(\tau')|^{1/2}$, we obtain

$$|\gamma^*\tau| = |R'^{-1}R| \le 2\frac{|R'|}{\det(R')}|R|.$$

Hence $\log^+|\gamma^*\tau| = O(\max\{\Lambda(\tau), \Lambda(\tau')\})$, and a similar bound holds in the case of $(\gamma^*\tau)^{-1}$.

It remains to bound $|\gamma|$. If c, d denote the two lower blocks of γ, then $\operatorname{Im}(\gamma^*\tau) = c\operatorname{Im}(\tau)$. Therefore $\log^+|c| = O(\max\{\Lambda(\tau), \Lambda(\tau')\})$, and

$$\log^+|d| \le \log^+(|c\tau| + |\gamma^*\tau|) = O\big(\max\{\Lambda(\tau), \Lambda(\tau')\}\big).$$

Finally, we bound the upper blocks a and b of γ in a similar way, using the relation $a\tau + b = \tau'(c\tau + d)$. □

Lemma 6.2.22. *There exists an absolute constant C such that the following holds. Let $\tau \in \mathbb{H}_2$ and $\varepsilon > 0$, and assume that the precision during algorithm 6.2.19 remains greater than $|\log\varepsilon| + 1$. Then the number of iterations is $O(\Lambda(\tau))$. Moreover, during the algorithm, the quantities $|\log(|\det(\gamma^*\tau)|)|$, $\Lambda(\tau')$ and $\log|\gamma|$ remain in $O(\Lambda(\tau))$.*

Proof. The number of iterations is $O(\Lambda(\tau))$ by [Str10, Prop. 5.16]: observe that [Str10, Lem. 5.14 and 5.15] still apply, because $\det\operatorname{Im}(\tau')$ is strictly increasing in algorithm 6.2.19. The proof of [Str10, Lem. 5.17] also applies

to algorithm 6.2.19 with slightly worse constants. This shows that $\log|\tau'|$ and $\log|\det(\gamma^*\tau)|$ remain in $O(\Lambda(\tau))$.

During the algorithm, we also have $\log^+ m_2(\tau') = O(\Lambda(\tau))$ by [Str10, Lem. 5.12]. Moreover $\det \mathrm{Im}(\tau') \geq \det \mathrm{Im}(\tau)$, so

$$m_1(\tau')^{-1} \leq \frac{m_2(\tau')}{\det \mathrm{Im}(\tau')} \leq \frac{m_2(\tau')}{\det \mathrm{Im}(\tau)} \leq \frac{4m_2(\tau')}{3m_1(\tau)^2}$$

by (6.7). Therefore we also have $\Lambda(\tau') = O(\Lambda(\tau))$. The remaining bounds follow from lemma 6.2.21. □

We are now ready to prove the correctness and running time of algorithm 6.2.19.

Proposition 6.2.23. *There is an absolute constant C such that the following holds. Let $\tau \in \mathbb{H}_2$ and $\varepsilon > 0$. Then, given an approximation of τ to precision $N \geq C(\Lambda(\tau) + |\log \varepsilon|)$ as input, algorithm 6.2.19 does not run out of precision, and computes a matrix $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\gamma\tau \in \mathcal{F}_2^\varepsilon$ and $\log|\gamma| = O(\Lambda(\tau))$. It costs $O(\mathcal{M}(N)N\log N)$ binary operations.*

Proof. By lemma 6.2.22, there exists an absolute constant C' such that $\log|\gamma| \leq C'\Lambda(\tau)$ during the execution of algorithm 6.2.19 as long as the absolute precision remains at least $|\log\varepsilon| + 1$. Therefore, if C is chosen appropriately, step 4 in algorithm 6.2.19 ensures that the absolute precision is at least $|\log\varepsilon| + 1$ at every step. Hence the estimate on $\log|\gamma|$ and $\Lambda(\tau')$ remains valid until the end of the algorithm, and we can perform the approximate Minkowski reductions using lemma 6.2.20.

By lemma 6.2.22, there are $O(\Lambda(\tau))$ steps in algorithm 6.2.19, and by lemma 6.2.20, each step costs $O(\mathcal{M}(N)\log N)$ binary operations. Hence the cost is overall $O(\mathcal{M}(N)N\log N)$ binary operations. When the algorithm stops, the absolute precision is still greater than $|\log\varepsilon|+1$, hence the final τ' belongs to $\mathcal{F}_2^\varepsilon$. □

Given $\tau' \in \mathcal{F}_2^\varepsilon$, we can increase the imaginary parts of the coefficients slightly to obtain $\tau'' \in \mathcal{F}_2'$ such that

$$|\tau'' - \tau'| \leq C\varepsilon\,|\tau'|$$

for some absolute constant C. Theta constants at τ'' can be computed in quasi-linear time in the precision by proposition 6.2.17; they approximate the value of theta constants at τ'.

Corollary 6.2.24 (Under hypothesis 6.2.15)**.** *There exist an algorithm and an absolute constant C such that the following holds. Let $\tau \in \mathbb{H}_2$ and $N \geq 1$. Then, given an approximation of τ to precision $N + C\Lambda(\tau)$, the algorithm computes*

1. *a matrix* $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ *such that* $\log|\gamma| = O(\Lambda(\tau))$,
2. *a matrix* $\tau'' \in \mathcal{F}_2$ *such that* τ'' *is an approximation of* $\gamma\tau$ *to precision* N,
3. *an approximation of squares of theta constants at* $\gamma\tau$ *to precision* N,

within

$$O\big(\mathcal{M}(\Lambda(\tau))\Lambda(\tau)\log\Lambda(\tau) + \mathcal{M}(N)\log N\big)$$

binary operations.

Proof. Fix $\varepsilon = 0.01$, for instance. First, we apply proposition 6.2.23 to compute $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\gamma\tau \in \mathcal{F}_2^{\varepsilon}$, using $O\big(\mathcal{M}(\Lambda(\tau))\Lambda(\tau)\log\Lambda(\tau)\big)$ binary operations. Then, we recompute $\tau' := \gamma\tau$ to high precision, and reduce it further if necessary to land in $\mathcal{F}_2^{\varepsilon'}$ where $\varepsilon' = 2^{-N}\exp(-C\Lambda(\tau))$ for some appropriate constant C. This costs $O(\mathcal{M}(N + \Lambda(\tau)))$ binary operations. Finally we compute τ'' which satisfies conditions (6.3) and is close to τ'; the matrix τ'' is still an approximation of $\gamma\tau$, with $O(\Lambda(\tau))$ bits of precision lost. We output theta constants at τ'' to precision $N + O(1)$, which can be computed in time $O(\mathcal{M}(N)\log N)$ by proposition 6.2.17. □

6.3 Evaluating Hilbert modular equations

Let F be a real quadraic field. In this final section, we detail a heuristic algorithm to evaluate Hilbert modular equations of level β and their derivatives, where $\beta \in \mathbb{Z}_F$ is a totally positive prime of norm ℓ, given Igusa invariants (j_1, j_2, j_3) in a number field L lying on the Humbert surface. We let $R \in \mathrm{GL}_2(\mathbb{R})$ be the matrix chosen as in §2.3.4, defining the Hilbert embedding $H_R\colon \mathbb{H}_1^2 \to \mathbb{H}_2$.

The algorithm can be adapted to the case of Hilbert modular equations in Gundlach invariants for $F = \mathbb{Q}(\sqrt{5})$, and the case of Siegel modular equations. In both of these cases, using complete denominators for modular equations as in §5.1.4 allows us to design an algorithm with provably correct output.

6.3.1 Outline of the algorithm

A high-level view on the algorithm was already given at the beginning of chapter 6. In every complex embedding μ of L, we compute a period matrix $\tau_0 \in \mathcal{F}_2$ with Igusa invariants $(\mu(j_1), \mu(j_2), \mu(j_3))$, and invert the Hilbert embedding to find $\tau \in \mathbb{H}_1^2$ such that $H_R(\tau)$ is $\mathrm{Sp}_4(\mathbb{Z})$-equivalent to τ_0. Then we evaluate the Hilbert modular equations at τ using the

analytic formula given by definition 2.4.6. By the expression of Igusa invariants in terms of theta constants given in §2.3.2, this can be done by computing theta constants at all the matrices $H_R(\frac{1}{\beta}\eta\tau) \in \mathbb{H}_2$, where η runs through the set C_β^σ defined in §5.1.4. Similarly, the derivatives of modular equations at $(\mu(j_1), \mu(j_2), \mu(j_3))$ can be obtained from derivatives of theta constants at the matrices $H_R(\frac{1}{\beta}\eta\tau)$. Finally, we try to recognize the results as elements of the number field L.

The necessary precision to perform this algebraic reconstruction is controlled by the height of the modular equations evaluated at (j_1, j_2, j_3), which is $O_F\big(\ell \log \ell + \ell\, h(j_1, j_2, j_3)\big)$ by the results of chapter 5. However, we have no general formula for the implied constant in terms of the discriminant of F. Therefore the algebraic reconstruction step can only be heuristic in general. In practice, one could work at a high enough precision N so that an algebraic reconstruction succeeds, and check that the result remains identical when the computations are done at precision $2N$.

The situation improves if we can build a complete denominator for the modular equations, as we did in §5.1.4 in the case of Siegel modular equations, and Hilbert modular equations in Gundlach invariants for $F = \mathbb{Q}(\sqrt{5})$. In this case, we separately evaluate numerators and denominators of the modular equations using analytic formulæ. Then we only have to recognize algebraic *integers* from their images under the complex embeddings of L, and we can apply the methods of §6.1.3 to obtain a provably correct output.

Assume, for instance, that we want to evaluate Hilbert modular equations of level β for $F = \mathbb{Q}(\sqrt{5})$ at a given pair of Gundlach invariants $(g_1, g_2) \in L$. We compute a point $\tau \in \mathbb{H}_1^2$ realizing these Gundlach invariants by inverting the Hilbert embedding as above. By eq. (5.3), the denominator $D_\beta(\tau)$ has the analytic expression

$$D_\beta(\tau) = 2^{w_\beta} \frac{G_2(\tau)^{2\lfloor w_\beta/6 \rfloor}}{F_{10}(\tau)^a G_2(\tau)^b} \prod_{\eta \in C_\beta^\sigma} (\eta^*\tau)^{-10} F_{10}(\tfrac{1}{\beta}\eta\tau), \tag{6.8}$$

where $w_\beta = 10(2\ell + 2)$, and the integers $a, b \geq 0$ are chosen such that

$$2\lfloor w_\beta/6 \rfloor + w_\beta = 10a + 2b \quad \text{with } 0 \leq b \leq 4.$$

Recall that for every $\tau = (\tau_1, \tau_2) \in \mathbb{H}_1^2$, for every $\eta = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in C_\beta$, and for every $\lambda \in F$, we write $\lambda\tau = (\lambda\tau_1, \overline{\lambda}\tau_2)$ and

$$\eta^*\tau = (c\tau_1 + d)(\overline{c}\tau_2 + \overline{d}), \quad (\eta\sigma)^*\tau = \eta^*\big((\tau_2, \tau_1)\big).$$

By definition of the Hilbert modular equations, we also have the formulæ:

$$\begin{aligned}
D_\beta(t)\Psi_{\beta,1}\big(\mu(g_1),\mu(g_2),Y\big) &= D_\beta(t)\prod_{\eta\in C_\beta^\sigma}\big(Y-g_1(\tfrac{1}{\beta}\eta t)\big)\\
&= 2^{w_\beta}\frac{G_2(t)^{2\lfloor w_\beta/6\rfloor}}{F_{10}(t)^a G_2(t)^b}\prod_{\eta\in C_\beta^\sigma}(\eta^* t)^{-10}\big(F_{10}(\tfrac{1}{\beta}\eta t)Y-G_2^5(\tfrac{1}{\beta}\eta t)\big), \quad \text{and}\\
D_\beta(t)\Psi_{\beta,2}\big(\mu(g_1),\mu(g_2),Y\big) &= D_\beta(t)\sum_{\eta\in C_\beta^\sigma} g_2(\tfrac{1}{\beta}\eta t)\prod_{\eta'\neq\eta}\big(Y-g_1(\tfrac{1}{\beta}\eta' t)\big)\\
&= 2^{w_\beta}\frac{G_2(t)^{2\lfloor w_\beta/6\rfloor}}{F_{10}(t)^a G_2(t)^b}\sum_{\eta\in C_\beta^\sigma}(\eta^* t)^{-10}G_2^2(\tfrac{1}{\beta}\eta t)F_6(\tfrac{1}{\beta}\eta t)\\
&\qquad\qquad\cdot\prod_{\eta'\neq\eta}(\eta'^* t)^{-10}\big(F_{10}(\tfrac{1}{\beta}\eta t)Y-G_2^5(\tfrac{1}{\beta}\eta t)\big).
\end{aligned} \tag{6.9}$$

By proposition 5.1.13, these quantities (6.8) and (6.9) are the values, in each complex embedding μ of L, of certain elements

$$D\in\mathbb{Z}[g_1,g_2], \quad \text{and} \quad P_1,P_2\in\mathbb{Z}[g_1,g_2,Y].$$

If g_1 and g_2 are integers in L, then D, P_1 and P_2 are elements of $\mathbb{Z}_L$ as well. In the general case, let $c\in L$ be such that both cg_1 and cg_2 are integers; then, by proposition 5.1.17, it is sufficient to multiply D, P_1, and P_2 by $c^{10(\ell+1)/3}$ to obtain algebraic integers. Finally, we output

$$\Psi_{\beta,1}(g_1,g_2,Y)=P_1/D \quad \text{and} \quad \Psi_{\beta,2}(g_1,g_2,Y)=P_2/D.$$

The same idea of using D_β to obtain algebraic integers works for derivatives of modular equations too. The proof of correctness follows two steps.

1. In the first step, we assume that we are given approximations of the values $\mu(j_k)$ for $1\le k\le 3$ in a fixed complex embedding μ of L. We analyze the cost and precision losses in the various steps of the algorithm: computing τ_0 (§6.3.2), inverting the Hilbert embedding (§6.3.3), reducing the matrices $H_R(\frac{1}{\beta}\eta\tau)$ to the fundamental domain $\mathcal{F}_2$, computing theta constants, and evaluating modular equations analytically (§6.3.4). Towards the end, we restrict to the formulæ (6.8) and (6.9) for the sake of brevity; otherwise, the computations apply to all cases of modular equations in Igusa invariants discussed above.

2. In the second step (§6.3.5), we focus on the case of Hilbert modular equations for $\mathbb{Q}(\sqrt{5})$ in Gundlach invariants. We consider each of

the two situations of §6.1.3 for the reconstruction of integers in L: in each situation, we compute the required precision to recognize D, P_1 and P_2 uniquely, and conclude on the cost of the whole algorithm. The computations in the case of Siegel modular equations are very similar, and we simply point out the differences in running time.

6.3.2 Computing period matrices

Let d_L be the degree of L over $\mathbb{Q}$, and fix a complex embedding μ of L. In this section, we investigate the first step of the algorithm to evaluate modular equations: given Igusa invariants $(j_1, j_2, j_3) \in L^3$, compute a period matrix $\tau \in \mathcal{F}_2$ with Igusa invariants $(\mu(j_1), \mu(j_2), \mu(j_3))$. We assume that $j_3 \neq 0$: otherwise τ is not uniquely determined.

During the algorithm, we will consider a finite family $\Theta(j_1, j_2, j_3)$ of algebraic numbers constructed from j_1, j_2 and j_3. More precisely we consider Θ as a finite family of polynomials $Q \in \mathbb{Q}[X_1, \dots, X_n, Y]$, and the algebraic numbers that we consider are constructed as roots of polynomials of the form $Q(j_1, j_2, j_3, x_4 \dots, x_n, Y)$, where $x_4, \dots, x_n$ are previously constructed elements of $\Theta(j_1, j_2, j_3)$. When presented in this way, Θ does not depend on L, j_1, j_2, or j_3. As a toy example, consider the family Θ consisting of the single polynomial $X_1 - Y^2$; then $\Theta(j_1, j_2, j_3) = \{\sqrt{j_1}\}$. We call Θ a *finite recipe of algebraic extensions*.

If H denotes the height of (j_1, j_2, j_3), then the height of all elements of $\Theta(j_1, j_2, j_3)$ is in $O_\Theta(H)$ by propositions 4.3.1 and 4.3.3.

For every complex embedding μ of L, we define $B_{\Theta,\mu} \geq 0$ as the minimal real number such that

$$\left|\log(|\widetilde{\mu}(\theta)|)\right| \leq B_{\Theta,\mu}$$

for each nonzero $\theta \in \Theta(j_1, j_2, j_3)$ and each extension $\widetilde{\mu}$ of μ to the number field $L(\Theta(j_1, j_2, j_3))$. We can take $B_{\Theta,\mu} = O_\Theta(d_L H)$; moreover the sum of the bounds $B_{\Theta,\mu}$ over all the complex embeddings of L is also $O_\Theta(d_L H)$. A typical example of how we use $B_{\Theta,\mu}$ is as follows.

Proposition 6.3.1. *There exist an algorithm and a finite recipe of algebraic extensions Θ such that the following holds. Let L be a number field, let $j_1, j_2, j_3 \in L$ such that $j_3 \neq 0$, let μ be a complex embedding of L, and define $B_{\Theta,\mu}$ as above. Let $N \geq 1$. Then, given approximations of $\mu(j_k)$ for $1 \leq k \leq 3$ to precision N, the algorithm computes a genus 2 hyperelliptic curve $\mathcal{C}$ over $\mathbb{C}$ with Igusa invariants $\mu(j_1), \mu(j_2), \mu(j_3)$ within $O\big(\mathcal{M}(N + B_{\Theta,\mu})\big)$ binary operations, with a precision loss of $O(B_{\Theta,\mu})$ bits.*

Proof. Use Mestre's algorithm [Mes91]. This algorithm involves $O(1)$ elementary operations with complex algebraic numbers constructed from

the $\mu(j_k)$ for $1 \leq k \leq 3$, hence the estimates on the running time and precision losses follow from proposition 6.1.2. □

We prove that any period matrix $\tau \in \mathcal{F}_2$ of $\mathcal{C}$ is bounded in terms of $B_{\Theta,\mu}$ for some Θ. This is done by looking at theta quotients at τ.

Lemma 6.3.2. *There exists a finite recipe of algebraic extensions Θ such that the following holds. Let $\mathcal{C}$ be as in proposition 6.3.1, and let $\tau \in \mathcal{F}_2$ be a period matrix of $\mathcal{C}$. Then we have*

$$|\tau| = O(B_{\Theta,\mu}).$$

Proof. By Thomae's formulæ [Mum84, Thm. IIIa.8.1], the theta quotients $\theta_j(\tau)/\theta_0(\tau)$ for $j \in [\![1, 15]\!]$ are algebraic numbers constructed from the coefficients of $\mathcal{C}$, and are nonzero for $j \in \{0, 1, 2, 3, 4, 6, 8, 12\}$. Therefore, we can choose Θ in such a way that

$$\left|\log\left(|\theta_j(\tau)/\theta_0(\tau)|\right)\right| \leq B_{\Theta,\mu} \quad \text{for } j \in \{4, 8\}.$$

Write $\tau = \left(\begin{smallmatrix} z_1 & z_3 \\ z_3 & z_2 \end{smallmatrix}\right)$. By lemma 6.2.10, the real numbers $\mathrm{Im}(z_1)$ and $\mathrm{Im}(z_2)$ are in $O(B_{\Theta,\mu})$ as well, hence also $|\mathrm{Im}(z_3)|$ because $\det \mathrm{Im}(\tau) > 0$. Since $|\mathrm{Re}(\tau)| \leq 1/2$, the result follows. □

Proposition 6.3.3 (Under hypothesis 6.2.15)**.** *There exist an algorithm and a finite recipe of algebraic extensions Θ such that the following holds. Let L be a number field, let $j_1, j_2, j_3 \in L$ be such that $j_3 \neq 0$, let μ be a complex embedding of L, and define $B_{\Theta,\mu}$ as above. Let $N \geq 1$. Then, given approximations of $\mu(j_k)$ for $1 \leq k \leq 3$ to precision N, the algorithm computes a matrix $\tau \in \mathcal{F}_2$ such that the Igusa invariants at τ are the $\mu(j_k)$ for $1 \leq k \leq 3$. The algorithm involves $O\big(\mathcal{M}(N + B_{\Theta,\mu}) \log(N + B_{\Theta,\mu})\big)$ binary operations, and a precision loss of $O(\log N + B_{\Theta,\mu} \log B_{\Theta,\mu})$ bits.*

Proof. First, we compute a complex curve $\mathcal{C}$ as in proposition 6.3.1. Then, by Thomae's formulæ, there is a finite number of possibilities for the values of squares of theta quotients at τ; one of them corresponds to an actual matrix $\tau \in \mathcal{F}_2$, and the others correspond to other elements in the orbit $\mathrm{Sp}_4(\mathbb{Z})\tau$. When we run the algorithm of proposition 6.2.14 on these inputs, we may discard all resulting period matrices that do not belong to $\mathcal{F}_2$. In order to distinguish between the remaining possible values of τ, it is usually enough to compute theta constants to precision $O(1)$ using the naive algorithm, and match with the input. In extreme cases, we may resort to computing Igusa invariants at all remaining possible values of τ to precision $O(N + B_\mu)$, using $O\big(\mathcal{M}(N + B_\mu) \log(N + B_\mu)\big)$ binary operations, by proposition 6.2.17; this is where we use hypothesis 6.2.15. □

6.3.3 Inverting the Hilbert embedding

Let F be a real quadratic field, and let $R \in \mathrm{GL}_2(\mathbb{R})$ be a matrix defining a Hilbert embedding H_R, as in §2.3.3. Recall that for every $\tau \in \mathbb{H}_2$, we denote by $0 < \lambda_1(\tau) \leq \lambda_2(\tau)$ the two eigenvalues of $\mathrm{Im}(\tau)$.

Lemma 6.3.4. *Let F and R be as above. Then there exists a constant $C > 0$ depending on F and R such that for every $\tau = (\tau_1, \tau_2) \in \mathbb{H}_1^2$, we have*

$$\begin{aligned} \tfrac{1}{C}\lambda_1(H_R(\tau)) \leq \min\{\mathrm{Im}(\tau_1), \mathrm{Im}(\tau_2)\} \leq C\lambda_1(H_R(\tau)), \\ \tfrac{1}{C}\lambda_2(H_R(\tau)) \leq \max\{\mathrm{Im}(\tau_1), \mathrm{Im}(\tau_2)\} \leq C\lambda_2(H_R(\tau)). \end{aligned}$$

Proof. Use the definition (2.12) of the Hilbert embedding. □

Proposition 6.3.5. *Let F be a real quadratic field, and let R be as above. Then there exist an algorithm, a constant $C > 0$, and a finite recipe of algebraic extensions Θ depending on F and R such that the following holds. Let L be a number field, let $j_1, j_2, j_3 \in L$ be such that $j_3 \neq 0$, let μ be a complex embedding of L, and define $B_{\Theta,\mu}$ as in §6.3.2. Let $\mathcal{C}$ be a genus 2 hyperelliptic curve over $\mathbb{C}$ with Igusa invariants $\mu(j_1), \mu(j_2), \mu(j_3)$, and assume that $\mathrm{Jac}(\mathcal{C})$ has real multiplication by $\mathbb{Z}_F$. Let $\tau_0 \in \mathcal{F}_2$ be a period matrix of $\mathcal{C}$. Then there exists $\tau = (\tau_1, \tau_2) \in \mathbb{H}_1^2$ such that $H_R(\tau) \in \mathbb{H}_2$ is a period matrix of $\mathcal{C}$, and*

$$|\log(\mathrm{Im}\,\tau_i)| \leq CB_{\Theta,\mu} \quad \text{for } i = 1, 2.$$

Moreover, given an approximation of τ_0 to precision $N + CB_{\Theta,\mu}$, the algorithm computes τ to precision N within $O_F\big(\mathcal{M}(N + B_{\Theta,\mu}) \log B_{\Theta,\mu}\big)$ binary operations.

Proof. By lemma 6.3.2, if Θ is well chosen, we have

$$|\tau_0| = O(B_{\Theta,\mu}).$$

The result would be obvious from lemma 6.3.4 if there existed $\tau \in \mathbb{H}_1^2$ such that $\tau_0 = H_R(\tau)$, but this is not always the case. In general, by [BW03, Lem. 4.1], there exist coprime integers a, b, c, d, and e such that

$$b^2 - 4ac - 4de = \Delta \quad \text{and} \quad az_1(\tau_0) + bz_3(\tau_0) + cz_2(\tau_0) + d\det(\tau_0) + e = 0, \tag{6.10}$$

where Δ denotes the discriminant of F. Finding the integers a, b, c, d, e will allow us to construct a conjugate of τ_0 under $\mathrm{Sp}_4(\mathbb{Z})$ which admits a preimage under H_R.

We claim that the heights of a, b, c, d, e are in $O_F(B_{\Theta,\mu})$ for some choice of Θ. We prove this by comparing the analytic and rational representations (see [BL04, §1.2]) of the endomorphism $\sqrt{\Delta}$ on the complex abelian variety

$$A(\tau_0) = \mathbb{C}^2/(\mathbb{Z}^2 \oplus \tau_0\mathbb{Z}^2).$$

The rational representation of every endomorphism f in the image of $\mathbb{Z}_F$ inside $\mathrm{End}(A(\tau_0))^\dagger$ is of the form

$$\rho_{R,\tau_0}(f) = \begin{pmatrix} n & ma & 0 & md \\ -mc & mb+n & -md & 0 \\ 0 & me & n & -mc \\ -me & 0 & ma & mb+n \end{pmatrix} \qquad \text{for some } m, n \in \mathbb{Z},$$

by [BW03, Cor. 4.2]. On the other hand, the analytic representation of $\sqrt{\Delta}$, denoted by $\rho_{A,\tau_0}(\sqrt{\Delta})$, can be computed as follows. Let $\omega = (\omega_1, \omega_2)$ be a basis of differential forms on $A(\tau_0)$ such that $\mathrm{Sym}^2(\omega)$ corresponds by the Kodaira–Spencer isomorphism to a deformation of $A(\tau_0)$ along the Humbert surface. Then, by proposition 3.3.10, the matrix of $(\sqrt{\Delta})^*$ in the basis ω is of the form

$$\pm \begin{pmatrix} \sqrt{\Delta} & 0 \\ 0 & -\sqrt{\Delta} \end{pmatrix}.$$

Algorithm 3.3.11 shows that such a basis ω exists; moreover the base change matrix m between (dz_1, dz_2) and ω can be chosen in such a way that

$$\log \max\{|m|, |m^{-1}|\} = O_F(B_{\Theta,\mu})$$

after extending Θ in a suitable way. This proves that the analytic representation of the endomorphism $\sqrt{\Delta}$ on $A(\tau_0)$ satisfies

$$\log^+ |\rho_{A,\tau_0}(\sqrt{\Delta})| = O_F(B_{\Theta,\mu}).$$

For every $f \in \mathrm{End}(A(\tau_0))$, the rational and analytic representations of f are related by the following formula [BL04, Rem. 8.14]:

$$\rho_{A,\tau_0}(f)\,(\tau_0\ I_2) = (\tau_0\ I_2)\,\rho_{R,\tau_0}(f).$$

Taking imaginary parts, we find that there exist $m, n \in \mathbb{Z}$ such that

$$\mathrm{Im}(\tau_0) \begin{pmatrix} n & ma \\ -mc & mb+n \end{pmatrix} = \mathrm{Im}\big(\rho_{A,\tau_0}(\sqrt{\Delta})\tau_0\big), \quad \text{and}$$
$$\mathrm{Im}(\tau_0) \begin{pmatrix} 0 & md \\ -md & 0 \end{pmatrix} = \mathrm{Im}\big(\rho_{A,\tau_0}(\sqrt{\Delta})\big).$$

Therefore the heights of a, b, c, d, m and n are in $O_F(B_{\Theta,\mu})$. The same is true for the integer e by eq. (6.10). This proves our claim.

The algorithm to compute τ is the following. We compute the integers a, b, c, d and e using $O_F(\mathcal{M}(B_{\Theta,\mu}) \log B_{\Theta,\mu})$ binary operations with a quasi-linear version of the LLL algorithm [NSV11, NS16], using eq. (6.10). Then we apply the algorithm of [BW03, Prop. 4.5] to compute $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\gamma\tau_0$ lies in the image of H_R; the matrix γ has a simple expression in terms of a, b, c, d, and e, hence we also have

$$\log|\gamma| = O_F(B_{\Theta,\mu}).$$

By proposition 6.2.9, we also have

$$\Lambda(\gamma\tau_0) = O_F(B_{\Theta,\mu}),$$

so the result follows from lemma 6.3.4. □

6.3.4 Analytic evaluation of modular equations

We return to the case of Hilbert modular equations in Gundlach invariants for $F = \mathbb{Q}(\sqrt{5})$. Let L be a number field, let μ be a complex embedding of L, and let $(g_1, g_2) \in L$. We keep the notation of §6.3.1. In the following proposition, we detail the algorithm to evaluate the numerator and denominator of Hilbert modular equations of level β at $(\mu(g_1), \mu(g_2))$. We assume that $g_1 \neq 0$, so that the associated Igusa invariants (j_1, j_2, j_3) satisfy $j_3 \neq 0$.

In order to avoid complicated expressions, we hide logarithmic factors in the $\widetilde{O}$ notation from now on. Actually $\widetilde{O}(T)$ is always $O(\mathcal{M}(T \log T) \log T)$.

Proposition 6.3.6 (Under hypothesis 6.2.15)**.** *There exist an algorithm, a finite recipe of algebraic extensions Θ, and a constant $C > 0$ such that the following holds. Let L be a number field, let μ be a complex embedding of L, let $(g_1, g_2) \in L^2$ such that $g_1 \neq 0$, and define $B_{\Theta,\mu}$ as in §6.3.2. Finally, let $N \geq C(B_{\Theta,\mu} \log B_{\Theta,\mu} + \log \ell)$.*

Then, given approximations of $\mu(g_1)$ and $\mu(g_2)$ to precision N, the algorithm computes $\mu\big(D_\beta(g_1, g_2)\big) \in \mathbb{C}$ and $\mu\big(D_\beta \Psi_{\beta,m}(g_1, g_2, Y)\big) \in \mathbb{C}[Y]$ for $m \in \{1, 2\}$ within $\widetilde{O}(\ell B_{\Theta,\mu}^2 + \ell N)$ binary operations, with a precision loss of $\widetilde{O}(\ell B_{\Theta,\mu} + \log N)$ bits.

Proof. We first compute the associated Igusa invariants $\mu(j_k)$ for $1 \leq k \leq 3$ using proposition 2.3.16. Note that $j_3 \neq 0$. Using proposition 6.3.3, we compute a matrix $\tau_0 \in \mathcal{F}_2$ having these Igusa invariants in $\widetilde{O}(N + B_{\Theta,\mu})$ binary operations, with a precision loss of $O(\log N + B_{\Theta,\mu} \log B_{\Theta,\mu})$ bits, for

a suitable choice of Θ. Then, using proposition 6.3.5, we compute $\tau \in \mathbb{H}_1^2$ such that τ has Gundlach invariants (g_1, g_2), and

$$|\log(\operatorname{Im} \tau_i)| = O_F(B_{\Theta,\mu}), \quad \text{for each } i \in \{1, 2\}.$$

This costs $\widetilde{O}(N + B_{\Theta,\mu})$ binary operations, with a precision loss of $O(B_{\Theta,\mu})$ bits. The next step is to compute the points $\frac{1}{\beta}\eta\tau \in \mathbb{H}_1^2$ for all $\eta \in C_\beta^\sigma$: this costs $\widetilde{O}\big(\ell(N + B_{\Theta,\mu})\big)$ binary operations, with a precision loss of $O(\log \ell)$ bits. By lemma 6.3.4, we have for every $\eta \in C_\beta^\sigma$:

$$\Lambda\big(H_R(\tfrac{1}{\beta}\eta\tau)\big) = O(B_{\Theta,\mu} + \log \ell).$$

Then, we reduce the matrices $H_R(\frac{1}{\beta}\eta\tau)$ to the fundamental domain using corollary 6.2.24: we compute matrices $\tau_\eta \in \mathcal{F}_2$ and $\gamma_\eta \in \mathrm{Sp}_4(\mathbb{Z})$ such that

$$\gamma_\eta(\tfrac{1}{\beta}\eta\tau) = \tau_\eta, \qquad \text{for every } \eta \in C_\beta^\sigma,$$

with a precision loss of $O(B_{\Theta,\mu} + \log \ell)$ bits. We also compute squares of theta constants at every τ_η. This can be done in $\widetilde{O}(\ell B_{\Theta,\mu}^2 + \ell N)$ binary operations. Moreover, we have

$$\log |\gamma_\eta| \in O(B_{\Theta,\mu} + \log \ell).$$

Definition 2.3.4 yields the values of the Siegel modular forms h_4, h_6, h_{10}, h_{12} at all the matrices τ_η using $O(\ell)$ binary operations, and a precision loss of $O(1)$ bits.

At the end, we evaluate $D_\beta(\tau)$ using eq. (6.8), and the equality

$$F_{10}(\tfrac{1}{\beta}\eta\tau) = \big(\det \gamma_\eta^*(\tfrac{1}{\beta}\eta\tau)\big)^{-10} F_{10}(\tau_\eta),$$

for every $\eta \in C_\beta^\sigma$. By lemma 6.2.21, $O\big(\ell(B_{\Theta,\mu} + \log \ell)\big)$ bits of precision are lost in this computation; the total cost of computing $D_\beta(t)$ is $\widetilde{O}\big(\ell(N + B_{\Theta,\mu})\big)$ binary operations.

The polynomials $D_\beta(\tau)\mu\big(\Psi_{\beta,m}(g_1, g_2, Y)\big)$ for $m \in \{1, 2\}$, given by the formula (6.9), are computed using product trees as in lemma 6.1.3. This also costs $\widetilde{O}\big(\ell(N + B_{\Theta,\mu})\big)$ binary operations, and adds a precision loss of $O\big(\ell(B_\sigma + \log \ell)\big)$ bits. We conclude by summing precision losses and binary costs of each step. □

In the case of Siegel modular equations of level ℓ, the complexity and precision loss estimates are similar to those in proposition 6.3.6, with each occurence of ℓ replaced by ℓ^3. A further remark is that in the Siegel case, we can avoid costly reductions to $\mathcal{F}_2$ using the matrices η^R from the proof of proposition 5.1.10. This is especially attractive since this reduction step is the only one whose cost is not quasi-linear in the output size.

6.3.5 Algebraic reconstruction

Recall that we fix $F = \mathbb{Q}(\sqrt{5})$. Once the Hilbert modular equations of level β and their denominators have been computed at (g_1, g_2) in every complex embedding of L, we recognize their coefficients as algebraic numbers using the methods of §6.1.3. We present two results, one in the case of lifts from a finite field, and one in the case of a general number field.

In the case of a finite field, we are given a prime power $q = p^d$, and a monic polynomial $P \in \mathbb{Z}[X]$ of degree d, irreducible modulo p. We let $M \geq 1$ be such that $\log |P| \leq M$. We assume that a black box provides us with approximations of the roots of P to any desired precision. Then we represent elements of $\mathbb{F}_q$ as elements of $\mathbb{F}_p[X]/(P)$.

Proposition 6.3.7 (Under the conditions of the previous paragraph, and hypothesis 6.2.15)**.** *There exists an algorithm such that the following holds: given $g_1, g_2 \in \mathbb{F}_q$ such that $g_1 \neq 0$ and $D_\beta(g_1, g_2) \neq 0$, the algorithm computes the polynomials $\Psi_{\beta,m}(g_1, g_2) \in \mathbb{F}_q[Y]$ for $m \in \{1, 2\}$ within*

$$\widetilde{O}(\ell d^2 \log^2 p + \ell d^4 M^2 + \ell^2 d \log p + \ell^2 d^2 M)$$

binary operations.

If $dM = O(\log p)$, and if moreover $\ell = O(\log q)$, then the cost estimate simplifies to $\widetilde{O}(\log^3 q)$ binary operations. If $q = p$ is prime (i.e. $d = 1$), then the cost estimate simplifies to $\widetilde{O}(\ell \log^2 p + \ell^2 \log p)$ binary operations. Theorem 6.0.1 stated at the beginning of this chapter is the analogue of proposition 6.3.7 in the case of $\mathbb{F}_p$ for Siegel modular equations, where we replace each occurence of ℓ by ℓ^3.

Proof. Let L be the number field $\mathbb{Q}[X]/(P)$, and let α be a root of P in L. We lift g_1 and g_2 to elements of $\mathbb{Z}[\alpha]$ in such a way that the height of their coefficients is bounded above by $\log p$. Then we have

$$h(\alpha) \leq M + \log 2, \quad \text{and}$$
$$\max\{h(g_1), h(g_2)\} \leq \log(p) + dh(\alpha) + \log(d) = O(dM + \log p).$$

Since D_β and the coefficients of $\Psi_{\beta,m}$ for $m \in \{1, 2\}$ are elements of $\mathbb{Z}[g_1, g_2]$ of degree $O(\ell)$ and height $O(\ell \log \ell)$ by theorem 5.0.1, the algebraic integers we have to recognize are all elements of $\mathbb{Z}[\alpha]$, and the height of their coefficients is $O(\ell \log \ell + \ell dM + \ell \log p)$. By proposition 6.1.5, we can recognize each coefficient within $\widetilde{O}(\ell d^2 M + \ell d \log p)$ binary operations, provided that its values in every complex embedding of L are computed to precision at least $C(\ell \log \ell + \ell dM + \ell \log p)$, where C is some absolute constant.

Let μ be a complex embedding of L, and start at precision $N \geq 1$. Then $\mu(g_1)$ and $\mu(g_2)$ are obtained by replacing α by one of the complex roots of P: this can be done within $\widetilde{O}(d(M+N))$ binary operations, and a precision loss of $O(dM + \log p)$ bits, via Horner's algorithm. Then we run the algorithm of proposition 6.3.6, for each complex embedding μ of L. It is enough to choose N in

$$\widetilde{O}(\ell dM + \ell \log p + \ell B_{\Theta,\mu}),$$

for a cost of $\widetilde{O}(\ell B_{\Theta,\mu}^2 + \ell N)$ binary operations. Since we have

$$\sum_{\mu} B_{\Theta,\mu} = O(d \log p + d^2 M),$$

and each $B_{\Theta,\mu}$ is in $O(d \log p + d^2 M)$, the total cost of analytic evaluations over all embeddings is

$$\widetilde{O}(\ell d^2 \log^2 p + \ell d^4 M^2 + \ell^2 d \log p + \ell^2 d^2 M)$$

binary operations, and dominates the cost of algebraic reconstruction. □

If $g_1, g_2 \in \mathbb{Z}$ are small integers, then the complexity of evaluating modular equations is quasi-linear in the output size.

Proposition 6.3.8 (Under hypothesis 6.2.15)**.** *There exists an algorithm such that the following holds. Given and $g_1, g_2 \in \mathbb{Z}$ such that*

$$\max\{|g_1|, |g_2|\} \in O(1) \quad \textit{and} \quad g_1 D_\beta(g_1, g_2) \neq 0 \bmod p,$$

the algorithm computes the polynomials $\Psi_{\beta,m}(g_1, g_2) \in \mathbb{Q}[Y]$ for $m \in \{1, 2\}$ using a total of $O(\mathcal{M}(\ell^2 \log \ell) \log \ell)$ binary operations.

Proof. In this case, we have $B_{\Theta,\mu} = O(1)$. It is sufficient to round the result of proposition 6.3.6 with $N = C\ell \log \ell$, where C is an absolute constant, to the nearest integers. □

From another point of view, the complexity of evaluating Hilbert modular equations for F over a number field L can be bounded in terms of the discriminant Δ_L of L and the height of the operands. We assume that an LLL-reduced basis of $\mathbb{Z}_L$ has been precomputed. Moreover, if m_L is the matrix defined in proposition 6.1.6, we assume that a black box provides us with the coefficients of m_L^{-1} to any desired precision.

Proposition 6.3.9 (Under the conditions from the previous paragraph, and hypothesis 6.2.15)**.** *There exists an algorithm such that the following holds. Let $H \geq 1$, and let $g_1, g_2 \in L$ given as quotients of integers of height at most H such that $g_1 D_\beta(g_1, g_2) \neq 0$. Then the algorithm computes $\Psi_{\beta,m}(g_1, g_2) \in L[Y]$ for $m \in \{1, 2\}$ within*

$$\widetilde{O}(d^2\ell \log \Delta_L + d^3\ell + \ell d^2 H^2 + \ell^2 d^2 H + d^4)$$

binary operations.

In the case $L = \mathbb{Q}$, the cost estimate simplifies to $\widetilde{O}(\ell H^2 + \ell^2 H)$ binary operations.

Proof. For simplicity, assume that g_1 and g_2 are actually integers: in the general case we multiply D_β by an appropriate power of a common denominator of g_1 and g_2 in $\mathbb{Z}_L$.

We know that $D_\beta(g_1, g_2)$ and the coefficients of $D_\beta \Psi_{\beta,k}(g_1, g_2)$ are polynomials in $\mathbb{Z}[g_1, g_2]$ of degree $O(\ell)$ and height $O(\ell \log \ell)$, by theorem 5.0.1. Therefore they are algebraic integers of height $\widetilde{O}(\ell H)$. By proposition 6.1.6, we can recognize each coefficient within $\widetilde{O}(d^2\ell H + d^2 \log \Delta_L + d^4)$ binary operations, provided that complex approximations are computed at a precision N high enough; it is enough to take N in $\widetilde{O}(\log \Delta_L + d\ell H + d^2)$.

In order to obtain these approximations, we run the algorithm of proposition 6.3.6 in each complex embedding μ of L. For each embedding μ, the starting precision is chosen in $\widetilde{O}(\log \Delta_L + d\ell H + \ell B_{\Theta,\mu} + d^2)$, for a suitable recipe of algebraic extensions Θ. Therefore the cost to compute the required complex approximations in the embedding μ is

$$\widetilde{O}(\ell B_{\Theta,\mu}^2 + \ell \log \Delta_L + d\ell^2 H + \ell^2 B_{\Theta,\mu} + \ell d^2)$$

binary operations. The sum of the bounds $B_{\Theta,\mu}$ is in $O(dH)$, as well as each individual $B_{\Theta,\mu}$. Therefore, the total cost of computing complex approximations in all embeddings is $\widetilde{O}(\ell d^2 H^2 + \ell d \log \Delta_L + \ell^2 d^2 H)$ binary operations. □

The complexity results of propositions 6.3.7 and 6.3.9 are not entirely satisfactory: the dependence on $\log p$ in the finite field case, and on H in the number field case, is quadratic. This comes from the fact that the reduction algorithm to the Siegel fundamental domain (proposition 6.2.23) is quasi-quadratic in $\Lambda(\tau)$, and not quasi-linear. Reduction to the fundamental domain for elements of $\mathbb{H}_2$ is essentially equivalent to lattice reduction for dimension 4 *symplectic* lattices; it is likely that this reduction can be done in quasi-linear time, as for Gauss's algorithm [Sch91, Yap92] and for

the LLL algorithm [NSV11, NS16]. Another possibility to improve the complexities in propositions 6.3.7 and 6.3.9 would be to try to adapt the reduction algorithm for the matrices $H_R(\frac{1}{\beta}\eta\tau) \in \mathbb{H}_2$ for $\eta \in C_\beta^\sigma$, using the known matrix $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\gamma H_R(\tau) \in \mathcal{F}_2$.

Chapter 7

Elkies's method for abelian surfaces

The previous chapters presented several results and algorithms about higher-dimensional modular equations, and in particular modular equations of Siegel and Hilbert type in dimension 2. We know how to use them to compute isogenies between abelian surfaces (chapter 3), we can bound their degrees and heights (chapter 5), and finally we have efficient algorithms to evaluate them (chapter 6). Thus, all the necessary ingredients to generalize Elkies's method to principally polarized abelian surfaces are gathered. This short chapter details the resulting algorithm, building upon previous works on Schoof's and Atkin's methods for abelian surfaces [GH00, GS12, GKS11, BGLG+16]. In particular, we obtain a heuristic point-counting algorithm for principally polarized abelian surfaces with real multiplication by a fixed quadratic field F which, up to constants depending on F, has the same asymptotic complexity as the SEA algorithm for elliptic curves.

7.1 The characteristic polynomial of Frobenius

Let p be a large prime, let q be a power of p, and let A be a principally polarized abelian surface over $\mathbb{F}_q$. Let π_A be the Frobenius endomorphism of A, and let $\chi_A \in \mathbb{Z}[X]$ be its characteristic polynomial. Let $\ell \neq p$ be a prime. Then χ_A is the characteristic polynomial of π_A as an endomorphism of the ℓ-adic Tate module $T_\ell(A)$. In particular, $\chi_A \bmod \ell$ is the characteristic polynomial of π_A as an endomorphism of $A[\ell] \simeq (\mathbb{Z}/\ell\mathbb{Z})^4$.

Our goal, as in Elkies's method for elliptic curves (§1.2.2), is to compute $\chi_A \bmod \ell$ using subgroups of $A[\ell]$ obtained as kernels of isogenies instead of endomorphisms, in order to reduce the degrees of the polynomi-

als involved. We consider two cases:

- In the *Siegel case*, we assume nothing a priori on the ring $\mathrm{End}(A)$.
- In the *Hilbert case*, we fix a real quadratic field F and assume the existence of a real multiplication embedding $\mathbb{Z}_F \hookrightarrow \mathrm{End}(A)^\dagger$ defined over $\mathbb{F}_q$, which is not necessarily explicit or computable.

In the Hilbert case, we further assume that ℓ splits in $\mathbb{Z}_F$ into a product of two principal ideals generated by totally positive elements $\beta, \overline{\beta} \in \mathbb{Z}_F$. By the Chebotarev density theorem, this condition holds for a constant proportion of ℓ's as q grows to infinity; hence this assumption does not affect the asymptotic complexity of Elkies's method. Moreover, we can choose β in such a way that $\mathrm{Tr}_{F/\mathbb{Q}}(\beta) \in O_F(\sqrt{\ell})$ by [GKS11, §3.3].

The properties of χ_A are intimately related with the Rosati dual $\pi_A^\dagger$ of π_A. For the moment, we do not make the assumption that A is simple or ordinary. We know that $\pi_A^\dagger \circ \pi_A = q$ in $\mathrm{End}(A)$ [Mum70, Thm. 4 p. 206], hence $\mathbb{Q}(\pi_A) \subset \mathrm{End}^0(A)$ is a commutative subring stable under $\dagger$. Moreover, the roots of χ_A in $\mathbb{C}$ are complex numbers of the form $\lambda_1, \lambda_2, q/\lambda_1, q/\lambda_2$ with $|\lambda_k|^2 = q$ for $k \in \{1, 2\}$. Therefore the characteristic polynomial of π_A takes the form

$$\chi_A = X^4 - s_1 X^3 + (s_2 + 2q)X^2 - q s_1 X + q^2$$

where we set

$$r_1 = \lambda + \frac{q}{\lambda_1}, \quad r_2 = \lambda_2 + \frac{q}{\lambda_2}, \quad s_1 = r_1 + r_2, \quad \text{and} \quad s_2 = r_1 r_2.$$

We have $|s_1| \leq 4\sqrt{q}$ and $|s_2| \leq 4q$; in fact, by the proof of [Rü90, Lem. 3.1], we also have the inequalities

$$(|s_1| - 4\sqrt{q})^2 \geq s_1^2 - 4s_2 \geq 0.$$

The endomorphism $\xi_A = \pi_A + \pi_A^\dagger$ satisfies $\xi_A^2 - s_1 \xi_A + s_2 = 0$, and π_A satisfies the following relation over $\mathbb{Z}[\xi_A]$:

$$\pi_A^2 - \xi_A \pi_A + q = 0.$$

Note that $\xi_A \in \mathrm{End}^0(A)^\dagger$. By [Mum70, Thm. 3 p. 202], this ring is either $\mathbb{Q}$, $\mathbb{Q} \times \mathbb{Q}$, or a real quadratic number field, and is equal to F in the Hilbert case. In the point counting context, we want to compute s_1 and s_2 mod ℓ using the action of π_A on kernels of isogenies. The methods in the Siegel and Hilbert cases differ.

7.2 The Siegel case

Recall that $A[\ell]$ is endowed with an alternating, nondegenerate bilinear form denoted by $(x, y) \mapsto \langle x, y \rangle \in \mathbb{Z}/\ell\mathbb{Z}$ induced by the principal polarization of A. For all $x, y \in A[\ell]$, we have

$$\langle \pi_A(x), \pi_A(y) \rangle = q \langle x, y \rangle . \tag{7.1}$$

We say that ℓ is an *Elkies prime* if there exists an ℓ-isogeny φ with domain A which is defined over $\mathbb{F}_q$; equivalently, ℓ is Elkies if $A[\ell]$ admits a dimension 2 isotropic subspace which is stable under π_A. Let us relate this property with the decomposition of χ_A mod ℓ.

If P is a monic polynomial of degree d whose constant term a_0 is invertible, then we define its q-reciprocal polynomial as

$$\widetilde{P} = \frac{1}{a_0} X^d P(q/X),$$

and we say that P is *q-reciprocal* if $\widetilde{P} = P$. If $d = 2$, this simply means that $a_0 = q$.

Proposition 7.2.1. *Assume that ℓ is Elkies, and let φ be an ℓ-isogeny with domain A defined over $\mathbb{F}_q$. Let P be the characteristic polynomial of π_A on $\ker \varphi$. Then $\chi_A = P\widetilde{P}$ mod ℓ.*

Proof. Choose a symplectic basis of $A[\ell]$ whose first two vectors generate $\ker \varphi$. By (7.1), the matrix of π_A in this basis takes the form

$$\begin{pmatrix} M & * \\ 0 & qM^{-t} \end{pmatrix},$$

and the characteristic polynomial of qM^{-t} is $\widetilde{P}$. □

Proposition 7.2.2. *Let ℓ be a prime, and assume that one of the following properties holds:*

1. *χ_A splits modulo ℓ as a product of the form $P\widetilde{P}$ where P and $\widetilde{P}$ are coprime; or*
2. *χ_A is totally split modulo ℓ.*

Then ℓ is Elkies.

Proof. In case 1, we write $a = P(\pi_A)$ and $\widetilde{a} = \widetilde{P}(\pi_A)$ as endomorphisms of $A[\ell]$. Then we have a decomposition of $A[\ell]$ as $\ker a \oplus \ker \widetilde{a}$, both subspaces being of dimension 2. Let us show that $\ker a$ is isotropic. By (7.1), $\widetilde{a}$ is the adjoint of a; hence

$$\langle \ker a, \ker a \rangle = \langle \operatorname{Im} \widetilde{a}, \operatorname{Im} \widetilde{a} \rangle = \langle A[\ell], \operatorname{Im}(a\widetilde{a}) \rangle = 0.$$

In case 2, if v is an eigenvector of π_A, then $v^\perp$ is still π_A-stable, hence there exists $w \in v^\perp$ such that $\langle v \rangle \oplus \langle w \rangle$ is a π_A-stable subspace of dimension 2. □

The splitting of χ_A modulo ℓ also determines the number of maximally isotropic subspaces. For instance, if χ_A is totally split over $\mathbb{F}_\ell$, and if the eigenvalues satisfy the non-equalities $r_1 \neq r_2$, $\lambda_1 \neq q/\lambda_1$ and $\lambda_2 \neq q/\lambda_2$, then there are exactly four of them inside $A[\ell]$.

To summarize, if ℓ is an Elkies prime, then we can recover $\chi_A \bmod \ell$ by computing the characteristic polynomial of π_A on the kernel of *one* ℓ-isogeny, by proposition 7.2.1. Heuristically, we expect that about half of the primes will be Elkies given the characterization of proposition 7.2.2, so that it is sufficient to consider Elkies primes $\ell = O(\log q)$ to complete the point counting algorithm.

7.3 The Hilbert case

In the Hilbert case, the subgroup $A[\ell]$ decomposes as the orthogonal direct sum $A[\beta] \oplus A[\overline{\beta}]$, and both subspaces are stable under π_A and $\pi_A^\dagger$. Since $A[\beta]$ and $A[\overline{\beta}]$ are not isotropic, the determinant of π_A on both of these subspaces is q. Hence we have a decomposition of $\chi_A \bmod \ell$ as

$$\chi_A = (X^2 - t_\beta X + q)(X^2 - t_{\overline{\beta}} X + q)$$

where t_β and $t_{\overline{\beta}}$ are the traces of π_A on $A[\beta]$ and $A[\overline{\beta}]$ respectively.

We say that β (resp. $\overline{\beta}$) is *Elkies* if π_A admits an eigenvector on $A[\beta]$ (resp. on $A[\overline{\beta}]$). Equivalently, β is Elkies if there exists a β-isogeny φ with domain A defined over $\mathbb{F}_q$. In this case, knowing the eigenvalue of π_A on $\ker \varphi$ is sufficient to reconstruct t_β.

However, β and $\overline{\beta}$ may not be simultaneously Elkies, hence reconstructing $\chi_A \bmod \ell$ is not always possible. Instead, we observe that $\xi_A \in \mathbb{Z}_F$ and $\mathbb{Z}_F/\beta\mathbb{Z}_F \simeq \mathbb{Z}/\ell\mathbb{Z}$; in this identification, we have $t_\beta = \xi_A \bmod \beta$. Instead of using the Chinese remainder theorem to recover the coefficients of χ_A as integers, we recover ξ_A as an element of $\mathbb{Z}_F$.

We view $\mathbb{Z}_F$ as a 2-dimensional lattice, endowed with the Euclidean distance given by the embedding $\mathbb{Z}_F \to \mathbb{C}^2$, $x \mapsto (x, \overline{x})$. Then $\xi_A \in \mathbb{Z}_F$ has length at most $(8q)^{1/2}$. We claim that ξ_A is uniquely determined as soon as we know the value of $\xi_A \bmod I$, where I is an ideal of norm strictly greater that $16q$. Indeed, if ξ, ξ' are equal mod I but distinct, then $N_{F/\mathbb{Q}}(\xi - \xi') \geq N(I)$; therefore

$$\|\xi - \xi'\| \geq \sqrt{2\left|N_{F/\mathbb{Q}}(\xi - \xi')\right|} > 2(8q)^{1/2}.$$

We can stop the point counting algorithm when we have collected the values of ξ_A modulo distinct totally positive primes $\beta_1, \ldots, \beta_N$ of $\mathbb{Z}_F$ such that $\prod_{i=1}^{N} N_{F/\mathbb{Q}}(\beta_i) > 16q$. Heuristically, we expect that a given β is Elkies with probability about one half, so that it is sufficient to consider Elkies primes β such that $N_{F/\mathbb{Q}}(\beta) = O(\log q)$ to complete the algorithm.

7.4 Computing the action of Frobenius on isogeny kernels

If ℓ or β is an Elkies prime, we still have to explain how to find a suitable β- or ℓ-isogeny φ with domain A defined over $\mathbb{F}_q$, and how to compute the characteristic polynomial of π_A on $\ker \varphi$. The first part, computing φ, is where we apply the results of the previous chapters. The second part involves polynomial systems and is more classical, being already studied in the context of Schoof's algorithm on hyperelliptic curves [GS12, Abe18].

We only give a sketch of the algorithms. For simplicity, we make the following assumptions:

- The base field $\mathbb{F}_q$ is prime, i.e. $q = p$; moreover $p > 8\ell + 7$ in the Siegel case, and $p > 4\,\mathrm{Tr}_{F/\mathbb{Q}}(\beta) + 7$ in the Hilbert case.
- A is the Jacobian of a hyperelliptic genus 2 curve $\mathcal{C}$ defined over $\mathbb{F}_p$, and has automorphism group $\{\pm 1\}$ over an algebraic closure of $\mathbb{F}_p$ (in particular A is absolutely simple).

Moreover, we assume that the various other genericity assumptions used in the algorithms of chapters 3 and 6 hold true. For instance, in the Siegel case, we assume in particular that the denominator D_ℓ of Siegel modular equations does not vanish at $(j_1(A), j_2(A), j_3(A))$. If p is very large, say of cryptographic size, and if $\mathcal{C}$ is chosen at random, then all these genericity assumptions hold with overwhelming probability.

The algorithm to compute φ can be sketched as follows; we recall that $\ell = O(\log p)$.

1. Evaluate the modular equations and their derivatives at A using the algorithms of chapter 6, assuming that hypothesis 6.2.15 holds; this costs $\widetilde{O}(\ell^5 \log^2 p)$ binary operations in the Siegel case, and $\widetilde{O}_F(\ell \log^2 p)$ binary operations in the Hilbert case. If the Igusa invariants of A can be written as quotients of integers of height $O(1)$, then the cost of evaluating modular equations drops to $\widetilde{O}(\ell^6)$ and $\widetilde{O}_F(\ell^2)$ binary operations respectively.

   In the Hilbert case, the rational reconstruction step is heuristic, unless one computes an explicit complete denominator for modular equations, as we did in the case of Gundlach invariants for $F = \mathbb{Q}(\sqrt{5})$ (§5.1.4). It is advisable to use modular equations in nonsymmetric invariants if possible, since they have smaller degrees and allow us to distinguish between β and $\overline{\beta}$-isogenies.

2. Compute roots of modular equations to find an isogenous abelian surface; this step costs $\widetilde{O}(\ell^3 \log^2 p)$ binary operations in the Siegel case, and $\widetilde{O}_F(\ell \log^2 p)$ binary operations in the Hilbert case.

3. Compute an explicit representation of the isogeny φ using the algorithm of chapter 3; this costs $\widetilde{O}(\ell \log p)$ binary operations in the Siegel case, and $\widetilde{O}_F(\sqrt{\ell} \log p)$ binary operations in the Hilbert case.

At the end, we have computed a point P of $\mathcal{C}$, possibly over an extension of $\mathbb{F}_p$, and the rational representation of the isogeny φ at the base point P in the sense of §3.4.1. Denote this rational representation by $\varphi_P \colon \mathcal{C} \to \mathcal{C}'^{<2>}$. Following the proof of [Sch95, Prop. 6.1], [BGLG+16, Prop. 3], one can show that the kernel of φ will be defined over $\mathbb{F}_p$ in generic cases.

It remains to explain how to compute the characteristic polynomial of π_A on $\ker \varphi$ using φ_P as input. For simplicity, assume that P is a Weierstrass point of $\mathcal{C}$, and that all the points of $A[\ell]\backslash\{0\}$ can be written in the form $[Q_1 + Q_2 - K]$ where Q_1, Q_2 are distinct points of $\mathcal{C}$ and K denotes the canonical divisor of $\mathcal{C}$. Then such a point lies in $\ker \varphi$ if and only if

$$\varphi_P(Q_1) = -\varphi_P(Q_2) \quad \text{in } \mathrm{Jac}(\mathcal{C}').$$

From this characterization, one derives an ideal of polynomial equations describing the points of $\ker(\varphi)\backslash\{0\}$ in $\mathcal{C}^{<2>}$. In order to compute with the formal point of $\ker(\varphi)\backslash\{0\}$, we compute a Gröbner basis of this ideal. Asymptotically, the most costly step in this procedure is to compute the resultant of certain bivariate polynomials of degree $O(d)$ over $\mathbb{F}_p$ [GS12, §3.1], where d is the degree of the rational fractions describing ϕ_P. This costs $\widetilde{O}(d^3 \log p)$ binary operations; in other words $\widetilde{O}(\ell^3 \log p)$ and $\widetilde{O}_F(\ell^{3/2} \log p)$ binary operations in the Siegel and Hilbert cases, respectively. The degrees of the

polynomials in the resulting Gröbner basis is $O(d^2)$.

The final step is to compute the characteristic polynomial of π_A on $\ker \varphi$. In the Siegel case, where we have to determine two elements of $\mathbb{Z}/\ell\mathbb{Z}$, we perform $O(\ell^2 + \log p)$ multiplications modulo polynomials of degree $O(d^2)$, for a total cost of $\widetilde{O}\big((\ell^2 + \log p)\ell^2 \log p\big)$ binary operations. In the Hilbert case, we only determine one element in $\mathbb{Z}/\ell\mathbb{Z}$, therefore the total cost is only $\widetilde{O}_F(\ell \log^2 p)$ binary operations.

We conclude by summing the complexities of all the steps in the point counting algorithm, and multiplying by the number $O(\log p)$ of auxiliary primes to consider:

Result. *We obtain an algorithm which, given a large prime p and given a principally polarized abelian surface $A = \mathrm{Jac}(\mathcal{C})$ over $\mathbb{F}_p$ without extra automorphisms, computes $\chi_A \in \mathbb{Z}[X]$. In the generic case, its heuristic cost is $\widetilde{O}(\log^8 p)$ binary operations, or $\widetilde{O}(\log^7 p)$ binary operations if the Igusa invariants of A are quotients of integers of height $O(1)$. If A has real multiplication by $\mathbb{Z}_F$ where F is a fixed real quadratic field, its heuristic cost is $\widetilde{O}_F(\log^4 p)$ binary operations.*

In the Siegel case, the complexity of Schoof's method is already $\widetilde{O}(\log^8 p)$ binary operations. It can be further decreased using recent algorithms for the computation of bivariate resultants [Vil18], although the practical effects of this improvement are not demonstrated. Therefore, Elkies's method is (heuristically) asymptotically faster only when A has small invariants. In the Hilbert case, the heuristic complexity of Elkies's method improves on Schoof's method: up to constant factors depending on F, we reach the same asymptotic complexity as the SEA algorithm for elliptic curves.

Research perspectives

To conclude the thesis, we sketch several directions of research that our results suggest. They concern either improvements or generalizations of our algorithms.

- First and foremost, releasing an implementation of the algorithms of chapters 3 and 6 is in order, and perhaps even a full implementation of Elkies's method over prime finite fields in generic cases. In particular, this would allow us to compare the cost of Elkies's and Schoof's methods in the Siegel case experimentally.
- In order to implement a proved algorithm to compute genus 2 theta constants in quasi-linear time, it would be nice to prove hypothesis 6.2.15 with explicit upper bounds on the necessary initial precision in the Newton iterations. This calculation is perhaps feasible, and would be interesting also in the genus 1 case.
- The complexity of the algorithm of chapter 6 to evaluate modular equations over number fields is not quasi-linear in the output size in general. In particular, the term H^2 in proposition 6.3.9 comes from the fact that reduction to the fundamental domain $\mathcal{F}_2 \subset \mathbb{H}_2$ is not done in quasi-linear time, as proposition 6.2.23 shows. As we remarked at the end of chapter 6, it should be possible to reformulate this problem in terms of lattice reduction for dimension 4 symplectic lattices, and to adapt existing quasi-linear lattice reduction algorithms [NSV11, NS16] to this setting.
- In order to assess the complexity of Elkies's method for abelian surfaces, it would be interesting to generalize the known results about the distribution of Elkies primes in the case of elliptic curves [SS14, SS15] to higher dimensions.
- The isogeny algorithm of chapter 3 could be applied to the construction of explicit families of abelian surfaces with real multiplication, besides the examples presented in [GKS11, §4].

- In order to evaluate modular equations over a finite field $\mathbb{F}_q$, lifting to a number field as in proposition 6.3.7 is not the most natural thing to do. It is interesting to ask if there exists an algorithm to evaluate modular equations in quasi-linear time over unramified extensions of $\mathbb{Q}_p$, using canonical lifts of abelian surfaces.
- When constructing complete denominators of modular equations, the structure of the corresponding rings of modular forms over $\mathbb{Z}$ plays a central role. This raises the question of finding an algorithm to compute the ring of Hilbert modular forms over $\mathbb{Z}$ for $\mathbb{Q}(\sqrt{\Delta})$, when Δ varies. The result for $\Delta = 29$ and $\Delta = 37$ has recently been computed [Wil20]. It would also be interesting to compute rings of Hilbert modular forms over $\mathbb{Z}$ for totally real cubic fields.
- Finally, beyond abelian surfaces, it would be interesting to extend Elkies's method for point counting to the case of principally polarized abelian threefolds with real multiplication over finite fields. We should obtain strong asymptotic improvements over Schoof's algorithm in this case. A first step in this direction would be to generalize the isogeny algorithm of chapter 3 to Jacobians of quartic curves, using the newly found correspondance between Siegel modular forms in dimension 3 and concomitants of ternary quartics [CFvdG20].

Bibliography

[Abe18] S. Abelard. *Counting points on hyperelliptic curves in large characteristic: algorithms and complexity.* PhD thesis, Université de Lorraine, 2018.

[AM93] A. O. L. Atkin and F. Morain. Elliptic curves and primality proving. *Math. Comp.*, 61:29–68, 1993.

[And17] Y. André. On the Kodaira–Spencer map of abelian schemes. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 17(4):1397–1416, 2017.

[BB66] W. L. Baily, Jr. and A. Borel. Compactification of arithmetic quotients of bounded symmetric domains. *Ann. of Math. (2)*, 84:442–528, 1966.

[BCG+17] A. Bostan, F. Chyzak, M. Giusti, R. Lebreton, G. Lecerf, B. Salvy, and É. Schost. *Algorithmes efficaces en calcul formel.* CreateSpace, 2017.

[BGLG+16] S. Ballentine, A. Guillevic, E. Lorenzo García, C. Martindale, M. Massierer, B. Smith, and J. Top. Isogenies for point counting on genus two hyperelliptic curves with maximal real multiplication. In *Algebraic Geometry for Coding Theory and Cryptography*, volume 9, pages 63–94, Los Angeles, 2016. Springer.

[BGP17] J. I. Burgos Gil and A. Pacetti. Hecke and Sturm bounds for Hilbert modular forms over real quadratic fields. *Math. Comp.*, 86(306):1949–1978, 2017.

[BJW17] E. H. Brooks, D. Jetchev, and B. Wesolowski. Isogeny graphs of ordinary abelian varieties. *Res. Number Theory*, 3:28, 2017.

[BL04] C. Birkenhake and H. Lange. *Complex abelian varieties.* Springer, second edition, 2004.

[BL09] R. Bröker and K. Lauter. Modular polynomials for genus 2. *LMS J. Comp. Math.*, 12:326–339, 2009.

[BLS12] R. Bröker, K. Lauter, and A. V. Sutherland. Modular polynomials via isogeny volcanoes. *Math. Comp.*, 81:1201–1231, 2012.

[BM88] J.-B. Bost and J.-F. Mestre. Moyenne arithmético-géométrique et périodes de courbes de genre 1 et 2. *Gaz. Math.*, 38:36–64, 1988.

[BMSS08] A. Bostan, F. Morain, B. Salvy, and É. Schost. Fast algorithms for computing isogenies between elliptic curves. *Math. Comp.*, 77(263):1755–1778, 2008.

[Bol87] O. Bolza. Darstellung der rationalen ganzen Invarianten der Binärform sechsten Grades durch die Nullwerthe der zugehörigen θ-Functionen. *Math. Ann.*, 30(4):478–495, 1887.

[Bor88] C. W. Borchardt. Theorie des arithmetisch-geometrisches Mittels aus vier Elementen. In *Gesammelte Werke*, pages 373–431. Reimer, 1888.

[Bou72] N. Bourbaki. *Groupes et algèbres de Lie. Chapitres II et III.* Hermann, 1972.

[Bou75] N. Bourbaki. *Groupes et algèbres de Lie. Chapitres VII et VIII.* Hermann, 1975.

[Bru08] J. H. Bruinier. Hilbert modular forms and their applications. In *The 1-2-3 of modular forms*, pages 105–179. Springer, 2008.

[BS10] R. Bröker and A. V. Sutherland. An explicit height bound for the classical modular polynomial. *Ramanujan J.*, 22(3):293–313, 2010.

[BS11] G. Bisson and A. V. Sutherland. Computing the endomorphism ring of an ordinary elliptic curve over a finite field. *J. Number Theory*, 113:815–831, 2011.

[BW03] C. Birkenhake and H. Wilhelm. Humbert surfaces and the Kummer plane. *Trans. Amer. Math. Soc.*, 335(5):1819–1841, 2003.

[Car86] H. Carayol. Sur la mauvaise réduction des courbes de Shimura. *Compositio Math.*, 59(2):151–230, 1986.

[CE15] J.-M. Couveignes and T. Ezome. Computing functions on Jacobians and their quotients. *Lond. Math. Soc. J. Comput. Math.*, 18(1):555–577, 2015.

[CFv17] F. Cléry, C. Faber, and G. van der Geer. Covariants of binary sextics and vector-valued Siegel modular forms of genus two. *Math. Ann.*, 369(3-4):1649–1669, 2017.

[CFvdG20] F. Cléry, C. Faber, and G. van der Geer. Concomitants of ternary quartics and vector-valued Siegel and Teichmüller modular forms of genus three. *Sel. Math.*, 26(55), 2020.

[Che51] C. Chevalley. Deux théorèmes d'arithmétique. *J. Math. Soc. Japan*, 3(1):36–44, 1951.

[Cle72] A. Clebsch. *Theorie der binären algebraischen Formen.* B. G. Teubner, 1872.

[CLG09] D. X. Charles, K. E. Lauter, and E. Z. Goren. Cryptographic hash functions from expander graphs. *J. Cryptol.*, 22:93–113, 2009.

[CLM+18] W. Castryck, T. Lange, C. Martindale, L. Panny, and J. Renes. CSIDH: an efficient post-quantum commutative group action. In *Advances in Cryptology – Asiacrypt 2018*, volume 3, pages 395–427, Brisbane, 2018. Springer.

[CMSV19] E. Costa, N. Mascot, J. Sijsling, and J. Voight. Rigorous computation of the endomorphism ring of a Jacobian. *Math. Comp.*, 88(317):1303–1339, 2019.

[Coh84] P. Cohen. On the coefficients of the transformation polynomials for the elliptic modular function. *Math. Proc. Cambridge Philos. Soc.*, 95(3):389–402, 1984.

[Cox84] D. A. Cox. The arithmetic-geometric mean of Gauss. *Enseign. Math.*, 30:275–330, 1984.

[Cox13] D. A. Cox. *Primes of the form $x^2 + ny^2$*. John Wiley & Sons, second edition, 2013.

[CR15] R. Cosset and D. Robert. Computing (ℓ, ℓ)-isogenies in polynomial time on Jacobians of genus 2 curves. *Math. Comp.*, 84(294):1953–1975, 2015.

[Del70] P. Deligne. Travaux de Shimura, 1970.

[dG19] L. de Feo and S.D. Galbraith. SeaSign: compact isogeny signatures from class group actions. In *Advances in Cryptology – Eurocrypt 2019*, volume 3, pages 759–789, Darmstadt, 2019. Springer.

[DH76] W. Diffie and M. Hellman. New directions in cryptography. *IEEE Trans. Inf. Theory*, 22(6):644–654, 1976.

[dJP11] L. de Feo, D. Jao, and J. Plût. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies. In *Post-Quantum Cryptography*, pages 19–34, Taipei, 2011. Springer.

[DJRV17] A. Dudeanu, D. Jetchev, D. Robert, and M. Vuille. Cyclic isogenies for abelian varieties with real multiplication. https://hal.inria.fr/hal-01629829v1, 2017.

[dKS18] L de Feo, J. Kieffer, and B. Smith. Towards practical key exchange from ordinary isogeny graphs. In *Advances in Cryptology – Asiacrypt 2018*, volume 3, pages 365–394, Brisbane, 2018. Springer.

[dMPS19] L. de Feo, S. Masson, C. Petit, and A. Sanso. Verifiable delay functions from supersingular isogenies and pairings. In *Advances in Cryptology – Asiacrypt 2019*, volume 1, pages 248–277, Kobe, 2019. Springer.

[DR73] P. Deligne and M. Rapoport. Les schémas de modules de courbes elliptiques. In *Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, 1972)*, pages 143–316. Springer, 1973.

[DS05] F. Diamond and J. Shurman. *A first course in modular forms*. Springer, 2005.

[Dup06] R. Dupont. *Moyenne arithmético-géométrique, suites de Borchardt et applications*. PhD thesis, École polytechnique, 2006.

[Dup11] R. Dupont. Fast evaluation of modular functions using Newton iterations and the AGM. *Math. Comp.*, 80(275):1823–1847, 2011.

[Elk98] N. D. Elkies. Elliptic and modular curves over finite fields and related computational issues. In *Computational perspectives on number theory (Chicago, 1995)*, volume 7, pages 21–76. Amer. Math. Soc., 1998.

[EM02] A. Enge and F. Morain. Comparing invariants for class polynomials of imaginary quadratic fields. In *Algorithmic Number Theory*, pages 252–266, Sydney, 2002. Springer.

[Eng09a] A. Enge. The complexity of class polynomial computation via floating point approximations. *Math. Comp.*, 78(266):1089–1107, 2009.

[Eng09b] A. Enge. Computing modular polynomials in quasi-linear time. *Math. Comp.*, 78(267):1809–1824, 2009.

[ET14] A. Enge and E. Thomé. CMH: Computation of Igusa class polynomials. https://gitlab.inria.fr/cmh/cmh/, 2014.

[Fal83] G. Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörper. *Invent. Math.*, 73(3):349–366, 1983.

[FC90] G. Faltings and C.-L. Chai. *Degeneration of abelian varieties*. Springer, 1990.

[FM02] M. Fouquet and F. Morain. Isogeny volcanoes and the SEA algorithm. In *Algorithmic Number Theory*, pages 276–291, Sydney, 2002. Springer.

[FSM19] E. Freitag and R. Salvati Manni. On the variety associated to the ring of theta constants in genus 3. *Amer. J. Math.*, 141(3):705–732, 2019.

[Gau68] C. F. Gauss. *Werke*. Dietrich, 1868.

[GH00] P. Gaudry and R. Harley. Counting points on hyperelliptic curves over finite fields. In *Algorithmic number theory*, pages 313–332, Leiden, 2000. Springer.

[GHK+06] P. Gaudry, T. Houtmann, D. Kohel, C. Ritzenthaler, and A. Weng. The 2-adic CM method for genus 2 curves with application to cryptography. In *Advances in Cryptology – Asiacrypt 2006*, pages 114–129, Shanghai, 2006. Springer.

[GHS02] S.D. Galbraith, F. Hess, and N. P. Smart. Extending the GHS Weil descent attack. In *Advances in Cryptology – Eurocrypt 2002*, pages 29–44, Amsterdam, 2002. Springer.

[GKS11] P. Gaudry, D. Kohel, and B. Smith. Counting points on genus 2 curves with real multiplication. In *Advances in Cryptology – Asiacrypt 2011*, pages 504–519, Seoul, 2011. Springer.

[GM17] A. Guillevic and F. Morain. Discrete logarithms. In *Guide to Pairing-Based Cryptography*. CRC Press, 2017.

[Got59] E. Gottschling. Explizite Bestimmung der Randflächen des Fundamentalbereiches der Modulgruppe zweiten Grades. *Math. Ann.*, 138:103–124, 1959.

[Gru10] D. Gruenewald. Computing Humbert surfaces and applications. In *Arithmetic, geometry, cryptography and coding theory 2009*, pages 59–69. Amer. Math. Soc., 2010.

[GS12] P. Gaudry and É. Schost. Genus 2 point counting over prime fields. *J. Symb. Comput.*, 47(4):368–400, 2012.

[Har85] M. Harris. Arithmetic vector bundles and automorphic forms on Shimura varieties. I. *Invent. Math.*, 82(1):151–189, 1985.

[Har86] M. Harris. Arithmetic vector bundles and automorphic forms on Shimura varieties. II. *Compositio Math.*, 60(3):323–378, 1986.

[HP17] P. Habegger and F. Pazuki. Bad reduction of genus 2 curves with CM Jacobian varieties. *Compos. Math.*, 153(12):2534–2576, 2017.

[HS00] M. Hindry and J. H. Silverman. *Diophantine geometry*. Springer, 2000.

[Ibu12] T. Ibukiyama. Vector-valued Siegel modular forms of symmetric tensor weight of small degrees. *Comment. Math. Univ. St. Pauli*, 61(1):51–75, 2012.

[Igu60] J.-I. Igusa. Arithmetic variety of moduli for genus two. *Ann. of Math. (2)*, 72:612–649, 1960.

[Igu62] J.-I. Igusa. On Siegel modular forms of genus two. *Amer. J. Math.*, 84:175–200, 1962.

[Igu64] J.-I. Igusa. On the graded ring of theta-constants. *Amer. J. Math.*, 86(1):219–246, 1964.

[Igu66] J.-I. Igusa. On the graded ring of theta-constants (II). *Amer. J. Math.*, 88(1):221–236, 1966.

[Igu72] J.-I. Igusa. *Theta functions*. Springer, 1972.

[Igu79] J.-I. Igusa. On the ring of modular forms of degree two over $\mathbb{Z}$. *Amer. J. Math.*, 101(1):149–183, 1979.

[Jar08] F. Jarvis. Higher genus arithmetic-geometric means. *Ramanujan J.*, 17(1):1–17, 2008.

[JL06] A. Joux and R. Lercier. Counting points on elliptic curves in medium characteristic, 2006.

[JMV09] D. Jao, S. D. Miller, and R. Venkatesan. Expander graphs based on GRH with an application to elliptic curve cryptography. *J. Number Theory*, 129:1491–1504, 2009.

[Joh17] F. Johansson. Arb: efficient arbitrary-precision midpoint-radius interval arithmetic. *IEEE Trans. Comput.*, 66(8):1281–1292, 2017.

[Kan19] E. Kani. Elliptic subcovers of a curve of genus 2. I. The isogeny defect. *Ann. Math. Qué.*, 43(2):281–303, 2019.

[Kie20a] J. Kieffer. Degree and height estimates for modular equations on PEL Shimura varieties. https://hal.archives-ouvertes.fr/hal-02436057v2, 2020.

[Kie20b] J. Kieffer. Evaluating modular polynomials in genus two. https://hal.inria.fr/hal-02971326v1, 2020.

[Kie20c] J. Kieffer. Height bounds for polynomial and rational fractions from their values. Preprint, 2020.

[Kie21] J. Kieffer. Sign choices in the AGM for genus two theta constants. https://hal.inria.fr/hal-02967220, to appear in *Publ. Math. Besançon*, 2021.

[Kli90] H. Klingen. *Introductory lectures on Siegel modular forms*. Cambridge University Press, 1990.

[Kob87] N. Koblitz. Elliptic curve cryptosystems. *Math. Comp.*, 48(177):203–209, 1987.

[Kob89] N. Koblitz. Hyperelliptic cryptosystems. *J. Cryptol.*, 1:139–150, 1989.

[Koh96] D. Kohel. *Endomorphism rings of elliptic curves over finite fields*. PhD thesis, University of California at Berkeley, 1996.

[KPR19] J. Kieffer, A. Page, and D. Robert. Computing isogenies from modular equations in genus two. https://hal.archives-ouvertes.fr/hal-02436133v2, 2019.

[Lab18] H. Labrande. Computing Jacobi's θ in quasi-linear time. *Math. Comp.*, 87:1479–1508, 2018.

[LLL82] A. K. Lenstra, H. W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982.

[LM00] G. Laumon and L. Moret-Bailly. *Champs algébriques.* Springer, 2000.

[LR12] D. Lubicz and D. Robert. Computing isogenies between abelian varieties. *Compos. Math.*, 148(05):1483–1515, 2012.

[LT16] H. Labrande and E. Thomé. Computing theta functions in quasi-linear time in genus 2 and above. In *Algorithmic Number Theory Symposium XII*, pages 163–177, Kaiserslautern, 2016. LMS J. Comp. Math 19.

[LY11] K. Lauter and T. Yang. Computing genus 2 curves from invariants on the Hilbert moduli space. *J. Number Theory*, 131(5):936–958, 2011.

[Mar18] C. Martindale. *Isogeny graphs, modular polynomials, and applications.* PhD thesis, Universiteit Leiden and Université de Bordeaux, 2018.

[Mar20] C. Martindale. Hilbert modular polynomials. *J. Number Theory*, 213:464–498, 2020.

[Mer74] F. Mertens. Ein Beitrag zur analytischen Zahlentheorie. *J. Reine Angew. Math.*, 78:46–62, 1874.

[Mes91] J.-F. Mestre. Construction de courbes de genre 2 à partir de leurs modules. In *Effective methods in algebraic geometry (Castiglioncello, 1990)*, page 313–334. Birkhäuser, 1991.

[MFK94] D. Mumford, J. Fogarty, and F. Kirwan. *Geometric invariant theory.* Springer, third edition, 1994.

[Mil] E. Milio. Database of modular polynomials. https://members.loria.fr/EMilio/modular-polynomials.

[Mil85] V. S. Miller. Uses of elliptic curves in cryptography. In *Advances in Cryptology – Crypto '85 Proceedings*, pages 417–426, Santa Barbara, 1985. Springer.

[Mil86a] J. S. Milne. Abelian varieties. In *Arithmetic geometry (Storrs, 1984)*, pages 103–150. Springer, 1986.

[Mil86b] J. S. Milne. Jacobian varieties. In *Arithmetic geometry (Storrs, 1984)*, pages 167–212. Springer, 1986.

[Mil90] J. S. Milne. Canonical models of (mixed) Shimura varieties and automorphic vector bundles. In *Automorphic forms, Shimura varieties, and L-functions (Ann Arbor, 1988)*, volume 1, page 283–414. Academic Press, 1990.

[Mil05] J. S. Milne. Introduction to Shimura varieties. In *Harmonic analysis, the trace formula, and Shimura varieties*, pages 265–378. Amer. Math. Soc., 2005.

[Mil15] E. Milio. A quasi-linear time algorithm for computing modular polynomials in dimension 2. *LMS J. Comput. Math.*, 18:603–632, 2015.

[Mil17] J. S. Milne. *Algebraic groups*. Cambridge University Press, 2017.

[MN19] P. Molin and C. Neurohr. Computing period matrices and the Abel-Jacobi map of superelliptic curves. *Math. Comp.*, 88(316):847–888, 2019.

[Mol18] P. Molin. Hcperiods: Period matrices and Abel-Jacobi maps of hyperelliptic and superperelliptic curves. https://github.com/pascalmolin/hcperiods, 2018.

[MR20] E. Milio and D. Robert. Modular polynomials on Hilbert surfaces. *J. Number Theory*, 216:403–459, 2020.

[Mum70] D. Mumford. *Abelian varieties*. Oxford University Press, 1970.

[Mum83] D. Mumford. *Tata lectures on theta. I.* Birkhäuser, 1983.

[Mum84] D. Mumford. *Tata lectures on theta. II.* Birkhäuser, 1984.

[Nag83] S. Nagaoka. On the ring of Hilbert modular forms over $\mathbb{Z}$. *J. Math. Soc. Japan*, 35(4):589–608, 1983.

[Noe15] E. Noether. Der Endlichkeitssatz der Invarianten endlicher Gruppen. *Math. Ann.*, 77(1):89–92, 1915.

[NS16] A. Neumaier and D. Stehlé. Faster LLL-type reduction of lattice bases. In *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation*, pages 373–380, Waterloo, 2016. ACM.

[NSV11] A. Novocin, D. Stehlé, and G. Villard. An LLL-reduction algorithm with quasi-linear time complexity. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, pages 403–412, San Jose, 2011. ACM.

[NV09] P. Nguyen and B. Vallée, editors. *The LLL algorithm. Survey and applications*. Springer, 2009.

[OU73] F. Oort and K. Ueno. Principally polarized abelian varieties of dimension 2 and 3 are Jacobian varieties. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 20:377–381, 1973.

[Paz12] F. Pazuki. Theta height and Faltings height. *Bull. Soc. Math. France*, 140(1):19–49, 2012.

[Paz19] F. Pazuki. Modular invariants and isogenies. *Int. J. Number Theory*, 15(3):569–584, 2019.

[PH78] S. Pohlig and M. Hellman. An improved algorithm for computing logarithms over GF(p) and its cryptographic significance. *IEEE Trans. Inf. Theory*, 24(1):106–110, 1978.

[Pil90] J. Pila. Frobenius maps of abelian varieties and finding roots of unity in finite fields. *Math. Comp.*, 55(192):745–763, 1990.

[Piz90] A. K. Pizer. Ramanujan graphs and Hecke operators. *Bull. Am. Math. Soc.*, 23(1):127–137, 1990.

[Pol75] J. M. Pollard. A Monte Carlo method for factorization. *BIT Numer. Math.*, 15(3):331–334, 1975.

[Pom87] C. Pomerance. Fast, rigorous factorization and discrete logarithm algorithms. *Discrete Algorithms Complex.*, pages 119–143, 1987.

[Rap78] M. Rapoport. Compactifications de l'espace de modules de Hilbert-Blumenthal. *Compositio Math.*, 36(3):255–335, 1978.

[Ros41] B. Rosser. Explicit bounds for some functions of prime numbers. *Am. J. Math.*, 63(1):211–232, 1941.

[Rü90] H.-G. Rück. Abelian surfaces and Jacobian varieties over finite fields. *Compos. Math.*, 76(3):351–366, 1990.

[Sch85] R. Schoof. Elliptic curves over finite fields and the computation of square roots mod p. *Math. Comp.*, 44(170):483–494, 1985.

[Sch91] A. Schönhage. Fast reduction and composition of binary quadratic forms. In *Proceedings of the 1991 International Symposium on Symbolic and Algebraic Computation*, pages 128–133, Bonn, 1991. ACM.

[Sch95] R. Schoof. Counting points on elliptic curves over finite fields. *J. Théorie Nr. Bordx.*, 7(1):219–254, 1995.

[Sho94] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, Santa Fe, 1994. IEEE.

[Sil09] J. H. Silverman. *The arithmetic of elliptic curves*. Springer, second edition, 2009.

[SS14] I. E. Shparlinski and A. V. Sutherland. On the distribution of Atkin and Elkies primes. *Found. Comput. Math.*, 14:285–297, 2014.

[SS15] I. E. Shparlinski and A. V. Sutherland. On the distribution of Atkin and Elkies primes for reductions of elliptic curves on average. *LMS J. Comput. Math.*, 18:308–322, 2015.

[Str10] M. Streng. *Complex multiplication of abelian surfaces*. PhD thesis, Universiteit Leiden, 2010.

[Str14] M. Streng. Computing Igusa class polynomials. *Math. Comp.*, 83:275–309, 2014.

[Sut] A. V. Sutherland. Database of modular polynomials of all levels up to 300 and prime levels up to 1000 for the j-function. https://math.mit.edu/~drew/ClassicalModPolys.html.

[Sut11] A. V. Sutherland. Computing Hilbert class polynomials with the Chinese remainder theorem. *Math. Comp.*, 80(273):501–538, 2011.

[Sut13] A. V. Sutherland. On the evaluation of modular polynomials. In *Proceedings of the Tenth Algorithmic Number Theory Symposium*, pages 531–555. Math. Sci. Publ., 2013.

[The18] The Stacks project authors. The Stacks Project. https://stacks.math.columbia.edu/, 2018.

[The19] The PARI group. Pari/GP version 2.11.0. http://pari.math.u-bordeaux.fr/, 2019.

[Tho70] J. Thomae. Beitrag zur Bestimmung von $\vartheta(0,0,\ldots 0)$ durch die Klassenmoduln algebraischer Functionen. *J. Reine Angew. Math.*, 71:201–222, 1870.

[van88] G. van der Geer. *Hilbert modular surfaces*. Springer, 1988.

[van08] G. van der Geer. Siegel modular forms and their applications. In *The 1-2-3 of modular forms*, pages 181–245. Springer, 2008.

[vG13] J. von zur Gathen and J. Gerhard. *Modern computer algebra*. Cambridge University Press, third edition, 2013.

[Vil18] G. Villard. On computing the resultant of generic bivariate polynomials. In *Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*, pages 391–398, New York, 2018. ACM.

[Vé71] J. Vélu. Isogénies entre courbes elliptiques. *C. R. Acad. Sci. Paris*, A273:238–241, 1971.

[Wat69] W. C. Waterhouse. Abelian varieties over finite fields. *Ann. scient. Éc. Norm. Sup.*, 4(2):521–560, 1969.

[Wil20] B. Williams. The rings of Hilbert modular forms for $\mathbb{Q}(\sqrt{29})$ and $\mathbb{Q}(\sqrt{37})$. *J. Algebra*, 559(1):679–711, 2020.

[Yap92] C. K. Yap. Fast unimodular reduction: planar integer lattices. In *Proceedings of the 33rd Annual Symposium on Foundations of Computer Science*, pages 436–446, Pittsburgh, 1992. IEEE.