UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

PhD thesis

Diophantine estimates on modular structures

Modular polynomials and transcendence questions on genus 2 curves

Desirée Gijón Gómez

Advisor: Fabien Pazuki

Submitted: September 30, 2025

This thesis has been submitted to the PhD School of The Faculty of Science, University of Copenhagen

Desirée Gijón Gómez
desireegngz@gmail.com
Department of Mathematical Sciences
University of Copenhagen
Universitetsparken 5, 2100 Copenhagen
Denmark

Thesis title	Diophantine estimates on modular structures
Supervisor	Professor Fabien Pazuki
Assessment committee	Professor Elisenda Feliu (chair) *University of Copenhagen* Professor Javier Fresán *Sorbonne Université IMJ-PRG* Directeur de recherche Damien Robert *INRIA-Université de Bordeaux*
Date of submission	September 30, 2025
Date of defense	November 7, 2025

Para mi padre

Abstract

This thesis is a study on Diophantine estimates and transcendence results for invariants of genus two curves; more precisely, on generalizations of the Stéphanois theorem. One of the main objects of interest is the elliptic modular polynomials Φ_N and its generalizations to genus two, governing the behavior of (specific type of) isogenies.

It has two parts, Part I with the preliminary literature and Part II, with our contributions. Part II contains three projects. The first article [BGP25], a joint project with Fabien Pazuki and Florian Breuer, gives explicit upper and lower bounds for the (naive) height of Φ_N, extending previous results to arbitrary level N. It also points out a connection between its growth in terms on N with other suitable notions of size for the classical modular curves $X_0(N)$.

The second article [Gij25], a single-authored project, is an analysis of the natural sources of exceptions for our natural statement of transcendence for genus two curves. The CM points are obvious candidates for such an exception, and we compare them with the other positive dimensional subvarieties that we found.

The last project, split into the two last chapters, gives a generalization of all the steps of the proof of the Stéphanois theorem, except for the last one, for the Igusa invariants of genus two curves. It is informed by the previous study of exceptions, and it is the most general result that one can attempt to prove for these invariants.

Resumé

Denne afhandling er en undersøgelse af diofantiske estimater og transcendensresultater for invarianter af kurver af genus to; mere præcist af generaliseringer af Stéphanois sætning. Et af de vigtigste objekter af interesse er de elliptiske modulære polynomier Φ_N og deres generaliseringer til genus to, som styrerkontrollerer opførslen af (en bestemt type af) isogenier.

Afhandlingen har to dele: Del I med den indledende litteratur, og Del II med vores bidrag. Del II indeholder tre projekter. Den første artikel [BGP25], et fælles projekt med Fabien Pazuki og Florian Breuer, giver eksplicitte øvre og nedre grænser for (den naive) højden (naiv) af Φ_N, og udvider tidligere resultater til vilkårlige niveauer N. Den peger også på en forbindelse mellem væksten i N og andre passende størrelsesbegreber for de klassiske modulære kurver $X_0(N)$.

Den anden artikel [Gij25], et eneenkeltforfatterprojekt, er en analyse af de naturlige kilder til undtagelser for vores naturlige formulering af transcendens for kurver af genus to. CM-punkterne er åbenlyse kandidater til en sådan undtagelse, og vi sammenligner dem med de andre positive-dimensionelle delmængdervarieteter, som vi fandt.

Det sidste projekt, fordelt på de to sidste kapitler, giver en generalisering af alle trin i beviset af Stéphanois sætning, undtagen det sidste, for Igusa-invarianterne af kurver af genus to. Det bygger på den tidligere undersøgelse af undtagelser og er det mest generelle resultat, man kan forsøge at bevise for disse invarianter.

Thesis Statement

This thesis includes material from the following:

- Chapter 5 is a reproduction of [BGP25], published work with Florian Breuer and Fabien Pazuki.
- Chapter 6 is an improvement on the arxiv preprint [Gij25].

Acknowledgements

A PhD thesis only has one author, but the whole PhD path involves a lot of people. Then, first of all, sorry for the ones I have forgotten to mention.

I thank my supervisor Fabien Pazuki for accepting me as a PhD student three years ago, and for all the patience, guidance, encouragement and help. You brought to me a topic of research that I have thoroughly enjoyed, and I have learned many equally interesting things along the way. There is more to being a supervisor than just the mathematics and there are more things to learn as a student than just the research, so thank you for your role on the mathematician I am right now. Thank you also for all the dinners in all the different locations!

I thank Daniel Bertrand for hosting me for three months of Spring 2024 at IMJ-PRG. It is due to your questions and suggestions that I get to have the results to defend here. Thank you also for the continued conversations after I came back to Copenhagen, right until the end of my PhD, and thank you for all your insightful remarks on the (several) drafts on the preprint.

I get to thank Fabien Pazuki again, together with Florian Breuer, as a collaborator now. Thank you for the invitation right at that moment to join: it has truly been a gift that has kept on giving, in very surprising ways.

I thank Damien Robert, Javier Fresán and Elisenda Feliu, for agreeing to be in the assessment committee. I also thank all the mathematical conversations with Damien Robert and Javier Fresán.

I thank the members of the Zoo office, where I have spend more time that on my own place: Tim *Korte* Berland, Fadi Mezher, Huaitao Gui and Joan Ferrer Rodríguez, and the honorary members Viggo, Velcro, the ducks and the Nisse. Thank you Tim, my Number Theory brother, office mate and friend. Thanks for all the math, all the questions, answers and confusions, thanks for all the riddles of the day, for all the danish traditions and for the Number Theory colloquium; and thanks for Viggo and thanks for inviting me to your wedding (and thanks for checking my danish abstract). Thank you Fadi, for all the (definitely too many) weekends at the department on our first year, and for this chaotic energy that reminds me of home. Thank you Huaitao, for all the thoughtful gifts, for having the most peaceful aura I have ever seen in a mathematician, and for the Buildings Reading group; y gracias Joan, que ojalá hubiera llegado antes, a quien paso la responsabilidad de mantener la presencia española en el sótano.

I thank the (past and present) Number Theory department in Copenhagen: Morten Risager, Ian Kiming, Jasmin Matz, Hélène Esnault, Adrien Morin, Lars Kühne, Adele Betina, and Alf Söderberg, who now gets to step up as the senior PhD student. I thank Lars for being my Master Thesis advisor in a wonderful project. I thank Adrien for all the math questions you answered, and for the Number Theory coffee meetings, for the trip on the N3 days at Tromsø, and for Velcro. I also thank the rest of the N3 network, and all the participants for the Number Theory colloquium: Anton Fehnker (welcome to math, little PhD brother!), Ida Kolmanic, Emil Rugaard Wieser, Hugo Valling Lauritsen and Corentin Bernet.

I thank the Copenhagen Topology department for and for all the mathematical life it maintains, and for accepting me as an infiltrated number theorist in some of their talks (and dinners), and for lunch every day. Thank you Nathalie Wahl and Jesper Grodal for your dedication to the center.

I thank all the (past and present) members in this wonderful community: Oscar Bendix Harr, Nena Batenburg, Thomas Jan Mikhail, Maxime Ramzi, Adriano Cordova Fedeli, Dustin Clausen, Jonathan Laurent Clivio, Branko Juran, Pierre Elis, Qingyuan Bai, Andrea Bianchi,

Robert Burklund, Jan Steinebrunner, Cecilie Olesen Recke, Azélie Picot, Isaac Moselle, Priya Kaveri, Florian Riedel, Ryomei Iwasa, Rodrigue Haya, Marie-Camille Delarue, Dani Kaufman, Thomas Blom, Ishan Levy, Marius Kjærsgaard, Kaif Hilman, Alexis Aumonier, Vignesh Subramanian, Erik Lindel, Philippe Vollmuth, Mateusz Kandybo, Mate Laszlo, Alexander Gjelsvik Ravnanger, Imma Gálvez-Carrillo.

I thank IMJ-PRG for its hospitality in Spring 2024. I thank for all the Number Theory, Algebraic Geometry and Arithmetic Geometry, and for the stimulating life it has for PhD students, which was the precise right thing for me at that time; and for all the people I met there: João Ruiz, Drimik Roy, Tangi Pasquer, Thiago Landim, Antoine Galet, Nastaran Einabadi, Matteo Verni, Gabriel Ribeiro, Steffano Aloé, Sacha Zakharov, Anna Roig Sanchis, Pietro Piccione, and far more other people. I also thank the IMB at Bordeaux for all the visits I have made there and conferences I have attended in the past three years.

I thank the invitations I had received for giving talks: Riccardo Pengo for its invitation to the Diophantine Geometry Oberseminar; the organizers from the Atelier ANR j-invariant at Clermont-Ferrand for their invitation to a contributed talk; the organizers of the N3 days at Kiel for their invitation to a contributed talk, and the organizers of AGCT25 at CIRM for their invitation to a poster presentation.

I thank all the number theorists I have met in these three years, that have answered any of my questions or have made any mathematical gathering better: Martín Azón, Jean Kieffer, Simon Kristensen, Philipp Habegger, Yunqing Tang, Marco Streng, Riccardo Pengo, Marc Hindry, Stéphane Fischler, Maiken Balman Gravgaard, John Voight, Sabrina Kunzweiler, Elisa Lorenzo García, Christophe Ritzenthaler, Francesco Campagna, Martin Oor, Christopher Daw, Peter Beelen, Joël Ouaknine, Elena Berardini, Yuri Bilu, Sara Checcoli, Boris Adamczewski, Emanuele Tron, Jerson Caro, Florian Tilliet, Lorenzo Andreaus, Gaétan Guillot, Guy Fowler, Manoy Trip.

Thank you Oscar for being one of these people that make everyone around them happier, and all the good luck to you and Bella with the baby (and Stockholm). Thank you Jean for, among other things, your help with my next postdoc. Thank you Martín, for bringing a little bit of home to every conference we run into. Thank you Nena, for being a good friend, for your advice, and for all the weekends at the department. Gracias Thomas, por esos paseos que tanta falta me hacían. Thank you Mette, for Viggo, for inviting me to your wedding and for all those wonderful dinners. Thank you Teresa, for all the invitations at Bordeaux. Thank you Elisabeth for the weekend at Tromsø, for the day at Dyrehaven, for the coffee meetings and for Velcro. Thank you Elena for inviting me to your home for the true italian experience of Sanremo. Thank you to my flatmate Abby, for making our place less empty.

I acknowledge the non-academic wives and girlfriends I have met, for all the challenges and hardship that comes your way.

And finally, I thank all my family and friends back home. Gracias a mi madre, por básica y absolutamente todo, y por toda la fortaleza que ha necesitado, y gracias a mi hermano, por lo que nos ha tenido que soportar, y por cuidar de ella. Gracias a mi abuela Mari y a mi abuelo Pedro, que no los veo tanto como les gustaría y sufren por tenerme tan lejos siempre. Gracias a toda mi familia, por todo lo que hemos pasado, y por todo el apoyo. Gracias a Jorge por estar ahí sufriendo el mismo camino que yo desde 2016 y por ser quien mejor me entiende ahora mismo. Gracias a José Antonio, que sigue siendo mi hermanito topólogico, tres días más pequeño que yo, aunque cada vez que nos vemos intenta convencerme de que me dedique a algo que dé dinero. Gracias a mis niñas Marta y María, que aunque se me dé fatal mantener el contacto han sido una presencia constante en mi vida desde siempre. Gracias a Miguel por

cuando estuvo, y por todo lo que aprendí de él y lo que aprendí de mí misma. Gracias a Rocío por los consejos que me cuesta escuchar, y por Bo.

Y finalmente y el más importante de todos, gracias a mi padre, que no está desde el 21 de enero de 2024, que tendría que estar aquí para verme doctorada, y que donde quiera que esté espero que mire hacia abajo y esté orgulloso de mí.

Introduction

This PhD project had Nesterenko's theorem [Nes96, Theorem 1] as a starting point, which states

$$\operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(q, E_2(q), E_4(q), E_6(q)) \geq 3,$$

for the q-expansion of the Eisenstein series, for *any value* $0 < |q| < 1$.

It is equivalent to a statement of algebraic independence on the j-invariant of elliptic curves and its derivatives of first and second order with respect to the q-variable. From this perspective, Nesterenko's result is an improvement (and historically that is what happened) of an algebraic independence statement for the j-invariant in the q-variable, previously known as Mahler-Manin conjecture, which in French is referred to as the Stéphanois theorem [BDGP96, Théorème]:

$$\operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(q, j(q)) \geq 1.$$

A natural question is what happens to the modular invariants for higher genus curves, for instance genus two curves. There have been other work done in the direction of different generalizations of Nesterenko's theorem, see [Fon23], but we take a different route in the thesis. The j-invariant admits a generalization to genus two curves, called the Igusa invariants (from [Igu60]). They are defined as functions on the indecomposable locus of the coarse moduli space of principally polarized abelian surfaces $\mathcal{A}_2$, evaluated at the point corresponding to the Jacobian variety of a genus two curve. They are also Siegel modular functions with respect to $\mathrm{Sp}_4(\mathbb{Z})$, a generalization of modular functions to higher dimensional analytic spaces.

They share other structural properties with the j-invariant (namely, they are the generators of the field of functions on $\mathcal{A}_2$). They also admit Fourier expansions, resulting from translation symmetries of the type $j(\tau + 1) = j(\tau)$.

Our goal was to generalize the strategy of the Stéphanois theorem to the Igusa invariants. As the $\boldsymbol{q}$-expansions of the Igusa invariants are functionally algebraically independent, by [BZ01, Theorem 2], the expectation is a conjecture

$$\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q})) \geq 3, \tag{1}$$

for $\boldsymbol{q} = (q_1, q_2, q_3)$ *generically*, outside of a set of exceptions to be determined.

The material on this thesis is divided into two parts. Part I includes the preliminaries for the material in Part II, which contains the new results. We have intended for Part II to be read as independently as possible.

Part I is divided into four chapters. In Chapter 1, we present the Igusa invariants, together with the geometric explanation to the locus where they are not defined. We state the necessary results on principally polarized abelian surfaces and Siegel modular forms, but it is not an exhaustive or self-included presentation. In Chapter 2, we introduce the Hilbert modular surfaces, with the solution to the moduli problem they provide, and some basic results about Hilbert modular forms. It closes with a finite degree map to the Siegel threefold $\mathcal{A}_2$, which allows us to pull back Siegel modular forms to (symmetric) Hilbert modular forms. The image of this map is a Humbert surface, and Humbert surfaces are discussed in Chapter 3. This chapter starts with a discussion on Humbert singular relations to define the Humbert surfaces $\mathcal{H}_\Delta$ and $\mathcal{G}_\Delta$. We then introduce a generalization given by the Humbert invariant and the generalized Humbert varieties, which come from [Kan94] and [Kan19]. Chapter 4 discusses the moduli problem of principally polarized abelian surfaces with quaternionic multiplication by hereditary orders in an indefinite rational quaternion algebra. The corresponding coarse

moduli spaces are Shimura curves and classical modular curves, and as it is the case of the Hilbert modular surfaces, they admit finite degree maps to $\mathcal{A}_2$. However, these maps are no longer canonical and consequently the quaternionic loci they describe in $\mathcal{A}_2$ has several (but finitely many) irreducible components. The analog of the explicit formula on Hilbert space was given by [Has95] and we present them together with the reinterpretation of [LY20] and [GY19] in terms of positive definite quadratic forms of a given discriminant *and genus*. They are realized as Humbert varieties. We finally present this embedding problem for modular curves as a solution given by [Kan16] to a different problem, which extends the previous morphisms to a general level N instead of a square-free level N.

Part II is also organized into four chapters, numbered from 5 to 8. Chapter 5 is a joint published paper with Fabien Pazuki and Florian Breuer, in [BGP25] (the only part of this thesis that is in collaboration). We provide explicit lower and upper bounds to $h(\Phi_N)$, the naive height of the classical modular polynomial, for *any* level N. The lower bound is the first to be published, to the best of our knowledge. The upper bound gives in particular explicit estimates of an asymptotic result from [Coh84]. Both bounds are optimal, up to the second term. In addition, we remark that there are relations with the size of the modular curve $X_0(N)$. Chapter 6 is [Gij25], with some modifications. It is a study on the possible sources of exceptions to (1), motivated by a question of Daniel Betrand. CM points are a natural first candidate, and we prove that they are indeed exceptions to (1), and compute the transcendence degree.[1] However, we found *positive* dimensional subvarieties of $\mathcal{A}_2$ where the "type of exceptionality" of each CM point is generic: we can more precisely associate one such variety to a given CM point. In particular, CM points are not the only exceptions; there exist larger subvarieties that are exceptions. It is stated in Theorem 6.6. The previous study of the exceptions suggests a differentiated behavior according to (some) special subvarieties of $\mathcal{A}_2$, in particular that $\operatorname{trdeg}\mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q})) \geq 1$ for *any* $\boldsymbol{q} = (q_1, q_2, q_3)$, which would be a direct generalization of the Stéphanois theorem in [BDGP96]. The last two chapters advance into a generalization of the steps from the strategy of the proof. Chapter 7 presents the "modular" proof of the Stéphanois theorem as in [Wal96],[2] with the modification that one can restrict oneself to prime degree isogenies instead of all isogenies. This modification is more suitable to our application to the Igusa invariants, as the generalizations of modular polynomials in the literature are only considered for prime levels. This chapter is also to be read as the one dimensional introduction to Chapter 8, as both part of the same project. In Chapter 8, we generalize all the steps from the modular proof of Stéphanois, except for the last one, hence we are only able to claim a partial Stéphanois.

Chapter 5 is one dimensional in nature, and focus on upper (and lower) bounds on the naive height for the classical modular polynomials. The link with the rest of the thesis is the modular polynomials themselves: they are fundamental in the Stéphanois theorem in [BDGP96], and our work on the Igusa invariants relies on modular polynomials in a higher-dimensional setting. We focus only on the Siegel (see [BL09] and [Mil15]) and Hilbert (see [MR20] or [Mar20], for a different approach) modular polynomials, but they can be defined in the more general setting of Shimura varieties of PEL type, see [Kie21] and [Kie22].

There is another link with the rest of the thesis. In Chapter 5, we considered all the cyclic isogenies of degree N, and during the proof we split them into two groups ("small d" and "large d"), and the ones for "small d" were easier to analyze, as they did not require reduction

[1] conditional to Schanuel's conjecture.
[2] after a suggestion of Daniel Bertrand.

modulo $\mathrm{SL}_2(\mathbb{Z})$. When N is a prime p, this first selected group only has one isogeny. More precisely, for $\tau \in \mathbb{H}$, these two groups of (cyclic) isogenies of degree p are given by

$$\{p\tau\}, \text{ and } \left\{\frac{\tau+b}{p}, \text{ for } 0 \leq b \leq p-1\right\}. \tag{2}$$

It is a known result that the above points equidistribute in the upper half-plane $\mathbb{H}$ with respect to the Poincaré metric, [CU04, Theorem 2.1]. With [Sage], we computed the representatives $\widetilde{\tau} \in \mathcal{F}$ in the standard fundamental domain and plotted them in Figure 0.1.

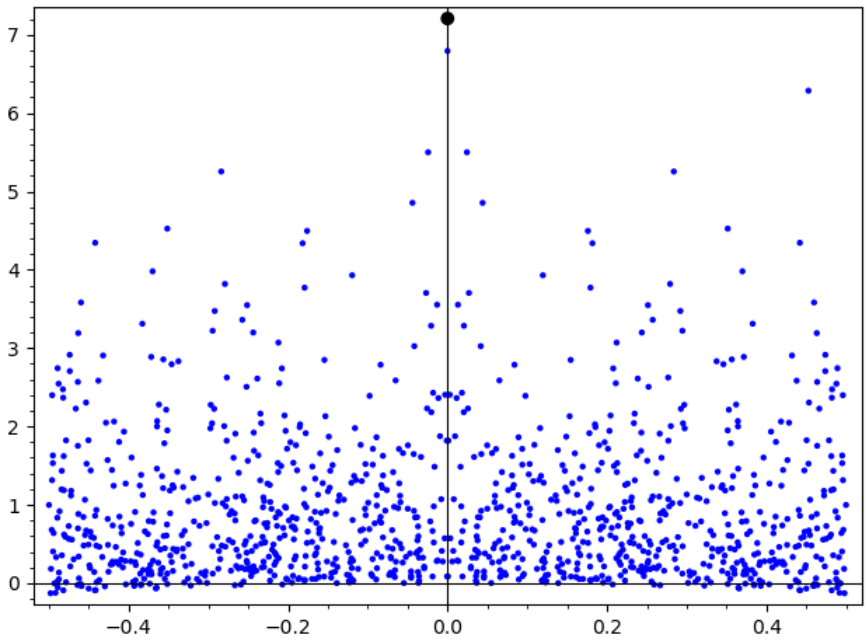


Figure 0.1: For $\tau = 1.23i$, and $p = 1097$, distribution of (2) in the standard fundamental domain $\mathcal{F}$. The imaginary part is in logarithmic scale. The black dot is the distinguished isogeny $p\tau$.

Hence, the isogeny given by $\boldsymbol{\tau} \to p\boldsymbol{\tau}$ is a distinguished one, because of the good properties it translates to a sequence of isogenies $\{n\boldsymbol{\tau}\}_{n\geq 0}$: for every given point $t \in \mathbb{H}$, the sequence always tends to the cusp at infinity. By the equidistribution result mentioned above, that is definitely not the case for any given sequence of isogenies, see Figure 0.2 for an example. The sequence $\{n\boldsymbol{\tau}\}_{n\geq 0}$ is still distinguished on the higher-dimensional setting.

A large part of this work is devoted to a transcendence proof. Here is a really brief sketch of how this type of strategy unfolds. Starting with numbers that are assumed algebraic, one constructs an auxiliary function that is

- "small" in a domain,
- "large" evaluated at the point,

and these upper and lower bounds contradict each other. The lower bounds are derived from the arithmetic information of the input (there are general constraints on how small an algebraic

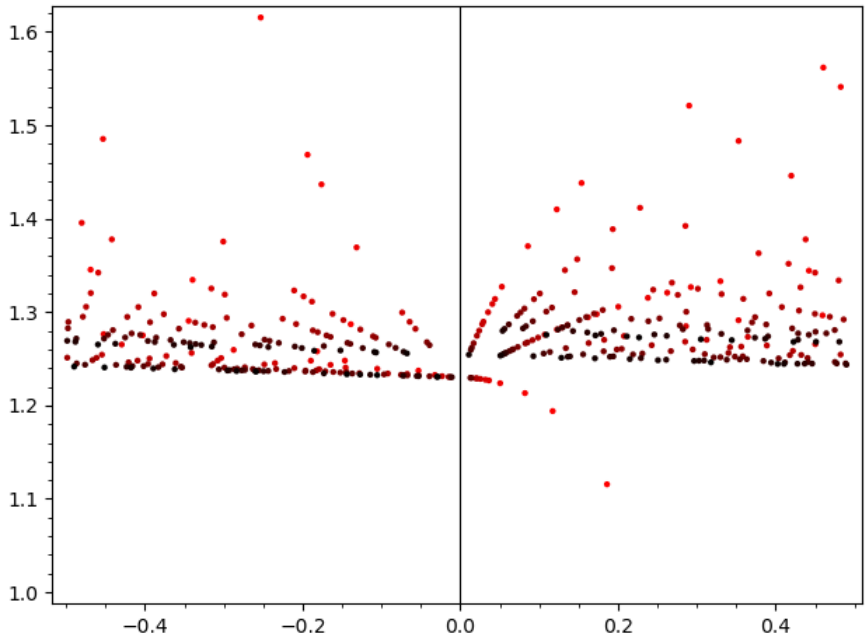


Figure 0.2: For $\tau = 1.23i$, it plots the sequence of isogenies (in increasing darkness of color) $\left\{\frac{\lfloor\sqrt{p}\rfloor+\tau}{p}\right\}$, for $\lfloor\cdot\rfloor$ the floor function and p varying in the 400 first prime numbers.

number can be, in terms of arithmetic data as the degree or height), by results loosely known as Liouville's inequalities, whereas the upper bounds come from complex analysis. Usually, one makes the auxiliary function $F(z)$ have a "high" order of vanishing at 0, let us say M, and by holomorphicity that high order of vanishing "propagates" and makes the function "small" on its domain: $F(z)/z^M$ is holomorphic, and by the maximum principle on that new function, $F(z) = O(z^M)$.

Some chapters have small appendices, and we end the thesis with a common bibliography section.

Contents

Part I

Preliminaries

El valor para marcharse, el miedo a llegar.
Llueve en el canal, la corriente enseña
el camino hacia el mar.
Todos duermen ya.
Dejarse llevar
suena demasiado bien
jugar al azar,
nunca saber dónde puedes terminar...
o empezar.

Copenhague, Vetusta Morla

Chapter 1

Invariants of genus two curves

The goal of this chapter is to present invariants for genus two curves that will help us study their moduli space, with the following properties:

- They admit both a definition in terms of the coefficients of a model of the curve, and as analytic function in a suitable moduli space.
- They are invariants of the isomorphism class (over an algebraically closed field) of the curve.
- They detect algebraicity of the curve.

As such, they generalize the j-invariant of elliptic curves. We will present them from three different perspectives, from algebraic, geometric and analytical point of view, respectively.

1.1 Algebra

On this section, we will focus mostly on describing the invariants algebraically. Fix now an algebraically closed field k of characteristic zero.

A smooth curve of genus two over k is hyperelliptic (see [CF96, Chapter 1, Section 3]), i.e. comes equipped with a ramified degree two cover to $\mathbb{P}^1$. Hence, it admits a model over k of the form $y^2 = p(x)$, with $p \in k[x]$ with all simple roots, as C is smooth. The covering $f : C \to \mathbb{P}^1$ is ramified at $n := \deg(P)$ points. By the Riemann-Hurwitz (see [Har77, Corollary IV.2.4]) formula

$$2 - 2g(C) = \underbrace{\deg(f)}_{=2}(2 - \underbrace{g(\mathbb{P}^1)}_{=0}) - \sum_{x \in C}(e_x - 1) = 4 - n,$$

where g is the genus of the curve and e_x means the ramification index at x. There are n ramification points and because the degree of the cover is 2, necessarily $e_x = 2$ in those points. Therefore, a hyperelliptic curve of genus g has $n = 2g + 2$ ramifications points. As one of those point could be the point at infinity in $\mathbb{P}^1$, if the curve is given by $y^2 = p(x)$, then $\deg(p) = 2g + 2$ or $2g + 1$, according to whether the cover ramifies at infinity or not. For our case of genus two, that means that the polynomial has degree 5 or 6.

That it is also the case with elliptic curves, they admit models defined by polynomials of degree 3 or 4. The standard choice in this case (Weierstrass models) is with degree three, and in this case the model is additionally smooth at infinity.[1] However, this cannot hold for any

[1] The degree four model is singular at the two points at infinity.

higher genus situation, as the genus-degree formula for the arithmetic genus of irreducible plane curves (see [Har77, Chapter V, Example 1.5.1]), predicts $g(C) = \frac{1}{2}(\deg(P)-1)(\deg(P)-2)$, hence the only instance of the arithmetic and smooth genus of C agreeing is for the degree 3 models of elliptic curves. In particular, the models that we can consider for genus two will always be singular at infinity.[2]

Hence, consider for our smooth genus two curve C with model given by $y^2 = p(x)$ with $\deg(p) = 6$, and set α_i, $i = 1, \ldots, 6$ for its (distinct) roots. We can define the following.

$$
\begin{aligned}
I_2(p) &= \sum(\alpha_1-\alpha_2)^2(\alpha_3-\alpha_4)^2(\alpha_5-\alpha_6)^2,\\
I_4(p) &= \sum(\alpha_1-\alpha_2)^2(\alpha_2-\alpha_3)^2(\alpha_3-\alpha_1)^2(\alpha_4-\alpha_5)^2(\alpha_5-\alpha_6)^2(\alpha_6-\alpha_4)^2,\\
I_6(p) &= \sum\prod_{l=1}^{3}\prod_{(i,j)\in C_l}(\alpha_i-\alpha_j)^2, \qquad (1.1)\\
\text{where } C_1 &= \{(1,2),(2,3),(3,1)\}, C_2 = \{(4,5),(5,6),(6,4)\}, C_3 = \{(1,4),(2,5),(3,6)\},\\
I_{10}(p) &= \prod(\alpha_i-\alpha_j)^2,
\end{aligned}
$$

where the sums range among all permutations of six elements. Note that I_{10} is the discriminant of p. As I_κ are symmetric polynomial functions on the roots, they are polynomial functions on the coefficients of f. Historically, they arise from the study of invariants of binary sextics during the 19th century and in [Igu60] they are linked to the study of genus two curves as more suitable analytic functions. They are jointly called *Igusa-Clebsch invariants.*

We note that in [Igu60] five weighted invariants are considered instead of four, so that the theory extends to fields of even characteristic.

Those are weighted invariants, with the weight given by the index in out notation. That means that under an isomorphism ϕ of the curve C,[3] $I_\kappa(\phi^* p) = (\det\phi)^\kappa I_\kappa(p)$. In more precise terms, they induce a map from $\mathcal{M}_2$, the moduli space of projective smooth genus two curves, to the weighted projective space $\mathbb{P}^{2,4,6,10}$, which restricts to $I_{10} \neq 0$.

If we want to consider instead *absolute* invariants on the isomorphism class of the curve, we have to make a non canonical choice of normalization. There are several choices in the literature for normalization of the invariants. As $I_{10} \neq 0$ if and only if C is a smooth curve, dividing by I_{10} is allowed. A standard choice of normalization[4] is

$$j_1(C) = \frac{I_2^5}{I_{10}},\ j_2(C) = \frac{I_2^3 I_4}{I_{10}},\ j_3(C) = \frac{I_2^2 I_6}{I_{10}}.$$

However, we will mostly work with the normalization (Igusa-Streng invariants) from [Str10]. Set $I_6' := 1/2(I_2 I_4 - 3I_6)$, and

$$\tilde{j_1}(C) = \frac{I_4 I_6'}{I_{10}},\ \tilde{j_2}(C) = \frac{I_2 I_4^2}{I_{10}},\ \tilde{j_3}(C) = \frac{I_4^5}{I_{10}^2}.$$

[2] for a smooth projective model, one can embed the curve in a *weighted* projective plane, as in [Gal12, Section 10.1].

[3] a linear change of coordinates in (x, z) for $\bar{p}(x, z)$ the homogenization of p.

[4] up to constants.

This second set of invariants have the advantage of having "smaller" denominators, as we will see. In any case, one can pass from one set to the other via:

$$\tilde{j}_1 = \frac{j_2(j_2 - 3j_3)}{j_1}, \qquad j_1 = \frac{\tilde{j}_2^5}{\tilde{j}_3^2},$$
$$\tilde{j}_2 = \frac{j_2^2}{j_1}, \qquad \text{and} \qquad j_2 = \frac{\tilde{j}_2^3}{\tilde{j}_3},$$
$$\tilde{j}_3 = \frac{j_2^5}{j_1^3}; \qquad j_3 = \frac{\tilde{j}_2^2(\tilde{j}_2 - 2\tilde{j}_1)}{3\tilde{j}_3}.$$

Notice that in both choices of absolute invariants, there is always a locus on $\mathcal{M}_2$ where they vanish simultaneously: the locus $I_2 = 0$, (resp. $I_4 = 0$) for the Igusa invariants (resp. Igusa-Streng invariants), or equivalently $j_1 = 0$ and $\tilde{j}_3 = 0$. That is an artificial side effect of the normalization, and one can choose another normalization for those situations, for example see [CQ05, before Lemma 1, page 74].

Finally, let us note that the Legendre normal form of an elliptic curve

$$y^2 = x(x-1)(x-\lambda),\ \lambda \in \mathbb{P}^1 \setminus \{0, 1, \infty\},$$

obtained by prescribing 3 out of the 4 ramification points to be $\{0, 1, \infty\}$ admits a natural generalization to genus 2 curves. It is called the *Rosenhain normal form* of the curve, for algebraically closed fields of characteristic not equal to two,

$$y^2 = x(x-1)(x-\lambda_1)(x-\lambda_2)(x-\lambda_3),\ \lambda_i \neq \lambda_j,\ \lambda_i \in \mathbb{P}^1 \setminus \{0, 1, \infty\},$$

and $(\lambda_1, \lambda_2, \lambda_3)$ are called the *Rosenhain* invariants. As with the elliptic λ-form, they are only invariants up to a level structure.

1.2 Geometry

The Jacobian of a curve C of genus g, $(J(C), \theta_C)$ is a principally polarized abelian variety (ppav) of dimension g, for more details see [BL04, Section 11] and [Mil86b, Section 1]. The assignment $\mathcal{M}_g \to \mathcal{A}_g$ is called the *Torelli map*.

Theorem 1.1 (Torelli)**.** *Over an algebraically closed field, the Torelli map $T : \mathcal{M}_g \hookrightarrow \mathcal{A}_g$ is injective.*

Proof. This is [Mil86b, Theorem 12.1] [BL04, Theorem 11.1.7] □

The closure of the image of the Torelli map is called *the Torelli locus*, and we call the image of the Torelli map the *open* Torelli locus. For genus $2, 3$ it follows from comparing the dimensions of $\mathcal{M}_g$ and $\mathcal{A}_g$ that the Torelli locus is the full $\mathcal{A}_g$ for $g = 2, 3$. The open Torelli locus in this low genus situation is also determined in the literature.

Definition 1.2. *Given a ppas $(A, \lambda) \in \mathcal{A}_g$, we say that it is* decomposable in $\mathcal{A}_g$ *if $(A, \lambda) \cong (A_1, \lambda_1) \times \ldots (A_r, \lambda_r)$, where $(A_i, \lambda_i) \in \mathcal{A}_{g_i}$, meaning that λ is the product polarization induced by the λ_i.*

Otherwise (A, λ) is called indecomposable*. We defined $\mathcal{A}_g^{ind} \subset \mathcal{A}_g$ as the locus of indecomposable ppav.*

Observe that $\mathcal{A}_g^{ind} = \mathcal{A}_g \setminus \sqcup_{0<k<g}(\mathcal{A}_k \times \mathcal{A}_{g-k})$, where $\mathcal{A}_k \times \mathcal{A}_{g-k}$ embeds naturally into $\mathcal{A}_g$ assigning the product polarization. For $g = 2$, then $\mathcal{A}_2^{ind} = \mathcal{A}_2 \setminus (\mathcal{A}_1 \times \mathcal{A}_1)$.

Theorem 1.3 (Weil)**.** *For $g = 2, 3$ the image of the Torelli morphism is exactly $\mathcal{A}_g^{ind}$. Equivalently, a principally polarized abelian surface (ppas) or threefold is either a Jacobian, or the principally polarized product of lower dimensional varieties.*

Proof. This can be found in [BL04, Corollary 11.8.2], and originally (for genus 2) comes from Weil's proof of the Torelli theorem in [Wei57, Satz 2]. □

Remark 1.4. *The Torelli morphism extents to the Deligne-Mumford compactification of $\mathcal{M}_g$ given by stable curves of compact type $\mathcal{M}_g^{ct}$, see [MO13, Section 1.3], but this extension fails to be injective at the boundary of $\mathcal{M}_g^{ct}$.*

For complex principally polarized abelian varieties, it is well known that its moduli space is given by $\mathrm{Sp}_{2g}(\mathbb{Z})\backslash\mathbb{H}_g$, see for example [BL04, Section 8]. The analytic space is the Siegel upper half-space of degree two:

$$\mathbb{H}_g = \{\boldsymbol{\tau} \in \mathrm{Mat}_g(\mathbb{C}) |\ \boldsymbol{\tau} = \boldsymbol{\tau}^t,\ \mathrm{Im}\,\boldsymbol{\tau} \text{ is positive definite}\}.$$

We denote the symplectic group

$$\mathrm{Sp}_{2g}(\mathbb{Z}) = \left\{ M \in \mathrm{Mat}_{2g}(\mathbb{Z}) |\ M^t J M = J, \text{ where } J = \begin{pmatrix} 0_g & I_g \\ -I_g & 0_g \end{pmatrix} \right\},$$

which acts on $\mathbb{H}_g$ by linear fractional transformations: for $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{Sp}_{2g}(\mathbb{Z})$, then

$$M\boldsymbol{\tau} = (\alpha\boldsymbol{\tau} + \beta)(\gamma\boldsymbol{\tau} + \delta)^{-1}.$$

More precisely, for $\boldsymbol{\tau} \in \mathbb{H}_g$, the associated ppas $(A_{\boldsymbol{\tau}}, H_{\boldsymbol{\tau}}) \in \mathcal{A}_g$ is parametrized by $\mathbb{C}^2 \big/ \boldsymbol{\tau}\mathbb{Z}^g \oplus \mathbb{Z}^g$, with Hermitian form is given by $(\mathrm{Im}\,\boldsymbol{\tau})^{-1}$ with respect to the standard basis of $\mathbb{C}^g$. Alternatively, the Riemann form on the basis $\boldsymbol{\tau}\mathbb{Z}^g \oplus \mathbb{Z}^g$ is on the standard symplectic form $\begin{pmatrix} 0_g & I_g \\ -I_g & 0_g \end{pmatrix}$.

For now on we fix $g = 2$. Then we set $\boldsymbol{\tau} = \begin{pmatrix} \tau_1 & \tau_2 \\ \tau_2 & \tau_3 \end{pmatrix}$, and observe that $\mathrm{Im}\,\boldsymbol{\tau}$ is positive definite if and only if $\tau_1, \tau_3 \in \mathbb{H}_1$ and $\mathrm{Im}(\tau_2)^2 < \mathrm{Im}\,\tau_1\,\mathrm{Im}\,\tau_3$.

1.3 Siegel modular forms

As it is standard in the classical modular forms, we define Siegel modular forms for finite index subgroups Γ of $\mathrm{Sp}_4(\mathbb{Z})$. There are other definitions for Siegel modular forms, that map from $\mathbb{H}_g \to V$ for V a finite-dimensional vector space, see [Gee08, Section 3]. The definition we are going to use is what is customary called *classical* or *scalar valued* Siegel modular form.

Definition 1.5. *Let Γ a finite index subgroup of $\mathrm{Sp}_4(\mathbb{Z})$, and $k \in \mathbb{Z}_{>0}$. Let $f : \mathbb{H}_2 \to \mathbb{C}$ a holomorphic function. We say that f is a* holomorphic Siegel modular form *of weight k with respect to Γ if for every $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma$ it holds*

$$f(M\boldsymbol{\tau}) = \det(\gamma\boldsymbol{\tau} + \delta)^k f(\boldsymbol{\tau}).$$

We call a quotient of two modular forms of the same weight for Γ *a* Siegel modular function *for* Γ.

We denote the (finite) $\mathbb{C}$*-vector space of Siegel modular forms of weight* k *with respect to* Γ *by* $A_k(\Gamma)$*, and set* $A(\Gamma) = \oplus_k A_k(\Gamma)$ *for the graded ring of Siegel modular forms with respect to* Γ*. We set the notation* $\mathbb{C}(\mathcal{A}_2)$ *for the field of functions of* $\mathcal{A}_2$*, that corresponds to the field of Siegel modular functions for* $\mathrm{Sp}_4(\mathbb{Z})$.

A holomorphic Siegel modular form f is in particular invariant under $\boldsymbol{\tau} \to \boldsymbol{\tau} + T$ for $T \in \mathrm{Sym}_2(\mathbb{Z})$, so admits a Fourier expansion of the form

$$f = \sum_{T \in \mathrm{Sym}_2^\vee(\mathbb{Z})} a_f(T) e^{i2\pi \,\mathrm{tr}(T\boldsymbol{\tau})}$$

for $\mathrm{Sym}_2^\vee(\mathbb{Z}) = \{M \in \mathrm{Sym}_2(\mathbb{Q}) : \mathrm{tr}(MT) \in \mathbb{Z} \text{ for all } M \in \mathrm{Sym}(\mathbb{Z})\}$, with Fourier coefficients given by

$$a_f(T) = \int_{\boldsymbol{x} \bmod 1} f(\boldsymbol{\tau}) e^{-i2\pi \,\mathrm{Tr}(T\boldsymbol{\tau})} d\boldsymbol{x},$$

where $\boldsymbol{\tau} = \boldsymbol{x} + i\boldsymbol{y}$. One can see that $\mathrm{Sym}_2^\vee(\mathbb{Z})$ consists of matrices M such that $2M$ has integral coefficients with the diagonal elements being even. These matrices are customary called *half integral symmetric*. In addition, by the invariance of unimodular transformations $\boldsymbol{\tau} \mapsto {}^t U \boldsymbol{\tau} U$ for $U \in \mathrm{GL}_2(\mathbb{Z})$, one has

$$a_f({}^t UTU) = a_f(T), \text{ for all } U \in \mathrm{GL}_2(\mathbb{Z}).$$

Remark 1.6. *In Subsection 8.1.1, we will interpret the abstract Fourier expansions as a power series with integral coefficients in terms of suitable* q*-variables.*

The first important difference with respect to the one dimensional case is that it is not necessary to impose boundness (or holomorphicity) conditions at infinity, as in the higher dimensional case those are automatic when f is holomorphic at $\mathbb{H}_2$ (and any $\mathbb{H}_g$), by what is called the Koecher principle.

Theorem 1.7 (Koecher)**.** *Let* f *a holomorphic Siegel modular form of weight* k*. Then* f *is bounded on any subset of* $\mathbb{H}_2$ *of the form* $\{\boldsymbol{\tau} \in \mathbb{H}_: \ \mathrm{Im}(\boldsymbol{\tau}) > cI_2\}$ *for fixed* $c > 0$*, where* $A > B$ *means* $(A - B)$ *is a positive definite matrix.*

In terms of the Fourier coefficients, it holds that for all T *not positive semi-definite,* $a_f(T) = 0$.

Hence the Fourier expansions of a Siegel modular form for $\mathrm{Sp}_4(\mathbb{Z})$ is indexed over $T \in \mathrm{Sym}_2^\vee(Z)$ such that T is positive semi-definite. We denote this subset by $\mathrm{Sym}_2^{\vee,+}(\mathbb{Z})$. Remark that if $T \in \mathrm{Sym}_2^{\vee,+}(\mathbb{Z})$ then either $\det(T) > 0$ and T is positive definite, or $\det(T) = 0$.

Proof. See [Gee08, Section 4, Theorem 1 and Theorem 2]. □

Definition 1.8. *Let* f *a Siegel modular form for* $\mathrm{Sp}_4(\mathbb{Z})$*. If* $a_f(N) = 0$ *for all* $N \in \mathrm{Sym}_2^{\vee,+}(\mathbb{Z})$ *with* $\det(N) = 0$*, we say that* f *is a* cusp form*. We denote the space of cusps forms of weight* k *by* $S_k(\mathrm{Sp}_4(\mathbb{Z}))$.

Remark 1.9. *In [Gee08, Section 5] there is an equivalent definition of cusp form using the Siegel operator.*

Proposition 1.10. *There are no non-trivial Siegel modular forms of negative weight.*

Remark 1.11. *There can be Siegel modular forms for* $\mathrm{Sp}_4(\mathbb{Z})$ *of odd weight (and in general for* $\mathrm{Sp}_{2g}(\mathbb{Z})$ *with* g *even).*

We can define analogs of the Eisenstein series and the theta constants.

Definition 1.12. *For* $k > 2$ *even we call the Eisenstein series*[5] *of weight* k

$$E_k(\boldsymbol{\tau}) = \sum_{\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbb{Z})\backslash \mathrm{Sp}_4(\mathbb{Z})} \det(\gamma\boldsymbol{\tau} + \delta)^{-k},$$

where the sum ranges over a set of representatives of $\mathrm{GL}_2(\mathbb{Z})\backslash \mathrm{Sp}_4(\mathbb{Z})$*, where* $\mathrm{GL}_2(\mathbb{Z})$ *is identified with the subgroup of unimodular transformations* $\left(\begin{smallmatrix} U & 0\\ 0 & {}^tU^{-1}\end{smallmatrix}\right) \in \mathrm{Sp}_4(\mathbb{Z})$ *acting by left multiplication.*

The Eisenstein series define a holomorphic Siegel modular form of weight k for $\mathrm{Sp}_4(\mathbb{Z})$, by [Gee08, Theorem 5.3].

We can also generalize the classical theta constants. We follow the conventions from [Str10, Section 7].

Definition 1.13. *We call an element* $c \in \{0, 1/2\}^4$ *a* theta characteristic. *We define the* theta constant *or* Theta Nullwerte of characteristic c *to be the function* $\theta[c] : \mathbb{H}_2 \to \mathbb{C}$ *defined by*

$$\theta[c](\boldsymbol{\tau}) = \sum_{n\in\mathbb{Z}^2} \exp\left(i\pi\left((n + c_a)\boldsymbol{\tau}\,{}^t(n + c_b) + 2(n + c_a)\,{}^tc_b\right)\right),$$

where for $c = (c_1, c_2, c_3, c_4)$ *we set* $c_a = (c_1, c_2)$ *and* $c_b = (c_3, c_4)$.

It follows that $\theta[c](\boldsymbol{\tau})$ is trivially zero if and only if $4c_a\,{}^tc_b \in \{0, 1\}$ is odd. We call the remaining ones *even* theta constants. Out of the 16 possible theta constants, there are 10 of them that are even, and they are given by:

$$\begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1/2\end{pmatrix}, \begin{pmatrix}0\\0\\1/2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1/2\\1/2\end{pmatrix}, \begin{pmatrix}1/2\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1/2\\0\\0\end{pmatrix}, \begin{pmatrix}1/2\\1/2\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1/2\\1/2\\0\end{pmatrix}, \begin{pmatrix}1/2\\0\\0\\1/2\end{pmatrix}, \begin{pmatrix}1/2\\1/2\\1/2\\1/2\end{pmatrix}.$$

We can also follow the convention by [Dup06], $c \mapsto 16c_2 + 8c_1 + 4c_4 + 2c_3$ is injective on the even theta characteristics and takes the values $\{0, 1, 2, 3, 4, 6, 8, 9, 12, 15\}$, so we could also use the notation $\theta_{16c_2+8c_1+4c_4+2c_3} = \theta[c]$.

They are holomorphic functions on $\mathbb{H}_2$ and Siegel modular forms[6] of weight $1/2$ for a suitable subgroup of $\mathrm{Sp}_4(\mathbb{Z})$.

Consider the following subgroups of $\mathrm{Sp}_4(\mathbb{Z})$:

$$\Gamma(n) = \ker\left(\mathrm{Sp}_4(\mathbb{Z}) \to \mathrm{Sp}_4\left(\mathbb{Z}/n\mathbb{Z}\right)\right),$$

$$\Gamma(n, 2n) = \left\{\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix} \in \Gamma(n),\ \alpha\,{}^t\beta \equiv \gamma\,{}^t\delta \equiv 0_2\ (2n)\right\}.$$

[5] On higher genus, one can even generalize them to Klingen Eisenstein series as in [Gee08, Section 5].

[6] with respect to a certain multiplier system.

Theorem 1.14. *The following properties hold.*

1. *(Igusa) We have* $A(\Gamma(4,8)) = \mathbb{C}[\theta_m\theta_n]$*, where we mean the ring generated by polynomials in all products of two theta constants.*

2. *We have* $A(\mathrm{Sp}_4(\mathbb{Z}))^{even} = \mathbb{C}[E_4, E_6, E_{10}, E_{12}]$*, for* $A(\mathrm{Sp}_4(\mathbb{Z}))^{even} = \sum_k A_{2k}\,\mathrm{Sp}_4(\mathbb{Z}))$*. There exists a Siegel modular cusp form of weight* 35*,* χ_{35}*, such that*

$$A(\mathrm{Sp}_4(\mathbb{Z})) = \mathbb{C}([E_4, E_6, E_{10}, E_{12}, \chi_{35})/\chi_{35}^2 = P$$

   for P *a specific isobaric polynomial in* E_4, E_6, E_{10}, E_{12}*.*

3. *(Thomae's formula, Rosenhain form) Let* $C \in \mathcal{M}_2$ *a genus two curve, and* $\boldsymbol{\tau} \in \mathbb{H}_2$ *a normalized period matrix for* $(\mathrm{Jac}(C), \theta) \in \mathcal{A}_2^{ind}$*. Then,* C *is isomorphic to* $y^2 = x(x-1)(x-\lambda_1)(x-\lambda_2)(x-\lambda_3)$*, for*[7]

$$\lambda_1(\boldsymbol{\tau}) = \left(\frac{\theta_{12}(\boldsymbol{\tau})\theta_4(\boldsymbol{\tau})}{\theta_8(\boldsymbol{\tau})\theta_0(\boldsymbol{\tau})}\right)^2,\ \lambda_2(\boldsymbol{\tau}) = \left(\frac{\theta_6(\boldsymbol{\tau})\theta_{12}(\boldsymbol{\tau})}{\theta_2(\boldsymbol{\tau})\theta_8(\boldsymbol{\tau})}\right)^2,\ \lambda_3(\boldsymbol{\tau}) = \left(\frac{\theta_6(\boldsymbol{\tau})\theta_4(\boldsymbol{\tau})}{\theta_2(\boldsymbol{\tau})\theta_0(\boldsymbol{\tau})}\right)^2.$$

Proof. See [Igu64, Theorem 1], [Igu62, Corollary to Theorem 2] and [Igu67, page 849]. The second part is attributed to Rosenhain by Igusa in [Igu60], and a proof can be found in [Gru08, Proposition 3.31]. □

Remark 1.15. *In [Igu62, page 179] it is given the Taylor expansions of* λ_i *around* $\tau_2 = 0$*. It follows that they can be holomorphically extended to diagonal matrices* $\boldsymbol{\tau} = \begin{pmatrix}\tau_1 & 0\\ 0 & \tau_2\end{pmatrix}$*, by* $\lambda_1(\boldsymbol{\tau}) = \lambda_2(\boldsymbol{\tau}) = \lambda_3(\boldsymbol{\tau}) = \lambda(\tau_1)$*, for the elliptic modular* λ*-function on* τ_1*.*

Finally, it follows that $\mathcal{A}_2$ is a *rational* threefold, meaning that its field of functions (of transcendence degree three) can be generated by three functions. In [Igu62, Section II], it is furthermore realized as birational to a weighted projective variety, and given as a suitable blow-up.[8] See also [Kli90, Proposition 11.3]. However, it is a low genus phenomena that the Siegle modular varieties $\mathcal{A}_g$ are rational, for $g \geq 7$ they are of general type, and $\mathcal{A}_2(n)$ is also of general type for $n \geq 4$, see [HS02, Theorem II.2.1].

1.3.1 From geometry to algebra

By Theorem 1.14, theta constants allow us to express the algebraic invariants in Equation (1.1) (invariants of the genus two curve in algebraic terms) as functions on $\mathcal{A}_2^{ind}$ (functions on its Jacobians), by given an inverse to Torelli's morphism.[9]

Therefore, they admits expressions in terms of theta constants. We follow the notation form [Str10]. Consider the set T of the 10 even theta characteristics, and define the following

[7] There are 720 choices for the Rosenhaim invariants, this is the choice of [Igu62].
[8] which is also the toroidal compactification of $\mathcal{A}_2$.
[9] According to [Str10, Remark 7.4], it was already done by Bolza and Spallek.

functions, originally from [Igu67, page 848]:

$$h_4 = \sum_{c\in T} \theta[c]^8,$$
$$h_6 = \sum_{\substack{b,c,d\in T \\ \text{syzygous}}} \pm(\theta[b]\theta[c]\theta[d])^4$$
$$h_{10} = \prod_{c\in T} \theta[c]^2,$$
$$h_{12} = \sum_{C\in S} \prod_{c\in T\setminus C} \theta[c]^4,$$

and see [Str10, Section 7.1] for where the sums are defined.

Proposition 1.16. *Recall the (algebraic) Igusa-Clebsch invariants $I_2, I_4, I_6, I_{10}, I_{12}$ from Equation (1.1). They define Siegel modular forms by evaluating them at $C : y^2 = x(x-1)(x-\lambda_1(\boldsymbol{\tau})(x-\lambda_2(\boldsymbol{\tau}))(x-\lambda_3(\boldsymbol{\tau}))$, that we denote with the same letter. Then if $\boldsymbol{\tau} \in \mathcal{A}_2^{ind}$, it holds*

$$I_2(\boldsymbol{\tau}) = \frac{h_{12}(\boldsymbol{\tau})}{h_{10}(\boldsymbol{\tau})},\ I_4(\boldsymbol{\tau}) = h_4(\boldsymbol{\tau}),\ I_6'(\boldsymbol{\tau}) = h_6(\boldsymbol{\tau}),\ I_{10}(\boldsymbol{\tau}) = h_{10}(\boldsymbol{\tau}),$$

and each I_κ are holomorphic Siegel modular forms of weight κ for $\mathrm{Sp}_4(\mathbb{Z})$.

Proof. This is [Igu67, page 848], in the form of [Str10, Lemma 7.3]. □

It is customary to use another set of generators for $A(\mathrm{Sp}_4(\mathbb{Z}))$, given by Eisenstein series and cusp forms. Considering the normalized cusp forms of Igusa [Igu62, page 195]:

$$\chi_{10} = \frac{-43867}{2^{12}\cdot 3^5\cdot 5^2\cdot 7\cdot 53}(E_4E_6 - E_{10})$$

and

$$\chi_{12} = \frac{131\cdot 593}{2^{13}\cdot 3^7\cdot 5^3\cdot 7^2\cdot 337}\left(3^2\cdot 7^2E_4^3 + 2\cdot 5^3E_6^2 - 691E_{12}\right).$$

One uses $E_4, E_6, \chi_{10}, \chi_{12}$ as generators instead. Remark that χ_{10} resembles the formula for the modular discriminant in terms of elliptic Eisenstein series. They can be related to the functions h_κ from above.

Proposition 1.17. *One has the following equalities for the Eisenstein series and cusp forms χ_{10} and χ_{12}.*

$$h_4 = 2^2E_4,\ h_6 = 2^2E_6,\ h_{10} = -2^{14}\chi_{10},\ h_{12} = 2^{17}3\chi_{12}.$$

The Fourier coefficients of $E_4, E_6, 4\chi_{10}, 12\chi_{12}$ are all integers, with each one of them having at least one Fourier coefficients being 1.

Proof. The first part is from [Igu67, page 848], and the second is [Str10, Appendix 1]. □

We write the Igusa and the Igusa-Streng invariants in terms of the standard generators as:

$$j_1 = 2\cdot 3^5\frac{\chi_{12}^5}{\chi_{10}^6},\ j_2 = \frac{3^3}{2^3}\frac{E_4\chi_{12}^3}{\chi_{10}^4},\ j_3 = \frac{3}{2^5}\frac{E_6\chi_{12}^2}{\chi_{10}^3} + \frac{3^2}{2^3}\frac{E_4\chi_{12}^3}{\chi_{10}^4},$$
$$\tilde{j_1} = \frac{1}{2^{10}}\frac{E_4E_6}{\chi_{10}},\ \tilde{j_2} = 2^7\cdot 3\frac{\chi_{12}E_4^6}{\chi_{10}},\ \tilde{j_3} = \frac{1}{2^{18}}\frac{E_4^5}{\chi_{10}^2}.$$

Remark 1.18. *Both set of invariants have denominators with powers of χ_{10}, and the Igusa-Streng invariants have them with lower powers.*

Proposition 1.19. *It follows that $\mathbb{C}(\mathcal{A}_2) = \mathbb{C}(j_1, j_2, j_3)$.*

Proof. It follows from [Igu62] that the field of functions is generated by three algebraically independent modular functions, and it follows form the formulas above that j_k are algebraically independent. □

By our geometric considerations, we expected those denominators: we constructed the invariants for the image of the Torelli morphism $\mathcal{A}_2^{ind}$, passing back to the invariants from $\mathcal{M}_2$. By Proposition 1.16 and Proposition 1.17 above, χ_{10} is a multiple of I_{10}, which does not vanish on $\mathcal{M}_2$ for that the curves are smooth. The complement of $\mathcal{A}_2^{ind}$ is $\mathcal{A}_1 \times \mathcal{A}_1$, identified with the canonical polarized product of elliptic curves. By Remark 1.4, they can be realized as generalized jacobians for curves in $\mathcal{M}_2^{ct}$, *but not in a unique way.* This is also corroborated analytically for our choice of Rosenhain invariants from Remark 1.15: they are defined, but they degenerate and become independent of τ_3, which would be absurd for invariants.

Reciprocally, on $\mathcal{A}_1 \times \mathcal{A}_1$, it follows that $\chi_{10} = 0$, by [Kli90, Proposition 9.2].

Chapter 2

Hilbert modular surfaces

This chapter is about Hilbert modular surfaces, from neither a self-contained nor exhaustive perspective. We present them as a solution to the moduli problem of abelian surfaces with real multiplication,[1] and then we present Hilbert modular forms. Finally, we consider the "forgetful functor" on our moduli space to map then into the Siegel threefold $\mathcal{A}_2$.

As we will see, complex tori with real multiplication are always polarizable, and hence the complex moduli space is in bijection with a quotient of a Cartesian product $\mathbb{H}^g$, where our group of transformations act component-wise. From this perspective, Hilbert modular forms are a more direct generalization of classical modular forms that the Siegel modular forms.

2.1 Note on conventions

This first section is a warning that there exists different conventions in the literature for the Hilbert modular group and its action on the Hilbert space. All of them are compatible (under some extra conditions) and solve the same moduli problem, but in our explicit manipulations these conventions do matter. Here we translate from one convention to another, that may be useful for the reader when navigating other different sources. After this section we will fix $\mathrm{SL}_2(\mathcal{O}_K)$ acting on $\mathbb{H}^2$ as a default.

Set K a real quadratic field over $\mathbb{Q}$ of discriminant Δ_K, $\widetilde{(\cdot)}$ for its non-trivial Galois automorphism, $\mathcal{O}_K$ its ring of integer, $\mathfrak{a}$ a fractional ideal, and $\mathfrak{a}^{-1}$ its inverse. Set also for $\mathfrak{a}$ a fractional ideal, $\mathfrak{a}^{++}$ the subset of totally positive elements. Then:

$$\mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(K)\ a, d \in \mathcal{O}_K,\ b \in \mathfrak{a}, c \in \mathfrak{a}^{-1} \right\} = \begin{pmatrix} \mathcal{O}_K & \mathfrak{a} \\ \mathfrak{a}^{-1} & \mathcal{O}_K \end{pmatrix} \cap \mathrm{SL}_2(K),$$

which is the group of K-linear isomorphisms $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of $K^2 \to K^2$ of determinant one that satisfy[2]

$$(\mathcal{O}_K \oplus \mathfrak{a}) \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a\mathcal{O}_K + c\mathfrak{a}) \oplus (b\mathcal{O}_K + d\mathfrak{a}) \subset \mathcal{O}_K \oplus \mathfrak{a}.$$

We set $\mathrm{SL}_2(\mathcal{O}_K) = \mathrm{SL}_2(\mathcal{O}_K \oplus \mathcal{O}_K)$, and $\partial_K = \sqrt{\Delta_K}\mathcal{O}_K$ and $\partial_K^{-1} = \frac{1}{\sqrt{\Delta_K}}\mathcal{O}_K$ for the different and codifferent of K, respectively. Remark that $\partial_K^{-1} = \mathcal{O}_K^\vee$, where for $\mathfrak{a}$ a fractional

[1] the details of the proof are omitted, for that in Section 4.1 in the next chapter we will give the proof for quaternionic multiplication, and the details are of very similar flavor.

[2] Some authors [Gee88] and [Bru08] follow the convention of transposing such matrices.

ideal in K, we denote the dual with respect to the trace as $\mathfrak{a}^\vee = \{\alpha \in K : \mathrm{tr}_{K|\mathbb{Q}}(\alpha\mathfrak{a}) \subset \mathbb{Z}\}$. Remark too that

$$\mathfrak{a}^\vee = \mathfrak{a}^{-1}\mathcal{O}_K^\vee = \mathfrak{a}^{-1}\partial_K^{-1} \tag{2.1}$$

We use the translation from [MR20, beginning of Section 2.2, Equations (4) and (5)] and [EK14, Section 3]. Roughly, we can think of either

1. the group $\mathrm{SL}_2(\mathcal{O}_K)$ acting on $\mathbb{H}^2$,
2. the group $\mathrm{SL}_2(\mathcal{O}_K)$ acting on $\mathbb{H} \times (-\mathbb{H})$,
3. or the group $\mathrm{SL}_2(\mathcal{O}_K \oplus \partial_K^{-1})$ acting on $\mathbb{H}^2$.

The action in any case is given by linear fractional transformations. As $K \otimes_{\mathbb{Q}} \mathbb{R} = \mathbb{R}^2$ for that K is totally real, there is an embedding of $\mathrm{SL}_2(K) \hookrightarrow (\mathrm{SL}_2(\mathbb{R}))^2$ acting on $\mathbb{C}^2$ by componentwise linear fractional transformations. In other words, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(K)$ acts on $\mathbb{C}^2$ by

$$(z_1, z_2) \mapsto \left(\frac{az_1 + b}{cz_2 + d}, \frac{\widetilde{a}z_2 + \widetilde{b}}{\widetilde{c}z_2 + \widetilde{d}} \right). \tag{2.2}$$

The three actions above are identified via the isomorphism, from number 2) to number 3)

$$\begin{aligned} \mathrm{SL}_2(\mathcal{O}_K) &\to \mathrm{SL}_2(\mathcal{O}_K \oplus \partial_K^{-1}) \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto \begin{pmatrix} a & \frac{b}{\sqrt{\Delta_K}} \\ c\sqrt{\Delta_K} & d \end{pmatrix} \\ \mathbb{H} \times (-\mathbb{H}) &\to \mathbb{H}^2 \\ (z_1, z_2) &\mapsto \left(\frac{z_1}{\sqrt{\Delta_K}}, \frac{-z_2}{\sqrt{\Delta_K}} \right). \end{aligned} \tag{2.3}$$

More precisely, we have the following commutative diagram ([EK14, Section 3 p. 2308])

$$\begin{array}{ccc} \mathrm{SL}_2(\mathcal{O}_K) \times (\mathbb{H} \times (-\mathbb{H})) & \longrightarrow & \mathbb{H} \times (-\mathbb{H}) \\ \downarrow & & \downarrow \\ \mathrm{SL}_2(\mathcal{O}_K \oplus \partial_K^{-1}) \times \mathbb{H}^2 & \longrightarrow & \mathbb{H}^2, \end{array}$$

where the horizontal maps are given by $(\gamma, \boldsymbol{z}) \mapsto \gamma\boldsymbol{z}$ as in (2.2) above, and the vertical maps are induced by the isomorphisms in Equation (2.3).

Furthermore, assume that K has a fundamental unit ε of norm -1, and set $\varepsilon > 0$. Then $\frac{\sqrt{\Delta_K}}{\varepsilon} \in \mathcal{O}_K^{++}$, then there is an identification 1) to 3)

$$\begin{aligned} \mathrm{SL}_2(\mathcal{O}_K) &\to \mathrm{SL}_2(\mathcal{O}_K \oplus \partial_K^{-1}) \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto \begin{pmatrix} a & b\frac{\varepsilon}{\sqrt{\Delta_K}} \\ c\frac{\sqrt{\Delta_K}}{\varepsilon} & d \end{pmatrix} \\ \mathbb{H}^2 &\to \mathbb{H}^2 \\ (z_1, z_2) &\mapsto \left(\frac{\varepsilon}{\sqrt{\Delta_K}} z_1, \frac{-\widetilde{\varepsilon}}{\sqrt{\Delta_K}} z_2 \right). \end{aligned} \tag{2.4}$$

One would naively want to think of convention number 1). However, the more natural one in the context of the moduli spaces (see Corollary 2.10) is number 3), which is only canonically translated to number 2) via dividing by $\sqrt{\Delta}$, of negative norm $-\Delta$, and that makes the change to $\mathbb{H} \times (-\mathbb{H})$. We would want to change $\sqrt{\Delta}$ so it is a totally positive element, so that the final conversion requires *units of norm* -1 in the field, which needs the simplifying assumption that the fundamental unit has negative norm.

The companion modular groups $\mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a})$ naturally show up in both the moduli space problem of abelian surfaces with real multiplication by $\mathcal{O}_K$, and in the Fourier expansions of Hilbert modular forms around cusps. From the point of view of Hilbert modular forms, they allow us to work "cusp-free": the Hilbert modular surface has a finite number of cusps, in bijection with the class group of K (see Lemma 2.15), and for a given Hilbert modular form, to study its behavior around a different cusp entails changing the modular group by the fractional ideal. We give more details in Subsection 2.3.2.

From a geometric perspective, they are also necessary to completely describe the moduli space of polarized abelian surfaces with real multiplication, as we will see in the next section. It is also the case that considering all groups $\mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a})$ is a natural consequence to working in the adelic language, as in [Gee88, Chapter I, Section 7].

2.2 Moduli space

We present the material of this section from [Gor02, Chapter 2, Section 2], [Gee88, Chapter IX, Section 1] and [EK14, Section 2].

The following construction holds more generally for g-dimensional abelian varieties with real multiplication by a totally real field of degree g, but we are only going to give the proofs in dimension two for simplicity.

Definition 2.1. *Consider a complex torus C of dimension two, K a real quadratic field, and $\mathcal{O} \subset K$ an order in K. We say that C* admits real multiplication by $\mathcal{O} \subset K$ *if there is an embedding of rings*

$$\mathcal{O} \hookrightarrow \mathrm{End}(C).$$

When $\mathcal{O} = \mathcal{O}_K$, the ring of integers of K we may say that C admits real multiplication by K*, or that it admits* maximal real multiplication by K*.*

Consider a principally polarized complex abelian surface (A, E), and set $\mathrm{End}^s(A, \lambda) \subset \mathrm{End}(A)$ *for the subring of symmetric endomorphism with respect to the Rosati involution. We say that A has* real multiplication by $\mathcal{O} \subset K$ *if*

$$\mathcal{O} \hookrightarrow \mathrm{End}^s(A, E).$$

We say that (A, E, ι) and (A', E', ι') are isomorphic as ppas with real multiplication by $\mathcal{O}$ if there is an isomorphism $\phi : (A, E) \to (A', E')$ such that ι and ι' are compatible, meaning that for $\phi^* : \mathrm{End}_0(A') \to \mathrm{End}_0(A)$, we have $\phi^* \circ \iota' = \iota$.

Remark 2.2. *Some authors, for example [Gee88, Chapter IX, Definition 1.3], follow another definition of real multiplication by K, requiring a $\mathbb{Q}$-algebra embedding $K \hookrightarrow \mathrm{End}_0(C)$, then setting $\mathcal{O} = K \cap \mathrm{End}(C)$, C has real multiplication by $\mathcal{O}$. We do not follow this convention here.*

Alternatively, for the notion of isomorphism, one could pass through the universal cover $\mathbb{C}^2$ and, via fixing the analytic and rational representations ρ_a and ρ_r and fix the action of K as in Equation (2.5) below. Then the compatibility condition is that the action of K commutes with $\rho_a(\phi)$ on C^2. See [Gor02, Section 2.1, Assertion (A) above Equation (2.62)].

Example 2.3. *Consider E an elliptic curve with* $\mathrm{End}(E) = \mathbb{Z}$*, and set $A = E^2$. We will see in Lemma 6.25 that, as a complex torus, it has RM by* every *quadratic field K.*

However, with the standard product principal polarization on E^2, the embeddings given in Lemma 6.25 only make it a RM abelian surface for the quadratic fields $K = \mathbb{Q}(\sqrt{d})$, d square free, such that d is the sum of two squares, which by Jacobi's two square theorem, make $d = 2^j p_1 \dots p_r$ for $j = 1, 2$ and $p_i \equiv 1 \pmod 4$. We will see in Example 4.24 that these are actually all of them.

We will see below that, unlike for general complex tori, that not always are abelian varieties, a complex torus with real multiplication by $\mathcal{O} \subset K$ is *always* polarizable. This also happens with complex torus admitting quaternionic and complex multiplication, see Remark 4.11. From [Cla03, Paragraph under Proposition 5.1, Section 0.3.2], where they are called *complex tori with sufficiently many endomorphisms*, they are always polarizable.[3]

The non-automatic detail is that the polarization is *principal*. It is a general fact of abelian varieties that any polarization is induced from a principal polarization via an isogeny [BL04, Proposition 4.1.2]. The endomorphism algebra is invariant under isogenies but that is not true for the endomorphism *ring*, so this new abelian surface will have real multiplication by a different order in the field.

Reciprocally, for a simple complex abelian variety (or over an algebraically closed field of characteristic zero), one can find isogenous abelian varieties A with $\mathrm{End}(A)$ isomorphic to any fixed order of the starting endomorphism algebra ([Rém17, Lemme 7.9]).[4] In the case of real multiplication, for simple abelian surfaces one can always consider an isogenous one with *maximal* real multiplication, but it may not admit a principal polarization.

These two facts suggest the following duality. For a given polarized abelian surface with real multiplication by an order $\mathcal{O} \subset K$, then *up to isogeny*:

- the polarization can be assumed principal, or
- the order can be assumed maximal.

In general these two conditions do not need to happen at the same time, there are examples of complex abelian varieties with neither of them, see [GR14a, Proposition 11.1].[5] Let us focus now on the moduli problem of complex *principally* polarized abelian surfaces (A, λ, ι) with *maximal* real multiplication. As a complex torus, $A = \mathbb{C}^2/\Lambda$, with an embedding $\iota : \mathcal{O}_K \hookrightarrow \mathrm{End}(A)$, we can consequently consider Λ as an $\mathcal{O}_K$-module.

Lemma 2.4. *With notation as above, Λ is a* projective *$\mathcal{O}_K$-module of rank two, and there exists fractional ideals $\mathfrak{a}$ and $\mathfrak{b}$ of K such that $\Lambda \cong \mathfrak{a} \oplus \mathfrak{b}$ as $\mathcal{O}_K$-modules. Furthermore,*

[3] This notion requires care, there exists *non-algebraizable* complex tori with "many" endomorphisms, but their endomorphism algebra are not Albert algebras, see [OZ95].

[4] and for non-necessarily simple, combining [Rém17, Théorème 1.2] and [GR14a, Théorème 1.6], there is an isogenous A maximal with $\mathrm{End}(A)$.

[5] remark that their construction of the counterexample does not work for genus 2, as there would be the exceptions b) and c) of [BL04, Theorem 9.9.1].

$\mathfrak{a} \oplus \mathfrak{b} \cong \mathcal{O}_K \oplus \mathfrak{ab}$ *and the isomorphism class of the module is completely determined by* $\mathfrak{ab}$ *in* $\mathrm{Cl}(K)$

Proof. As $\operatorname{rank}_{\mathbb{Z}} \Lambda = 4$ and $\operatorname{rank}_{\mathbb{Z}} \mathcal{O}_K = 2$, it follows that the $\mathcal{O}_K$-rank of Λ is two. As Λ is torsion-free as an $\mathcal{O}_K$-module, and $\mathcal{O}_K$ is a Dedekind domain, Λ is flat. Hence, it is a flat and finitely generated module over a Noetherian ring, so it is projective.

The rest of the statement follows from [FT93, Chapter II, Section 4, Theorem 13]. □

In light of this result, let us know consider $\mathfrak{a}, \mathfrak{b}$ fractional ideals of K and set $\Lambda = \mathfrak{a} \oplus \mathfrak{b}$. We define an alternating bilinear form on $\Lambda \times \Lambda$ by restricting a multiple of the standard alternating bilinear form on $K^2 \times K^2 \to K$ given by the determinant, and then considering the field trace to $\mathbb{Q}$.

In details, consider for $r \in K^\times$,

$$\begin{aligned} E_r : \Lambda \times \Lambda &\to \mathbb{Q} \\ (a_1, b_1) \times (a_2, b_2) &\mapsto \operatorname{tr}_{K|\mathbb{Q}} \left(r \det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \right) = \operatorname{tr}(r(a_1 b_2 - a_2 b_1)). \end{aligned}$$

For given $\boldsymbol{z} = (z_1, z_2) \in \mathbb{H}^2$, define

$$\begin{aligned} \phi_{\boldsymbol{z}} : K \oplus K &\to \mathbb{C}^2 \\ (l, k) &\mapsto l\boldsymbol{z} + k = (lz_1 + k, \widetilde{l} z_2 + \widetilde{k}), \end{aligned}$$

and

$$\Lambda_{\boldsymbol{z}} = \phi_{\boldsymbol{z}}(\Lambda) = \phi_{\boldsymbol{z}}(\mathfrak{a} \oplus \mathfrak{b}) = \{(a_1 z_1 + b_1, \widetilde{a_2} z_2 + \widetilde{b_2}), a_1, a_2 \in \mathfrak{a}, b_1, b_2 \in \mathfrak{b}\} \subset \mathbb{C}^2.$$

We observe that $\Lambda_{\boldsymbol{z}}$ is an $\mathcal{O}_K$-module, via the module structure of Λ as an $\mathcal{O}_K$-module and the map $\phi_{\boldsymbol{z}}$. Therefore, $\phi_{\boldsymbol{z}}$ induces a ring embedding $\mathcal{O}_K \to \operatorname{End}(A_{\boldsymbol{z}})$. More precisely, $\iota_{\boldsymbol{z}} : \mathcal{O}_K \to \operatorname{End}(A_{\boldsymbol{z}}), l \mapsto \iota_{\boldsymbol{z}}(l)$ is given by

$$\begin{aligned} \iota_{\boldsymbol{z}}(l) : \mathbb{C}^2 &\to \mathbb{C}^2 \\ (u_1, u_2) &\mapsto (lu_1, \widetilde{l} u_2), \end{aligned} \tag{2.5}$$

and it follows that $\iota_{\boldsymbol{z}}(l)(\Lambda_{\boldsymbol{z}}) \subset \Lambda_{\boldsymbol{z}}$, hence $\iota_{\boldsymbol{z}}(l) \in \operatorname{End}(A_{\boldsymbol{z}})$.

We also define

$$\begin{aligned} E_{r,\boldsymbol{z}} : \Lambda_{\boldsymbol{z}} \times \Lambda_{\boldsymbol{z}} &\to \mathbb{Q} \\ \phi_{\boldsymbol{z}}(a_1, b_1) \times \phi_{\boldsymbol{z}}(a_2, b_2) &\mapsto E_r((a_1, b_1), a_2, b_2)). \end{aligned}$$

Extend $E_{r,\boldsymbol{z}}$ $\mathbb{R}$-linearly to $\mathbb{C}^2 \to \mathbb{R}$ and set $H_{r,\boldsymbol{z}}(u, v) = E_{r,z}(iu, w)) + iE_{r,z}(u, v)$ for $u, v \in \mathbb{C}^2$ the associated hermitian form.

Proposition 2.5. *The assignment* E_r *above is an alternating bilinear form and it holds that*

1. $E_r(\Lambda \times \Lambda) \subset \mathbb{Z}$ *if and only if* $r \in (\partial_K \mathfrak{ab})^{-1} = (\mathfrak{ab})^\vee$. *In that case, the determinant* $\det(E_r) = N_{K|Q}(r\partial_K \mathfrak{ab})^2$, *for the ideal norm.*

2. *The hermitian form $H_{r,\boldsymbol{z}}$ can be written in coordinates, for $u, v \in \mathbb{C}^2$:*

$$H_{r,\boldsymbol{z}}((u_1,u_2),(v_1,v_2)) = \frac{u_1\overline{v_1}r}{\mathrm{Im}(z_1)} + \frac{u_2\overline{v_2}\widetilde{r}}{\mathrm{Im}(z_2)}$$

Then $H_{r,z}$ is positive definite if and only if r is totally positive.

3. *The embedding ι from Equation (2.5) restricts to $\iota : \mathcal{O}_K \hookrightarrow \mathrm{End}^s(A, H_{r,\boldsymbol{z}})$*

Proof. The first two items come from [Gor02, Lemma 2.8 and Lemma 2.9], and we only give the proof of item 1. Remark that $(\partial_K \mathfrak{a}\mathfrak{b})^{-1} = (\mathfrak{a}\mathfrak{b})^\vee$ follows from (2.1) above, and then the equivalence is tautological. For the determinant computation, notice first that it follows from the first part that $r\partial_K\mathfrak{a}\mathfrak{b} \subset \mathcal{O}_K$, so its ideal norm is its index in $\mathcal{O}_K$ and in particular is a positive integer. Set $\Lambda \subset \Lambda_r^* \subset K$ the dual of Λ with respect to E_r, *i.e.* $\Lambda_r^* = \{\alpha \in K \oplus K |\, E_r(\alpha, \Lambda) \subset \mathbb{Z}\}$, then we have

$$\Lambda_r^* = \frac{1}{r}\mathfrak{b}^\vee \oplus \frac{1}{r}\mathfrak{a}^\vee = (r\partial_k\mathfrak{b})^{-1} \oplus (r\partial_k\mathfrak{a})^{-1}\,.$$

By the theory of elementary divisors applied to $\Lambda_r^* \supset \Lambda_r$, there is a basis of Λ such that E_r admits a matrix $\begin{pmatrix} 0 & D \\ -D & 0 \end{pmatrix}$, with D a diagonal matrix with the elementary divisors of $\Lambda_r^* \supset \Lambda$, taking the determinant it follows: $\det E_r = (\det D)^2 = [\Lambda_r^* : \Lambda]$. For the index computations,

$$\left[(r\partial_k\mathfrak{b})^{-1} \oplus (r\partial_k\mathfrak{a})^{-1} : \mathfrak{a} \oplus \mathfrak{b}\right] = \left[(r\partial_k\mathfrak{b})^{-1} : \mathfrak{a}\right]\left[(r\partial_k\mathfrak{a})^{-1} : \mathfrak{b}\right] = \left[(r\partial_k\mathfrak{a}\mathfrak{b})^{-1} : \mathcal{O}_K\right]^2 = [\mathcal{O}_K : r\partial_k\mathfrak{a}\mathfrak{b}]^2\,.$$

For item 3, by [BL04, Proposition 5.1.1], one needs check that for $l \in \mathcal{O}_K$, $\iota_{\boldsymbol{z}}(l)$ is self-adjoint with respect to $H_{r,\boldsymbol{z}}$, *i.e.* $H_{r,\boldsymbol{z}}(u, \iota_{\boldsymbol{z}}(l)v) = H_{r,\boldsymbol{z}}(\iota_{\boldsymbol{z}}(l)u, v)$, which follows directly from the definition of $\iota_{\boldsymbol{z}}(l)$ in Equation (2.5) and Item 2. □

We take a small detour to notice that one can also think of the (non necessarily principal) polarizations of the complex torus $\mathbb{C}^2/\Lambda_{\boldsymbol{z}}$ as being parametrized by $(\mathfrak{a}\mathfrak{b})^{\vee,++}$. This is a notion for abelian surface with real multiplication by $\mathcal{O}_K$, called the *polarization module*.

Consider the Néron-Severi group $\mathrm{NS}(A)$. For the following definition, we recall that a polarization is equivalently an isogeny $f : A \to \widehat{A}$ to the dual abelian variety such that $f = \hat{f}$ under the identification $\hat{\widehat{A}} \cong A$. Observe that if A has real multiplication by $\mathcal{O}_K$, then $\widehat{A}$ does too, simply by setting $\widehat{\iota} : \mathcal{O}_K \to \mathrm{End}(\widehat{A})$ as $\widehat{\iota}(k) = \widehat{\iota(k)}$.

Therefore, it makes sense to consider isogenies $f : A \to \widehat{A}$ that furthermore are $\mathcal{O}_K$-linear.

Definition 2.6. *Let A be an abelian variety with real multiplication by $\mathcal{O}_K$. Then we define* the polarization module[6] *as*

$$P(A) = \{f : A \to \hat{A},\ \hat{f} = f, f \text{ is } \mathcal{O}_K\text{-linear}\} \subset \mathrm{NS}(A),$$

and the positive cone:

$$P(A)^{++} = \{f \in P(A),\ f \text{ is a polarization}\} \subset \mathrm{NS}(A)^{++},$$

and the pair $(P(A), P(A)^{++})$ is called the ordered polarization module.

[6] this notions depend on the embedding ι, as it induces the $\mathcal{O}_K$-module structure in the first place, but we omit it form the notation for readability.

Proposition 2.7. *Let A an abelian surface with real multiplication by $\mathcal{O}_K$.*

- *$P(A)$ is a projective $\mathcal{O}_K$-module of rank 1.*
- *$P(A)^+$ is non-empty, it is a positive cone in $P(A)$, and generates $P(A)$ as an $\mathcal{O}_K$-module.*

Proof. This is [Gor02, Lemma 2.13] □

Corollary 2.8. *The polarization module of $A_{\boldsymbol{z}}$ is $(\mathfrak{c}, \mathfrak{c}^{++})$ where $\mathfrak{c} = (\partial_K \mathfrak{a}\mathfrak{b})^{-1}$. $E_{r,\boldsymbol{z}}$ induces a principal polarization if and only if $\mathfrak{a}\mathfrak{b} = \partial_K^{-1}$ in the narrow class group $\mathrm{Cl}(K)^+$.*

In our notion of isomorphism in the moduli space of abelian surfaces with RM, one can either ask for the isomorphism to respect the *fixed* polarization, or only to leave the polarization module invariant.[7] The second choice being weaker, the corresponding moduli space gets mod out by $\mathrm{GL}_2(\mathcal{O}_K \oplus \mathfrak{a})^{++}$, of invertible matrices with $\det \in (\mathcal{O}_K^*)^{++}$ instead of $\mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a})$.

Compare it too with the following result from [BL04, Theorem 5.2.4 and Remark 5.2.5]: under the abelian group isomorphism $\mathrm{NS}_0(A)$ and $\mathrm{End}_0^s(A)$, there is a *set isomorphism* between polarizations (of degree d) and totally positive elements (of analytic norm d). This is not a contradiction with this result, as the positive cone of the polarization module only considers polarizations that are $\mathcal{O}_K$-linear.

Theorem 2.9. *Set $[\mathfrak{a}] \in \mathrm{Cl}(K)^+$ for representatives of the narrow class group of K. Then we have the bijections:*

$$\left\{ \begin{array}{l} (A,\iota) \text{ ab. surf. with RM by } \mathcal{O}_K \\ \text{and polarization module } (\mathfrak{c},\mathfrak{c}^{++}) \end{array} \right\} \Big/ \underset{(\mathfrak{c},\mathfrak{c}^{++})}{\text{isom. pres.}} \longrightarrow \mathrm{GL}_2(\mathcal{O}_K \oplus \mathfrak{a})^{++} \backslash \mathbb{H}^2,$$

for $\mathfrak{c} = (\partial_K \mathfrak{a})^{-1}$, and

$$\left\{ \begin{array}{l} (A,\iota) \text{ ab. surf. with RM by } \mathcal{O}_K \\ \text{and polarization module } (\mathfrak{c},\mathfrak{c}^{++}) \end{array} \right\} \Big/ \underset{\text{the polarization}}{\text{isom. pres.}} \longrightarrow \mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a}) \backslash \mathbb{H}^2.$$

Therefore, we have

$$\left\{ \begin{array}{l} \text{isomorphism classes of } (A,\iota) \\ \text{ab. surf. with RM by } \mathcal{O}_K \end{array} \right\} \cong \bigsqcup_{(\mathfrak{c},\mathfrak{c}^{++})} \mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a}) \backslash \mathbb{H}^2.$$

Proof. See [Gor02, Theorem 2.17 and Corollary 2.19]. □

Corollary 2.10. *We have the following bijection*

$$\left\{ \begin{array}{l} \text{isomorphism classes of } (A,\iota) \\ \text{ppas with RM by } \mathcal{O}_K \end{array} \right\} \cong \mathrm{SL}_2(\mathcal{O}_K \oplus \partial_K^{-1}) \backslash \mathbb{H}^2.$$

Proof. Follows from Theorem 2.9 and Corollary 2.8. □

[7] observe that $(\mathfrak{a}, \mathfrak{a}^{++})$ admits an action by $(\mathcal{O}_K^*)^{++}$.

We remark that in Theorem 2.9, the isomorphism notions are for abelian surfaces with RM by $\mathcal{O}_K$, meaning that the isomorphisms are compatible with the respective embeddings $\iota : \mathcal{O}_K \hookrightarrow \mathrm{End}^s(A)$. Then, even for the case of principal polarization $\mathcal{O}_K \oplus \partial_K^{-1}$, the isomorphism class of the ppas $A_{\boldsymbol{z}} \in \mathcal{A}_2$ it is not the same that the one on the moduli spaces above.

We actually have a natural involution on the analytic spaces $\mathbb{H}^2 \hookrightarrow \mathbb{H}^2$, $(z_1, z_2) \mapsto (z_2, z_1)$, given by

$$\begin{aligned} \mathbb{H}^2 \times \mathrm{SL}_2(\mathcal{O}_K \oplus \partial_K^{-1}) &\to \mathbb{H}^2 \times \mathrm{SL}_2(\mathcal{O}_K \oplus \partial_K^{-1}) \\ ((z_1, z_2), \gamma) &\mapsto ((z_2, z_1), \widetilde{\gamma}), \end{aligned}$$

where $\widetilde{\gamma}$ means the real conjugation applied component-wise to the matrix. This induces a well-defined involution in the Hilbert modular surface, but it is not trivial: no element in the Hilbert modular group permutes the variables (z_1, z_2).

However, one can see from the explicit assignments that the corresponding principally polarized abelian surfaces $A_{(z_1,z_2)}$ and $A_{(z_2,z_1)}$ are isomorphic in $\mathcal{A}_2$. The obstruction to be $\mathcal{O}_K$-isomorphic is that, if we set $\iota : \mathcal{O}_K \hookrightarrow \mathrm{End}^s(A_{(z_1,z_2)})$ for the real multiplication structure, then one can check that the induced one by the isomorphism on $A_{(z_2,z_1)}$ is $\widetilde{\iota} := \iota(\widetilde{\cdot})$. From the Definition 2.1, if we set $f : A_{(z_1,z_2)} \to A_{(z_2,z_1)}$ for the isomorphism, the compatibility condition translates to $f\iota(l) = \iota(\widetilde{l})$, which is not verified. Indeed, as maps on $\mathbb{C}^2 \to \mathbb{C}^2$, for given $l \in \mathcal{O}_K$,

$$f \circ \iota(l)(u_1, u_2) = (\widetilde{l}u_2, lu_1) \neq (\widetilde{l}u_1, lu_2) = \iota(\widetilde{l}).$$

We see more details about this in Subsection 2.3.1.

2.3 Hilbert modular forms

The results from the above section are true in any genus g: a complex torus with real multiplication by a totally real field of degree g is always polarizable, and for maximal real multiplication we can parametrize the moduli space of principally polarized abelian varieties of dimension g by a quotient of $\mathbb{H}^g$ by a suitable discrete subgroup of $(\mathrm{SL}_2(\mathbb{R}))^g$. Being able to consider *all* complex tori, the functions and vector bundles on this complex space are simply functions on $\mathbb{H}^g$, satisfying symmetry properties coordinate-wise. In a way, if one where to ask for a generalization of classical modular forms on $\mathbb{H}$, Hilbert modular forms are more direct ones than Siegel modular forms.

The material form this section comes from [Bru08] and [Gee88]. We set again $g = 2$.

Definition 2.11. *Let $\Gamma \subset \mathrm{SL}_2(K)$ commensurable with $\mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a})$. Let $f : \mathbb{H}^2 \to \mathbb{C}$ a meromorphic function. We say it is a* meromorphic Hilbert modular form *of weight $\boldsymbol{k} = (k_1, k_2) \in \mathbb{Z}^2$ for Γ if for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$,*

$$f(\gamma \boldsymbol{z}) = (cz_1 + d)^{k_1} (\widetilde{c}z_2 + \widetilde{d})^{k_2} f(\boldsymbol{z}).$$

In addition, if $k_1 = k_2$ we say that f has parallel *weight, and if $f(z_1, z_2) = f(z_2, z_1)$, we say that it is* symmetric.

If f is furthermore holomorphic as a complex function on $\mathbb{H}^2$, we say that f is a holomorphic *Hilbert modular form.*

Given two Hilbert modular forms of the same weight, we call its quotient a Hilbert modular function, *and if they are both symmetric, we call the quotient symmetric too.*

Remark 2.12. *Note that if f as above is symmetric, then it has parallel weight.*

Analogously to Siegel modular forms, there is a Koecher's principle, see Theorem 2.18 below, so one does not need to impose boundedness conditions at infinity as an extra condition for holomorphicity.

2.3.1 Symmetric Hilbert modular surfaces

It is natural to consider the symmetry condition $f(z_1, z_2) = f(z_2, z_1)$ for Hilbert modular forms, so in the definition we have singled them out. Note again that it is not given by any $\gamma \in \Gamma$, as the linear fractional transformations act independently in z_1, z_2.

Let us denote by σ the involution $\sigma(z_1, z_2) = (z_2, z_1)$. One could also extend $\mathrm{SL}_2(\mathcal{O}_K)$ to $\mathrm{SL}_2(\mathcal{O}_K) \rtimes \langle\sigma\rangle = \mathrm{SL}_2(\mathcal{O}_K) \cup \mathrm{SL}_2(\mathcal{O}_K)\sigma$, and consider the corresponding analytic space. It is called *a symmetric Hilbert surface*, for that its field of functions is given by symmetric Hilbert modular functions. From a moduli problem perspective, it mods out by the canonical involution we saw at the end of Section 2.2.

Observe that for Siegel modular forms with respect to $\mathrm{Sp}_4(\mathbb{Z})$, there are fractional linear transformations that permute the variables. For example, taking $U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and the unimodular transformation $\begin{pmatrix} U & 0 \\ 0 & {}^tU^{-1} \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z})$, it acts as $\boldsymbol{\tau} \mapsto U\boldsymbol{\tau}\,{}^tU$, hence as $\begin{pmatrix} \tau_1 & \tau_2 \\ \tau_2 & \tau_3 \end{pmatrix} \mapsto \begin{pmatrix} \tau_3 & \tau_2 \\ \tau_2 & \tau_1 \end{pmatrix}$. We will see in Proposition 2.26 that it is possible to pull-back Siegel modular forms to Hilbert modular forms, via a finite degree map from the Hilbert modular surface to a subvariety of $\mathcal{A}_2$, and this pull back of the Siegel modular form will have to be symmetric.[8]

Remark 2.13. *While $\mathcal{A}_2$ is a rational variety, which imply that its field of functions (of transcendence degree three) admits three algebraically independent generators (the Igusa invariant), it is hardly ever the case that Hilbert surfaces are rational surfaces.*[9] *However, as noted in [MR20],*[10] symmetric *Hilbert surfaces are rational more often.*

Theorem 2.14. *Symmetric and non-symmetric Hilbert modular surfaces are of general type except for finitely many discriminants.*

- *The Hilbert modular surface of discriminant Δ is rational for $\Delta = 5, 8, 12, 13, 17$.*
- *The symmetric Hilbert modular surface (equivalently, the Humbert surface) is rational for all fundamental discriminants*[11] *$1 < \Delta < 100$.*

Proof. This is [HZ77, Theorem 3],[12] and [EK14, second to last paragraph in page 2298]. For the second result, it is not claim that the list is complete.

For more details for other Hilbert surfaces, see also [Gee88, Theorem VII.3.3]. □

[8] It is true for $\mathcal{A}_2$, it may not be for covers by subgroups of $\mathrm{Sp}_4(\mathbb{Z})$.
[9] Except for finitely many discriminants, a Hilbert modular surface is of general type.
[10] and we thanked Damien Robert for pointing it out to us.
[11] there are thirty of them.
[12] it is not Theorem 2 in *loc. cit.*, as it is a different Hilbert modular surface.

2.3.2 Cusps of the Hilbert modular surface

The complex manifolds $\mathrm{SL}_2(\mathcal{O}_K)\backslash\mathbb{H}^2$ are not compact, as the one dimensional modular curves, and they also admit a compactification by adding finitely many points called *cusps*. They are defined analogously as the extension of the action of $\mathrm{SL}_2(\mathcal{O}_K)$ on $\mathbb{H}^2$ as in (2.2) above, to $\mathbb{P}^1(K) = K \cup \{\infty\}$. For $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(K)$, we define

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \frac{\alpha}{\beta} = \frac{a\frac{\alpha}{\beta} + b}{c\frac{\alpha}{\beta} + d} = \frac{a\alpha + b\beta}{c\alpha + d\beta}.$$

We call a *cusp of* $\mathrm{SL}_2(\mathcal{O}_K)$ to one of the orbits of $\mathrm{SL}_2(\mathcal{O}_K)$ in $\mathbb{P}^1(K)$, and set $\infty = (1:0) \in \mathbb{P}^1(K)$ and denote its orbit *the cusp at infinity*.

Remark that $(\alpha : \beta) \in \mathbb{P}^1(K)$ is defined up to scalar multiplication $r \in K^*$, so the fractional ideal $\alpha\mathcal{O}_K + \beta\mathcal{O}_K$ is only well defined in the class group $\mathrm{Cl}(K)$.

Lemma 2.15. *The assignment*

$$\begin{aligned} \mathrm{SL}_2(\mathcal{O}_K)\backslash\mathbb{P}^1(K) &\to \mathrm{Cl}(K) \\ (\alpha : \beta) &\mapsto [\alpha\mathcal{O}_K + \beta\mathcal{O}_K] \end{aligned}$$

is a bijection.

Remark 2.16. *In the modular case* $\mathrm{SL}_2(\mathbb{Z})$ *acts* transitively *on* $\mathbb{P}^1(\mathbb{Q})$*, we note that the argument over the integers do not generalize to* $\mathcal{O}_K$*. For* $\mathrm{SL}_2(\mathbb{Z})$ *it is enough to show that every cusp is in the* $\mathrm{SL}_2(\mathbb{Z})$*-orbit of* ∞*, or that for* $a/b \in \mathbb{Q}$ *with* $(a, b) = 1$*, one needs to find* $c, d \in \mathbb{Z}$ *with*

$$ad - bc = 1,$$

so that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ *and* $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \infty = a/c$*. Over the integers the existence of* $c, d \in \mathbb{Z}$ *as above is proven by Bézout's lemma. It generalizes to principal domains, so to Dedekind domains* $\mathcal{O}_K$ *with* $h(K) = 1$*. Lemma 2.15 hence proves that this is the only case where the action of* $\mathrm{SL}_2(K)$ *is transitive over* $\mathbb{P}^1(K)$*.*

Proof. We give the proof from [Bru08, Lemma 1.3] and [Gee88, Chapter I, Proposition (1.1)]. For well-definedness of the assignment, if $(\gamma : \delta) = (a\alpha + b\beta : c\alpha + d)$ for $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathcal{O}_K)$, take $r \in K^*$ so that $\gamma = r(a\alpha + b\beta)$ and $\delta = r(c\alpha + d)$, then

$$\gamma\mathcal{O}_K + \delta\mathcal{O}_K = r\left((a\alpha + b\beta)\mathcal{O}_K + (c\alpha + d)\mathcal{O}_K\right) = r\left(\alpha\mathcal{O}_K + \beta\mathcal{O}_K\right),$$

where the last equality follows for that $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible in $\mathrm{SL}_2(\mathcal{O}_K)$, so $(a\alpha + b\beta), (c\alpha + d)$ is another $\mathcal{O}_K$ basis of the fractional ideal $\alpha\mathcal{O}_K + \beta\mathcal{O}_K$. Hence, they are the same ideal class in $\mathrm{Cl}(K)$.

Surjectivity follows from the known fact (see for example [Mar18, Theorem 17]) that for every fractional ideal $\mathfrak{a}$ in $\mathcal{O}_K$ there exists $\alpha, \beta \in \mathcal{O}_K$ with $\mathfrak{a} = \alpha\mathcal{O}_K + \beta\mathcal{O}_K$.

For injectivity for $(\alpha : \beta), (\gamma : \delta) \in \mathbb{P}^1(K)$, without loss of generalization, assume that $\alpha\mathcal{O}_K + \beta\mathcal{O}_K = \gamma\mathcal{O}_K + \delta\mathcal{O}_K$ and set any one of them as $\mathfrak{a}$. As $1 \in \mathfrak{a}\mathfrak{a}^{-1}$, there exists $\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \tilde{\delta} \in \mathfrak{a}^{-1}$ with

$$\begin{aligned} \alpha\tilde{\beta} - \beta\tilde{\alpha} &= 1 \\ \gamma\tilde{\delta} - \delta\tilde{\gamma} &= 1, \end{aligned}$$

so that the matrices

$$M = \begin{pmatrix} \alpha & \tilde{\alpha} \\ \beta & \tilde{\beta} \end{pmatrix}, \; N = \begin{pmatrix} \gamma & \tilde{\gamma} \\ \delta & \tilde{\delta} \end{pmatrix} \in \begin{pmatrix} \mathfrak{a} & \mathfrak{a}^{-1} \\ \mathfrak{a} & \mathfrak{a}^{-1} \end{pmatrix} \cap \mathrm{SL}_2(K)$$

send ∞ to α/β and γ/δ, respectively. Notice that M, N do not have coefficients in $\mathrm{SL}_2(\mathcal{O}_K)$. However, consider MN^{-1}, it sends γ/δ to α/β and it belongs to $\mathrm{SL}_2(\mathcal{O}_K)$: as N has determinant 1, its inverse is its adjoint, which belongs to $\begin{pmatrix} \mathfrak{a}^{-1} & \mathfrak{a}^{-1} \\ \mathfrak{a} & \mathfrak{a} \end{pmatrix}$, hence

$$MN^{-1} \in \begin{pmatrix} \mathfrak{a} & \mathfrak{a}^{-1} \\ \mathfrak{a} & \mathfrak{a}^{-1} \end{pmatrix} \begin{pmatrix} \mathfrak{a}^{-1} & \mathfrak{a}^{-1} \\ \mathfrak{a} & \mathfrak{a} \end{pmatrix} \subset \mathrm{SL}_2(\mathcal{O}_K).$$

□

As part of the proof above, for a given cusp in $\mathbb{P}^1(K)$, $\sigma := (\alpha : \beta)$, setting $\mathfrak{a} = \alpha\mathcal{O}_K + \beta\mathcal{O}_K \in \mathrm{Cl}(K)$, the matrix

$$M_\sigma = \begin{pmatrix} \alpha & \tilde{\alpha} \\ \beta & \tilde{\beta} \end{pmatrix},$$

sends ∞ to α/β, where $\tilde{\alpha}, \tilde{\beta} \in \mathfrak{a}^{-1}$ and satisfy $1 = \alpha\tilde{\beta} - \beta\tilde{\alpha}$, which does not necessarily lie in $\mathcal{O}_K$. To work cusp-free for Hilbert modular forms, one could then fix the cusp at ∞ and consider instead all subgroups, for $\mathfrak{a} \in \mathrm{Cl}(K)$:

$$M_\sigma^{-1} \mathrm{SL}_2(\mathcal{O}_K) M_\sigma \subset \begin{pmatrix} \mathfrak{a}^{-1} & \mathfrak{a}^{-1} \\ \mathfrak{a} & \mathfrak{a} \end{pmatrix} \begin{pmatrix} \mathfrak{a} & \mathfrak{a}^{-1} \\ \mathfrak{a} & \mathfrak{a}^{-1} \end{pmatrix} \subset \begin{pmatrix} \mathcal{O}_K & (\mathfrak{a}^2)^{-1} \\ \mathfrak{a}^2 & \mathcal{O}_K \end{pmatrix} \tag{2.6}$$

and it holds that $M_\sigma^{-1} \mathrm{SL}_2(\mathcal{O}_K) M_\sigma = \mathrm{SL}_2(\mathcal{O}_K \oplus (\mathfrak{a}^2)^{-1})$. This is another context in which the companion modular groups show up naturally.

The analytic manifold $\mathrm{SL}_2(\mathcal{O}_K)\backslash\mathbb{H}^2$ admits a compactification by adding the (finitely many) cusps, as the extended $(\mathbb{H}^2)^* = \mathbb{H}^2 \cup \mathbb{P}^1(K)$, and as $\mathrm{SL}_2(\mathcal{O}_K)\backslash(\mathbb{H}^2)^*$. This is actually the Bailey-Borel compactification of the Hilbert surface ([Bru08, Theorem 1.11] and [Fre90, Theorem II.4.1]), but in contrast to the one dimensional case, the cusps are *singular*.[13] A minimal resolution of singularities was constructed in [Hir71].

Finally, the isotropy group of the cusp at ∞ in $\mathrm{SL}_2(\mathcal{O}_K \oplus (\mathfrak{a}^2)^{-1})$ is given by

$$\left\{ \begin{pmatrix} u & \mu \\ 0 & u^{-1} \end{pmatrix} : u \in (\mathcal{O}_K)^*, \; \mu \in (\mathfrak{a}^2)^{-1} \right\}. \tag{2.7}$$

[13] there are also quotient singularities at the elliptic points, as it happens in the elliptic case too.

2.3.3 Fourier expansion at the cusp at ∞

We are going to present the Fourier expansions of holomorphic Hilbert modular forms at the cusps of $\mathrm{SL}_2(\mathcal{O}_K)$. By Equation (2.6), it is equivalent as to describe the Fourier expansion at ∞ with respect to all companion groups $\mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a})$.

We follow [Bru08, Section 1.3] and [Gee88, Chapter 6]. Consider f a holomorphic Hilbert modular form for $\Gamma \subset \mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a})$. Setting $M \subset (\mathfrak{a}^2)^{-1}$, then as in Equation (2.7), the isotropy group of the cusp at ∞ for Γ is of the form

$$G(M,V) := \left\{ \begin{pmatrix} u & \mu \\ 0 & u^{-1} \end{pmatrix} : u \in V,\ \mu \in M \right\},$$

for $M \subset K$ some $\mathbb{Z}$-module of rank two and $V \subset (\mathcal{O}_K)^*$ a finite index subgroup that acts on M.

Definition 2.17. *In general, for a given cusp and Γ, the pair (M,V) describing the isotropy group as above is called* the type of the cusp *(in [Gee88, End of Section 1.1, page 7]).*

As f is a Hilbert modular form for Γ, then $f(\boldsymbol{z}+\mu) = f(\boldsymbol{z})$ for all $\mu \in M$. Therefore, f admits a uniformly convergent Fourier series (see [Fre90, Lemma I.4.1]),

$$f = \sum_{\omega \in M^\vee} a_f(\omega) e^{i2\pi \operatorname{tr}(\omega \boldsymbol{z})},$$

where $\operatorname{tr}(\omega \boldsymbol{z}) := \omega z_1 + \widetilde{\omega} z_2$ and $M^\vee = \{\lambda \in K | \operatorname{tr}_{K|\mathbb{Q}}(\lambda\mu) \in \mathbb{Z}, \text{ for all } \mu \in M\}$ is the dual lattice of M with respect to $\operatorname{tr}_{K|\mathbb{Q}}$. The Fourier coefficients are given by

$$a_f(\omega) = \frac{1}{\operatorname{vol}(\mathbb{R}^2/M)} \int_{\mathbb{R}^2/M} f(\boldsymbol{z}) e^{-i2\pi \operatorname{tr}(\omega \boldsymbol{z})} dx_1 dx_2, \tag{2.8}$$

where $M \hookrightarrow \mathbb{R}^2$ via the two real embeddings of K.

As in the case of Siegel modular forms in Theorem 1.7, we have the following version of the Koecher principle.[14]

Theorem 2.18. *Let $f : \mathbb{H}^2 \to \mathbb{C}$ a* holomorphic *Hilbert modular form for Γ of weight $\boldsymbol{k} = (k_1, k_2)$, and let (M,V) the type of the cusp at ∞ of Γ. Then*

- *for all $\omega \in \mathcal{M}^\vee$ and $u \in V$, $a_f(u^2\omega) = u^{k_1}\widetilde{u}^{k_2} a_f(\omega)$;*
- *and if $a_f(\omega) \neq 0$ then either $\omega = 0$ or $\omega \gg 0$.*

Therefore, f admits a Fourier expansion of the form

$$f = a_f(0) + \sum_{\omega \in M^{\vee,++}} a_f(\omega) e^{i2\pi \operatorname{tr}(\omega \boldsymbol{z})}.$$

Proof. This is [Bru08, Theorem 1.20]. □

[14] In the Hilbert literature, it seems to be called the Götzky-Koecher principle. By [Gee08, Section 4, Theorem 2], it was proved first by Götzky for Hilbert modular forms and by Koecher in the Siegel case.

Definition 2.19. *Let f a holomorphic Hilbert modular form as above, and consider its Fourier coefficients at one of the cusps. If $a_f(0) = 0$, we say that f vanishes at that cusp. If f vanishes at all cusps, we say that f is a* cusp *form.*

The Koecher principle has similar consequences as Proposition 1.10. In particular, we have the following.

Corollary 2.20. *Let f holomorphic Hilbert modular of weight $\boldsymbol{k} = (k_1, k_2)$ If $k_1 \neq k_2$ then f is a cusp form.*

Proof. This is [Bru08, Proposition 1.23]. □

It also holds a Hecke type bound for the Fourier coefficients of a modular form, and hence the sequence of Fourier coefficients has polynomial growth in terms of the field norm $N = N_{K|\mathbb{Q}}$.

Theorem 2.21 (Hecke's bound)**.** *Let f be a holomorphic Hilbert modular form of* paralel *weight $k = k_1 = k_2$ for $\mathrm{SL}_2(\mathcal{O}_K)$. Consider its Fourier expansion at the cusp at ∞,*

$$f = \sum_{\omega \in M^\vee} a_f(\omega) e^{i2\pi \operatorname{tr}(\omega \boldsymbol{z})}.$$

- *Assume f is a cusp form. Then the function $f(\boldsymbol{z}) (\operatorname{Im} z_1 \operatorname{Im} z_2)^{k/2}$ is Γ-invariant and bounded in $\mathbb{H}^2$, and we have $a_f(\omega) = O(N(\omega)^k/2)$, where the implicit constant can be taken as $\operatorname{vol}(R^2/M) \max_{\mathbb{H}^2} f(\boldsymbol{z}) (\operatorname{Im} z_1 \operatorname{Im} z_2)^{k/2}$.*
- *In general, if f is not a cusp form, then $a_f(\omega) = O(N(\omega)^{k-1})$.*

Proof. This is [Bru08, Proposition 1.42] or [Gee88, Proposition I.6.6]. The proof for cusp forms is analogous to the Hecke proof, and for general forms one proves the bounds for Eisenstein series and uses a decomposition of the finite vector space of Hilbert modular forms of parallel weight (k, k) as the direct sum of the the space of cusps forms and the space of Eisenstein series (see [Bru08, Theorem 1.35 and Theorem 1.37]). □

2.4 Modular embeddings of Hilbert surfaces into the Siegel threefold

In this section, we describe mappings from Hilbert surfaces into the Siegel threefold. Their images define subvarieties of $\mathcal{A}_2$ called *Humbert surfaces*, which will be the object of study of Chapter 3.

Remark 2.22. *These maps are called* modular embeddings *in [Gee08] and [Gor02], but they are* not *embeddings in general. There is a general notion of "modular embeddings" in the literature, see [CW90, Theorem 1], for Fuchsian groups Γ and embedding into the uniformizing space of abelian varieties with generalized complex multiplication. The terminology is then used as complex analytic embeddings $\mathbb{H} \hookrightarrow \mathbb{H}^r$ (or $\mathbb{H}_2$ in our case), that induces a finite degree map between the analytic quotients $\Gamma\backslash\mathbb{H}$ and the corresponding one in the codomain, that it is an algebraic map of the corresponding algebraic varieties.*

A very general description is given at [Gee88, end of Section IV.1, pages 209-210][15] and at [Gee82, page 327-328],[16] but we present the more explicit description from [LY11, Section 3] and [MR20, Section 2.3]. For simplicity, we assume there exists $\varepsilon \in K^*$ with $N(\varepsilon) = -1$, and set $\varepsilon > 0$, and we are going to fix the Hilbert group as $\mathrm{SL}_2(\mathcal{O}_K)$. Then by Corollary 2.10 and Equation (2.3), we can think of $\mathrm{SL}_2(\mathcal{O}_K)\backslash\mathbb{H}^2$ as the moduli space of principally polarized abelian surfaces with RM by $\mathcal{O}_K$.

By the moduli interpretation, we have the forgetful functor into $\mathcal{A}_2$, and we want to realize it as an finite degree map

$$\mathrm{SL}_2(\mathcal{O}_K)\backslash\mathbb{H}^2 \to \mathrm{Sp}_4(\mathbb{Z})\backslash\mathbb{H}_2, \tag{2.9}$$

by means of compatible embeddings $\mathbb{H}^2 \hookrightarrow \mathbb{H}_2$ and $\mathrm{SL}_2(\mathcal{O}_K) \hookrightarrow \mathrm{Sp}_4(\mathbb{Z})$. This kind of maps are broadly called in the literature *modular embeddings*, and we will discuss them in more details for the quaternionic multiplication and the Shimura curves case in Section 4.2.

Define $^* : K \hookrightarrow \mathrm{Mat}_2(K)$ by $a^* = \mathrm{diag}(a, \widetilde{a})$, and $^* : \mathbb{H}^2 \hookrightarrow \mathbb{H}_2$ by $\boldsymbol{z}^* = \mathrm{diag}(z_1, z_2)$. Consider also $^* : \mathrm{Mat}_2(K) \hookrightarrow \mathrm{Mat}_4(K)$ by $\gamma = (\gamma_{ij}) \mapsto \gamma^* = (\gamma_{ij}^*)$.

Fix $\{e_1, e_2\}$ a $\mathbb{Z}$-basis of $\mathcal{O}_K$, define $R = \begin{pmatrix} e_1 & e_2 \\ \widetilde{e_1} & \widetilde{e_2} \end{pmatrix}$ and

$$S = \mathrm{diag}({}^tR, R^{-1})\mathrm{diag}\left(I_2, \left(\frac{\sqrt{\Delta}}{\varepsilon}\right)^*\right).$$

Theorem 2.23. *The following embeddings are well defined*

$$\begin{aligned} \phi_{e_1,e_2} : \mathbb{H}^2 &\hookrightarrow \mathbb{H}_2 \\ \boldsymbol{z} &\mapsto {}^tR\left(\frac{\varepsilon}{\sqrt{\Delta}}\boldsymbol{z}\right)^* R, \\ \varphi_{e_1,e_2} : \mathrm{SL}_2(K) &\hookrightarrow \mathrm{Sp}_4(\mathbb{Q}) \\ \gamma &\mapsto S\gamma^* S^{-1}; \end{aligned}$$

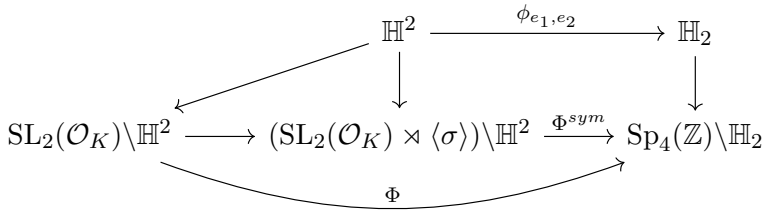
and satisfy the subsequent properties

- *we have a restriction to*
$$\varphi_{e_1,e_2} : \mathrm{SL}_2(\mathcal{O}_K) \hookrightarrow \mathrm{Sp}_4(\mathbb{Z}).$$
- *They are compatible with the respective actions of* $\mathrm{SL}_2(\mathcal{O}_K)$ *on* $\mathbb{H}^2$ *and* $\mathrm{Sp}_4(\mathbb{Z})$ *on* $\mathbb{H}_2$*, hence they induce a map as in in Equation (2.9), that we denote* Φ_{e_1,e_2}*.*
- *If* $\{f_1, f_2\}$ *is another* $\mathbb{Z}$*-basis of* $\mathcal{O}_K$*, then there exists* $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ *such that for all* $\boldsymbol{z} \in \mathbb{H}^2$*,* $\phi_{f_1,f_2}(\boldsymbol{z}) = \gamma\phi_{e_1,e_2}(\boldsymbol{z})$*. In particular, the map*
$$\Phi : \mathrm{SL}_2(\mathcal{O}_K)\backslash\mathbb{H}^2 \to \mathrm{Sp}_4(\mathbb{Z})\backslash\mathbb{H}_2$$
is independent of the choice of $\mathbb{Z}$*-basis of* $\mathcal{O}_K$*.*

[15] into the moduli space of polarized abelian surfaces with a polarization of type $(1, t)$, for $\mathcal{O}_K$ the maximal order and all companion groups.

[16] for general orders in a real quadratic field, for the companion group of principal polarizations.

- *Denote* $\sigma(z_1, z_2) = (z_2, z_1)$*, then there exists* $M_\sigma \in \mathrm{Sp}_4(\mathbb{Z})$ *such that* $\phi_{e_1,e_2}(\sigma \boldsymbol{z}) = M_\sigma \phi_{e_1,e_2}(\boldsymbol{z})$*. Therefore,* ϕ_{e_1,e_2} *and* φ_{e_1,e_2} *also defines a map form the symmetric Hilbert modular surface* $(\mathrm{SL}_2(\mathcal{O}_K) \rtimes \langle\sigma\rangle)\backslash\mathbb{H}^2$*, that we call* Φ^{sym}*. More precisely, we have the following commutative diagram*



where all unnamed maps are the natural ones.

Taking the standard $\mathbb{Z}$*-basis of* $\mathcal{O}_K$ *given by* $\{1, \omega\}$ *where* $\omega = \begin{cases} \sqrt{\Delta}/2, & \text{if } \Delta \equiv 0(4), \\ (1+\sqrt{\Delta})/2, & \text{if } \Delta \equiv 1(4) \end{cases}$*.*
Set too $(k, l) = \begin{cases} \left(\frac{\Delta}{4}, 0\right) & \text{if } \Delta \equiv 0(4), \\ \left(\frac{\Delta-1}{4}, 1\right), & \text{if } \Delta \equiv 1(4) \end{cases}$*, so that* $\Delta = 4k+l$ *with* $l \in \{0, 1\}$ *and* $x^2 - lx - k = 0$ *is the minimal equation for* ω*. Then the maps are given explicitly by*

$$\phi_{1,\omega}(z_1, z_2) = \begin{pmatrix} \frac{\varepsilon z_1 - \widetilde{\varepsilon} z_2}{\sqrt{\Delta}} & \frac{\varepsilon\omega z_1 - \widetilde{\varepsilon}\widetilde{\omega} z_2}{\sqrt{\Delta}} \\ \frac{\varepsilon\omega z_1 - \widetilde{\varepsilon}\widetilde{\omega} z_2}{\sqrt{\Delta}} & \frac{\varepsilon\omega^2 z_1 - \widetilde{\varepsilon}\widetilde{\omega}^2 z_2}{\sqrt{\Delta}} \end{pmatrix}$$

$$\phi_{1,\omega}\begin{pmatrix} a_1 + a_2\omega & b_1 + b_2\omega \\ c_1 + c_2\omega & d_1 + d_2\omega \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & b_2 & b_1 + lb_2 \\ ka_2 & a_1 + la_2 & b_1 + lb_2 & kb_2 + l(b_1 + b_2) \\ kc_2 - lc_1 & c_1 & d_1 & kd_2 \\ c_1 & c_2 & d_2 & d_1 + ld_2 \end{pmatrix}$$

Furthermore, for $\boldsymbol{\tau} = \begin{pmatrix} \tau_1 & \tau_2 \\ t_2 & \tau_3 \end{pmatrix} \in \mathrm{Im}\,\phi_{1,\omega}$*, it holds*

$$\begin{cases} \frac{\Delta-1}{4}\tau_1 + \tau_2 - \tau_3 = 0, & \text{if } \Delta \equiv 1(4), \\ \frac{\Delta}{4}\tau_1 - \tau_3 = 0, & \text{if } \Delta \equiv 0(4). \end{cases} \tag{2.10}$$

Proof. The details are in [LY11, Proposition 3.1], with the reformulation from [MR20, Proposition 2.3 and Equations (7), (10)], and [Run99, proof of Lemma 4] for the embedding $\mathrm{SL}_2(\mathcal{O}_K) \hookrightarrow \mathrm{Sp}_4(\mathbb{Z})$. Note that via Equation (2.4), one reformulates to an analogous statement for the Hilbert group $\mathrm{SL}_2(\mathcal{O}_K \oplus \partial_K^{-1})$, for that this is the one naturally associated with the moduli problem of abelian surfaces with RM by $\mathcal{O}_K$ and *principal* polarization. □

The equation (2.10) is a particular case of a Humbert singular relations, that will be the object of study of the next chapter. We will also see that the image of the map ϕ in $\mathcal{A}_2$ defines a subvariety of $\mathcal{A}_2$ called the Humbert surface of (the fundamental) discriminant Δ_K.

Proposition 2.24. *The map* ϕ^{sym} *as in Theorem 2.23 is a birational map onto the Humbert surface of discriminant* Δ_K*.*

Proof. This is [MR20, Proposition 2.8]. □

Remark 2.25. *We have stated the result for the full Hilbert modular group, but there are analogous results by passing through compatible subgroups of* $\mathrm{SL}_2(\mathcal{O}_K)$ *and* $\mathrm{Sp}_4(\mathbb{Z})$ *and in terms of covers of Hilbert surfaces into covers of* $\mathcal{A}_2$*, see [MR20, Section 2.5].*

2.4.1 Pull-back of modular forms

We close this chapter by observing that the map from Theorem 2.23 induces a pull-back map from the respective field of functions, and from the modular forms of fixed weight.

As we mentioned already, the pullback is to the symmetric Hilbert surface. This pullback can also be determined in terms of the Fourier expansions and we can give formula for the Fourier coefficients.

Proposition 2.26. *Assume* K *has a fundamental unit of norm* -1 *and set* ε *such a unit, with* $\varepsilon > 0$*. Set* $\{e_1, e_2\}$ *a* $\mathbb{Z}$*-basis for* $\mathcal{O}_K$ *and* ϕ_{e_1,e_2} *from Theorem 2.23. Consider* f *a holomorphic Siegel modular form of weight* k *for* $\mathrm{Sp}_4(\mathbb{Z})$*. Then* $g = \phi^*_{e_1,e_2} f$ *defines a symmetric holomorphic Hilbert modular form for* $\mathrm{SL}_2(\mathcal{O}_K)$ *with (paralel) weight* k*.*

Consider the Fourier expansion of f *at* ∞*.*

$$f = \sum_{T \in \mathrm{Sym}_2(\mathbb{Z})^\vee, T \geq 0} a_f(T) e^{i2\pi \operatorname{tr}(T\boldsymbol{\tau})},$$

then the Fourier expansion of g *at the cusp at* ∞ $a_g(0) + \sum_{\omega \in \partial_K^{-1,++}} a_g(\omega) e^{i2\pi \operatorname{tr}(\omega \boldsymbol{z})}$ *is given by* $a_g(0) = a_f(0_2)$ *and for* $\omega \in \partial_K^{-1,++}$*, setting* $t = \frac{\sqrt{\Delta_K}}{\varepsilon}\omega \in \mathcal{O}_K^{++}$*, then*

$$a_g(\omega) = \sum_{\substack{T \in \mathrm{Sym}_2(\mathbb{Z})^\vee, T \geq 0 \\ Q_T(e_1,e_2) = t}} a_f(T),$$

where $Q_T(x, y)$ *is the associated binary quadratic form for* T*.*

Proof. For the first part, from Theorem 2.23, we only need to check that if we set $\begin{pmatrix} * & * \\ C(\phi(\gamma)) & D(\phi(\gamma)) \end{pmatrix} = \phi(\gamma)$ for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathcal{O}_K)$, then $\det(C(\gamma)\phi(\boldsymbol{z}) + D(\gamma)) = (cz_1 + d)(\widetilde{c}z_2 + \widetilde{d})$. This follows from the same matrix computations as in the proof of [LY11, Proposition 3.1].

The second part is [LY11, Proposition 3.2]. Recall that for a given holomorphic Hilbert modular form g with respect to $\mathrm{SL}_2(\mathcal{O}_K)$, its Fourier expansion at the cusp at ∞ is given by $a_g(0) + \sum_{\omega \in (\partial_K)^{-1,++}} a_g(\omega) e^{i2\pi \operatorname{tr}(\omega \boldsymbol{z})}$. Observe that a $\mathbb{Z}$-basis of ∂_K^{-1} is given by $\{e_1/\sqrt{\Delta_K}, e_2/\sqrt{\Delta_K}\}$, and that $\omega \in \partial_K^{-1,++}$ if and only if $t := \frac{\sqrt{\Delta_K}}{\varepsilon}\omega \in \mathcal{O}_K^{++}$.

By the map ϕ_{e_1,e_2} from Theorem 2.23, a Siegel modular form is indexed in $T \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+}$. Then, restricting to the $\operatorname{Im} \phi_{e_1,e_2} \in \mathbb{H}_2$, for $R = \begin{pmatrix} e_1 & e_2 \\ \widetilde{e_1} & \widetilde{e_2} \end{pmatrix}$,

$$\begin{aligned} \operatorname{tr}(T\boldsymbol{\tau}) = \operatorname{tr}(T\phi_{e_1,e_2}\boldsymbol{z}) &= \operatorname{tr}\left(T\, {}^tR \begin{pmatrix} \frac{\varepsilon}{\sqrt{\Delta}} z_1 & 0 \\ 0 & -\frac{\widetilde{\varepsilon}}{\sqrt{\Delta}} z_2 \end{pmatrix} R \right) = \operatorname{tr}\left(RT\, {}^tR \begin{pmatrix} \frac{\varepsilon}{\sqrt{\Delta}} z_1 & 0 \\ 0 & -\frac{\widetilde{\varepsilon}}{\sqrt{\Delta}} z_2 \end{pmatrix} \right) \\ &= \begin{pmatrix} e_1 & e_2 \end{pmatrix} T \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \frac{\varepsilon}{\sqrt{\Delta}} z_1 + \begin{pmatrix} \widetilde{e_1} & \widetilde{e_2} \end{pmatrix} T \begin{pmatrix} \widetilde{e_1} \\ \widetilde{e_2} \end{pmatrix} \frac{-\widetilde{\varepsilon}}{\sqrt{\Delta}} z_2 = t\frac{\varepsilon}{\sqrt{\Delta}} z_1 + \widetilde{t}\frac{\widetilde{\varepsilon}}{-\sqrt{\Delta}} z_2 = \operatorname{tr}(\omega \boldsymbol{z}) \end{aligned}$$

for $t = Q_T(e_1, e_2)$ where Q_T is the associated binary quadratic form to T and $\omega = \frac{\varepsilon}{\sqrt{\Delta}} t$, as above. □

Chapter 3

Humbert singular relations

In this chapter, we present an alternative solution to the real multiplication moduli problem *directly on* $\mathcal{A}_2$. Then we recover the surfaces birational to the symmetric Hilbert varieties from the end of the previous chapter, and generalize them to not maximal real multiplication, on one hand, and to loci of points in $\mathcal{A}_2$ isogenous to product of elliptic curves. In a way, this last example is a degeneration to Hilbert modular surfaces, and includes $\mathcal{A}_1 \times \mathcal{A}_1 \subset \mathcal{A}_2$.

We have a first section presenting the Humbert singular relations, that describe the before mentioned loci. Then, we describe generalizations of these loci via the refined Humbert invariant and generalized Humbert varieties. For the former, the original definition comes from [Kan94], but we present an equivalent one by [GRV24], for any dimension and any characteristic. For the latter, in this chapter we simply present them and recopilate some general results from [Kan19] and [Kan16]. We will close the chapter with an alternative proof of Appendix 6.5.3.

At the end of Chapter 4, we will recover the notion of Humbert curves to discuss Shimura and modular curves on $\mathcal{A}_2$ on the same footing, and to relate them to the loci of quaternionic multiplication in $\mathcal{A}_2$.

3.1 Endomorphisms of abelian surfaces and Humbert singular relations

The material of this section is a reproduction of [Gij25, Section 4], which in turn comes from [BW03], a modern exposition on the original papers [Hum01], and [LY20, Section 3].

Consider a principally polarized abelian surface $(A_{\boldsymbol{\tau}}, E)$ for $\boldsymbol{\tau} \in \mathbb{H}_2$, $E = \begin{pmatrix} 0 & I_2 \\ -I_2 & 0 \end{pmatrix}$ the associated Riemann form (alternatively, with Hermitian form $H = (\operatorname{Im} \boldsymbol{\tau})^{-1}$). Setting the lattice $\Lambda_{\boldsymbol{\tau}} := (\boldsymbol{\tau}\ I_2)\mathbb{Z}^4 \subset \mathbb{C}^2$, as a complex torus $A_{\boldsymbol{\tau}}$ is identified with $\mathbb{C}^2/\Lambda_{\boldsymbol{\tau}}$.

Consider furthermore the endomorphism ring (as a complex torus) $\operatorname{End}(A_{\boldsymbol{\tau}})$ and the endomorphism algebra $\operatorname{End}^0(A_{\boldsymbol{\tau}}) := \operatorname{End}(A_{\boldsymbol{\tau}}) \otimes \mathbb{Q}$.

By [BL04, Prop 1.2.1], the elements of $\operatorname{End}(A_{\boldsymbol{\tau}})$ correspond to $\mathbb{C}$-linear maps $F : \mathbb{C}^2 \to \mathbb{C}^2$ such that $F(\Lambda_{\boldsymbol{\tau}}) \subset \Lambda_{\boldsymbol{\tau}}$. We have therefore two representations for $\operatorname{End}(A)$: considering the linear map $\mathbb{C}^2 \to \mathbb{C}^2$ induces the *analytical representation* $\rho_{a,\boldsymbol{\tau}} : \operatorname{End}(A_{\boldsymbol{\tau}}) \to \operatorname{Mat}_2(\mathbb{C})$, and the linear map on the lattice $\Lambda \to \Lambda$ induces the *rational representation* $\rho_{r,\boldsymbol{\tau}} : \operatorname{End}(A_{\boldsymbol{\tau}}) \to \operatorname{Mat}_4(\mathbb{Z})$. The compatibility condition that for any $f \in \operatorname{End}(A_{\boldsymbol{\tau}})$ $\rho_{a,\boldsymbol{\tau}}(f)(\Lambda) \subset \Lambda$ implies for

any $f \in \mathrm{End}(A_{\boldsymbol{\tau}})$:

$$\rho_{a,\boldsymbol{\tau}}(f)(\boldsymbol{\tau}\, I_2) = (\boldsymbol{\tau}\, I_2)\rho_{r,\boldsymbol{\tau}}(f). \tag{3.1}$$

Remark 3.1. *Setting* $\rho_{r,\boldsymbol{\tau}}(f) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *for matrices in* $\mathrm{Mat}_2(\mathbb{Z})$ *then this is equivalent to*

$$(\rho_{a,\boldsymbol{\tau}}(f)\boldsymbol{\tau} \quad \rho_{a,\boldsymbol{\tau}}(f)) = (\boldsymbol{\tau}A + C \quad \boldsymbol{\tau}B + D)$$

and solving for $\rho_{a,\boldsymbol{\tau}}(f)\boldsymbol{\tau}$ *yields the equation*

$$(\boldsymbol{\tau}A + C) = (\boldsymbol{\tau}B + D)\,\boldsymbol{\tau}. \tag{3.2}$$

The fact that $A_{\boldsymbol{\tau}}$ *admits a group structure implies that* $\mathbb{Z} \subset \mathrm{End}(A_{\boldsymbol{\tau}}$*, which identifies with matrices* nI_4*,* $n \in \mathbb{Z}$ *under the rational representation. Note that in this case Equation* (3.2) *verifies trivially as* $A = D = nI_2$ *for some* $n \in \mathbb{Z}$ *and* $B = C = 0$*.*

Generically in $\mathbb{H}_2$*, we observe that* $\mathrm{End}(A) = \mathbb{Z}$ *for an abelian variety, as Equation* (3.2) *yields restrictions on the period matrix.*

Furthermore, the two representations are related as by [BL04, Prop. 1.2.3]: the extended rational representation:

$$\rho_r \otimes 1 : (\mathrm{End}(A) \otimes \mathbb{Q}) \otimes \mathbb{C} \to \mathrm{End}_{\mathbb{C}}(\Lambda \otimes \mathbb{C}) \cong \mathrm{End}_{\mathbb{C}}(\mathbb{C}^2 \times \mathbb{C}^2)$$

satisfies $\rho_r \otimes 1 \cong \rho_a \otimes \overline{\rho_a}$ (where $\overline{\rho_a}$ means the complex conjugated representation to ρ_a).

Note that these representations depends on the choice of period matrix $\boldsymbol{\tau} \in \mathbb{H}_2$. If $M \in \mathrm{Sp}_4(\mathbb{Z})$ then

$$\rho_{r,M\boldsymbol{\tau}} = {}^tM^{-1}\rho_{r,\boldsymbol{\tau}}\,{}^tM, \tag{3.3}$$

as ${}^tM$ is the rational representation on the isomorphism of the complex tori $A_{\boldsymbol{\tau}}$ and $A_{\boldsymbol{\tau}'}$. Likewise, if $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ the analytic representation of said isomorphism is ${}^t(\gamma\tau + \delta)$, hence

$$\rho_{a,M\boldsymbol{\tau}} = {}^t(\gamma\tau + \delta)^{-1}\rho_{a,\boldsymbol{\tau}}\,{}^t(\gamma\tau + \delta). \tag{3.4}$$

The principal polarization induces an anti-involution on $\mathrm{End}(A)$ called the *Rosati involution,* $' : \mathrm{End}(A) \to \mathrm{End}(A)$ corresponding to the adjoint operator to the Riemann form E. That implies ([BL04][Prop. 5.1.1]) for all $\lambda, \mu \in \Lambda$, $f \in \mathrm{End}(A_{\boldsymbol{\tau}})$

$$E(\rho_{r,\boldsymbol{\tau}}(f)(\lambda), \mu) = E(\lambda, \rho_{r,\boldsymbol{\tau}(f')}(\mu))$$

or, in term of matrices,

$${}^t\rho_{r,\boldsymbol{\tau}}(f)\begin{pmatrix} 0 & I_2 \\ -I_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & I_2 \\ -I_2 & 0 \end{pmatrix}\rho_{r,\boldsymbol{\tau}}(f'). \tag{3.5}$$

Likewise, for the analytic representation, as $H = (\mathrm{Im}\,\boldsymbol{\tau})^{-1}$, we have

$$\overline{\rho_{a,\boldsymbol{\tau}}(f')}\,\mathrm{Im}\,\boldsymbol{\tau} = \mathrm{Im}\,\boldsymbol{\tau}\;{}^t\rho_{a,\boldsymbol{\tau}}(f), \tag{3.6}$$

where $\bar{\cdot}$ indicates complex conjugation.

We call $f \in \mathrm{End}(A)$ *symmetric* if $f' = f$. We denote $\mathrm{End}^s(f) \subset \mathrm{End}(A)$ the subring of symmetric endomorphisms. We have the following result ((from [BW03, Lemma 4.1]) for symmetric endomorphisms.

Lemma 3.2. *Consider an endomorphism $f \in \mathrm{End}(A_{\boldsymbol{\tau}})$, with rational representation given by $\rho_{r,\boldsymbol{\tau}}(f) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. Denote $A = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$, $b = B_{12}$, and $c = C_{12}$. For the following statements, we have (1) $\iff$ (2), and any of them implies (3).*

1. *The endomorphism f is symmetric.*
2. *The matrices B and D are antisymmetric and $D = {}^tA$: more precisely it follows that*
$$B = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & c \\ -c & 0 \end{pmatrix}, D = {}^tA.$$
3. *If $\boldsymbol{\tau} = \begin{pmatrix} \tau_1 & \tau_2 \\ \tau_2 & \tau_3 \end{pmatrix}$, then it satisfies*
$$a_2\tau_1 + (a_4 - a_1)\tau_2 - a_3\tau_3 + b\left(\tau_2^2 - \tau_1\tau_3\right) + c = 0.$$

In particular, for non-trivial equations, $\{1, \tau_1, \tau_2, \tau_3, \det(\boldsymbol{\tau})\}$ are $\mathbb{Q}$-linearly dependent.

Proof. The equivalence (1) $\iff$ (2) comes from unwinding (3.5), which gives
$$\begin{pmatrix} -{}^tC & {}^tA \\ -{}^tD & {}^tB \end{pmatrix} = \begin{pmatrix} C & D \\ -A & -B \end{pmatrix},$$
hence $D = {}^tA$ and B and C are antisymmetric. This equation is equivalent to $f = f'$.

Now (2) $\implies$ (3) is a consequence of (3.2), that is equivalent to
$$(\boldsymbol{\tau}A + C) = \left(\boldsymbol{\tau}B + {}^tA\right)\boldsymbol{\tau},$$
$$\boldsymbol{\tau}B\boldsymbol{\tau} - C = (\boldsymbol{\tau}A - {}^t(\boldsymbol{\tau}A)),$$
and as the three matrices involved are antisymmetric, the matrix equation above holds if and only if it holds for the off diagonal element, which is the equation of the statement. □

Remark 3.3. *As before, $\mathbb{Z} \subset \mathrm{End}(A)$ and actually $\mathbb{Z} \subset \mathrm{End}^s(A)$, but this yields trivial equations for $\boldsymbol{\tau}$. In addition, note that if $f \in \mathrm{End}^s(A)$, then $n + mf \in \mathrm{End}^s(A)$, with rational representation $nI_4 + m\rho_{r,\boldsymbol{\tau}}(f)$.*

Remark that on the linear equation in Lemma 3.2 we can take $a_1 = 0$, as that means considering $f - a_1 \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ instead of f, with rational representation $\rho_{r,\boldsymbol{\tau}}(f) - a_1 I_4$. Likewise, can also asume $gcd(a_2, a_3, a_4, b, c) = 1$: otherwise, it means that for $r = gcd(a_2, a_3, a_4, b, c)$, then $\rho_{r,\boldsymbol{\tau}}(f) = rM$ for $M \in \mathrm{Mat}_4(\mathbb{Z})$, which means that $f = rg$ for $g \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$.

For given $(a, b, c, d, e) \in \mathbb{Z}^5$, one can define the following matrix $R_{(a,b,c,d,e)} = \begin{pmatrix} A & B \\ C & {}^tA \end{pmatrix}$ where
$$A = \begin{pmatrix} 0 & a \\ -c & b \end{pmatrix}, B = \begin{pmatrix} 0 & d \\ -d & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & e \\ -e & 0 \end{pmatrix}.$$

Lemma 3.4. *If we have $(a, b, c, d, e) \in \mathbb{Z}^5$ such that*
$$a\tau_1 + b\tau_2 + c\tau_3 + d(\tau^2 - \tau_1\tau_3) + e = 0,$$
then there exists $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ with $\rho_{r,\boldsymbol{\tau}}(f) = R_{(a,b,c,d,e)}$.

Proof. By the argument at the end of the proof of Lemma 3.2, the equation in the statement is in fact equivalent to

$$(\boldsymbol{\tau} A + C) = \left(\boldsymbol{\tau} B + {}^{t}A\right) \boldsymbol{\tau},$$

which is equivalent to (3.2), and to the compatibility condition (3.1). Hence, there exists an endomorphism $f \in \mathrm{End}(A_{\boldsymbol{\tau}})$ with such a rational representation. It is symmetric by Lemma 3.2. □

Remark 3.5. *In the proof of Lemma 3.2 and Lemma 3.4, the only instance where we use that the abelian variety has dimension $g = 2$ is that the matrix equation involves antisymmetric matrices in $\mathrm{Mat}_2(\mathbb{C})$, so it is a subspace of dimension one. That is why Humbert singular relations only arise naturally in this dimension.*

In conclusion, the existence of $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ with $f \notin \mathbb{Z}$ is determined, via $\rho_{r,\boldsymbol{\tau}}(f)$, by certain quadratic relations for the components of the period matrix $\boldsymbol{\tau}$.

Definition 3.6. *We call* Humbert singular relation *(HSR for short) the following equation for $a, b, c, d, e \in \mathbb{Z}$,*

$$a\tau_1 + b\tau_2 + c\tau_3 + d(\tau_2^2 - \tau_1\tau_3) + e = 0.$$

If in addition $\gcd(a, b, c, d, e) = 1$, we say that the relation is primitive.

Let us consider now the analytic representation $\rho_{a,\boldsymbol{\tau}}(f)$ of a symmetric endomorphism, and its characteristic polynomial as a matrix in $\mathrm{Mat}_2(\mathbb{C})$:

$$P_f(t) := \det(tI_2 - \rho_{a,\boldsymbol{\tau}}(f)) = t^2 - \mathrm{tr}(\rho_{a,\boldsymbol{\tau}}(f))t + \det(\rho_{a,\boldsymbol{\tau}}(f)). \tag{3.7}$$

By (3.1) and Lemma 3.2, for $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$,

$$\rho_{a,\boldsymbol{\tau}}(f) = \boldsymbol{\tau} B + {}^{t}A = \begin{pmatrix} -d\tau_2 & d\tau_1 - c \\ -d\tau_3 + a & d\tau_2 + b \end{pmatrix}.$$

Hence, it follows (using the singular relation solved by $\boldsymbol{\tau}$):

$$\begin{aligned} \mathrm{tr}(\rho_{a,\boldsymbol{\tau}}(f)) &= b, \\ \det(\rho_{a,\boldsymbol{\tau}}(f)) &= ac + de. \end{aligned}$$

Therefore, $P_f(t) = t^2 - bt + (ac + de)$. More is true, $\rho_{a,\boldsymbol{\tau}}$ induces a ring isomorphism ([BW03, Corollary 4.4])

$$\{n + mf,\ n, m \in \mathbb{Z}\} \to {}^{\mathbb{Z}[t]}\!\!\diagup\!\!{}_{\left(t^2 - bt + ac + de\right)}.$$

Note that $\Delta = \Delta(f) := \mathrm{Disc}(P_f) = b^2 - 4(ac + de)$. Remark that $\Delta \equiv 0, 1 \bmod 4$.

Lemma 3.7. *If $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ then $\Delta(f) \geq 0$, and $\Delta(f) = 0$ if and only if $f \in \mathbb{Z}$.*

Proof. (From [BW03, Proposition 4.7]). By (3.6), it follows that $\rho_{a,\boldsymbol{\tau}}(f)\,\mathrm{Im}\,\boldsymbol{\tau}$ is a Hermitian matrix. Consider λ_k an eigenvalue of $\rho_a(f)$ with eigenvector v_k, for $k = 1, 2$. Because f is symmetric,

$$\lambda_k H(v_k, v_k) = H(\rho_a(f)v_k, v_k) = H(v_k, \rho_a(f)v_k) = \overline{\lambda_k} H(v_k, v_k),$$

therefore λ_k is real, and is λ_k is a root of P_f, it follows that $\Delta(f) \geq 0$ and $\Delta(f) = (\lambda_1 - \lambda_2)^2$. It can only be 0 if the two eigenvalues of $\rho_a(f)$ coincide and are integers, say $m \in \mathbb{Z}$, as $P_f(t) \in \mathbb{Z}[t]$. Therefore, $\rho_a(f) - mI_2 = \rho_a(f - m)$ is a nilpotent matrix (with exponent two, necessarily, for it is a 2×2 matrix). Then $g = f - m \in \mathrm{End}^s(A)$ and $gg' = g^2 = 0$, but the Rosati involution is a positive definite involution, which implies $\mathrm{tr}(gg') = 0$ if and only if $g = 0$. For a more elementary argument, one can check that the following linear algebra problem on $\mathrm{Mat}_2(\mathbb{C})$

$$\begin{cases} Z^2 = 0 \\ Z \operatorname{Im}(\boldsymbol{\tau}) \text{ is Hermitian} \end{cases}$$

only has as solution $Z = 0$, and then $\rho_a(f - m) = 0$, and $f \in \mathbb{Z}$. A possible proof goes by inspection: in dimension two a nilpotent matrix is characterized by trace and determinant zero, so there exists $z, u, v \in \mathbb{C}$ with $-z^2 = uv$ such that

$$Z = \begin{pmatrix} z & u \\ v & -z \end{pmatrix}.$$

One can impose the Hermitian conditions on

$$\begin{pmatrix} z & u \\ v & -z \end{pmatrix} \begin{pmatrix} \operatorname{Im}(\tau_1) & \operatorname{Im}(\tau_2) \\ \operatorname{Im}(\tau_2) & \operatorname{Im}(\tau_3) \end{pmatrix} = \begin{pmatrix} z \operatorname{Im}(\tau_1) + u \operatorname{Im}(\tau_2) & z \operatorname{Im}(\tau_2) + u \operatorname{Im}(\tau_3) \\ v \operatorname{Im}(\tau_1) - z \operatorname{Im}(\tau_2) & v \operatorname{Im}(\tau_2) - z \operatorname{Im}(\tau_3) \end{pmatrix},$$

resulting on the equations (where we set $y_i = \operatorname{Im}(\tau_i)$, remark that $y_1, y_3 > 0$):

$$\begin{aligned} zy_1 + uy_2 &= \bar{z}y_1 + \bar{u}y_2, \\ vy_2 - zy_3 &= \bar{v}y_2 - \bar{z}y_3, \\ vy_1 - zy_2 &= \bar{z}y_1 + \bar{u}y_3, \end{aligned}$$

equivalent to

$$\operatorname{Im}(z)y_1 = -\operatorname{Im}(u)y_2, \tag{3.8}$$

$$\operatorname{Im}(z)y_3 = \operatorname{Im}(v)y_2, \tag{3.9}$$

$$v = \frac{2\operatorname{Re}(z)y_2 + \bar{u}y_3}{y_1}. \tag{3.10}$$

Note that (3.10) implies

$$y_1 \operatorname{Im}(v) = -y_3 \operatorname{Im}(u). \tag{3.11}$$

As $-z^2 = uv$, it follows from (3.10) that

$$y_1 z^2 + 2y_2 u \operatorname{Re}(z) + y_3 |u|^2 = 0, \tag{3.12}$$

which implies, as $\operatorname{Re}(z^2) = \operatorname{Re}(z)^2 - \operatorname{Im}(z)^2$

$$y_1(\operatorname{Re}(z)^2 - \operatorname{Im}(z)^2) + 2y_2 \operatorname{Re}(u) \operatorname{Re}(z) + y_3(\operatorname{Re}(u)^2 + \operatorname{Im}(u)^2) = 0,$$

or equivalently,

$$y_1 \operatorname{Re}(z)^2 + 2y_2 \operatorname{Re}(u) \operatorname{Re}(z) + y_3 \operatorname{Re}(u)^2 = y_1 \operatorname{Im}(z)^2 - y_3 \operatorname{Im}(u)^2. \tag{3.13}$$

We rewrite the right hand side. Together with (3.8) and (3.9),

$$y_1 \operatorname{Im}(z)^2 - y_3 \operatorname{Im}(u)^2 = -\operatorname{Im}(z)\operatorname{Im}(u)y_2 + y_1 \operatorname{Im}(v)\operatorname{Im}(u) = \operatorname{Im}(u)\left(-y_2 \operatorname{Im}(z) + y_1 \operatorname{Im}(v)\right)$$
$$= \operatorname{Im}(u)\left(\frac{-\operatorname{Im}(v)y_2^2}{y_3} + y_1 \operatorname{Im}(v)\right) = \operatorname{Im}(u)\operatorname{Im}(v)\frac{y_1y_3 - y_2^2}{y_3} \leq 0,$$

because $\operatorname{Im}(u)\operatorname{Im}(v) \leq 0$ by (3.11), $\det(\operatorname{Im}\boldsymbol{\tau}) > 0$ and $y_3 > 0$. But on the other hand,

$$y_1 \operatorname{Re}(z)^2 + 2y_2 \operatorname{Re}(u)\operatorname{Re}(z) + y_3 \operatorname{Re}(u)^2 = (\operatorname{Re}(z)\ \operatorname{Re}(u))\begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix}\begin{pmatrix} \operatorname{Re}(z) \\ \operatorname{Re}(u) \end{pmatrix} \geq 0,$$

as $\operatorname{Im}(\boldsymbol{\tau})$ is positive definite. Therefore, the only possibility for Equation (3.13) to hold is if $\operatorname{Re}(z) = \operatorname{Re}(u) = 0$ and $\operatorname{Im}(u)\operatorname{Im}(v) = 0$. By (3.10) we have $0 = \operatorname{Im}(u) = \operatorname{Im}(v)$, and by (3.8) $\operatorname{Im}(z) = 0$. Hence $z, u, v \in \mathbb{R}$ and we have $z = u = 0$, so by (3.10) we conclude $v = 0$. □

Remark 3.8. *We remark that the end of the proof of Lemma 3.7 implies that there are no* symmetric *nilpotent endomorphisms. When $A_{\boldsymbol{\tau}} \simeq E^2$, the endomorphism algebra is a matrix algebra and contains nilpotent elements, therefore they cannot be symmetric with respect to the Rosati involution.*

Compare this with [BL04, Section 5.3] (for any dimension g): there exist symmetric idempotents*, the norm-endomorphisms detecting abelian subvarieties. Roughly speaking, they are constructed as "projections" $f_X = f \in \operatorname{End}^s(A)$ to a subvariety $X \subset A$ such that $f|_X = e(X)_{[X]}$, for $e(X)$ the exponent of the induced polarization. These endomorphisms are symmetric with respect to the Rosati involution and satisfy $f^2 = e(X)f$. One can further consider the (canonical) idempotent $e(X)^{-1}f \in \operatorname{End}_0^s(A)$. By [BL04, Theorem 5.3.2], the symmetric idempotents of $\operatorname{End}_0(A)$ are in bijection with the abelian subvarieties of A.*

In particular, for ppas, it follows that for $f \in \operatorname{End}^s(A)$, then either f is invertible in $\operatorname{End}_0(A)$ or there exists $\mu \in \mathbb{Z}$ such that $f - \mu$ is a multiple of an idempotent element in $\operatorname{End}_0(A)$.

After Lemma 3.7, we can consider the following definition. Remark that $\mathcal{H}_0 = \mathcal{A}_2$, so we do not consider it.

Definition 3.9. *For $\Delta > 0$, $\Delta \equiv 0, 1 \bmod 4$, we define the following loci in $\mathcal{A}_2$:*

$$\mathcal{G}_\Delta := \{A \in \mathcal{A}_2 |\ \textit{there exists } f \in \operatorname{End}^s(A) \textit{ with } \Delta(f) = \Delta\} \subset \mathcal{A}_2,$$
$$\mathcal{H}_\Delta := \{A \in \mathcal{A}_2 |\ \textit{there exists } f \in \operatorname{End}^s(A) \text{ primitive } \textit{with } \Delta(f) = \Delta\} \subset \mathcal{A}_2,$$

where we call $f \in \operatorname{End}(A)$ primitive *if there not exists $m \in \mathbb{Z}$ with $f = mg$ with $g \in \operatorname{End}(A)$. We call $\mathcal{H}_\Delta$ the* Humbert surface *of discriminant Δ.*[1]

Note that $\mathcal{G}_{m^2\Delta}$ is well defined for any $m \in \mathbb{Z}$ and $\mathcal{G}_\Delta \subset \mathcal{G}_{m^2\Delta}$. In addition, we have

$$\mathcal{G}_\Delta = \bigcup_{m^2|\Delta} \mathcal{H}_{\frac{\Delta}{m^2}}$$

[1] We saw the notation $\mathcal{G}_\Delta$ in [Gee88, Equation above Proposition IV.2.3, page 211], but it does not appear to have a proper name.

Lemma 6.21 will imply that the discriminant of a HSR is invariant under $\mathrm{Sp}_4(\mathbb{Z})$-action, so the corresponding preimages in $\mathbb{H}_2$ are

$$\{\boldsymbol{\tau} \in \mathbb{H}_2 : \text{ there exists a HSR } l \text{ with } \Delta(l) = \Delta \text{ such that } \boldsymbol{\tau} \text{ satisfies } l\},$$
$$\{\boldsymbol{\tau} \in \mathbb{H}_2 : \text{ there exists a } \textit{primitive} \text{ HSR } l \text{ with } \Delta(l) = \Delta \text{ such that } \boldsymbol{\tau} \text{ satisfies } l\}.$$

If $A \in \mathcal{G}_\Delta$ there exists $m^2|\Delta$ such that $A \in \mathcal{H}_{\Delta'}$ with $\Delta' = \frac{\Delta}{m^2}$.

We distinguish the following cases in terms of this discriminant, which is the content of [BW03, Proposition 4.8, Proposition 4.9].

Proposition 3.10. *Suppose that $\boldsymbol{\tau}$ satisfies a primitive HSR with discriminant Δ. Then the following holds:*

- *If Δ is a fundamental discriminant, then there exists an embedding of the ring of integers*

$$\mathcal{O}_{\mathbb{Q}(\sqrt{\Delta})} \hookrightarrow \mathrm{End}^s(A).$$

  More generally, if Δ is the discriminant of an order $\mathcal{O}$ in a real quadratic field, then there exists an embedding

$$\mathcal{O} \hookrightarrow \mathrm{End}^s(A).$$

  In both cases, there is an embedding of a real quadratic field $K \hookrightarrow \mathrm{End}_0^s(A)$.

- *If $\Delta = \delta^2$, then A is not simple. More precisely, there exists an isogeny of degree δ^2:*

$$(E_1 \times E_2, p_1^*\mathcal{O}_{E_1}(\delta) \otimes p_2^*\mathcal{O}_{E_2}(\delta)) \to (A, H).$$

Corollary 3.11. *Taking $\delta = 1$ above, it follows that $\mathcal{A}_1 \times \mathcal{A}_1 = \mathcal{H}_1$.*

We will proof in the next subsection that for choose a basis for $\rho_{r,\tau}(f)$ in order to obtain a symmetric (non trivial) endomorphism of the following special shape:

$$\begin{pmatrix} A & 0 \\ 0 & {}^tA \end{pmatrix}, \; A = \begin{pmatrix} 0 & a \\ -1 & b \end{pmatrix}$$

with

$$(a, b) = \begin{cases} \left(\frac{-\Delta}{4}, 0\right) & \text{if } \Delta \equiv 0 \bmod 4, \\ \left(\frac{1-\Delta}{4}, 1\right) & \text{if } \Delta \equiv 1 \bmod 4. \end{cases}$$

Therefore, as per Lemma 3.2, given $\boldsymbol{\tau} \in \mathbb{H}_2$ satisfying a HSR of discriminant Δ, we have a distinguished one, which we call the *normalized* HSR of discriminant Δ.

Proposition 3.12 (Humbert's lemma)**.** *For $\boldsymbol{\tau} \in \mathbb{H}_2$, if $A_{\boldsymbol{\tau}}$ satisfies a Humbert singular relation of discriminant Δ, then there exists $M \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\boldsymbol{\tau}' = M\boldsymbol{\tau}$ solves:*

$$\frac{-\Delta}{4}\tau_1' + \tau_3' = 0 \text{ if } \Delta \equiv 0 \text{ mod } 4,$$
$$\frac{1-\Delta}{4}\tau_1' + \tau_2' + \tau_3' = 0 \text{ if } \Delta \equiv 1 \text{ mod } 4.$$

We will say more on how Humbert singular relations behave under $\mathrm{Sp}_4(\mathbb{Z})$-action in the next section.

Remark that as a corollary, we have the following.

Corollary 3.13. *For any discriminant Δ, $\mathcal{H}_\Delta$ is an irreducible surface, and the number of connected components of $\mathcal{G}_\Delta$ is $\#\{m \in \mathbb{Z}_{>0},\ m^2|\Delta\}$.*

Remark 3.14. *For the more general definition of a Humbert surface in [Gee88, Chapter IX, Section 2], that is is defined a subvariety of the moduli of $(1,t)$-polarized abelian surfaces, for $t > 1$, it is not always the case that $\mathcal{H}_\Delta$ is irreducible, see [Gee88, Chapter IX, Theorem 2.4].*

3.1.1 Proof of Humbert's lemma

We study in more detail the $\mathrm{Sp}_4(\mathbb{Z})$-action on a given HSR. Remark that for $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z})$, its action on $\mathbb{H}_2$ is given by fractional linear transformations:

$$M\boldsymbol{\tau} := (\alpha\boldsymbol{\tau} + \beta)(\gamma\boldsymbol{\tau} + \delta)^{-1}.$$

and by Equation (3.2), Lemma 3.2, and Lemma 3.4, given $\boldsymbol{\tau} \in \mathbb{H}_2$ satisfying

$$a\tau_1 + b\tau_2 + c\tau_3 + d(\tau_2^2 - \tau_1\tau_3) + e = 0,$$

then we consider $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ with $\rho_{r,\boldsymbol{\tau}}(f) = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 0 & a \\ -c & b \end{pmatrix} & \begin{pmatrix} 0 & d \\ -d & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & e \\ -e & 0 \end{pmatrix} & {}^tA \end{pmatrix}$. Then $\boldsymbol{\tau}' = M\boldsymbol{\tau}$ solves a Humbert singular relation associated with the matrix ${}^tM^{-1}\rho_{r,\boldsymbol{\tau}}(f)$. Let us study the effect on the singular relations by families of elements in $\mathrm{Sp}_4(\mathbb{Z})$. We eventually be interested in the effect they have on the quadratic term.

By [Kli90, Chapter I, Section 3, Proposition 6] we have two alternative sets of generators for $\mathrm{Sp}_4(\mathbb{Z})$.

Proposition 3.15. *The modular group $\mathrm{Sp}_{2n}(\mathbb{Z})$ is generated by either:*

- *The union of $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$ and the translations $\begin{pmatrix} I_2 & T \\ 0 & I_2 \end{pmatrix}$, for T in any generator set for $\mathrm{Sym}_n(\mathbb{Z})$.*
- *The union of the set of the subgroup of unimodular transformations*

$$U = \left\{ \begin{pmatrix} U & 0 \\ 0 & {}^tU^{-1} \end{pmatrix},\ U \in \mathrm{GL}_n(\mathbb{Z}) \right\},$$

and the n subgroups composed by "diagonal" transformations of $\mathrm{SL}_2(\mathbb{Z})$ given by $\alpha \to m_\alpha^i$ for $i = 1, \ldots, n$, where the four elements of α are sent to the elements in positions $(i,i), (i,i+n), (i+n,i)$ and $(i+n,i+n)$, and the elements of I_{2n} elsewhere

In the case of $\mathrm{Sp}_4(\mathbb{Z})$ these embeddings are given by

$$\begin{gathered}\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{Sp}_4(\mathbb{Z}) \\ \begin{pmatrix}\alpha_{1,1} & \alpha_{2,1}\\ \alpha_{1,2} & \alpha_{2,2}\end{pmatrix} \mapsto \begin{pmatrix}\alpha_{1,1} & 0 & \alpha_{2,1} & 0\\ 0 & 1 & 0 & 0\\ \alpha_{1,2} & 0 & \alpha_{2,2} & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\end{gathered}$$

and

$$\begin{gathered}\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{Sp}_4(\mathbb{Z}) \\ \begin{pmatrix}\alpha_{1,1} & \alpha_{2,1}\\ \alpha_{1,2} & \alpha_{2,2}\end{pmatrix} \mapsto \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & \alpha_{1,1} & 0 & \alpha_{2,1}\\ 0 & 0 & 1 & 0\\ 0 & \alpha_{1,2} & 0 & \alpha_{2,2}\end{pmatrix}.\end{gathered}$$

Notice that these are the generators of the "diagonal transformations":

$$\begin{pmatrix}D_1 & D_2\\ D_3 & D_4\end{pmatrix},$$

for D_i diagonal matrices. One can check that a matrix of this form belongs to $\mathrm{Sp}_4(\mathbb{Z})$ if and only if, writing $D_i = \mathrm{diag}(\alpha_i, \beta_i)$ for $i = 1, \dots, 4$, it holds $D_1 D_4 - D_2 D_3 = I_2$, or equivalently

$$\alpha_1\alpha_4 - \alpha_2\alpha_3 = \beta_1\beta_4 - \beta_2\beta_3 = 1.$$

Therefore, only if and only if it is in the image of the morphisms:

$$\begin{gathered}\Phi : \mathrm{SL}_2(\mathbb{Z}) \times \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{Sp}_4(\mathbb{Z}) \\ \begin{pmatrix}\alpha_1 & \alpha_2\\ \alpha_3 & \alpha_4\end{pmatrix}, \begin{pmatrix}\beta_1 & \beta_2\\ \beta_3 & \beta_4\end{pmatrix} \mapsto \begin{pmatrix}\begin{pmatrix}\alpha_1 & \\ & \beta_1\end{pmatrix} & \begin{pmatrix}\alpha_2 & \\ & \beta_2\end{pmatrix}\\ \begin{pmatrix}\alpha_3 & \\ & \beta_3\end{pmatrix} & \begin{pmatrix}\alpha_4 & \\ & \beta_4\end{pmatrix}\end{pmatrix}\end{gathered}$$

Notice that Φ is a group homomorphism, so $\Phi(\alpha, \beta) = \Phi(\alpha, I_2)\Phi(I_2, \beta)$, which are precisely the matrices given in the above Proposition.

One can likewise check that matrices of the form $\Phi(\alpha, I_2)$ act on $\mathbb{H}_2$ almost as a one-dimensional fractional transformation for τ_1:

$$\boldsymbol{\tau} = \begin{pmatrix}\tau_1 & \tau_2\\ \tau_3 & \tau_4\end{pmatrix} \mapsto \begin{pmatrix}\frac{\alpha_1\tau_1+\alpha_2}{\alpha_3\tau_1+\alpha_4} & \frac{\tau_2}{\alpha_3\tau_1 1+\alpha_4}\\ \frac{\tau_2}{\alpha_3\tau_1 1+\alpha_4} & \tau_3 - \frac{\alpha_3\tau_2^2}{\alpha_3\tau_1+\alpha_4}\end{pmatrix}$$

and analogously for $\Phi(I_2, \beta)$.

We therefore have the following four families of matrices to consider.

- The inverse $J = \begin{pmatrix}0 & I_2\\ -I_2 & 0\end{pmatrix}$, acting as $\boldsymbol{\tau} \mapsto -\boldsymbol{\tau}^{-1}$.
- The translations $\begin{pmatrix}I_2 & T\\ 0 & I_2\end{pmatrix}$, for $T \in \mathrm{Sym}_2(\mathbb{Z})$, acting as $\boldsymbol{\tau} \mapsto \boldsymbol{\tau} + T$.

- The unimodular transformations $\begin{pmatrix} U & 0 \\ 0 & {}^tU^{-1} \end{pmatrix}$ for $U \in \mathrm{GL}_2(\mathbb{Z})$, acting as $\boldsymbol{\tau} \mapsto U\boldsymbol{\tau}\,{}^tU$.

- The diagonal transformations $\Phi(\alpha, \beta)$ with $\alpha, \beta \in \mathrm{SL}_2(\mathbb{Z})$, acting as $\boldsymbol{\tau} \to (D_1\boldsymbol{\tau} + D_2)(D_3\boldsymbol{\tau} + D_4)^{-1}$ for $D_i = \mathrm{diag}(\alpha_i, \beta_i)$ for $i = 1, \ldots, 4$.

Proposition 3.16. *Consider $\boldsymbol{\tau} \in \mathbb{H}_2$ satisfying a Humbert singular relation (a, b, c, d, e). Consider $\boldsymbol{\tau}' = M\boldsymbol{\tau}$*

- *for $M = J$, then $\boldsymbol{\tau}'$ satisfies the Humbert singular relation given by*

$$(-c, b, -a, -e, -d).$$

- *for M a translation by $T = \begin{pmatrix} n_1 & n_2 \\ n_2 & n_3 \end{pmatrix}$, then $\boldsymbol{\tau}'$ satisfies the Humbert singular relation given by*

$$(a + dn_3, b - 2dn_2, c + dn_1, d, e - (an_1 + bn_2 + cn_3 + d\det T),$$

in particular, if $d = 0$ then $(a, b, c, 0, e - (an_1 + bn_2 + cn_3);)$

- *for M a unimodular transformation, then $\boldsymbol{\tau}'$ satisfies the Humbert singular relation given by*

$$(q, s - p, -r, \pm d, \pm e)$$

where $\begin{pmatrix} p & q \\ r & s \end{pmatrix} = {}^tU^{-1} \begin{pmatrix} 0 & a \\ -c & b \end{pmatrix} {}^tU$, so $b = s + p$ and $ac = ps - qr$;

- *for $M = \begin{pmatrix} D_1 & D_2 \\ D_3 & D_4 \end{pmatrix}$, then $\boldsymbol{\tau}'$ satisfies a Humbert singular relation associated to (a', b', c', d', e') where*

$$\begin{aligned} d' &= a\alpha_4\beta_3 - c\alpha_3\beta_4 + d\alpha_4\beta_4 - e\alpha_3\beta_3, \\ e' &= -a\alpha_2\beta_1 + c\alpha_1\beta_2 - d\alpha_2\beta_2 + e\alpha_1\beta_1. \end{aligned}$$

Proof. The proof follows simply by computations of ${}^tM^{-1}\rho_{r,\boldsymbol{\tau}}(f)\,{}^tM$, and general properties of symplectic matrices. For easy computation, note that if $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z})$, then

$$M^{-1} = \begin{pmatrix} {}^t\delta & -{}^t\beta \\ -{}^t\gamma & {}^t\alpha \end{pmatrix}.$$

We also use that an antisymmetric matrix in dimension two is completely determined by its off diagonal element, which in determine up to sign by the determinant of the matrix.

Set $A = \begin{pmatrix} 0 & a \\ -c & b \end{pmatrix}$, $B = \begin{pmatrix} 0 & d \\ -d & 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & e \\ -e & 0 \end{pmatrix}$, and $R = \begin{pmatrix} A & B \\ C & {}^tA \end{pmatrix}$.

Then the computations ${}^tM^{-1}R\,{}^tM$:

$$
\begin{gathered}
{}^tJR\,{}^tJ = \begin{pmatrix} {}^tA & -C \\ -B & A \end{pmatrix} \\
\begin{pmatrix} I_2 & 0 \\ -T & I \end{pmatrix}\begin{pmatrix} A & B \\ C & {}^tA \end{pmatrix}\begin{pmatrix} I_2 & 0 \\ T & I_2 \end{pmatrix} = \begin{pmatrix} I_2 & 0 \\ -T & I \end{pmatrix}\begin{pmatrix} A+BT & B \\ C+{}^tAT & {}^tA \end{pmatrix} = \begin{pmatrix} A+BT & B \\ -TA-TBT+C+{}^tAT & -BT+{}^tA \end{pmatrix} \\
\begin{pmatrix} {}^tU^{-1} & 0 \\ 0 & U \end{pmatrix}\begin{pmatrix} A & B \\ C & {}^tA \end{pmatrix}\begin{pmatrix} {}^tU & 0 \\ 0 & U^{-1} \end{pmatrix} = \begin{pmatrix} {}^tU^{-1}A\,{}^tU & {}^tU^{-1}BU^{-1} \\ UC\,{}^tU & U\,{}^tAU^{-1} \end{pmatrix} \\
\begin{pmatrix} D_4 & -D_3 \\ -D_2 & D_1 \end{pmatrix}\begin{pmatrix} A & B \\ C & {}^tA \end{pmatrix}\begin{pmatrix} D_1 & D_3 \\ D_2 & D_4 \end{pmatrix} = \begin{pmatrix} D_4 & -D_3 \\ -D_2 & D_1 \end{pmatrix}\begin{pmatrix} AD_1+BD_2 & AD_3+BD_4 \\ CD_1+{}^tAD_2 & CD_3+{}^tAD_4 \end{pmatrix}.
\end{gathered}
$$

Item 1 now follows immediately. Item 2 follows from the fact that $-TA-TBT+C+{}^tAT = C+({}^tAT-TA)+(-TBT)$ is antisymmetric matrix. Now

$$
TA = \begin{pmatrix} n_1 & n_2 \\ n_2 & n_3 \end{pmatrix}\begin{pmatrix} 0 & a \\ -c & b \end{pmatrix} = \begin{pmatrix} * & an_1+bn_2 \\ -n_3c & * \end{pmatrix},
$$

hence $({}^tAT-TA)_{1,2} = -n_3c-(an_1+bn_2)$. In addition, $\det(-TBT) = -\det(T)^2d^2$, so $(-TBT)_{1,2} = \pm\det(T)d$, but to determine the sign on has to compute the product, getting the negative sign. Therefore, $(-TA-TBT+C+{}^tAT)_{12} = e-(an_1+bn_2+cn_3+d\det(T))$.

For Item 3, we use again properties of antisymmetric matrices, noticing that $\det(U) = \pm 1$.

In Item 4, we only compute the product that we have for C' :

$$
\begin{aligned}
C' = D_4AD_3 + D_4BD_4 - D_3CD_3 - D_3\,{}^tAD_4 &= (D_4AD_3 - D_3\,{}^tAD_4) + D_4BD_4 - D_3CD_3 \\
&= \begin{pmatrix} 0 & d' \\ -d' & 0 \end{pmatrix}
\end{aligned}
$$

with $d' = a\alpha_4\beta_3 - c\alpha_3\beta_4 + d\alpha_4\beta_4 - 3\alpha_3\beta_3$, that follows from

$$
D_4AD_3 = \begin{pmatrix} * & a\alpha_4\beta_3 \\ -\alpha_3\beta_4c & * \end{pmatrix},
$$

and

$$
(D_4BD_4)_{1,2} = d\alpha_4\beta_4.
$$

And analogously for B', it follows from

$$
B' = -(D_2AD_1 - D_1\,{}^tAD_1) - D_2BD_2 + D_1\,{}^tAD_2
$$

□

We remark what happens if $d = 0$ in the Proposition 3.16.

Corollary 3.17. *Consider $\boldsymbol{\tau} \in \mathbb{H}_2$ satisfying a* linear *Humbert singular relation given by $l = (a, b, c, 0, e)$. Then if $M \in \mathrm{Sp}_4(\mathbb{Z})$ is one of the following:*

- *a translation by $T \in \mathrm{Sym}_2(\mathbb{Z})$,*
- *a unimodular transformation,*

or composition of such matrices, then $M\boldsymbol{\tau}$ satisfies a linear Humbert singular relation. If in addition l is homogeneous ($e = 0$), then J is also included in the list.

Proof of Humbert's lemma. We sketch the algorithm in [BW03, Proposition 4.5] and [Run99, Theorem 2], with our general considerations from Proposition 3.16. We only present the steps until a homogeneous linear relation $e = d = 0$. First we will make the following observations about the diagonal transformations from 3.16, as those are the one with more freedom to transform d. Start with $l = (a, b, c, d, e)$, to trasnform it to a linear HSR via diagonal transformations, we would like to solve for $\alpha_4, \beta_4, \alpha_3, \beta_3$ such that

$$0 = a\alpha_4\beta_3 - c\alpha_3\beta_4 + d\alpha_4\beta_4 - e\alpha_3\beta_3.$$

Remark then that α_i, β_i for $i = 1, 2$ can be determined via the Bézout's lemma, for that $\alpha_1\alpha_4 - \alpha_2\alpha_3 = \beta_1\beta_4 - \beta_2\beta_3 = 1$.

Assume $\alpha_4 = 0$, then it reduces to

$$0 = -\alpha_3(c\beta_4 + e\beta_3),$$

hence we simply require $c\beta_4 + e\beta_3 = 0$. So if $g = \gcd(c, e)$, set $\alpha_4 = 0$, $\beta_3 = -c/g$ and $\beta_4 = e/g$. We express such a matrix as

$$\begin{pmatrix} \operatorname{diag}(*, *) & \operatorname{diag}(*, *) \\ \operatorname{diag}(*, \frac{-c}{\gcd(c,e)}) & \operatorname{diag}(0, \frac{e}{\gcd(c,e)}) \end{pmatrix}, \tag{3.14}$$

where * means that the rest of parameters can be recovered from $1 = -\alpha_2\alpha_3 = \beta_1 e/g + \beta_2 c/g$.

Assume now what happens with $l = (a, b, c, 0, e)$. To preserve its linearity, we argue as above, but in this case we can also make $e' = 0$, so we want to solve simultaneously

$$\begin{aligned} 0 &= a\alpha_4\beta_3 - c\alpha_3\beta_4 - e\alpha_3\beta_3 \\ 0 &- a\alpha_2\beta_1 + c\alpha_1\beta_2 + e\alpha_1\beta_1 \end{aligned}$$

if we assume $\alpha_4 = 0$ again, as $1 = \alpha_1\alpha_4 - \alpha_2\alpha_3 = -\alpha_2\alpha_3$, we can assume $\alpha_2 = 1$ and $\alpha_3 = -1$, hence

$$\begin{aligned} 0 &= (c\beta_4 + e\beta_3) \\ 0 &= -a\beta_1 + c\alpha_1\beta_2 + e\alpha_1\beta_1, \end{aligned}$$

so if $g = \gcd(c, e)$ then and if we consider r, s such that $rc + se = \gcd(c, e)$, then we can set $\beta_3 = -c/g$, $\beta_4 = e/g$, $\beta_1 = s$, $\beta_2 = r$, which in our system above implies

$$0 = -as + cr\alpha_1 + es\alpha_1 = -as + \alpha_1(es + cr) = -as + \alpha_1 g,$$

or $g\alpha_1 = as$. If we can assume $g|a$, we can solve for α_1. We express such a matrix as

$$\begin{pmatrix} \operatorname{diag}(*\frac{a}{\gcd(c,e)}, *) & \operatorname{diag}(*, *) \\ \operatorname{diag}(*, \frac{-c}{\gcd(c,e)}) & \operatorname{diag}(0, \frac{e}{\gcd(c,e)}) \end{pmatrix}, \tag{3.15}$$

where * means that those coefficients are solved from $1 = -\alpha_2\alpha_3 = \beta_1\beta_4 - \beta_2\beta_3$.

This extra technical condition $\gcd(c, e)|a$ can be assured for linear HSR via translations of the form $T = \begin{pmatrix} n & 0 \\ 0 & 0 \end{pmatrix}$. Then $e \mapsto e - na \in \mathbb{Z}\gcd(e, a)$, and by Dirichlet's theorem on

primes of arithmetic progressions, there exists a prime number p with $p > |c|$ such that $e - na = \gcd(e, a)p$. Under this translation

$$(a, b, c, 0, e) \mapsto (a, b, c, \underbrace{e - na}_{:=e'}),$$

Hence $\gcd(c, e') = \gcd(c, \gcd(e, a)p) | \gcd(e, a) | a$, because $p \not| c$ by construction.

In conclusion, the steps of the algorithm are as follows: starting with $l = (a, b, c, d, e)$ a Humbert singular relation,

1. Use a diagonal transformation as in Equation (3.14) to get $(a_1, b_1, c_1, 0, e_1)$.
2. Use a translation by $T = \begin{pmatrix} n & 0 \\ 0 & 0 \end{pmatrix}$ to transform it to $(a_1, b_1, c_1, 0, e_2)$ such that $\gcd(c_1, e_2) | a_1$.
3. Use a diagonal transformation as in Equation (3.15) to transform it to $(a_2, b_2, c_2, 0, 0)$.

□

Remark 3.18. *In some special cases, one could design a simpler algorithm that the one in the proof above, assuming that in Steps 1 and 3 one has solution to the corresponding Diophantine equation.*

Proof. By our proposition above, we can transform the HSR (a, b, c, d, e) using J and translations by a matrix T as follows.

1. Use a translations T_1 to get $(a_1, b_1, c_1, d, 0)$.
2. Use J and get $(-c_1, b_1, -a_1, 0, -d)$.
3. Use a translation T_2 to get $(a_2, b_2, c_2, 0, 0)$.

That relies of being able to solve for $n_1, n_2, n_3 \in \mathbb{Z}$ the equation

$$an_1 + bn_2 + cn_3 + d(n_1 n_3 - n_2^2) = e. \tag{3.16}$$

For $d = 0$ there always exists such (n_1, n_2, n_3) if and only if $\gcd(a, b, c) = 1$. For $d \neq 0$, we assume that such an integral solution exists, which is equivalent to the following. For readability, we change the notation to the variables to (x, y, z). We follow the general solution by Lagrange [Lag68] to the general quadratic binary Diophantine equation, see [SSW08, Introduction].[2] We can then rewrite as

$$\begin{aligned}
d(xz - y^2) + ax + by + cz &= e \\
x \underbrace{(dz + a)}_{:=z'} - dy^2 + by + cz &= e \\
dxz' - d^2y^2 + dby + cz' - ac &= de \\
z' \underbrace{(dx + c)}_{:=x'} - d^2y^2 + dby - ac &= de \\
z'x' &= dy \underbrace{(dy - b)}_{:=y'} + de + ac \\
z'x' &= (y' + b)y' + de + ac.
\end{aligned}$$

[2] we saw it here

The last equation has infinitely many integers solutions: for any given $y' \in \mathbb{Z}$, take (z', x') pair of divisors of $(y'+b)y' + de + ac$. To recover solutions in (x, y, z), we furthermore need to impose the congruences $x' \equiv c$, $y' \equiv -b$ and $z' \equiv a \bmod d$. Taking congruences mod d, it holds that $x'z' \equiv ac \pmod{d}$. □

3.1.2 The Néron-Severi group of an abelian surface

We finish this section with the relation between $\mathrm{End}^s(A)$ and the Néron-Severi group $\mathrm{NS}(A)$, the group of divisors modulo algebraic equivalence, or the image of the first Chern class $\mathrm{Pic}(A) \to H^2(A, \mathbb{Z})$. For a line bundle L on A and for any $x \in A$, the line bundle $t_x^* L \otimes L^{-1}$ has first Chern class zero, so it belongs to $\mathrm{Pic}^0(A) = \hat{A}$. This induces a map $\phi_L : A \to \hat{A}$ given by $x \mapsto t_x^* L \otimes L^{-1}$, that only depends on the first Chern class of L in $\mathrm{NS}(A)$. For a principal polarization H_0, it is moreover an isomorphism $\phi_0 : A \to \hat{A}$. We can consider the assignment:

$$\begin{aligned} \mathrm{NS}(A) &\to \mathrm{End}(A) \\ L &\mapsto \phi_0^{-1} \circ \phi_L. \end{aligned}$$

In fact, it is compatible with the Rosati involution, restricting to $\mathrm{NS}(A) \to \mathrm{End}^s(A)$. More can be said:

Proposition 3.19. *The assignment*

$$\begin{aligned} \mathrm{NS}(A) &\to \mathrm{End}^s(A) \\ L &\mapsto \phi_0^{-1} \circ \phi_L \end{aligned}$$

is a group isomorphism.

Proof. This is [BL04, Proposition 5.2.1]. □

This actually generalizes to an isomorphism of $\mathbb{Q}$-vector spaces for H_0 not necessarily principal. Conversely, we can think of $\mathrm{End}^s(A)$ as parameterizing other line bundles in A, in particular, other (principal) polarizations: by [BL04, Theorem 5.2.4], there is a bijection between the polarizations of degree d in $\mathrm{NS}(A)$ and the totally positive endomorphisms in $\mathrm{End}^s(A)$ of norm d.

3.2 The refined Humbert invariant

In Subsection 6.4.1, we will consider the positive definite lattice of Humbert singular relations. The definitions, as presented, only work for principally polarized abelian *surfaces*. As complementary to that, we present here a generalization for every genus and every characteristic. [3] This is known as the Humbert refined invariant and it is originally defined by Kani in the series of papers [Kan94], [Kan16] and [Kan19], but our presentation here will be more influenced by the reinterpretation in [GRV24, Section 2].

We will see in Subsection 6.4.1 that the lattice of singular relations $\mathcal{L}$ from Definition 6.18 will fit into a the following short exact sequence (Equation (6.4)

$$0 \to \mathbb{Z} \to \mathrm{End}^s(A_{\boldsymbol{\tau}}) \to \mathcal{L}_{\boldsymbol{\tau}} \to 0. \tag{3.17}$$

[3] However we only prove statements for characteristic zero.

By Proposition 3.19, for a principally polarized abelian variety, $\mathrm{NS}(A) \cong \mathrm{End}^s(A)$. Hence, it makes sense to define a generalization of the quadratic form Δ directly on either one of them.[4]

Kani's original definiton is defined on the Néron-Severi group of the abelian variety, and the quadratic form is defined in terms of intersection pairing of divisors. In [GRV24, Section 2], it is defined directly on $\mathrm{End}^s(A)$, and this is the approach we present here.

Consider $A \in \mathcal{A}_g$ a complex[5] principally polarized abelian variety of dimension g. Then for any $f \in \mathrm{End}_0(A)$, we can consider the characteristic polynomial of $\rho_r(f) \in \mathcal{M}_{2g}(\mathbb{Q})$

$$P_f^r(t) = \det(tI_{2g} - \rho_r(f)) \in \mathbb{Q}[t],$$

a rational polynomial of degree $2g$. Likewise, with the analytic representation, $\rho_a(f)$ and $P_f^a(t) = \det(tI_g - \rho_a(f)) \in \mathbb{C}[t]$. Remark that P^a is the polynomial we used in (3.7) above in the $g = 2$ case.

We furthermore define $\mathrm{tr}_r(f) = \mathrm{tr}(\rho_r(f))$ and $\mathrm{tr}_a(f) = \mathrm{tr}(\rho_a(f))$. Notice then that they are terms appearing in the $2g-1$ and $g-1$ term, respectively, in the corresponding characteristic polynomials.

We collect the following properties of these polynomials.

Proposition 3.20. *For any $f \in \mathrm{End}_0(A)$:*

- $P_f^r(t) = P_f^a(t)\overline{P_f^a(t)}$.
- *For $' : \mathrm{End}_0(A) \to \mathrm{End}_0(A)$ the Rosati involution, $P_{f'}^a(t) = \overline{P_f^a(t)}$.*
- *For every $n \in \mathbb{Z}$, $P_f^r(n) = \deg(n_A - f)$,[6] where $n_A \in \mathrm{End}(A)$ is multiplication-by-n endomorphism.*
- *If $f \in \mathrm{End}(A)$ then $P_f^r(t) \in \mathbb{Z}[t]$, and if $f \in \mathrm{End}_0^s(A)$ then $P_f^a(t)$ has real coefficients and $P_f^r(t) = (P_f^a(t))^2$.*
- $\mathrm{tr}_r(f) = 2\,\mathrm{Re}(\mathrm{tr}_a(f))$.
- *We have a positive definite symmetric bilinear form on $\mathrm{End}_0(A)$ given by $(f, g) \mapsto \mathrm{tr}_r(f'g)$.*

Proof. This is [BL04, Proposition 5.1.2, Corollary 5.1.3, Lemma 5.1.4, Theorem 5.1.8]. □

In particular, there is an integral positive definite quadratic form $\mathrm{End}^s(A) \to \mathbb{Z}$ given by $f \mapsto \mathrm{tr}_r(f^2)$. As what happens with our analysis of the Humbert singular relations, we would want to mod out the effect of the trivial symmetric endomorphisms $\mathbb{Z} \subset \mathrm{End}^s(A)$. For that, observe that for $n \in \mathbb{Z}$ $\mathrm{tr}_r(n) = 2gn$ and $\mathrm{tr}_r(n^2) = (\mathrm{tr}_r(n))^2$, so one can consider the following definition.

[4] more precisely, it descents to a quotient by a subgroup.

[5] the following construction does work over fields of any characteristic, though we only present the characteristic zero case here.

[6] This the more general definition of the rational characteristic polynomial in any characteristic.

Definition 3.21. ([Kan94, Section 3] and [GRV24, Definition 2.1]) *Consider (A, E) a principally polarized abelian surfaces and* $\mathrm{End}^s(A) \subset \mathrm{End}(A)$ *the subring of symmetric endomorphisms with respect to the Rosati involution. We define the* refined Humbert invariant

$$\begin{aligned} q_{A,E} : \mathrm{End}^s(A) &\to \mathbb{Z} \\ f &\mapsto \frac{1}{4}\left(2g\,\mathrm{tr}_r(f^2) - (\mathrm{tr}_r(f))^2\right) \end{aligned} \tag{3.18}$$

It takes integral values by [GRV24, Remark 2.7]. We gather the following results.

Proposition 3.22. *Let (A, E) be a complex ppas and $q = q_{A,E}$ its refined Humbert invariant as above. Then*

1. *For any $f \in \mathrm{End}^s(A)$, $q_{A,E}(f) = g\,\mathrm{tr}_a(f^2) - (\mathrm{tr}_a(f))^2$.*
2. *For any $n \in \mathbb{Z}$ and $f \in \mathrm{End}^s(A)$, $q(f + n_A) = q(f)$. Furthermore, $q(f) \geq 0$, $q(f) = 0$ if and only if $f \in \mathbb{Z}$. Hence, it induces a well defined positive definite quadratic form in the quotient*
$$\mathcal{L} := \mathrm{End}^s(A)/\mathbb{Z}$$
3. *If $g = 2$, it agrees with the positive quadratic form from Subsection 6.4.1, $\Delta(f) = \mathrm{Disc}(P_f^a) = (\mathrm{tr}(\rho_a(f))^2 - 4\det(\rho_a(f))$.*

Proof. For item 1), by Proposition 3.20, as $f \in \mathrm{End}^s(A)$ then $\mathrm{tr}_r(f) = 2\,\mathrm{tr}_a(f)$, and the formula follows. For item 2), $q(f + n) = q(f)$ is a simple computation using trace properties:

$$2g\,\mathrm{tr}_r((f+n)^2) - (\mathrm{tr}_r(f+n))^2 = 2g\left(\mathrm{tr}_r(f^2) + 2gn^2 + 2n\,\mathrm{tr}_r(f)\right) - (\mathrm{tr}_r(f) + 2gn)^2 = 2g\,\mathrm{tr}_r(f^2) - (\mathrm{tr}_r(f))^2.$$

The rest is is [GRV24, Remark 2.2]: setting $h := 2gf - \mathrm{tr}_r(f)$, then $\mathrm{tr}_r(h) = 0$ and

$$4g^2 q(f) = q(h) = \frac{g}{2}\,\mathrm{tr}_r(h^2) \geq 0,$$

where the last inequality comes from the quadratic form $\mathrm{tr}_r(f'g)$ being positive definite. Remark too that then $q(f) = 0$ if and only if $2gf - \mathrm{tr}_r(f) = 0$, and that it is equivalent to $f \in \mathbb{Z}$.

For item 3), it follows because $2\,\mathrm{tr}(A^2) - (\mathrm{tr}(A))^2 = (\mathrm{tr}(A))^2 - 4\det(A)$ holds for any matrix $A \in \mathrm{Mat}_2(\mathbb{C})$. Indeed, it is equivalent to $(\mathrm{tr}(A))^2 - \mathrm{tr}(A^2) = 2\det(A)$. As $A \in \mathrm{Mat}_2(\mathbb{C})$ satisfies $A^2 - \mathrm{tr}(A)A + \det(A)I_2 = 0$, one takes traces in $\det(A)I_2 = \mathrm{tr}(A)A - A^2$. □

3.3 Generalized Humbert varieties

Likewise, one can generalize the concept of Humbert surfaces of discriminant Δ in Definition 3.9 to a Humbert variety associated to a quadratic form. This definition comes from [Kan16, Section 3].

First recall the following notion for quadratic forms in several variables from [AB04, Definition 3.5 and Definition 3.12]

Definition 3.23. *Consider a f, g two quadratic forms over $\mathbb{Z}$ of n and r variables respectively, with $r \leq n$. Consider $M(f) \in \mathrm{Sym}_n(\mathbb{Z}[1/2])$, $M(g) \in \mathrm{Sym}_r(\mathbb{Z}[1/2])$ their corresponding matrix representation. We say that f* represents *g (over Z) and we write $f \to g$, if there exists $P \in \mathrm{Mat}_{n\times r}(\mathbb{Z})$ of full rank r, that we call a* representation, *such that*

$$^tPM(f)P = M(g).$$

If it holds that $\gcd\{det(M)\} = 1$ for M ranging over all $r \times t$ minors of P, we say that the representation is primitive. We say that f represents g primitively *if there exist one primitive representation, and we write $f \to^{pr} g$.*

We notice that specializing $r = 1$, we recover the usual notions of a quadratic form representing an integer N (primitively), identifying N with the unary quadratic form Nx^2.

Definition 3.24. *Let q a positive definite r-ary integral quadratic form $r = 1, 2, 3$. We define the following locus of A_2, that we call* Humbert variety *associated with q:*

$$\mathcal{H}(q) = \{(A, E) \in A_2 : q_{A,E} \text{ primitively represents } q\} \subset A_2,$$

where $q_{A,E}$ is the refined Humbert invariant associated to (A, E) as in Definition 3.21.

Equivalently [Kan16, pg.5 paragraph above Remark 8], $(A, E) \in \mathcal{H}(q)$ if and only if there exists an injective homomorphism $f : \mathbb{Z}^r \hookrightarrow \mathcal{L} = {}^{\mathrm{End}^s(A)}\!/_{\mathbb{Z}}$ with ${}^{\mathcal{L}}\!/_{\mathrm{Im}\, f}$ torsion free such that $q = q_{A,E} \circ f$.

Remark 3.25. *For unary quadratic forms, it follows that $\mathcal{H}(\Delta x^2) = \mathcal{H}_\Delta$. We also have $\mathcal{H}(q) \subset \mathcal{H}_\Delta$ if and only $q \to^{pr} \Delta$.*

It follows that $\dim \mathcal{H}(q) = 3-r$, so that is why we restrict $r = 1, 2, 3$ in the above definition. For ternary quadratic forms, it will follow from Theorem 6.28 that $\mathcal{H}(q)$ is a union of isotypic CM points,[7] and they have been studied recently in [Kir22] and [GRV24].

For binary quadratic forms, we will discuss them in detail in Subsection 4.1.2 and Section 4.2. Here, we just say that among the special curves of $\mathcal{A}_2$ as listed in [DO21, Table 1], their irreducible components correspond to Shimura and modular curves. We do not recover the third type $Q \times CM$, because again by Theorem 6.28, $\mathrm{rank}\, \mathcal{L} = 1$ generically on these curves, and the refined Humbert invariant is unary.

In [Kan19, Section 2] it is furthermore proven that this locus defines closed subsets of $\mathcal{A}_2$, so they are a finite union of curves, which can be either Shimura or modular curves.

Remark 3.26. *In general, $\mathcal{H}(q)$ is not irreducible.*

Describing Shimura and modular curves in $\mathcal{A}_2$ conjunctly as $\mathcal{H}(q)$ for suitable binary quadratic forms q allows us to describe an intersection of Humbert surfaces $\mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2}$, when $\Delta_1 \neq \Delta_2$. This problem was previously studied in [HM95] and [Run99], and more recently in [Gru08].

First, let us observe a natural restriction on the values represented by the quadratic forms.

Lemma 3.27. *Assume that q is a positive definite quadratic form such that $\mathcal{H}(q) \neq \emptyset$. Then for any $n \in \mathbb{Z}_{>0}$ such that $q \to n$, n is a discriminant, i.e. $n \equiv 0, 1 \pmod 4$. We shorthand this property by $f \equiv 0, 1 \pmod 4$.*

[7] here we mean that they correspond to an abelian surface that is CM isotypic.

Proof. It follows from the definition that $\mathcal{H}(q) \subset \mathcal{H}_\Delta$ if and only if q represents Δ primitively, or $q \to^{pr} \Delta$. Hence, as $\mathcal{H}_\Delta \neq \emptyset$ if and only if $\Delta > 0$ and Δ is a discriminant (*i.e.* $\Delta \equiv 0, 1 \pmod 4$) by Lemma 3.7, it follows that $H(q) \neq \emptyset$ implies that the integers primitively represented by q are discriminants, and the conclusion follows. □

Proposition 3.28. *For $\Delta_1 \neq \Delta_2 \in \mathbb{Z}_{>0}$ discriminants, it follows*

$$\mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2} = \bigcup_{\substack{q \to^{pr} \Delta_1 \\ q \to^{pr} \Delta_2}} \mathcal{H}(q),$$

when q ranges through the finite union of $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of positive definite binary quadratic forms q with $q \equiv 0, 1 \pmod 4$ that primitively represent both Δ_1 and Δ_2.

Proof. This is [Kan19, Proposition 12] and by [Kan19, Remark 13], the discriminant of a such a quadratic form q is bounded by $4\Delta_1\Delta_2$. □

Notice that such an intersection has a priori Shimura and modular curves. The rest of the work of [Kan19] focuses in the case of $\Delta_1 = N^2$, an in that case necessarily the intersection only has modular curves, for that it is contained in $\mathcal{H}_{N^2}$, and the generic points of Shimura curves correspond, in particular, to simple surfaces (generically). We will resume in Subsection 4.2.

In [Run99] the case of general Δ_i is considered and the following results are obtained.

Proposition 3.29. *Assume $\mathcal{H}(q) \neq \emptyset$ and q in the conditions of Theorem 4.18 1. Then $\mathcal{H}(q) \subset \mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2}$ if and only if q is in the $\mathrm{GL}_2(\mathbb{Z})$-equivalence class of a form given by $\Delta_1 x^2 + *xy + \Delta_2 y^2$.*

If q is primitive, then $\mathcal{H}(q)$ is irreducible. Furthermore, $\mathcal{H}(q)$ admits a "Humbert's presentation" in $\mathbb{H}^2$ by the intersection of the following Humbert singular relations

$$l_1 = (k, l, -1, 0, 0) : k\tau_1 + l\tau_2 - \tau_3 = 0,$$
$$l_2 = (a, t, 0, c, 1) : a\tau_1 + t\tau_2 + c(\tau_2^2 - \tau_1\tau_3) + 1 = 0,$$

where l_1 is the normalized Humbert singular relation for Δ_1 and a, t, c satisfy $\Delta_2 = t^2 - 4c$.

In particular, if $\gcd(\Delta_1, \Delta_2) = 1$ then $\mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2}$ contains all $\mathcal{H}(\Delta_1 x^2 + axy + \Delta_2 y^2)$ for $0 \leq a$ such that $a^2 \leq \Delta_1\Delta_2$, and they are irreducible.

Proof. This [Run99, Corollary 9, Theorem 10, Corollary 15] □

3.3.1 An alternative proof of Appendix 6.5.3

We give now an alternative proof of Appendix 6.5.3.

Theorem 3.30. *Let $A \in \mathcal{A}_2$ and assume it has a normalized period matrix $\boldsymbol{\tau} \in \mathbb{H}_2$ such that $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}^{lin} = 2$. Then $\mathrm{End}^0(A) \cong \mathrm{Mat}_2(\mathbb{Q})$.*

Proof. Alternatively, if $A \in \mathcal{A}_2$ is a generic point in a Shimura curve, then $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}^{lin} = 1$ for any period matrix $\boldsymbol{\tau}$ of A.

The following suggestion comes originally from Daniel Bertand. Observe first that a Shimura curve in $\mathcal{A}_2$ is compact. Consider the sequence of matrices $\{n\boldsymbol{\tau}\}_{n \geq 0}$. They define (n, n)-isogenous abelian surfaces to $A_{n\boldsymbol{\tau}}$ and they tend to the boundary of $\mathcal{A}_2$. Hence, they eventually leave the Shimura curve.

Consider l_1 and l_2 two *linear* and primitive Humbert singular relations that generate $\mathcal{L}_{\boldsymbol{\tau}}^{lin}$. Write $l_i = (a_i, b_i, c_i, 0, e_i) \in \mathbb{Z}$ so that

$$a_i\tau_1 + b_i\tau_2 + c_i\tau_3 + e_i = 0,$$

and consider too $\Delta_i := \Delta(l_i) = (b_i)^2 - 4a_ic_i$ for $i = 1, 2$. It then follows that $A_{\boldsymbol{\tau}} \in \mathcal{H}_{\Delta_1} \cup \mathcal{H}_{\Delta_2}$.

Consider $n\boldsymbol{\tau}$ for $n \in \mathbb{Z}_{>0}$. It then solves

$$a_i(n\tau_1) + b_i(n\tau_2) + c_i(n\tau_3) + ne_i = 0,$$

which has the same discriminant Δ_i and for almost all n, $\gcd(a_i, b_i, c_i, ne_i) = 1$. In other words, as $\boldsymbol{\tau}$ solves a *linear* HSR, then $n\boldsymbol{\tau}$ parametrizes an isogenous variety $A_{n\boldsymbol{\tau}}$ which stays in the same Humbert surface. Therefore, the sequence $\{A_{n\boldsymbol{\tau}}\}_{n\in\mathbb{Z}_{>0}} \subset \mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2}$.

But by [Kan19, Proposition 12] above, such an intersection is a *finite* union of Shimura and modular curves. But if $\mathrm{End}^0(A_{\boldsymbol{\tau}})$ is a division algebra, as End^0 is invariant by isogeny, the same is true for $\mathrm{End}^0(A_{n\boldsymbol{\tau}})$. This implies that if $A_{\boldsymbol{\tau}}$ belongs to a Shimura curve, the sequence $\{A_{n\boldsymbol{\tau}}\}_{n\in\mathbb{Z}_{>0}}$ stays in the union of the Shimura curves in $\mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2}$, but this is a finite union of compact curves, while the sequence $\{A_{n\boldsymbol{\tau}}\}_{n\in\mathbb{Z}_{>0}}$ goes to the boundary of $\mathcal{A}_2$. This is a contradiction, hence $A_{\boldsymbol{\tau}}$ was in a modular curve, as those are the non-compact ones. □

Chapter 4

Shimura and modular curves in the Siegel threefold

This chapter has two different sections. In the first section, we present the construction of the Shimura curves as solutions to the moduli problem of quaternionic multiplication.[1] We present the problem of mapping them into $\mathcal{A}_2$, with a solution given by [Has95], and the complications derived with the non-existence of a canonical embedding. We reformulate this problem as in [LY20] and [GY19], who associated $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of positive definite quadratic forms to each embedding, or to each connected component of the quaternionic loci in $\mathcal{A}_2$, and derive formulas for counting them in terms of class numbers.

In this constructions, the classical modular curves $X_0(N)$ are realized as Shimura curves associated with $D = 1$. In the second section, we present a solution in [Kan16] to a different problem, the moduli problem of Jacobians isomorphic to products of elliptic curves, that describes the same loci as the quaternionic loci, without the restriction of square-free level on $X_0(N)$. Finally, we recall the language of generalized Humbert varieties from the end of Chapter 3, and translate the previous results in those terms.

4.1 Quaternionic multiplication: moduli space and quaternion modular embeddings

In this section we recall the classical construction (going back to Shimura in [Shi63]) of Shimura curves as a solution to a moduli space problem. We present the whole proof of this construction in the case of *hereditary* orders (subcase of Eichler orders) in a indefinite quaternion algebra over $\mathbb{Q}$. We then state the rest of the theory in this context, though it is known to experts that it is generalizable to Eichler orders.

We remark that there are other modular interpretations to the Shimura curves $X_0^D(N)$ ([AB04, Section 2.4]), that mimic the modular interpretation of the classical modular curves $X_0(N)$ (the data of an elliptic curve with a cyclic subgroup of N-torsion points of order N, or a cyclic isogeny of degree N). We do not use that modular interpretation, but for a comparison of the moduli problems, see [Cla03, Section 0.3.2, Proposition 53].

[1] On higher level N, the moduli problem is not quite the standard one associated with the modular curves $X_0^D(N)$

We follow [Voi21, Section 43.6] and [LY20, Section 2]. As the construction is fairly technical, we do not present all the details here, but we intent to point out the steps where the hereditary condition is used. In general terms, it is on the one hand to guarantee the polarization to be *principal*, and to have surjectivity of the asginment.

Let B a quaternion algebra over $\mathbb{Q}$ of discriminant D. We say that B is *indefinite* if it does not ramify at the infinity place, i.e. $B_\infty := B \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathrm{Mat}_2(\mathbb{R})$. Remark that the discriminant D of a indefinite quaternion algebra is a product of an *even* number of primes,[2] and that $D = 1$ if and only if $B \cong \mathrm{Mat}_2(\mathbb{Q})$.

Definition 4.1. *Let B an indefinite quaternion algebra over $\mathbb{Q}$ with discriminant $D \in \mathbb{Z}_{>0}$. We say that $\mathcal{O} \subset B$ is an* Eichler *order if it is the intersection of two maximal orders.*

They admit a local description ("locally maximal at every prime except for finitely many"), and we have the following equivalent characterization of Eichler orders.

Proposition 4.2. *An Eichler order $\mathcal{O} \subset B$ as above is completely determined by its* level *$N \in \mathbb{Z}$ with $(N, D) = 1$. They admit the local description:*

- *for every $p^e || N$, $\mathcal{O}_p$ is isomorphic to*

$$\begin{pmatrix} \mathbb{Z}_p & \mathbb{Z}_p \\ p^e\mathbb{Z}_p & \mathbb{Z}_p \end{pmatrix} \subset \mathrm{Mat}_2(\mathbb{Z}_p);$$

- *for every $p|D$, $\mathcal{O}_p$ is the unique discrete valuation ring in the division algebra B_p;*
- *for every other prime p, $\mathcal{O}_p$ is isomorphic to* $\mathrm{Mat}_2(\mathbb{Z}_p)$.

Proof. The local characterization follows from [Voi21, Section 23.4]. For Eichler orders, local isomorphisms lift to a global isomorphism by [Voi21, Theorem 28.2.11],[3] which are given by conjugation by a suitable element. Therefore, an Eichler order is characterized, up to conjugation, by its level. □

We focus on the Eichler orders such that the level is *square-free*. Those are furthermore *hereditary* orders, by [Voi21, §23.3.1 v)].

Definition 4.3. *In the same context as above, an order $\mathcal{O} \subset B$ is called* hereditary *if every (left/right) ideal $I \subset \mathcal{O}$ is projective as a (left/right) $\mathcal{O}$-module. Equivalently, if every (left/right) fractional $\mathcal{O}$-ideal is invertible.*

It can be proven that the distinction left/right is unnecessary ([Voi21, §21.4.2]). As B is an indefinite quaternion algebra, $B_\infty := B \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathrm{Mat}_2(\mathbb{R})$, we may fix an embedding

$$\iota : B \hookrightarrow \mathrm{Mat}_2(\mathbb{R}).$$

[2] A quaternion algebra over $\mathbb{Q}$ ramifies at a finite *even* number of places, and its discriminant coincides with the product of the (finite) primes where it ramifies.

[3] Eichler orders in *indefinite* quaternion algebras have *class* and *type* number 1.

Remark 4.4. *Such an embedding is unique up to conjugation by a matrix in* $\mathrm{GL}_2(\mathbb{R})$ *by the Skolem-Noether theorem (see [Voi21, Theorem 7.7.1]), so we could even set it as the following one. Assume* B *is given with the description* $\left(\frac{a,b}{\mathbb{Q}}\right)$ *as a quaternion* $\mathbb{Q}$*-algebra. Then* B *is indefinite if and only if either* a *or* b *are positive by [Voi21, Exercise 2.4], so we assume* $a > 0$*. Then we have the* $\mathbb{Q}$*-algebra embedding* $B \to \mathrm{Mat}_2(\mathbb{Q}(\sqrt{a})) \subset \mathrm{Mat}_2(\mathbb{R})$ *given by*

$$x_1 + x_2 I + x_3 J + x_4 IJ \mapsto \begin{pmatrix} x_1 + x_2\sqrt{a} & x_3 + x_4\sqrt{a} \\ b(x_3 - x_4\sqrt{a}) & x_1 - x_2\sqrt{a} \end{pmatrix}.$$

Consider the subgroup $\mathcal{O}^1$ of elements on quaternionic norm 1, $\iota(\mathcal{O}^1)$ as a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$, and set[4]

$$\Gamma := {}^{\iota(\mathcal{O}^1)}\!/_{\{\pm I_2\}} \subset \mathrm{PSL}_2(\mathbb{R}). \tag{4.1}$$

4.1.1 Principally polarized abelian surfaces with QM

Consider first a principally polarized abelian surface A with $\mathrm{End}(A) \cong \mathcal{O}$. The polarization induces a Rosati involution $'$ on $\mathrm{End}(A)$, that extends to $\mathrm{End}_0(A) := \mathrm{End}(A) \otimes_{\mathbb{Z}} \mathbb{Q} \cong D$. Hence via ι and extending scalars to $\mathbb{R}$, the Rosati involution $'$ induces a positive involution on $D \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathrm{Mat}_2(\mathbb{R})$. The positive involutions in $\mathrm{Mat}_2(\mathbb{R})$ are characterized as follows.

Lemma 4.5. *All positive involutions on* $\mathrm{Mat}_2(\mathbb{R})$ *are given by* $\mu \in \mathrm{GL}_2(\mathbb{R})$ *such that* $\mu^2 < 0$ [5] *via*

$$f \mapsto f' = \mu^{-1} \bar{f} \mu,$$

where $f \to \bar{f}$ *is the standard involution on* $\mathrm{Mat}_2(\mathbb{R})$ *given by adjugation of matrices. Such an element* μ *is unique up to scalar multiple.*

Proof. Let us observe first that one positive involution in $\mathrm{Mat}_2(\mathbb{R})$ is given by $A \mapsto {}^tA$. In terms of the standard involution, it gets written as

$$A \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1} \overline{A} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

By an application of the Skolem-Noether theorem (see [Voi21, Proposition 8.4.7 and Example 8.4.15]), any other positive involution in $M_2(\mathbb{R})$ has the form $A \mapsto \mu^{-1}\, {}^tA\mu$ with $\mu \in \mathrm{Sym}_2(\mathbb{R})$ positive definite. Suppose that there is another ν such that $\mu^{-1}\, {}^tA\mu = \nu^{-1}\, {}^tA\nu$, which implies that $(\mu\nu^{-1})^{-1}\, {}^tA(\mu\nu^{-1}) = {}^tA$, or $(\mu\nu^{-1})$ is a central element in $\mathrm{Mat}_2(\mathbb{R})$, so it is a scalar.

Note that μ positive definite implies $\mu = {}^t\mu$ and $\overline{\mu}\mu = \det(\mu) I_2$ with $\det(\mu) > 0$. Hence we can rewrite them to $A \mapsto \mu^{-1} j^{-1} \overline{A} j \mu$, with $j = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{-1}$. Now $\mu = {}^t\mu = j^{-1}\overline{\mu} j = -j\mu j$, hence $\det(\mu) I_2 = \overline{\mu}\mu = j\mu j^{-1}\mu = -(j\mu)^2$. Identifying $\mathbb{R}$ with $\mathbb{R} I_2$ in $\mathrm{Mat}_2(\mathbb{R})$, note that $(j\mu)^2 = -\det \mu I_2 < 0$, as we wanted. □

As our positive involution in $B \otimes \mathbb{R}$ was induced by a positive involution in $\mathrm{End}(A)$, the element μ which parametrizes it also belongs to $\mathrm{End}(A)$. Remark that $\mu^2 < 0$ is equivalent to

[4] Up to conjugation, Γ only depends on D and N.

[5] via the characteristic polynomial, equivalent to $\mathrm{tr}(\mu) = 0$ and $\det(\mu) > 0$.

$\mathrm{tr}(\mu) = 0$ and $n(\mu) > 0$, where tr, $\bar{\cdot}$ and n are the quaternionic trace, conjugation and norm, respectively, induced by $\mathrm{End}(A) \cong \mathcal{O}$.

We will see (the discussion right below together with Remark 4.11) that the fact that the polarization of A is *principal* forces

$$\begin{aligned} &\mu^2 = -DN, \\ &\mu^{-1} \in \mathrm{End}(A)^\vee := \{\alpha \in \mathrm{End}(A) |\ \mathrm{tr}(\alpha\overline{\beta}) \in \mathbb{Z} \text{ for all } \beta \in \mathrm{End}(A)\}. \end{aligned} \tag{4.2}$$

On the other direction, let us consider $\mathcal{O} \subset B$ an hereditary order of level N (square-free) and D the discriminant of B. For $\tau \in \mathbb{H}$, set $v_\tau = \begin{pmatrix} \tau \\ 1 \end{pmatrix}$. To construct principally polarized abelian surfaces with QM by $\mathcal{O}$, first we assign the complex torus. Defined the lattice in $\mathbb{C}^2$

$$\Lambda_\tau := \iota(\mathcal{O})v_t \subset \mathbb{C}^2.$$

As $\mathrm{rank}_{\mathbb{Z}} \mathcal{O} = \dim_{\mathbb{Q}} B = 4$, we have that Λ_τ is a lattice of the correct rank, and $A_\tau = {}^{\mathbb{C}^2}\!/_{\Lambda_\tau}$ is a complex torus of dimension two.

For the polarization on A_τ assume $\mu \in \mathcal{O}$ such that $\mu^2 = -DN$ and $\mu^{-1} \in \mathcal{O}^\vee$, as in Equation (4.2) above, and define

$$\begin{aligned} E_\tau : \Lambda_\tau \times \Lambda_\tau &\to \mathbb{Z} \\ (\iota(\alpha)v_\tau, \iota(\beta)v_\tau) &\mapsto \mathrm{tr}(\mu^{-1}\alpha\overline{\beta}). \end{aligned}$$

Observe that μ (more precisely μ^{-1}) is recovered as setting $\beta = 1$ above: it is the unique element (via non-degeneracy of the trace) on $\mathcal{O}^\vee$ such that $E_\tau(\iota(\alpha)v_\tau, v_\tau) = \mathrm{tr}(\mu^{-1}\alpha)$.

It is a technical computation to verify that E_τ is a Riemann form ([Voi21, Lemma 43.6.16.]). There is a priori a sign ambiguity (either E_τ or $-E_\tau$ induces a Hermitian positive definite form) and by [LY20, Lemma 3] it is resolved by setting

$$\iota(\mu) = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}), \text{ with } c > 0.$$

We do single out the computation of E_τ inducing a *principal* polarization in A_τ ([Voi21, Lemma 43.6.22.]) for hereditary orders, *i.e.* that E is unimodular: for α_i $i = 1, \dots, 4$ an integral basis of $\mathcal{O}$,[6] the determinant of the following matrix is one:

$$\begin{aligned} &\det((E(\iota(\alpha_i)v_\tau, \iota(\alpha_j)v_t)_{i,j}) = \det((\mathrm{tr}(\mu^{-1}\alpha_i\overline{\alpha_j})_{i,j}) \\ &= \det(L_{\mu^{-1}}) \det(\mathrm{tr}(\alpha_i\overline{\alpha_j})_{i,j}) \text{ where } L_x(y) = xy \\ &= n(\mu^{-1})^2 \underbrace{\det(\mathrm{tr}(\alpha_i\overline{\alpha_j})_{i,j})}_{=:\mathrm{disc}(\mathcal{O})} \text{ as } \det L_x = (n(x))^2 \text{ [Voi21, Remark 3.3.8]} \\ &= \frac{1}{(DN)^2}(DN)^2 = 1 \text{ as } \mathrm{disc}(\mathcal{O}) = (DN)^2 \text{ by [Voi21, Equation (15.1.2), §23.4.19]}. \end{aligned}$$

Finally, by construction we have that Λ_τ is an $\mathcal{O}$-module (via ι), hence we have a ring embedding $\mathcal{O} \to \mathrm{End}(A_\tau)$. Furthermore, μ was chosen at the beginning, by Lemma 4.5, so

[6] We can assume $\alpha_1 = 1$ and that α_i, $i \neq 1$ has trace zero, so that $-\det(\mathrm{tr}(x_i x_j)_{i,j}) = \det(\mathrm{tr}(x_i\overline{x_j})_{i,j})$.

that (A_τ, E_τ) satisfies the following compatibility condition with respect to μ:

$$\begin{array}{ccc} B & \xrightarrow{\iota_\tau} & \mathrm{End}_0(A_\tau) \\ \downarrow{\scriptstyle f^*=\mu^{-1}\bar{f}\mu} & & \downarrow{\scriptstyle \prime} \\ B & \xrightarrow{\iota_\tau} & \mathrm{End}_0(A_\tau) \end{array}$$

Compare with the notion of a ppas with RM from Definition 2.1: there is also a compatibility condition imposed regarding the Rosati involution. In that case, for the only positive involution in a real quadratic field is the identity, the Rosati involution has to restrict to the identity in the image of $\mathcal{O}_K$ (alternatively, the ring embedding maps to the symmetric subring $\mathrm{End}^s(A)$). In the case of quaternionic multiplication, the compatibility condition is agreeing with a fixed positive involution in $\mathcal{O}$, which it is not unique. Hence one could be more detailed in the definition and consider ppas with QM by $(\mathcal{O}, ^*)$, for * a fixed positive involution in $\mathcal{O}$.

Definition 4.6 ([Rot04b])**.** *Consider $\mathcal{O} \subset B$ a hereditary order of level N in a indefinite quaternion algebra over $\mathbb{Q}$ with discriminant D, and consider μ as in Equation (4.2) above. Consider a ppas (A, E). We say that (A, E, ι) has QM by $(\mathcal{O}, \mu)$ if there exists a ring embedding $\iota : \mathcal{O} \hookrightarrow \mathrm{End}(A)$, such that the restriction of the Rosati involution to $\mathcal{O}$ agrees with $f \mapsto \mu^{-1}\overline{f}\mu$.*

We say that (A, E, ι) and (A', E', ι') are isomorphic as $(\mathcal{O}, \mu)$-ppas if there is an isomorphism $\phi : (A, E) \to (A', E')$ such that ι and ι' are compatible, meaning that for $\phi^* : \mathrm{End}_0(A') \to \mathrm{End}_0(A)$, we have $\phi^* \circ \iota' = \iota$.

We state the following main result of this section.

Theorem 4.7. *Consider the moduli space $X_0^D(N) := \{(A, E)$ principally polarized complex abelian surfaces with QM by $(\mathcal{O}, \mu)\}$, up $(\mathcal{O}, \mu)$-isomorphism, and Γ as in Equation (4.1). The assignment*

$$\begin{aligned} \mathbb{H} &\to X_0^D(N) \\ \tau &\mapsto (A_\tau, E_\tau) \end{aligned}$$

induces a bijection $\Gamma\backslash\mathbb{H} \to X_0^D(N)$.

Remark 4.8. *When $D = 1$, the Eichler orders of level N are all isomorphic to $\begin{pmatrix} \mathbb{Z} & \mathbb{Z} \\ N\mathbb{Z} & \mathbb{Z} \end{pmatrix}$, and then Γ is precisely the Hecke congruence subgroup $\Gamma_0(N)$, hence for $D = 1$, on the analytic side one recovers the classical modular curves, with a different modular interpretation. We will discuss that in more detail in Section 4.2.*

Remark 4.9. *Different choices of μ yield abstract isomorphisms $X_0^D(N)(\mu) \cong X_0^D(N)(\mu')$ so they are dropped from the notation. But there is no* canonical *choice of μ. If one would want to consider the forgetful functor*

$$X_0^D(N) \to \mathcal{A}_2,$$

different choices of μ give rise to different mappings, with different images in $\mathcal{A}_2$.

They are not an embedding either, as the notion of isomorphism in the moduli space $X_0^D(N)$ is finer than the one in $\mathcal{A}_2$, for that it requires to respect the distinguished element μ. In

particular, those mappings $X_0^D(N) \to \mathcal{A}_2$ factor through the quotient by a subgroup of the group of Atkin-Lehner involutions (see [GY19, Theorem A (4)], [Rot04a, Corollary 4.5] for maximal orders). This subgroup is generically of order two, generated by one involution, and it can be of order four, with a second involution called a twisting involution *if the quaternion algebra B admits* twisting *elements ([Rot04a, definition after Proposition 4.1.]).*

Finally, these quotients of $X_0^D(N)$ are birational to its image in $\mathcal{A}_2$ ([Rot04b, Theorem 3.5] for maximal orders), being the Heegner points in the Shimura curve the obstruction.

Proof of Theorem 4.7. We have to prove the following:

1. Well-definedness and injectivity.
2. Existence of $\mu \in \mathcal{O}$ with the desired properties, namely
   a) $\mu^2 = -DN$,
   b) $\mu^{-1} \in \mathcal{O}^\vee$,
   c) $c > 0$ in $\iota(\mu) = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$
3. Surjectivity.

The last two items are those that explicitly use properties of μ. For 2a, it is enough to show that there is an embedding of orders[7] $\phi : \mathbb{Z}_K \hookrightarrow \mathcal{O}$, for $K = \mathbb{Q}(\sqrt{-DN})$ (remark that N is square free as we assumed $\mathcal{O}$ hereditary) and $\mathbb{Z}_K$ its ring of integers. We can prove more: such an embedding exists and it is *optimal* (*i.e.* if we extend ϕ to a $\mathbb{Q}$-algebra embedding $K \to B$, then $\phi(K) \cap \mathcal{O} = \phi(\mathbb{Z}_K)$). The theory of optimal embeddings allows us to count the number of such embeddings up to conjugation by elements in $\mathcal{O}^\times$. More precisely, by [Voi21, Example 30.7.5], together with the fact that as $\mathcal{O}$ is Eichler order in an *indefinite* quaternion algebra over $\mathbb{Q}$,[8] we have the following formula for the number of embeddings in terms of the local embedding numbers:

$$\begin{aligned} &\#\{\mathcal{O}^\times\text{-conjugacy classes of optimal } \mathbb{Z}_K \hookrightarrow \mathcal{O}\} \\ &= h(\mathbb{Z}_K) \prod_{p|D} \left(1 - \left(\frac{K}{p}\right)\right) \prod_{p|N} \left(1 + \left(\frac{K}{p}\right)\right), \end{aligned} \tag{4.3}$$

where $h(\mathbb{Z}_K)$ is the usual class number, and the splitting symbol is defined by

$$\left(\frac{K}{p}\right) := \begin{cases} 1, & \text{if } K_p \simeq \mathbb{Q}_p \times \mathbb{Q}_p \text{ is split;} \\ 0, & \text{if } K_p \supseteq \mathbb{Q}_p \text{ is a ramified field extension;} \\ -1 & \text{if } K_p \supseteq \mathbb{Q}_p \text{ is an unramified field extension.} \end{cases}$$

Then it is enough to show that the right hand side in Equation (4.3) does not vanish. Assume by contradiction that:

- For some $p|D$, $\mathbb{Q}_p(\sqrt{-DN}) \simeq \mathbb{Q}_p \times \mathbb{Q}_p$, but that cannot happen for $\mathbb{Q}_p(\sqrt{-DN})$ is a field.[9]

[7] only in this section we will use this notation for the ring of integers of a number field.
[8] then the class number of $\mathcal{O}$ is one and the sum in the formula is just one term.
[9] there is no square root of r in $\mathbb{Q}_p$ if $v_p(r) = 1$.

- For some $p|N$, $\mathbb{Q}_p(\sqrt{-DN}) \supseteq \mathbb{Q}_p$ is *unramified*. This cannot happen, as for every p, there is a unique unramified quadratic extension of $\mathbb{Q}_p$ and it is known ([Voi21, §13.1 under Equation (13.1.2)]) that such extension is given by $\mathbb{Q}_2(\sqrt{-3})$ for $p = 2$ and otherwise by $\mathbb{Q}_p(\sqrt{e})$ for e a quadratic non-residue modulo p. But DN is not a quadratic non-residue modulo p, as $p|N$.

Therefore, the number of embeddings in Equation (4.3) is non-zero, so there exists $\mu \in \mathcal{O}$ such that $\mu^2 = -DN$.

It is a consequence of $\mu^2 = -DN$ that $\mu^{-1} \in \mathcal{O}^\vee$ [Voi21, Lemma 43.6.7]. Finally, we require a choice of sign so that the Hermitian form is positive definite, and that choice of sign is checked in [LY20, Lemma 3], by evaluating the Hermitian form at one carefully chosen element.

For surjectivity of the assignment, assume (A, E) a principally polarized abelian surface with QM by $(\mathcal{O}, \mu)$. Consider the complex torus $A = \mathbb{C}^2/\Lambda$. It then follows that Λ has a natural left $\mathcal{O}$-module structure, and likewise $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$ is a B-module. As $\mathbb{Q}$-vector spaces $\dim_{\mathbb{Q}}(\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}) = \operatorname{rank}_{\mathbb{Z}} \Lambda = 4 = \dim_{\mathbb{Q}}(B)$, so $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$ is a B-module of rank one. If B is a division algebra (equivalently $D > 1$) then $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$ is a free B-module: there exists $w \in \mathbb{C}^2$ such that $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q} = Bw$, and so $\Lambda = Iw$ for some left $\mathcal{O}$-ideal (a lattice $I \subset B$ with left order containing $\mathcal{O}$). In the case of $B = \mathrm{Mat}_2(\mathbb{Q})$ the same conclusion holds using that B is a simple $\mathbb{Q}$-algebra.

We need more of I, we need it to be *principal* as an $\mathcal{O}$-ideal. As $\mathcal{O}$ is hereditary, I is invertible, and by [Voi21, Main Theorem 16.1.3], this is equivalent to being locally principal. In turn, as $\mathcal{O}$ is Eichler in an indefinite quaternion algebra, it has class number one by [Voi21, Theorem 28.2.11], so I is a principal $\mathcal{O}$-ideal. Changing w if necessary, we have that $\Lambda = \iota(\mathcal{O})w$ for $w \in \mathbb{C}^2$.

We have Λ in the form that we need, except for having $w = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ in the form $\begin{pmatrix} \tau \\ 1 \end{pmatrix}$ for $\tau \in \mathbb{H}$, but that comes from the theory of elementary divisors for the period lattice of a polarized abelian variety. The last thing to check it is that the given principal polarization E agrees with the one induced by E_τ. In this context, it all stems form the fact that they induced Rosati involutions are prescribed to agree on $\mathcal{O}$ ([Voi21, Lemma 43.6.23]).

First, to check that E_τ induces a Rosati involution that agrees with $f \mapsto \mu^{-1}\bar{f}\mu$ on $\mathcal{O}$, remark that by definition of the Rosati involution, for $f \in \mathrm{End}(A_{\boldsymbol{\tau}})$, f' is uniquely determined by the adjoint condition $E_\tau(x, fy) = E_\tau(f'x, y)$, for all $x, y \in \Lambda_\tau$. Hence, it is enough to check that for $\iota(f) \in \iota_\tau(\mathcal{O})$, and $\alpha, \beta \in \mathcal{O}$,

$$E_\tau(\iota(\alpha)v_\tau, \iota f \iota(\beta)v_\tau) = E_\tau(\iota(\mu^{-1}\bar{f}\mu)\iota(\alpha)v_\tau, \iota(\beta)v_\tau),$$

and that is a straightforward computation unravelling the definitions.

Second, there is a unique principal polarization compatible with μ: for another such $\tilde{E}$, compatibility with μ implies in particular that (the extension to $\mathbb{R}$-scalars of) the induced Rosati involution restricts to[10] $\iota(B) \subset \mathrm{End}_0(A_\tau)$, and as $\iota(B) \cong \mathrm{Mat}_2(\mathbb{R})$ can argue as the beginning, as in Lemma 4.5, there will be $\nu \in \mathcal{O}$ with such that the Rosati involution induced by $\tilde{E}$ agrees with $f \mapsto \nu^{-1}\bar{f}\nu$. But by Lemma 4.5, it follows that $\mu = \lambda\nu$ for some scalar λ, which implies that $E_\tau = \lambda\tilde{E}$, but as both E_τ and $\tilde{E}$ are principal, necessarily $\mu = \nu$. □

[10] this is automatic in the generic case $\mathrm{End}_0(A) \cong B$.

Remark 4.10. *Observe that in the generic case* $\mathrm{End}(A) \cong \mathcal{O}$ *(equivalently, when* A *is simple), by our argument above Lemma 4.5, there always exists* $\mu \in \mathcal{O}$ *such that* A *has QM by* $(\mathcal{O}, \mu)$*. However, it can happen, when* $\mathrm{End}_0(A)$ *is strictly larger than* B *(i.e. when* A *is CM isotypical) that* A *is principally polarized, we have an embedding* $\mathcal{O} \hookrightarrow \mathrm{End}(A)$*, but that* A *does not have QM by any* $\mu \in \mathcal{O}$*. That comes from the fact that there are more positive involutions of* $\mathrm{End}_0(A) \otimes \mathbb{R} \cong \mathrm{Mat}_2(\mathbb{C})$*, and that they do not leave* $B \otimes \mathbb{R} \cong \mathrm{Mat}_2(\mathbb{R}) \subset \mathrm{Mat}_2(\mathbb{C})$ *invariant. For an example, see [LY20, Remark 1, Example 32].*

Remark 4.11. *As it happens analogously with abelian surfaces with real multiplication, a complex torus* $\mathbb{C}^2/\Lambda$ *with "enough" endomorphisms (as in* $\mathbb{Q}$*-algebra embedding* $B \to \mathrm{End}(\mathbb{C}^2/\Lambda) \otimes \mathbb{Q}$*) admits automatically a polarization. We cannot guarantee a* principal *polarization for any given order* $\mathcal{O} \subset B$*, but we can for hereditary ones. Furthermore, if the Rosati involution is determined on* $\mathcal{O}$*, then the principal polarization is unique.*

Recall that in the case of real multiplication, the birational map from Theorem 2.23 from the symmetric Hilbert modular surface to the Humbert surface is indeed canonical, and by Corollary 3.13, the Humbert surface is irreducible. We notice that the difference with the QM case comes from the fact that the Rosati involution always acts as the identity of the totally real fields embedded in the endomorphism algebra of a polarized abelian variety. Therefore, the only positive involution to consider on a totally real field is the identity.

4.1.2 Quaternion modular embeddings

As said above, there are mappings from the Shimura curves $X_0^D(N) \to \mathcal{A}_2$ (depending on the distinguished element μ). In [Has95], they are written down explicitly as embeddings $\mathbb{H} \to \mathbb{H}_2$.

Remark 4.12. *These maps are called "quaternion modular embeddings" in classical literature and we keep it for historical reasons and for coherence with the Hilbert case, but they do not seem to be used in recent literature. As with the morphisms from Section 2.4, and recall Remark 2.22, they are not embeddings of* $X_0^D(N)$ *into* A_2*: the factor through an extension of* Γ *by (one or two) Atkin-Lehner involution, and on that quotient they induced a birational morphism into the image.*

The dependence on μ was translated as follows: $B = \mathbb{Q}1 + \mathbb{Q}I + \mathbb{Q}J + \mathbb{Q}IJ$ was recovered as a quaternion algebra as $\left(\frac{-DN,p}{\mathbb{Q}}\right)$ for p a prime number solving some explicit congruence relations. Then, as $\mathcal{O}$ is uniquely determined up to conjugation, one could specify $\mathcal{O}$ by a $\mathbb{Z}$-basis, such that $I \in \mathcal{O}$, and one could fix $\rho = I$ in the explicit calculations of the Riemann form and period matrix for (A_τ, E_τ). A priori it might look like a canonical choice of μ, but the overall effect it is that the embedding depends explicitly on the prime p.

In [LY20], for maximal orders, and later in [GY19] for hereditary orders, the dependence of the prime is reformulated as dependence on positive definite quadratic forms that represent p, associated with the pair (D, N). Alternatively, consider the following loci in $\mathcal{A}_2$.

Definition 4.13. *Consider an indefinite quaternion algebra* B *over* $\mathbb{Q}$ *of discriminant* D *and a hereditary order* $\mathcal{O} \subset B$ *of level* N*. We define* the quaternionic locus of QM by $\mathcal{O}$ *as*

$$\mathcal{Q}_{D,N} = \cup_\mu \mathfrak{X}_\mu, \tag{4.4}$$

where $\mu \in \mathcal{O}$ *runs for the allowed elements as in Equation (4.2) above, and* $\mathfrak{X}_\mu$ *is the image of* $X_0^D(N)(\mu)$ *into* $\mathcal{A}_2$ *after the forgetful functor.*

Remark 4.14. *By Remark 4.10, it contains in particular all $A \in \mathcal{A}_2$ such that* $\mathrm{End}(A) \cong \mathcal{O}$. *Hence, if one would consider the locus $\mathcal{Q}'_{D,N}$ of $A \in \mathcal{A}_2$ with an embedding $\mathcal{O} \hookrightarrow \mathrm{End}(A)$, without further conditions on the Rosati involution, then $\mathcal{Q}'_{D,N} \setminus \mathcal{Q}_{D,N}$ is a discrete union of CM points.*

These loci for $N = 1$ (for maximal orders, but in more general settings that abelian surfaces) were first study in [Rot03b], [Rot04b] and [Rot04a], as part of his PhD thesis [Rot03a]. The results in [LY20] and [GY19] extend them to refine the count of the number of components of $\mathcal{Q}_{D,N}$.

Let us explain first how to associate a positive definite quadratic form to a ppas (A, E) with QM by $(\mathcal{O}, \mu)$, with $\mathcal{O}$ hereditary order of level N in a indefinite quaternion algebra B over $\mathbb{Q}$ of discriminant D. Assume that $\mathrm{End}_0(A) \cong B$, we are going to construct a positive definite lattice in $\mathcal{O}$, isometric to $(\mathcal{L}_{\boldsymbol{\tau}}, \Delta)$, for $\boldsymbol{\tau}$ a normalized period matrix for (A, E). Remark that we have the quaternionic trace and the norm in B, $\mathrm{tr} : B \to \mathbb{Q}$, $\alpha \mapsto \alpha + \bar{\alpha}$ and $\mathrm{nr} : B \to \mathbb{Q}$, $\alpha \mapsto \alpha\bar{\alpha}$, and that any $\alpha \in \mathbb{Q}$ solves the monic polynomial

$$x^2 - \mathrm{tr}(\alpha)x + \mathrm{nr}(\alpha),$$

and if $\alpha \in \mathcal{O}$ this polynomial is furthermore in $\mathbb{Z}[x]$ by [Voi21, Corollary 10.3.3 and Corollary 10.3.6].

We first note that we have a natural candidate for quadratic form in B, given by nr.

Lemma 4.15. *For B and indefinite quaternion algebra over $\mathbb{Q}$, there exists a two dimensional $\mathbb{Q}$-vector subspace where* nr *is negative definite.*

Proof. If we have a presentation of $B = \mathbb{Q}1 + \mathbb{Q}I + \mathbb{Q}J + \mathbb{Q}IJ$ by $\left(\frac{a,b}{\mathbb{Q}}\right)$, then [Voi21, Equation (4.1.1)]

$$\mathrm{nr}(t + xI + yJ + zIJ) = t^2 - ax^2 - by^2 + abz^2.$$

As B is indefinite, $B \otimes \mathbb{R} \cong \mathrm{Mat}_2(\mathbb{R})$, there are non-zero elements with vanishing norm, and over $\mathbb{R}$ that can only happen if one of a, b is positive (without loss of generalization, assume $a > 0$). In particular, the quadratic form is *indefinite* over $\mathbb{R}$, and it is in any case not definite over $\mathbb{Q}$. But if we restrict to a suitable two dimensional vector space, namely $\mathbb{Q}I + \mathbb{Q}J$ or $\mathbb{Q}I + \mathbb{Q}IJ$, depending on the sign of b, we get a binary negative definite quadratic form. □

We now follow [LY20, Section 3], [GY19, Section 2.2] and [Run99, Section 6]. As (A, E) has QM by $(\mathcal{O}, \mu)$, the Rosati involution is given by $\alpha \mapsto \mu^{-1}\bar{\alpha}\mu$ on $\mathcal{O}$. Consider $\alpha \in \mathcal{O}$ is Rosati invariant, meaning $\alpha = \mu^{-1}\bar{\alpha}\mu$, and equivalently

$$\mu\alpha = \bar{\alpha}\mu = -\bar{\alpha}\bar{\mu} = -\overline{\mu\alpha},$$

as $\mathrm{tr}\,\mu = 0$. Hence, the Rosati invariant elements on $\mathcal{O}$ are exactly,

$$\mu^{\perp} := \{\beta \in \mathcal{O} |\ \mathrm{tr}(\mu\beta) = 0\} \subset \mathcal{O}.$$

It is a lattice of rank 3, as $\mu^{\perp} \otimes \mathbb{Q}$ is a $\mathbb{Q}$-vector space $\mathbb{Q}$-isomorphic (via μ) to the subspace of traceless quaternions $\{\alpha \in B | \mathrm{tr}(\alpha) = 0\}$. Notice that $\mathbb{Z} \subset \mu^{\perp}$ and that for any $x \in \mu^{\perp}$, $y := 2x - \mathrm{tr}(x) \in \mu^{\perp}$ and $\mathrm{tr}(y) = 0$. We consider then the distinguish sublattice of the traceless quaternions:

$$\{1, \mu\}^{\perp} := \{\beta \in \mathcal{O} |\ \mathrm{tr}(\beta) = 0,\ \mathrm{tr}(\mu\beta) = 0\} = \{\beta \in \mathcal{O}, \beta + \bar{\beta} = 0,\ \mu\beta + \beta\mu = 0\} \subset \mathcal{O}.$$

The quadratic form we consider is given by the discriminant of the polynomial $x^2 - \mathrm{tr}(\alpha)x + \mathrm{nr}(\alpha)$,

$$\begin{aligned} q : \mu^\perp &\to \mathbb{Z} \\ \alpha &\mapsto \mathrm{tr}(\alpha)^2 - 4\,\mathrm{nr}(\alpha). \end{aligned}$$

It will be shown in Proposition 4.16 below that q induces a quadratic form on ${}^{\mu^\perp}\!/_{\mathbb{Z}}$, that we will identify with $\{1,\mu\}^\perp$ via $x \mapsto 2x - \mathrm{tr}(x)$. If $\mathrm{tr}(\alpha) = 0$ then $q(\alpha) = -4\,\mathrm{nr}(\alpha)$, so $\{1,\mu\}^\perp$ is a two dimensional lattice where the restriction of nr is negative definite.

We can more canonically consider $(\alpha,\beta) \mapsto -\mathrm{tr}(\alpha\bar{\beta})$ as the bilinear quadratic form on

$$L_\mu := (2\mathcal{O} + \mathbb{Z}) \cap \{1,\mu\}^\perp.$$

Remark that this bilinear form is not the associated with $-\,\mathrm{nr}$, but with $-2\,\mathrm{nr}$, but for computations it is more natural to work with it, as $\det(-2\,\mathrm{nr}) = -\,\mathrm{disc}(\mathrm{nr})$, and we are eventually more interested in the quadratic form. Compare with Proposition 3.22.

Proposition 4.16. *We have the following properties.*

1. *For any $\alpha \in \mu^\perp$ and any $n \in \mathbb{Z}$, $q(\alpha + n) = q(\alpha)$.*
2. *For any $\alpha \in \mu^\perp$, $q(\alpha) \equiv 0, 1 \bmod 4$, $q(\alpha) \geq 0$ and $q(\alpha) = 0$ if and only if $\alpha \in \mathbb{Z}$.*
3. *The form q induces a positive definite quadratic form on ${}^{\mu^\perp}\!/_{\mathbb{Z}}$, and $({}^{\mu^\perp}\!/_{\mathbb{Z}}, q)$ is isometric to $(L_\mu, -\,\mathrm{nr}(\cdot))$.*

Proof. Item 1) follows from a calculation. For item 2), $q(\alpha) \equiv 0, 1 \bmod 4$ follows from the definition. With respect to positive definitiveness, we tensor with $\mathbb{R}$ and argue over $\mathrm{Mat}_2(\mathbb{R})$, where by Lemma 4.5, the positive involutions were analogously given by $A \to {}^tA$ and conjugation by positive definite symmetric matrices. As tr and nr are given by the standard trace and determinant on A, it is equivalent to argue that the invariant matrices by a positive involution of $\mathrm{Mat}_2(\mathbb{R})$ have real eigenvalues. For $A = {}^tA$ it is clear, for a general positive involution one argues as in Lemma 3.7. Simmilarly, if $q(\alpha) = 0$, then its polynomial ramifies as $(\alpha - n)^2 = 0$ for some $n \in \mathbb{Z}$, and one concludes as in Lemma 3.7 that nilpotent elements cannot be invariant under a positive involution because otherwise $0 < \mathrm{tr}(\alpha\alpha^*) = \mathrm{tr}(\alpha^2) = 0$.

The isometry is given by [LY20, Lemma 17], it is the inverse to $\alpha \in \mu^\perp, \alpha \mapsto 2\alpha - \mathrm{tr}(\alpha) \in (2\mathcal{O} + \mathbb{Z}) \cap \{1,\mu\}^\perp = L_\mu$, and it follows that

$$4q(\alpha) = q(2\alpha) = \mathrm{disc}(2\alpha - \mathrm{tr}(\alpha)) = -4\,\mathrm{nr}(2\alpha - \mathrm{tr}(\alpha)),$$

so $q(\alpha) = -\,\mathrm{nr}(2\alpha - \mathrm{tr}(\alpha))$. □

Corollary 4.17. *We have that the quadratic form q above represents squares if and only if $D = 1$.*

Proof. Let $D = 1$, so $B \simeq \mathrm{Mat}_2(\mathbb{Q})$. Then B does have elements of norm zero. But in the proof of the proposition above, there cannot be nilpotent elements on $\mu^\perp \subset \mathcal{O}$. There can still be elements of norm zero on $\mu^\perp$, recall remark 3.8, but in particular for those element $\mathrm{tr} \neq 0$. Hence, for such an element α, $q(\alpha) = \mathrm{tr}(\alpha)^2$ is a perfect square.

Reciprocally, if for an element α we have $q(\alpha) = k^2$ for $k \in \mathbb{Z}$, then its polynomial $x^2 - \mathrm{tr}(\alpha)x + \mathrm{nr}(\alpha) \in \mathbb{Z}[x]$ has integer roots $(x - n_1)(x - n_2)$. If we consider $\beta := \alpha - n_1$, then $\beta(\beta - (n_1 + n_2)) = 0$ so β is a zero divisor and $\mathrm{nr}(\beta) = 0$. □

Now we give an important result from [LY20] and [GY19], clarifying the link between the quadratic form above and the distinct Shimura curves $\mathfrak{X}_\mu$.

Theorem 4.18. *Fix now $\{\alpha, \beta\}$ a $\mathbb{Z}$-basis of ${}^{\mu^\perp}\!/_{\mathbb{Z}}$, and the binary positive definite quadratic from $Q_\mu(x,y) = q(x\alpha+y\beta) := ax^2+bxy+cy^2$. Furthermore, consider[11] its $\mathrm{GL}_2(\mathbb{Z})$-equivalence class $[Q_\mu]_{\mathrm{GL}}$.*

1. *It verifies the following properties.*
   - *The discriminant of $Q_\mu(x,y)$ as a quadratic form is*
   $$b^2 - 4ac = -16DN.$$
   - *Consider $n \in \mathbb{Z}_{>0}$ represented by Q_μ. Then*
     - *$n \equiv 0, 1 \ mod\ 4$.*
     - *The quaternion algebra $\left(\frac{-DN,n}{\mathbb{Q}}\right)$ has discriminant D (equivalently, it is isomorphic over $\mathbb{Q}$ to B).*
2. *(Supposing N square-free) Assume Q satisfying the properties above, then there exists $\mu \in \mathcal{O}$ such that Q_μ is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to Q.*
3. *(Supposing N square-free) There is a bijection between:*
$$\{\text{components of } \mathcal{Q}_{D,N}\} \leftrightarrow \{\text{ quadratic forms satisfying 1)}\}/_{\mathrm{GL}_2(\mathbb{Z})}$$
$$\mathfrak{X}_\mu \mapsto Q_\mu$$

In the language of Section 3.3, for N square-free,

$$\mathcal{Q}_{D,N} = \bigcup_q \mathcal{H}(q),$$

where q ranges through the $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of the quadratic forms in 1. The Humbert varieties $\mathcal{H}(q)$ are furthermore irreducible and non-empty.

Remark 4.19. *As part of the proof of [LY20, Lemma 18], for $\{\alpha, \beta\}$ $\mathbb{Z}$-basis[12] of ${}^{\mu^\perp}\!/_{\mathbb{Z}}$, then $\alpha\beta - \beta\alpha = \pm\mu$ and the order $\mathcal{O}$ admits a $\mathbb{Z}$-basis given by*

$$\mathcal{O} = \mathbb{Z} + \mathbb{Z}\alpha + \mathbb{Z}\beta + \mathbb{Z}\alpha\beta,$$

but observe that α, β are not given as standard generators of a quaternion algebra, as by the proof of [LY20, Lemma 18] $\alpha\beta \neq -\beta\alpha$ in general.

Remark 4.20. *Let us rewrite the condition $\mathrm{disc}\left(\frac{-DN,n}{\mathbb{Q}}\right) = D$. It can be expressed in terms of Hilbert symbols as $(-DN, n)_p = -1$ for the primes p with $p|D$, and $(-DN, n)_q = 1$ for any other prime, including the prime at ∞.[13] Assume now that $\gcd(n, -16DN) = 1$ and n is*

[11] remark this is well-defined and independent of the choice of basis.

[12] a specific choice of representative of α, β depending on the parity of their traces.

[13] This is simply that the quaternion algebra is indefinite, and it always holds because $-DN$ and n have different signs.

primitively represented by Q. By [Cox22, Lemma 2.5], $-16DN$ (hence $-DN$) is a quadratic residue modulo n and for any p odd prime:

$$\left(\frac{-DN}{p}\right)^{v_p(n)} = 1$$

Remark that by [Ser73, Chapter III, 1.2, Theorem 1], Hilbert symbols can be computed in terms of Legendre symbols. More precisely, for p an odd prime and $a, b \in \mathbb{Q}$ written as $a = p^\alpha u$ and $b = p^\beta v$ with $v_p(u) = v_p(v) = 0$, it follows that

$$(a,b)_p = (-1)^{\alpha\beta\frac{p-1}{2}} \left(\frac{u}{p}\right)^\beta \left(\frac{v}{p}\right)^\alpha,$$

Therefore, for p an odd prime and $p = \infty$,

$$(-DN, n)_p = \begin{cases} -1, & \text{if } p = \infty, \\ 1, & \text{if } p \nmid DNn, \\ \left(\frac{-DN}{p}\right)^{v_p(n)} = 1, & \text{if } p | n, \\ \left(\frac{n}{p}\right), & \text{if } p | DN \end{cases}$$

and for $p = 2$ there is a similar formula, or it can be deduced from the product formula of the Hilbert symbol. Hence, the only non automatic conditions are for the quadratic form Q_μ are the last one for $p|DN$.

Proof. This is [LY20, Lemma 18, Lemma 20, Lemma 21] (for the case $N = 1$) and [GY19, Theorem A] claims and extension to the case N square-free. We give a sketch of the relevant proofs.

For item 1, one passes to the isometric positive definite lattice $(L_\mu, -\operatorname{nr}(\cdot))$. For the discriminant of the quadratic form, it is (changing sign) the determinant of $(L_\mu, -\operatorname{tr}(\cdot, \bar{\cdot}))$, and it can be computed locally at every prime. It also follows that every integer represented by Q is congruent to $0, 1 \bmod 4$. Set now $n = Q(x_0, y_0)$ and consider $\gamma \in L_\mu$ corresponding to $(x_0\alpha + y_0\beta) + \mathbb{Z} \in {}^{\mu^\perp}\!/_{\mathbb{Z}}$, hence $n = -\operatorname{nr}(\gamma) = \gamma^2$ and $\mu\alpha + \alpha\mu = 0$. That implies that the $\mathbb{Q}$-subalgebra of B, $\mathbb{Q}[\mu, \gamma] \subset B$ generated by μ and α satisfies

$$\mu^2 = -DN,\ \gamma^2 = n,\ \mu\alpha = -\alpha\mu,$$

so it is isomorphic to the quaternion algebra $\left(\frac{-DN,n}{\mathbb{Q}}\right)$, and as in particular $4 = \operatorname{rank}_{\mathbb{Q}} B = \operatorname{rank}_{\mathbb{Q}}(\mathbb{Q}[\mu, \gamma])$, it holds that $\mathbb{Q}[\mu, \gamma] = B$. Therefore, $\left(\frac{-DN,n}{\mathbb{Q}}\right)$ has discriminant D.

Item 2) is [LY20, Lemma 21]. Given a quadratic form $Q(x, y) = ax^2 + bxy + cy^2$ satisfying the properties in item 1), set the algebra over $\mathbb{Q}$ of rank 4, $B' = \mathbb{Q} + \mathbb{Q}X + \mathbb{Q}Y + \mathbb{Q}XY$ with multiplication rule:

$$X^2 = a,\ Y^2 = c,\ XY + YX = b.$$

We identify it with a more standard presentation of a quaternion algebra via:

$$I := \frac{XY - YX}{4} = \frac{2YX - b}{4},$$
$$J := Y,$$

with inverse

$$\begin{aligned} Y &:= J \\ X &:= \frac{4I+b}{2X} = \frac{(b+4I)J}{2a} \end{aligned} \tag{4.5}$$

and one checks that $I^2 = (b^2-4ac)/16 = -DN$ and $IJ = -JI$, so this is the quaternion algebra $\left(\frac{-DN,a}{\mathbb{Q}}\right) \simeq B$. Setting $\mu' := I$, it follows that $\mu'^2 = -DN$. On the other hand, consider α, β defined by:

$$\alpha := \begin{cases} \frac{X}{2} & \text{if } a \equiv 0 \bmod 4, \\ \frac{1+X}{2} & \text{if } a \equiv 1 \bmod 4; \end{cases} \quad \beta := \begin{cases} \frac{Y}{2} & \text{if } c \equiv 0 \bmod 4, \\ \frac{1+Y}{2} & \text{if } c \equiv 1 \bmod 4. \end{cases} \tag{4.6}$$

One can check the following by computation ([LY20, Equation (11)]):

- The element $\alpha\beta + \beta\alpha \in \mathbb{Z} + \mathbb{Z}\alpha + \mathbb{Z}\beta + \mathbb{Z}\alpha\beta$. In particular the same is true for $\beta\alpha$, so $\mathcal{O}' := \mathbb{Z} + \mathbb{Z}\alpha + \mathbb{Z}\beta + \mathbb{Z}\alpha\beta$ is an order in B.
- The determinant of the Gram matrix of $\mathrm{tr}(\cdot, \bar{\cdot})$ with respect to $\{1, \alpha, \beta, \alpha\beta\}$ is $(DN)^2$, hence $\mathrm{disc}(\mathcal{O}') = DN$, so $\mathcal{O}'$ is conjugate to $\mathcal{O}$ as both are Eichler orders in B of level N.
- One checks that $\alpha\beta - \beta\alpha$ commutes with μ', so it belongs to $\mathbb{Q}(\mu')$, and furthermore $\alpha\beta - \beta\alpha = \pm\mu'$. In particular $\mu' \in \mathcal{O}'$.
- The lattice $L_{\mu'} := (\mathbb{Z}+2\mathcal{O}') \cap (\mu')^\perp$ is given by $L'_\mu = \mathbb{Z}\tilde{I} + \mathbb{Z}\tilde{J}$ and $-\mathrm{nr}(x\tilde{I}+y\tilde{J}) = Q(x,y)$. Therefore $\{\alpha, \beta\}$ is a $\mathbb{Z}$-basis for $\mu'/\mathbb{Z}$.

In conclusion, setting μ as the conjugate element to μ' under $\mathcal{O} \simeq \mathcal{O}'$ and (after maybe considering $-\mu'$) is an allowed element in $\mathcal{O}$ and we can consider the Shimura curve $\mathfrak{X}_\mu$ with its associated quadratic form Q_μ, and it follows that $[Q_\mu]_{\mathrm{GL}} = [Q]_{\mathrm{GL}}$.

Item 3) is [LY20, Lemma 21]. □

Lemma 4.21. *Consider one of the quadratic forms $Q(x,y) = ax^2 + bxy + cy^2$ in Theorem 4.18, and assume N square free, then either Q is primitive (meaning $\gcd(a,b,c) = 1$) or $Q = 4Q'$ for Q' a primitive quadratic form of discriminant $-DN$ only in the case of $DN \equiv 3 \bmod 4$.*

Proof. This is [LY20, Lemma 23]. □

For simplicity, for the rest of the section we consider the first case.

We circle back to our original goal of writing explicit quaternion modular embeddings. For that, we have to write down a choice for μ and an explicit symplectic $\mathbb{Z}$-basis $\gamma_1, \dots, \gamma_4$ for $\mathcal{O}$ with respect to E_μ, $(u,v) \mapsto \mathrm{tr}(\mu^{-1}u\bar{v})$. With that choice, and recalling the embedding $\iota : B \hookrightarrow \mathrm{Mat}_2(\mathbb{R})$, and $v_t = \begin{pmatrix} \tau \\ 1 \end{pmatrix}$ for $\tau \in \mathbb{H}$, then $(A_\tau, E_\mu) \in \mathfrak{X}_\mu$ has a (big) period matrix

$$\Pi(\tau) = (\iota(\gamma_1)v_\tau, \dots \iota(\gamma_4)v_t) = (\Omega_1(\tau), \Omega_2(\tau)),$$

so we can consider a normalized period matrix given by $\Omega_2(\tau)^{-1}\Omega_1(\tau)$.

First, we need an explicit symplectic $\mathbb{Z}$-basis of the order $\mathcal{O}$. Let us first consider the problem of finding an explicit integral basis of the order $\mathcal{O}$. This was first stated in [Ibu82], but we deduce these formulas from the data of a primitive quadratic form Q and the prime p.

Consider p a prime number, represented by Q such that $p \nmid DN$. Then by Theorem 4.18, $p \equiv 1 \bmod 4$ and $B \cong \left(\frac{-DN,p}{\mathbb{Q}}\right)$. Write I, J, IJ for its standard generators as a quaternion algebra:

$$I^2 = -DN,\ J^2 = p,\ IJ = -JI.$$

We want to write μ and $\mathcal{O}$ in terms of this data. We can set $\mu = I$, let us focus of setting a $\mathbb{Z}$-base for the order.

Recall the formulas in the proof of Theorem 4.18, so that we could write $\mathcal{O} = \mathbb{Z} + \mathbb{Z}\alpha + \mathbb{Z}\beta + \mathbb{Z}\alpha\beta$. We have formulas for α and β in Equation (4.6), and we could use them via Equation (4.5), but for that we need to consider the term b, writing Q as $ax^2 + bxy + cy^2$. To avoid fixing a specific form for Q, we are going to "recover" b from the data that Q has discriminant $-16DN$ and represents p (primitively, for that p is prime). By [Cox22, Lemma 2.3], Q is $\mathrm{SL}_2(\mathbb{Z})$-equivalent to a quadratic form of the form $px^2 + bxy + cy^2$, and we would like to solve for b and c in $b^2 - 4pc = -16DN$ (hence $\left(\frac{-DN}{p}\right) = 1$). As $p \nmid DN$, let us assume $DN | b, c$, and further set that they are divisible by 4. Write $b = s4DN$ and $c = t4DN$. Hence the discriminant equation reads $(s4DN)^2 - 16ptDN = -16DN$, which implies

$$s^2DN - pt = -1,$$

or, equivalently, $s^2DN + 1$ is divisible by p. Observe that s is necessarily even. There exists a solution for such s because $-DN$ is a quadratic residue modulo p: then its multiplicative inverse in $\mathbb{F}_p$ for (that we write $(-DN)^{-1}$) is also a quadratic residue, so there exists x such that $x^2 \equiv (-DN)^{-1} \bmod p$, or $x^2 + (-DN)^{-1} \equiv 0 \bmod p$, hence $x^2DN + 1 \equiv 0 \bmod p$. Applying formulas (4.5) and (4.6) one gets:

$$\alpha = \frac{1+J}{2},$$
$$\beta = \frac{(sDN+I)J}{p},$$

and for the product $\alpha\beta$,

$$\alpha\beta = \frac{1}{2p}(1+J)(sDN+I)J = \frac{sDNJ + IJ - psDN - pI}{2p}.$$

This last generator of $\mathbb{Z} + \mathbb{Z}\alpha + \mathbb{Z}\beta + \mathbb{Z}\alpha\beta$ has a more complicated expression, but it can be traded as a generator for $\frac{I(1+J)}{2}$.

This expressions for Eichler orders come from [Has95], which are a generalization of [Ibu82]. Note that they are not a symplectic basis for E_I, see [Has95, Lemma 2.4].

We can now state the following result for quaternion modular embeddings: originally from [Has95] and explained in terms of $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of quadratic forms in [LY20] and [GY19].

Theorem 4.22. *Consider Q, p, s, t as above. Then the order $\mathcal{O} \subset B$ given by $\mathbb{Z}e_1 + \mathbb{Z}e_2 + Ze_3 + \mathbb{Z}e_4$ with*

$$e_1 = 1, \quad \alpha_2 = \frac{1+J}{2}, \quad e_3 = \frac{I(1+J)}{2}, \quad e_4 = \frac{sDNJ + IJ}{p},$$

is an Eichler order of level N in B. Then $I \in \mathcal{O}$ and with the choice of $\mu = I$, $\mu^{\perp}/\mathbb{Z}$ is identified with $\mathbb{Z}e_2 + \mathbb{Z}e_4$, and $\mathbb{Q}_\mu(x,y) = \mathrm{disc}(xe_2 + ye_4)$ is given by

$$Q_\mu(x,y) = px^2 + 4sDNxy + 4tDNy^2.$$

Furthermore, a symplectic basis $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ of $\mathcal{O}$ for the alternating bilinear form $(v,w) \mapsto \mathrm{tr}(\mu^{-1}v\overline{w})$ is given by

$$\gamma_1 = e_3 - \frac{p-1}{2}e_4, \quad \gamma_2 = -sDNe_1 - e_4, \quad \gamma_3 = e_1, \quad \gamma_4 = e_2,$$

and with respect to this choice of symplectic basis, the normalized period matrix of (A_z, E_μ) is given by

$$\frac{1}{pz}\begin{pmatrix} -(\widetilde{\varepsilon})^2 + \frac{(p-1)sDN}{2}z + DN\varepsilon^2 z^2 & \widetilde{\varepsilon} - (p-1)sDNz - DN\varepsilon z^2 \\ \widetilde{\varepsilon} - (p-1)sDNz - DN\varepsilon z^2 & -1 - 2sDNz + DNz^2 \end{pmatrix},$$

for $\varepsilon = \frac{1+\sqrt{p}}{2}$ and $\widetilde{\varepsilon} = \frac{1-\sqrt{p}}{2}$, with singular relations (corresponding to e_2 and e_4) given by

$$\left(1, 1, \frac{1-p}{4}, 0, 0\right), \ \left(0, 2sDN, 0, 1, DN(s^2DN - t)\right).$$

Proof. This is the first case of [GY19, Lemma 4 and Lemma 5], originally proven in [Has95, Theorem 2.2, Theorem 3.5]. □

Remark 4.23. *This embedding is not in the "Humbert presentation" of Proposition 3.29.*

4.1.2.1 On the number of components of the quaternionic loci

Following [Rot04b], [LY20] and [GY19], we provide more details on the quaternionic loci $\mathcal{Q}_{D,N}$ defined in Equation (4.4).

By Theorem 4.18, the number of components of $\mathcal{Q}_{D,N}$ is the number of $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of binary quadratic forms satisfying the properties of said result. In [LY20, Theorem 1], for the case $N = 1$, a formula for the number of components in terms of class numbers. Remark that the class number of a discriminant is the number of primitive positive definite quadratic forms of said discriminant modulo $\mathrm{SL}_2(\mathbb{Z})$-equivalence, and Theorem 4.7 is stated in terms of $\mathrm{GL}_2(\mathbb{Z})$-equivalence.

In Theorem 4.18, starting with a quadratic form $Q(x,y)$ of discriminant $-16D_0$ for D_0 a square-free number, one could ask in more general terms how to assign the pair (D,N). In the context of the theorem, that from the condition that for every a represented by Q, all quaternion algebras $\left(\frac{-D_0,a}{\mathbb{Q}}\right)$ have the same discriminant D, a factor of D_0. There is another way to phrase it, as in [GY19, Introduction, pg 2], by considering the *genus* of the quadratic form. Two positive definite quadratic forms of the same (negative) discriminant N are said to belong to the same genus if they represent the same values[14] in $\left(\mathbb{Z}/N\mathbb{Z}\right)^*$, and every genus contains the same (finite) number of $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes by [Cox22, Corollary 3.14].

[14] Equivalently, by [Cox22, Theorem 3.21] if they are equivalent over the p-adic integers $\mathbb{Z}_p$ for every prime p, or equivalent over $\mathbb{Q}$ with a matrix with denominators prime to $2N$.

Assume that $Q(x,y) = ax^2+bxy+cy^2$ is a positive definite quadratic form of discriminant $-16D_0$ for D_0 a square-free number, and that every number represented by Q is congruent to $0,1 \bmod 4$. That property, together with D_0 being square free, implies by Lemma 4.21 that Q is either primitive or $Q = 4Q'$ for Q' a quadratic form of discriminant $-D_0$ (necessarily primitive), and that can only happen if $D_0 \equiv 3 \bmod 4$. As we did above, we are going to give details only in the former case and one may check [GY19, Introduction, page 2] for the latter.

Set C_{-16D_0} for the class group of primitive binary quadratic forms of discriminant $-16D_0$, and write $D_0 = p_1 \cdots p_k$. By [Cox22, Proposition 3.11, Theorem 3.15], there are 2^{k+1} genera of forms of discriminant $-16D_0$, and they are described as follows. Consider the following characters on C_{-16D_0}

$$\chi_{-4} : C_{-16D_0} \to \{\pm 1\} \qquad \chi_{p_j} : C_{-16D_0} \to \{\pm 1\}$$
$$P \mapsto \left(\frac{-4}{n}\right) = (-1)^{(n-1)/2}; \qquad P \mapsto \begin{cases} \left(\frac{n}{p_j}\right) & \text{if } p_j \text{ odd} \\ \left(\frac{8}{n}\right) = (-1)^{(n^2-1)/8} & \text{if } p_j = 2; \end{cases}$$

where n is any positive integer represented by P with $\gcd(n, 2D_0) = 1$ and $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol. In the proof of [Cox22, Theorem 3.15] (remark we are in the two last cases of the table), these are the assigned characters to describe the genera of C_{-16D_0}. Furthermore, it also holds that either $\chi_{-4}\chi_{p_1} \cdots \chi_{p_k}$ or $\chi_{p_1} \cdots \chi_{p_k}$ is the trivial character, depending on if the odd part of D_0 is congruent to 1 or 3 mod 4.

For our quadratic form Q, it follows that $\chi_{-4}(Q) = 1$, so in any of the cases above, necessarily $\chi_{p_j}(Q) = -1$ for an *even* number of primes. Define now:

$$D_Q := \prod_{p_j, \chi_{p_j}(Q) = -1} p_j, \; N_Q := \prod_{p_j, \chi_{p_j}(Q) = 1} p_j.$$

Then $D_Q N_Q = D_0$ and D_Q has an *even* number of prime factors, so it is the discriminant of an indefinite quaternion algebra over $\mathbb{Q}$, and can consider $\mathcal{O}$ an hereditary order of level N.

In conclusion, the pair (D, N) is determined by the genus of the quadratic form. Observe that $D = 1$ if and only if the quadratic form belongs to the principal genus.

Reciprocally, by Remark 4.20, $(-DN, n)_{p_k} = \chi_{p_k}(Q)$. Therefore, the quadratic forms we consider form a full genus in this case. In the case $D_0 \equiv 3 \bmod 4$, one considers both C_{-16D_0} and C_{-D_0} and one genus in each discriminant, see [LY20, Lemma 24].

Example 4.24. *For any N, consider the quadratic form $Q(x,y) = x^2 + 4Ny^2$, which its $\mathrm{SL}_2(\mathbb{Z})$-equivalence class is the principal class of C_{-16N}. We will see in the next subsection that this is the quadratic form associated to $\mu = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$ and the diagonal map of $X_0(N) \to \mathcal{A}_2$ given by $\tau \mapsto \begin{pmatrix} \tau & 0 \\ 0 & N\tau \end{pmatrix}$. In particular, for $N = 1$, $x^2 + 4y^2$ is the quadratic form associated to (generic) self-products $A = E^2$ with the product polarization. By Remark 3.25, A admits maximal real multiplication by a quadratic field K if and only $\Delta(K)$, the discriminant of the quadratic field, can be expressed as a sum $x_0^2 + (2y_0)^2$. Hence, the quadratic fields we described in Example 2.3 are all the quadratics fields such that A has* maximal *real multiplication by.*

4.2 Modular curves and Jacobians isomorphic to products of elliptic curves

Following the previous section, when $D = 1$ we recover modular embeddings for the classical modular curves $X_0(N)$ for N square-free. In this section, we are going to present an alternative approach to describe the associated quadratic forms, with the advantage that works for any N, not necessarily square-free. We follow [Kan16].

Before that, note that Theorem 4.22 also applies when $D = 1$, but in this specific case one can directly give different embeddings that are *linear*. We will state them in Lemma 6.31, following the examples in [LY20, Section 6].

In Subsection 6.5.2.2, when N is square-free, we will explicitly construct *all* such embeddings $X_0(N) \to \mathcal{Q}_{1,N}$, as they are indexed in every allowed $\mu \in \begin{pmatrix} \mathbb{Z} & \mathbb{Z} \\ N\mathbb{Z} & \mathbb{Z} \end{pmatrix}$, but we cannot conclude that for general N. We also compute the quadratic form Δ associated with $\mathcal{L}_{\boldsymbol{\tau}}$, for $\boldsymbol{\tau}$ in the image of those embeddings, hence the quadratic form associated with each allowed μ..

These quadratic forms are precisely the ones given in [Kan16] by an alternative approach, for *general* N. That allowed us to reframe our embeddings in this more general set up. We do not claim is it the full $\mathcal{Q}_{1,N}$ for general N.

We present now the problem of [Kan16]. This is an extended version of [Gij25, Section 5.2.2].

In concerns the following problem. From Theorem 1.3 it follows that for any $C \in \mathcal{M}_2$ and its Jacobian $(\mathrm{Jac}(C), \theta)$, and any E, E' elliptic curves

$$(\mathrm{Jac}(C), \theta) \ncong (E \times E', \lambda_E \otimes \lambda_{E'}),$$

as ppas, where $\lambda_E \oplus \lambda_{E'}$ is the canonical induced principal polarization on $E \times E'$. But it does not impede $\mathrm{Jac}(C)$ to be isomorphic to $E \times E'$ as *unpolarized* abelian varieties. Equivalently, there could be *another* principal polarization $\tilde{\lambda}$ on $E \times E'$ such that

$$(\mathrm{Jac}(C), \theta) \cong (E \times E', \tilde{\lambda}), \tag{4.7}$$

Equation (4.7) imposes conditions on E, E' and C. Before stating them, we set some facts about NS, End and End^s for a product $E \times E'$ equipped with the product polarization, that we extracted from [Kan16, Appendix].

Consider (E_1, λ_1) and (E_2, λ_2) and the maps $p_i : E_1 \times E_2 \to E_i$, $\iota_i : E_i \to E_1 \times E_2$ for $i = 1, 2$.

Lemma 4.25. *There is an isomorphism*

$$\hat{p_1} \times \hat{p_2} : \hat{E_1} \times \hat{E_2} \to \widehat{E_1 \times E_2}$$

and

$$\begin{aligned} \mathrm{End}(E_1 \times E_2) &\to (\mathrm{Hom}(E_i, E_j))_{i,j} \\ f &\mapsto (p_j f \iota_i)_{i,j}. \end{aligned}$$

The Rosati involution corresponding to $\lambda_1 \otimes \lambda_2$ is $(f_{ij})_{ij} \mapsto {}^t((\lambda_i)^{-1}\widehat{f_{ij}}\lambda_j)$, i.e.

$$\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix} \to \begin{pmatrix} \lambda_1^{-1}\hat{f_{11}}\lambda_1 & \lambda_2^{-1}\hat{f_{21}}\lambda_1 \\ \lambda_1^{-1}\hat{f_{12}}\lambda_2 & \lambda_2^{-1}\hat{f_{22}}\lambda_2 \end{pmatrix};$$

so we identify

$$\mathrm{End}^s(E_1\times E_2)=\left\{\begin{pmatrix}\alpha_{11} & \alpha_{12}\\ \lambda_2^{-1}\hat{\alpha_{12}}\lambda_1 & \alpha_{22}\end{pmatrix}:\ \alpha_{ii}\in\mathrm{End}^s(E_i,\lambda_i),\ \alpha_{12}\in\mathrm{Hom}(E_1,E_2)\right\}.$$

Furthermore, $\mathrm{End}^{s,++}(E_1\times E_2,\lambda_1\otimes\lambda_2)$ *corresponds to*

$$\left\{\begin{pmatrix}[a]_{E_1} & \alpha_{12}\\ \lambda_1^{-1}\hat{\alpha_{12}}\lambda_2 & [b]_{E_2}\end{pmatrix}:\ a,b\in\mathbb{Z}_{>0},\ \alpha_{12}\in\mathrm{Hom}(E_1,E_2),\ ab-\deg(\alpha_{12})>0\right\}.$$

Proof. This is the content of [Kan16, Paragraph after Equation (64) in pg.35, Proposition 61, Corollary 25]. Remark that [Kan16, Corollary 25] applies also if the elliptic curves have CM, for that $\mathrm{End}^s(E)\cong\mathbb{Z}$. □

Proposition 4.26. *Assume* C,E,E' *satisfying* (4.7) *above, and set* $F:Jac(C)\to E\times E'$ *for the isomorphism. Then*

- *The elliptic curves* E,E' *are isogenous. Let* $\phi:E\to E'$ *be an isogeny between them.*
- *The principal polarization on* $E\times E'$ *depends on* $\deg\phi$ *as follows. Consider* $\lambda_E\otimes\lambda_{E'}$ *the canonical product principal polarization on* $E\times E'$*. Then there exists* $a,b\in\mathbb{Z}_{>0}$ *with*

$$ab-\deg\phi=1, \tag{4.8}$$

such that for

$$f:=\begin{pmatrix}[a]_E & \phi\\ \lambda_E^{-1}\hat{\phi}\lambda_{E'} & [b]_{E'}\end{pmatrix}\in\mathrm{End}(E\times E'),$$

then $\tilde{\lambda}:=(\lambda_E\otimes\lambda_{E'})\circ f$ *is a principal polarization with*

$$\theta=F^*\tilde{\lambda}$$

Proof. This is [Kan16, Proposition 26], where the isogeny can furthermore be taken cyclic. □

Observe that an isogeny between elliptic curves $\phi':E\to E'$ is always of the form $\phi'=k\phi$ with ϕ cyclic. Then the condition (4.8) for $\phi'=k\phi$:

$$ab-k^2\deg(\phi)=1. \tag{4.9}$$

In the generic case of (4.7) that $\mathrm{End}(E)=\mathbb{Z}$, by Proposition 4.26 E and E' are isogenous and E' does not have CM either. Therefore $\mathrm{Hom}(E,E')\cong\mathbb{Z}$, so there exists an isogeny ϕ such that $\mathrm{Hom}(E,E')=\mathbb{Z}\phi$. Such an isogeny is *cyclic* and of *minimal degree*, suggesting the following definition.

Definition 4.27 ([Kan16])**.** *Let* $C\in\mathcal{M}_2$ *and consider* $(\mathrm{Jac}(C),\theta)\in\mathcal{A}_2^{ind}$*, and* $N\in\mathbb{Z}_{>0}$*. We say that* $\mathrm{Jac}(C)$ is a Jacobian of type N[15] *if*

$$\mathrm{Jac}(C)\cong E_1\times E_2,$$

and there exists a cyclic *isogeny* $\phi:E_1\to E_2$ *with* $\deg\phi=N$*.*

Furthermore, we define the following locus in $\mathcal{A}_2^{ind}$ *for* $N\in\mathbb{Z}_{>0}$*:*

$$K(N)=\{\mathrm{Jac}(C)\in\mathcal{A}_2^{ind}|\ \mathrm{Jac}(C)\ \textit{has type}\ N\}.$$

[15] the original terminology in [Kan16] is "curve of type N", for that it is a property of the curve C, but it is more natural to us to thing of C as a point in the moduli space, so we change the notation for Jacobian of type N.

By [Kan16, Theorem 1], these loci are one-dimensional, covered by finitely many irreducible components which are the image of modular embeddings of $X_0(N)$. More precisely, in the language of Section 3.3, $K(N) = H(q)$ with q a quadratic form of type N, as defined below.

Definition 4.28. ([Kan16, Definition pg.6]) *Let $N \in \mathbb{Z}_{>0}$, and let $f(x,y) = ax^2 + bxy + cy^2$ an integral binary quadratic form. We say that f is* of type N *if*

1. *for any $x, y \in \mathbb{Z}$, $f(x,y) \equiv 0, 1 \pmod 4$.*
2. *its discriminant $\Delta(f) = b^2 - 4ac = -16N$,*
3. *There exits $\delta \in \mathbb{Z}_{>0}$ with $\gcd(\delta, N) = 1$ such that $f \to^{pr} \delta^2$, meaning that f represents δ^2 primitively.*

The first condition in the definition is natural by Lemma 3.27, and in terms of characters and genera as in Subsection 4.1.2.1, it means $\chi_{-4}(f) = 1$. The last condition implies that (when $\gcd(a,b,c) = 1$), f belongs to the principal genus of C_{-16N}.

It will also follow form Proposition 4.29 below that either f is primitive or $f = 4g$ for g a primitive form. Compare it with Lemma 4.21 in the case of quaternionic multiplication, but that one is only under the condition of N square free.

More can be said, following [Kan16, Section 5].

Proposition 4.29. *Let $N \in \mathbb{Z}_{>0}$ and f an integral binary form. Then the following are equivalent.*

1. *f has type N as in Definition 4.28 above.*
2. *either f is primitive, or $f = 4g$ for g primitive (and in that case $N \equiv 3 \pmod 4$), and f (resp. g) belongs to the principal genus of C_{-16N} (resp. C_{-4N}).*
3. *f is $\mathrm{SL}_2(\mathbb{Z})$-equivalent to one of the following forms. Consider the parameter set*

$$P(N) = \{(a,b,k) \in \mathbb{Z}_{>0}^2 \times \mathbb{Z} : \ ab - k^2 N = 1\}, \tag{4.10}$$

and for any $(a,b,k) \in P(N)$, set

$$f_{(a,b,k)} := a^2x^2 + 2kN(ab+2) + b^2N(ab+3)y^2. \tag{4.11}$$

Proof. This is [Kan16, Theorem 13]. □

Remark 4.30. *There are trivial solutions to the parameter set $(1,1,0)$, which give $f_{1,1,0} = x^2 + 4Ny^2$, which is the $\mathrm{SL}_2(\mathbb{Z})$-equivalence class the principal form in C_{-16N}. More generally, there are trivial solutions parametrized by $(1, 1+k^2N, k)$ for any $k \in \mathbb{Z}$, and all $f_{(1,1+k^2N,k)}$ are $\mathrm{SL}_2(\mathbb{Z})$-equivalent to the principal form, for that they represent 1.*

Finally, from Equation (4.10) and Equation (4.9) it is natural to think there is a link, which we specify in Theorem 4.31 below, describing $K(N)$ as an union of $H(q)$ for q quadratic forms of type N. Before that, observe that our locus $K(N)$ was defined inside $\mathcal{A}_2^{ind} = \mathcal{A}_2 \backslash (\mathcal{A}_1 \times \mathcal{A}_2) = \mathcal{A}_2 \setminus \mathcal{H}_1$. Therefore, we should discard quadratic forms that represent 1, so the class of $x^2 + 4Ny^2$.

Theorem 4.31. *Consider $N \in \mathbb{Z}_{>0}$, then*

$$K(N) = \bigcup_q \mathcal{H}^{ind}(q) \subset \mathcal{A}_2^{ind},$$

where $\mathcal{H}^{ind}(q) = \mathcal{H}(q) \setminus \mathcal{H}_1$, and q ranges through the $\mathrm{GL}_2(\mathbb{Z})$-equivalence class of quadratic forms of type N, minus $[x^2 + 4Ny^2]_{\mathrm{GL}_2(\mathbb{Z})}$.

For such a q as above, $\mathcal{H}(q) \neq \emptyset$ and $\mathcal{H}(q)$ is irreducible: more precisely, for every $(a, b, k) \in P(N)$, there exists a proper morphism $X_0(N) \to \mathcal{A}_2$ with either degree 2 or degree 4 such that its image is $\mathcal{H}(f_{a,b,k})$. This morphism induces a birational isomorphim with $X_0(N)/w_N$ for w_N the Fricke involution,[16] or a degree four quotient of $X_0(N)$.

For N square free, it follows that $K(N) = \mathcal{Q}_{1,N}$.

Proof. This is [Kan16, Theorem 12, Theorem 31 and Proposition 45]. The morphism can be seen as an algebraic counterpart to Theorem 4.7 for $D = 1$ and general N. We will realize them as quaternion modular embeddings in Lemma 6.31. □

There are other results about irreducibility and non-triviality for $\mathcal{H}(q)$ that we collect here. This is not exhaustive.

Proposition 4.32. *Let q a positive definite quadratic form with $q \equiv 0, 1 \pmod 4$.*

- (Modular curve case) *Assume that q represents a square primitively (not necessarily coprime with its discriminant). Then $\mathcal{H}(q) \neq \emptyset$, ([Kan19, Theorem 1]), and it may not be irreducible ([Kan19, Theorem 4 b)]), but there are bounds for the number of irreducible components. If q is primitive then $H(q)$ is irreducible. ([Kan19, Corollary 5])*
- (Shimura curve case) *Assume q is in the conditions of 4.18 1, in particular* $\mathrm{disc}(q) = -16DN$ *with D an product of an even number of primes, and* $\gcd(D, N) = 1$.
  - *Assume N square-free, then $\mathcal{H}(q) \neq \emptyset$ and it is irreducible. If N is not square-free, then $\mathcal{H}(q)$ may not be irreducible ([LY20, Remark 3]), see [Run99, Example 13].*
  - *If q is primitive then $\mathcal{H}(q)$ is irreducible [Run99, Theorem 10].*

From Proposition 3.28, Theorem 4.18 and 4.31 we have several loci in $\mathcal{A}_2$ that admit description in terms of unions of generalized Humbert curves. We compare them here. In the intersection $\mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2}$ the corresponding quadratic forms may have different discriminants, while in both $\mathcal{Q}_{D,N}$ and $K(N)$ is the same discriminant. If one of the Δ_i is a perfect square, then $\mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2}$ can only contain modular curves, whereas $\mathcal{H}_{\Delta_1} \cap \mathcal{H}_{\Delta_2}$ in for non squares Δ_i contains both types of curves

For D square-free and (general) N we have

$$\mathcal{Q}_{D,N} \subseteq \bigcup_q \mathcal{H}(q),$$

for q ranging through $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of quadratic forms as in Theorem 4.18 1), and $\mathcal{H}(q)$ may not be irreducible.

[16] a distinguished Atkin-Lehner involution of the modular curve.

For D, N square-free, we have

$$\mathcal{Q}_{D,N} = \bigcup_q \mathcal{H}(q),$$

for q ranging through the same set, and every $\mathcal{H}(q)$ is *irreducible* and non-empty.

For $D = 1$ and N square-free, we have

$$K(N) = \mathcal{Q}_{1,N} \setminus \mathcal{H}_1.$$

For general N, we have $K(N) \subset \mathcal{Q}_{1,N} \setminus \mathcal{H}_1$, but more precisely, consider

$$\widetilde{\mathcal{Q}}_N := \bigcup_{\tilde{\mu}} X_{\tilde{\mu}} \subset \mathcal{Q}_{1,N}$$

the subset given by the embeddings from Lemma 6.31, then

$$K(N) = \widetilde{\mathcal{Q}}_N \setminus \mathcal{H}_1.$$

Finally, all possible CM points that occur in $\cup_N \mathcal{Q}_{1,N}$ belong already to $\cup_N \widetilde{\mathcal{Q}}_N$. More precisely, set $\mathfrak{CM}$ the subset of $\cup_N \mathcal{Q}_{1,N}$ consisting of all CM points that occur in any of $\mathcal{Q}_{1,N}$. Alternatively, $\mathfrak{CM}$ is the set of all CM isotypic abelian surfaces in $\mathcal{A}_2$. Analogously, set $\widetilde{\mathfrak{CM}} \subset \cup_N \widetilde{\mathcal{Q}}_N$. Then $\widetilde{\mathfrak{CM}}$ is the set of all abelian surfaces A that are isomorphic as unpolarized abelian varieties to $E_1 \times E_2$, with E_i isogenous CM elliptic curves. Then, by the theorem of Shioda and Mitani [BL04, Corollary 10.6.3], $\widetilde{\mathfrak{CM}} = \mathfrak{CM}$.

Part II

Results

Sácame de aquí, mi corazón,
si no logran encontrarnos,
ya sólo seremos munición
disparada contra el fango
una bala que no sabe que ha fallado.

Figurantes, Vetusta Morla

Chapter 5

Explicit bounds on the coefficients of modular polynomials and the size of $X_0(N)$

Abstract. We give explicit upper and lower bounds on the size of the coefficients of the modular polynomials Φ_N for the elliptic j-function. These bounds make explicit the best previously known asymptotic bounds. We then give an explicit version of Silverman's Hecke points estimates. Finally, we give an asymptotic comparison between the Faltings height of the modular curve $X_0(N)$ and the height of the modular polynomial Φ_N. [1]

Keywords: Modular polynomials, modular curves, elliptic curves, heights.
Mathematics Subject Classification: 11F32, 11G05, 11G50, 14G40.

5.1 Introduction

Modular curves play a central role in modern arithmetic questions. They are a key feature in the solution of famous diophantine equations, in the study of the Mordell-Weil group of elliptic curves (both for the torsion subgroup and for the Birch and Swinnerton-Dyer conjecture) and in isogeny-based cryptography. It is thus useful to be able to represent these curves explicitly and to estimate how complicated their models are.

A classical way to estimate complexity of models is via height theory. For any non-zero polynomial P in one or more variables and integer coefficients we define its *height* to be

$$h(P) := \log \max |c|, \quad \text{where } c \text{ ranges over all coefficients of } P.$$

Let N be a positive integer and denote by $\Phi_N = \Phi_N(X, Y) \in \mathbb{Z}[X, Y]$ the modular polynomial for the elliptic j-function. It vanishes at pairs of j-invariants of elliptic curves

[1] The authors thank Pascal Autissier, Joe Silverman, and Emmanuel Ullmo, for conversations around this topic at the occasion of the Hindry 65 conference in Bordeaux. They also thank Autissier for comments on an earlier version of the text. They thank Riccardo Pengo and Paolo Dolce for providing the reference [DM24]. They thank the referee for constructive feedback. The authors were supported by the IRN GandA (CNRS). The first author is supported by the Alexander-von-Humboldt Foundation. memorThe third author is supported by ANR-20-CE40-0003 Jinvariant.

linked by a cyclic N-isogeny, see [Lan87, Chapter 5]. The equation $\Phi_N(X,Y)=0$ is a plane affine integral model for the modular curve $X_0(N)$ (but not in general a smooth model).

Paula Cohen Tretkoff [Coh84] proved that when N tends to $+\infty$

$$h(\Phi_N) = 6\psi(N)\big[\log N - 2\kappa_N + O(1)\big], \tag{5.1}$$

where

$$\psi(N) = N\prod_{p\mid N}\left(1+\frac{1}{p}\right) \quad \text{and} \quad \kappa_N = \sum_{p\mid N}\frac{\log p}{p}.$$

Work of Autissier [Aut03] and Breuer-Pazuki [BP24] show that one may profitably replace κ_N with λ_N, where

$$\lambda_N := \sum_{p^n\|N}\frac{p^n-1}{p^{n-1}(p^2-1)}\log p.$$

The terms κ_N and λ_N are compared in [BP24], which leads in particular to

$$h(\Phi_N) = 6\psi(N)\big[\log N - 2\lambda_N + O(1)\big]. \tag{5.2}$$

Numerical computations as reported in [BP24] suggest that the bounded term implied by the $O(1)$ in (5.2) is smaller than the one in (5.1).

As modular polynomials have various cryptographic or algorithmic applications, it is useful to obtain explicit bounds on the $O(1)$ term. In [BS10], Bröker and Sutherland obtained asymptotically optimal bounds in the case where N is prime, and Pazuki [Paz19a] provided explicit bounds in the case of general N, but these were not quite asymptotically optimal.

The most recent work providing an explicit upper bound is [BP24], where the first and third authors proved that for any $N\geq 2$,

$$h(\Phi_N) \leq 6\psi(N)\big[\log N - 2\lambda_N + \log\log N + 4.436\big]. \tag{5.3}$$

The term $\log\log N$ was superfluous, an unfortunate artifact of the method used in [BP24]. A natural idea to try to remove it is via equidistribution results. However, that would be at the cost of losing the explicit nature of the upper bound, hence jeopardizing our other efforts. We are nevertheless now able to remove the $\log\log N$ in the following theorem, where we provide both explicit upper and lower bounds.

Theorem 5.1. *Let $N\geq 1$. The height of the modular polynomial $\Phi_N(X,Y)$ is bounded by*

$$6\psi(N)\big[\log N - 2\lambda_N - 0.0351\big] \leq h(\Phi_N) \leq 6\psi(N)\big[\log N - 2\lambda_N + 9.5387\big].$$

The main new idea to improve the upper bound comes from technical inequalities involving Farey sequences. We use both reduced and non-reduced elements in the upper half plane in the key equation (5.5), which help us obtain better estimates of the Mahler measures at play.

To obtain the lower bound, we use a specialization trick to reduce the calculations to the Mahler measure of the one-variable polynomial $\Phi_N(X,0) = \Phi(X,j(\rho))$, where we can use explicit complex multiplication properties.

After presenting some preliminaries in Section 5.2, we prove the upper bound of Theorem 5.1 in Section 5.3. We prove the lower bound of Theorem 5.1 in Section 5.4.

As a corollary to Theorem 5.1, we add the following explicit result on Hecke points in Section 5.5, giving an explicit version of a result of Silverman [Sil90]. For any elliptic curve E defined over $\overline{\mathbb{Q}}$, for any $N \geq 2$ and any cyclic subgroup $C \subset E(\overline{\mathbb{Q}})$ of order N, denote by $j_{E/C}$ the j-invariant of the isogenous elliptic curve E/C.

Theorem 5.2. *Let E be an elliptic curve defined over $\overline{\mathbb{Q}}$ with j-invariant j_E. Let $h_\infty(\cdot)$ denote the absolute logarithmic Weil height. For any $N \geq 2$, one has*

(a)
$$h_\infty(j_E) - \frac{1}{\psi(N)} \sum_{\substack{C \text{ cyclic} \\ \#C=N}} h_\infty(j_{E/C}) \geq -\frac{h(\Phi_N)}{\psi(N)} - \frac{2\log(\psi(N)+1)}{\psi(N)} \geq -6\log N + 12\lambda_N - 58.34.$$

(b)
$$h_\infty(j_E) - \frac{1}{\psi(N)} \sum_{\substack{C \text{ cyclic} \\ \#C=N}} h_\infty(j_{E/C}) \leq 6.67 + 6\min\left\{0, \log\left(1+h_\infty(j_E)\right) - \log N + 2\lambda_N + 0.25\right\}.$$

The proof is given in Section 5.5. It combines Silverman's method, Mahler measure estimates, and the explicit bounds from Theorem 5.1.

The height of Φ_N is a way to measure the size of the curve $X_0(N)$. But there are other ways of measuring the size of $X_0(N)$: the Faltings height of the curve, the Faltings height of its Jacobian $J_0(N)$, the height of a Hecke correspondence with respect to a carefully chosen metrized line bundle, the self-intersection of the Arakelov canonical sheaf, are all used in the literature. One could even think of the size of classical Heegner points on the modular Jacobian as a way to measure the complexity of $J_0(N)$, hence of $X_0(N)$. So what is the size of $X_0(N)$? We gather in the following theorem some asymptotic results that are easy to derive from the existing literature, and which explain that the height of Φ_N, despite being elementary, captures some of this deeper information.

Theorem 5.3. *We have the following properties.*

(a) Let h_{Falt} denote the stable Faltings height as recalled in Definition 5.21. For any integer $N \geq 1$, one has the equality $h_{\mathrm{Falt}}(X_0(N)) = h_{\mathrm{Falt}}(J_0(N))$. Then when N is square-free and coprime to 6 and tends to infinity, one has
$$h_{\mathrm{Falt}}(X_0(N)) \sim \frac{1}{6^3} h(\Phi_N).$$

(b) Let T_N be the Hecke correspondence in $\mathbb{P}^1 \times \mathbb{P}^1$ and let $\hat{\mathcal{L}}$ be the associated metrized line bundle as given by Autissier in [Aut03]. Then when N tends to infinity one has
$$h_{\hat{\mathcal{L}}}(T_N) \sim 2h(\Phi_N).$$

(c) Let k be a quadratic field of discriminant D_k, of class number h_k, with $2u_k$ roots of unity. Assume $D_k < 0$, $D_k = 1 (\mathrm{mod}\, 4)$, and consider $x_{D_k} \in X_0(N)$ the related Heegner point for each compatible N. It gives rise to a cycle $c_{D_k} = (x_{D_k}) - (\infty) \in J_0(N)$. Then when N tends to infinity, one has
$$\hat{h}_{J_0(N)}(c_{D_k}) \sim \frac{h_k u_k}{6\psi(N)} h(\Phi_N).$$

(d) Let $\overline{\omega}^2$ denote the self-intersection of the Arakelov canonical sheaf of the minimal regular model of $X_0(N)$, for any $N \geq 2$ coprime to 6. Then when N tends to infinity, one has

$$\overline{\omega}^2 \sim \frac{1}{24} h(\Phi_N).$$

We prove Theorem 5.3 in Section 5.6.

5.2 Preliminaries

Denote by $\mathbb{H} = \{\tau \in \mathbb{C} \,:\, \mathrm{Im}(\tau) > 0\}$ the upper half-plane, on which $\mathrm{SL}_2(\mathbb{Z})$ acts via fractional linear transformations. A fundamental domain for this action is

$$\mathcal{F} = \left\{ \tau \in \mathbb{H} \,:\, |\tau| \geq 1,\ -\frac{1}{2} < \mathrm{Re}\,\tau \leq \frac{1}{2} \text{ and } \mathrm{Re}\,\tau \geq 0 \text{ if } |\tau| = 1 \right\}.$$

For any $\tau \in \mathbb{H}$, we denote by $\tilde{\tau} \in \mathcal{F}$ the unique representative in this fundamental domain of the $\mathrm{SL}_2(\mathbb{Z})$-orbit of τ.

The j-function $j : \mathbb{H} \to \mathbb{C}$ is $\mathrm{SL}_2(\mathbb{Z})$-invariant, and satisfies a q-expansion of the form

$$j(\tau) = q^{-1} + 744 + 196884q + \cdots, \quad \text{where} \quad q = e^{2\pi i \tau}.$$

We will also consider the modular discriminant function $\Delta : \mathbb{H} \to \mathbb{C}$, which is a cusp form of weight 12 for $\mathrm{SL}_2(\mathbb{Z})$, and we choose to normalise it such that its q-expansion is

$$\Delta(\tau) = q \prod_{n=1}^{\infty} (1-q^n)^{24} = q - 24q^2 + 252q^3 + \cdots. \tag{5.4}$$

This modular form plays a key role in this paper. Let us start by computing in the next lemma two special values which will be used in the sequel.

Lemma 5.4. *Let Δ be the discriminant modular form, normalized as in (5.4). We have*

(a) $\Delta(\rho) = -\dfrac{3^3}{(2\pi)^{24}} \Gamma\left(\dfrac{1}{3}\right)^{36}$, *where $\rho = e^{\frac{i\pi}{3}}$ and Γ stands for Euler's Gamma function, and*

(b) $\Delta(i) = \dfrac{1}{2^{24}\pi^{18}} \Gamma\left(\dfrac{1}{4}\right)^{24}$.

Proof. Let us start with (a). We have classically $(2\pi)^{12}\Delta(\rho) = g_2(\rho)^3 - 27 g_3(\rho)^2$, with g_2, g_3 the normalized Eisenstein series, and $g_2(\rho) = 0$ is a direct computation. For the value $g_3(\rho)$, we work with the elliptic curve in complex Weierstrass form $y^2 = 4x^3 - 4$, which has period lattice $\Lambda = \omega\mathbb{Z} + \rho\omega\mathbb{Z}$, with period

$$\omega = 2\int_1^{+\infty} \frac{dt}{\sqrt{4t^3-4}} = \int_1^{+\infty} \frac{dt}{\sqrt{t^3-1}} = \frac{1}{3} B\left(\frac{1}{6}, \frac{1}{2}\right) = \frac{\Gamma(\frac{1}{3})^3}{2^{\frac{4}{3}}\pi},$$

where $B(.,.)$ is the Euler B function, as classically defined for any complex numbers z_1, z_2 with positive real part by

$$B(z_1, z_2) := \int_0^1 t^{z_1-1}(1-t)^{z_2-1} dt = \frac{\Gamma(z_1)\Gamma(z_2)}{\Gamma(z_1+z_2)}.$$

Writing in generic Weierstrass form $4x^3-4=4x^3-g_2x-g_3$, we simply read off $g_2(\Lambda)=0$ and $g_3(\Lambda)=4$. We can now compute $g_3(\Lambda)=\omega^{-6}g_3(\mathbb{Z}+\rho\mathbb{Z})=\omega^{-6}g_3(\rho)$, hence $g_3(\rho)^2=4^2\omega^{12}$, which gives the claim for $\Delta(\rho)$.

We treat part (b) similarly: $(2\pi)^{12}\Delta(i)=g_2(i)^3-27g_3(i)^2$, and $g_3(i)=0$ is a direct computation. For the value $g_2(i)$, we work with the elliptic curve in complex Weierstrass form $y^2=4x^3-4x$, which has period lattice $\Lambda=\omega_0\mathbb{Z}+i\omega_0\mathbb{Z}$, with period

$$\omega_0=2\int_1^{+\infty}\frac{dt}{\sqrt{4t^3-4t}}=\int_1^{+\infty}\frac{dt}{\sqrt{t^3-t}}=\frac{1}{2}B\left(\frac{1}{4},\frac{1}{2}\right)=\frac{\Gamma(\frac{1}{4})^2}{2^{\frac{3}{2}}\pi^{\frac{1}{2}}}.$$

Writing in generic Weierstrass form $4x^3-4x=4x^3-g_2x-g_3$, we read off $g_3(\Lambda)=0$ and $g_2(\Lambda)=4$. We can now compute $g_2(\Lambda)=\omega_0^{-4}g_2(\mathbb{Z}+i\mathbb{Z})=\omega_0^{-4}g_2(i)$, hence $g_2(i)^3=4^3\omega_0^{12}$, which gives the claim for $\Delta(i)$. □

Remark 5.5. *From old work of Hurwitz, one can also derive another expression of $\Delta(i)$ using another period. From equation (7) page 201 of [Hur98] we get*

$$\Delta(i)=\frac{2^{18}}{(2\pi)^{12}}\left(\int_0^1\frac{dt}{\sqrt{1-t^4}}\right)^{12}.$$

Our first analytical tool is the following result, which is a refinement of (3.18) of [Paz19a].

Lemma 5.6. *Let $f(\tau)=\log\max\left\{|\Delta(\tau)|,|j(\tau)\Delta(\tau)|\right\}$. Then for all $\tau\in\mathcal{F}$,*

$$-5.5335<f(\tau)\leq f(i)=\log\left(\frac{3^3}{2^{18}\pi^{18}}\Gamma\left(\frac{1}{4}\right)^{24}\right)<1.1266.$$

Proof. We have

$$j(\tau)=\frac{g_2(\tau)^3}{(2\pi)^{12}\Delta(\tau)},$$

where $g_2(\tau)$ is again the normalized Eisenstein series of weight 4. Thus

$$f(\tau)=\begin{cases}\log|\Delta(\tau)| & \text{if } |j(\tau)|<1\\ 3\log|g_2(\tau)|-12\log(2\pi) & \text{if } |j(\tau)|\geq 1.\end{cases}$$

The boundary of $\mathcal{F}$ consists of a circular arc C from ρ to ρ^2, where $\rho=e^{\frac{i\pi}{3}}$, as well as the two vertical half-lines L from ρ to $i\infty$ and L' from ρ^2 to $i\infty$.

Since $j(\tau)$ has simple zeros at ρ and ρ^2, and no other zeroes near $\mathcal{F}$, one finds that $|j(\tau)|\leq 1$ in small neighbourhoods of these two points. Their intersection with $\mathcal{F}$ consists of two connected components, $D\cup D'=\{\tau\in\mathcal{F}\ :\ |j(\tau)|\leq 1\}$, where $\rho\in D$ and $\rho^2\in D'$.

By the Maximum Modulus Principle, $f(\tau)$ attains its extrema either at the cusp $i\infty$ or on the boundary components L', C, L, ∂D and $\partial D'$.

Using SageMath [Sage], we computed f restricted to these boundary components. The computations can be found on the GitHub repository `https://github.com/florianbreuer/ModularPolynomials`.

The results are symmetric around the imaginary axis, so Figure 5.1 shows the plot of $f(\tau)$ for τ on the contour from i via ρ to $i\infty$, as well as on ∂D.

We find that f attains its maximum at $f(i)<1.1266$ and its minimum >-5.5335 where ∂D meets L. At the cusp, $f(i\infty)=0$, which lies between these two extreme values. The formula for $f(i)$ comes directly from Lemma 5.4.

□

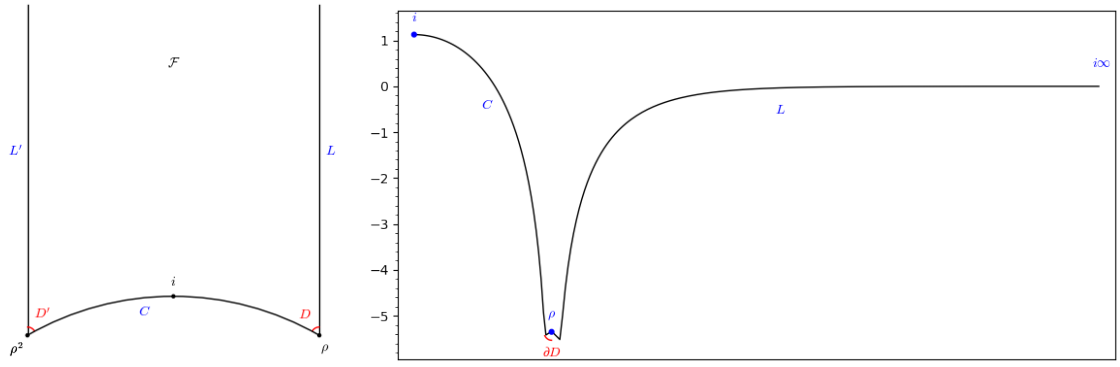


Figure 5.1: *(Left)* The fundamental domain $\mathcal{F}$. *(Right)* Plot of $f(\tau)$ for $\tau \in C \cup L \cup \partial D$ and $\mathrm{Re}(\tau) \geq 0$.

Remark 5.7. *Repeating the computations in [BP24] using the upper bound in Lemma 5.6 instead of [BP24, (13)], we obtain in the following corollary a slight improvement on the constant in (5.3).*

Corollary 5.8. *Let $N \geq 2$. the height of the modular polynomial $\Phi_N(X, Y)$ is bounded by*

$$h(\Phi_N) \leq 6\psi(N)\big[\log N - 2\lambda_N + \log\log N + 4.238\big].$$

□

5.3 Proof of the upper bound in Theorem 5.1

5.3.1 Strategy of proof.

Let us start by denoting, for $N \geq 1$,

$$C_N = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \ : \ a, b, d \in \mathbb{Z},\ ad = N,\ a \geq 1,\ 0 \leq b \leq d-1,\ \gcd(a,b,d) = 1 \right\}.$$

We have

$$\#C_N = \sum_{d|N} \sum_{\substack{0 \leq b < d \\ (b,r)=1}} 1 = \sum_{d|N} \frac{d\varphi(r)}{r} = \psi(N),$$

where we denote the gcd $r = (d, \frac{N}{d})$ for each d.

The relevance of the matrices in C_N is the following. For each $\gamma \in C_N$ and $\tau \in \mathbb{H}$, define

$$\tau_\gamma := \gamma(\tau) = \frac{a_\gamma \tau + b_\gamma}{d_\gamma}.$$

Then the elliptic curves $\mathbb{C}/\tau\mathbb{Z} + \mathbb{Z}$ and $\mathbb{C}/\tau_\gamma\mathbb{Z} + \mathbb{Z}$ are linked by a cyclic isogeny of degree N. Conversely, up to isomorphism, all cyclic N-isogenies from $\mathbb{C}/\tau\mathbb{Z} + \mathbb{Z}$ are obtained this way. In particular, the modular polynomial $\Phi_N(X, Y)$ satisfies

$$\Phi_N(X, j(\tau)) = \prod_{\gamma \in C_N} (X - j(\tau_\gamma)).$$

An interpolation argument (see Lemma 5.14) allows us to estimate the height of $\Phi_N(X,Y)$ in terms of the heights of the specialised polynomials $\Phi_N(X,j(\tau))$ for suitable values of $\tau\in\mathbb{H}$. These, in turn, are related to their logarithmic Mahler measures:

$$S_N(\tau) := m\big(\Phi_N(X,j(\tau))\big) = \sum_{\gamma\in C_N} \log\max\big\{1,|j(\tau_\gamma)|\big\}.$$

This Mahler measure is now our top priority. Let us work on the formula defining $S_N(\tau)$ and start with equation (14) from [BP24], valid for any $N\geq 1$ and $\tau\in\mathbb{H}$:

$$S_N(\tau) = \sum_{\gamma\in C_N} \log\max\{|\Delta(\tilde\tau_\gamma)|,|j(\tau_\gamma)\Delta(\tilde\tau_\gamma)|\}+6\sum_{\gamma\in C_N}\big[\log\operatorname{Im}\tilde\tau_\gamma-\log\operatorname{Im}\tau_\gamma\big]-\psi(N)\log|\Delta(\tau)|. \tag{5.5}$$

Recall that here $\tilde\tau_\gamma\in\mathcal{F}$ denotes the representative of τ_γ in the fundamental domain $\mathcal{F}$. We invoke [Aut03, Lemme 2.3]:

$$\sum_{\gamma\in C_N}\log\frac{d_\gamma}{a_\gamma} = \psi(N)(\log N - 2\lambda_N), \tag{5.6}$$

which combined with

$$\operatorname{Im}\tau_\gamma = \operatorname{Im}\left(\frac{a_\gamma\tau+b_\gamma}{d_\gamma}\right) = \frac{a_\gamma}{d_\gamma}\operatorname{Im}\tau$$

gives

$$-\sum_{\gamma\in C_N}\log\operatorname{Im}\tau_\gamma = \psi(N)\big(\log N - 2\lambda_N - \log\operatorname{Im}\tau\big). \tag{5.7}$$

Inject equality (5.7) in equation (5.5) and use the upper bound from Lemma 5.6 (note that $j(\tau_\gamma)=j(\tilde\tau_\gamma)$) to get:

$$\begin{aligned} S_N(\tau) &= \sum_{\gamma\in C_N}\log\max\{|\Delta(\tilde\tau_\gamma)|,|j(\tilde\tau_\gamma)\Delta(\tilde\tau_\gamma)|\} + 6\psi(N)\big[\log N - 2\lambda_N\big] \qquad (5.8)\\ &\quad + 6\sum_{\gamma\in C_N}\log\operatorname{Im}\tilde\tau_\gamma - \psi(N)\log\big[|\Delta(\tau)|(\operatorname{Im}\tau)^6\big]\\ &\leq 6\psi(N)\big[\log N - 2\lambda_N + 0.1878\big] + 6\sum_{\gamma\in C_N}\log\operatorname{Im}\tilde\tau_\gamma - \psi(N)\log\big[|\Delta(\tau)|(\operatorname{Im}\tau)^6\big]. \qquad (5.9)\end{aligned}$$

Our strategy is to set $\tau=iy$ with $y\geq 1$ and obtain an explicit upper bound for the sum $\sum_{\gamma\in C_N}\log\operatorname{Im}\tilde\tau_\gamma$, which we will decompose into a sum with large d and a sum with small d:

$$\sum_{\gamma\in C_N}\log\operatorname{Im}\tilde\tau_\gamma = \sum_{\substack{\gamma\in C_N\\ d_\gamma\geq\sqrt{Ny}}}\log\operatorname{Im}\tilde\tau_\gamma + \sum_{\substack{\gamma\in C_N\\ d_\gamma<\sqrt{Ny}}}\log\operatorname{Im}\tilde\tau_\gamma. \tag{5.10}$$

Our strategy is inspired by [Coh84], where the author uses a similar decomposition.

5.3.2 Large d

Consider $\gamma \in C_N$ for which $d = d_\gamma \geq \sqrt{Ny}$. As in [Coh84], we will approximate $\tilde{\tau}_\gamma$ with a representative $\hat{\tau}_\gamma \in \mathrm{SL}_2(\mathbb{Z})\tau_\gamma$ satisfying $\operatorname{Im}\hat{\tau}_\gamma \geq \frac{1}{2}$.

We start with the following lemma, which relies on the Farey sequence of order M.

Lemma 5.9. *Let $M \geq 1$ be an integer. Then one can express the interval*

$$I_M = \left[\frac{1}{M+1}, \frac{M+2}{M+1}\right) = \bigcup_{k=1}^{M} \bigcup_{\substack{h=1 \\ (h,k)=1}}^{k} I_M\left(\frac{h}{k}\right),$$

as a disjoint union of intervals $I_M\left(\frac{h}{k}\right)$ of the form $[\rho_1, \rho_2)$ containing $\frac{h}{k}$ and such that

$$\frac{1}{2Mk} \leq \frac{h}{k} - \rho_1 \leq \frac{1}{(M+1)k},$$

$$\frac{1}{2Mk} \leq \rho_2 - \frac{h}{k} \leq \frac{1}{(M+1)k}.$$

Proof. This is [Coh84, Lemma 3]. □

Recall that $\tau = iy$ with $y \geq 1$ and $d \geq \sqrt{Ny}$. Then we set

$$M := \left\lfloor \frac{d}{\sqrt{Ny}} \right\rfloor \geq 1.$$

Let $\gamma = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in C_N$. If $\frac{b}{d} \in [0, \frac{1}{M+1})$ then we replace b by $b + d$; this merely has the effect of replacing τ_γ by $\tau_\gamma + 1$, which is in the same $\mathrm{SL}_2(\mathbb{Z})$-orbit.

Next, choose a matrix $\delta = \begin{pmatrix} s & u \\ k & -h \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ for which $\frac{b}{d} \in I_M(\frac{h}{k})$ and define

$$\hat{\tau}_\gamma := \delta(\gamma(\tau)).$$

The entries s and u may be chosen in such a way (multiplying δ by a suitable translation matrix) that $-\frac{1}{2} < \operatorname{Re}(\hat{\tau}_\gamma) \leq \frac{1}{2}$.

Lemma 5.10. *The elements $\hat{\tau}_\gamma$ constructed above satisfy the following estimates.*

(a) $\operatorname{Im}\hat{\tau}_\gamma \geq \dfrac{1}{2}$,

(b) $\log \operatorname{Im}\hat{\tau}_\gamma \leq \log \dfrac{d^2}{Nyk^2}$, *and*

(c) $\log \operatorname{Im}\tilde{\tau}_\gamma \leq \log \operatorname{Im}\hat{\tau}_\gamma + \log 4$.

Proof. We compute

$$\operatorname{Im}\hat{\tau}_\gamma = \frac{Ny}{d^2k^2} \cdot \frac{1}{\left(\frac{Ny}{d^2}\right)^2 + \left(\frac{b}{d} - \frac{h}{k}\right)^2} = \frac{d^2}{Nyk^2} \cdot \frac{1}{1 + \frac{\left(\frac{b}{d} - \frac{h}{k}\right)^2}{\left(\frac{Ny}{d^2}\right)^2}}. \tag{5.11}$$

and so

$$\log \operatorname{Im} \hat{\tau}_\gamma = \log \frac{d^2}{Nyk^2} - \log \left(1 + \frac{\left(\frac{b}{d} - \frac{h}{k}\right)^2}{\left(\frac{Ny}{d^2}\right)^2} \right),$$

It follows that

$$\log \operatorname{Im} \hat{\tau}_\gamma \leq \log \frac{d^2}{Nyk^2}.$$

We also have $|\frac{b}{d} - \frac{h}{k}| \leq \frac{\sqrt{Ny}}{dk}$, so

$$0 \leq \frac{\left(\frac{b}{d} - \frac{h}{k}\right)^2}{\left(\frac{Ny}{d^2}\right)^2} \leq \frac{Ny}{d^2k^2} \cdot \frac{d^4}{N^2y^2} = \frac{d^2}{Nyk^2}.$$

Furthermore, as $\frac{d^2}{Nyk^2} \geq \frac{M^2}{k^2} \geq 1$, we also find in equation (5.11)

$$\operatorname{Im} \hat{\tau}_\gamma \geq \frac{1}{2}.$$

Finally, combining this with $-\frac{1}{2} < \operatorname{Re} \hat{\tau}_\gamma < \frac{1}{2}$ it follows that

$$\hat{\tau}_\gamma \in \mathcal{F} \cup S\mathcal{F} \cup ST^{-1}\mathcal{F} \cup ST\mathcal{F},$$

where

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

are the standard generators of $\mathrm{SL}_2(\mathbb{Z})$. In particular, we find

$$\log \operatorname{Im} \tilde{\tau}_\gamma \leq \log \operatorname{Im} \hat{\tau}_\gamma + \log 4.$$

□

Now we estimate the sum in (5.10) for those $\gamma \in C_N$ with $d \geq \sqrt{Ny}$. We note that

$$2 \log M \leq 2 \log d - \log(Ny).$$

Lemma 5.11. *Suppose $\tau = iy$ with $y \geq 1$ and $N \geq 1$. Then*

$$\sum_{\substack{\gamma \in C_N \\ d_\gamma \geq \sqrt{Ny}}} \log \operatorname{Im} \tilde{\tau}_\gamma \leq \left(4.75 + 3.5 \log 2 + \frac{0.5 + \log 2}{2\sqrt{N}} \right) \psi(N).$$

Proof. Let us start with

$$\sum_{\substack{\gamma \in C_N \\ d_\gamma \geq \sqrt{Ny}}} \log \operatorname{Im} \tilde{\tau}_\gamma \leq \sum_{\substack{d \mid N \\ d \geq \sqrt{Ny}}} \sum_{\substack{0 \leq b < d \\ (b,r)=1}} \left(\log \operatorname{Im} \hat{\tau}_\gamma + \log 4 \right) \tag{5.12}$$

$$\leq \left(\sum_{\substack{d \mid N \\ d \geq \sqrt{Ny}}} \sum_{\substack{0 \leq b < d \\ (b,r)=1}} \log \operatorname{Im} \hat{\tau}_\gamma \right) + \log(4)\psi(N),$$

as the number of terms in the sum is bounded by $\#C_N = \psi(N)$. By Lemma 5.11,

$$\sum_{\substack{d|N \\ d\geq\sqrt{Ny}}} \sum_{\substack{0\leq b<d \\ (b,r)=1}} \log \operatorname{Im} \hat{\tau}_\gamma \leq \sum_d \sum_b \log \frac{d^2}{Nyk^2} \tag{5.13}$$

$$= \sum_d \sum_{k=1}^{M} \underbrace{\sum_{h=1}^{k} \sum_{\frac{b}{d}\in I_M(\frac{h}{k})}}_{(i)} \left[2\log\frac{d}{k} - \log(Ny)\right].$$

Let us bound the number of terms in the sum (i) above. By Lemma 5.9, the length of $I_M(\frac{h}{k})$ is bounded by $\frac{2}{(M+1)k}$, hence the number of terms in the inner sum is bounded by the number of integers b with $(b,r)=1$ in $dI_M(\frac{h}{k})$, for fixed k and $h = 1, \dots k$. For an interval of length r, we have $\varphi(r)$ integers coprime with r. Therefore[2], for fixed d and k,

$$\#\left\{\frac{b}{d} \in I_M(\frac{h}{k}),\ 0\leq b<d,\ (b,r)=1\right\} \leq \varphi(r)\left\lceil \frac{1}{r}\frac{2d}{(M+1)k}\right\rceil \leq \varphi(r)\left(\frac{2d}{(M+1)kr}+1\right). \tag{5.14}$$

Summing over $1 \leq h \leq k$, we bound the number of terms in the sum (i) by

$$\frac{2d\varphi(r)}{(M+1)r} + k\varphi(r).$$

Hence we split the sum in Equation (5.13) in two:

$$\sum_{\substack{d|N \\ d\geq\sqrt{Ny}}} \sum_{\substack{0\leq b<d \\ (b,r)=1}} \log \operatorname{Im} \hat{\tau}_\gamma \tag{5.15}$$

$$\leq \sum_d \sum_{k=1}^{M} \frac{2d\varphi(r)}{(M+1)r}\left[2\log\frac{d}{k} - \log(Ny)\right] + \sum_d \sum_{k=1}^{M} k\varphi(r)\left[2\log\frac{d}{k} - \log(Ny)\right]$$

[2]The referee pointed out the stronger upper bound $\frac{\varphi(r)}{r}\frac{2d}{(M+1)k} + \varphi(r)(1-\frac{\varphi(r)-1}{r})$, which didn't lead to manageable expressions.

We deal with the first sum in the right hand side of inequality (5.15).

$$\begin{aligned}
&\sum_d \sum_{k=1}^{M} \frac{2d\varphi(r)}{(M+1)r}\left[2\log\frac{d}{k} - \log(Ny)\right] \qquad (5.16)\\
&= \sum_d \frac{2d\varphi(r)}{(M+1)r}\left[2M\log d - 2\sum_{k=1}^{M}\log k - M\log(Ny)\right]\\
&\le \sum_d \frac{2d\varphi(r)}{r}\frac{M}{M+1}\left[2 + \log\frac{d^2}{M^2Ny} - \frac{1}{6M(M+1)} - 2\frac{\log(\sqrt{2\pi M})}{M}\right] \quad (ii)\\
&\le \sum_d \frac{2d\varphi(r)}{r}\frac{M}{M+1}\left[2 + 2\log\frac{M+1}{M} - \frac{1}{6M(M+1)} - \frac{\log(2\pi M)}{M}\right] \quad (iii)\\
&\le \sum_{d|N} \frac{2d\varphi(r)}{r} 2 \quad (iv)\\
&= 4\psi(N),
\end{aligned}$$

where to reach (ii) we used

$\sum_{k=1}^{M}\log k = \log(M!)$ and by [Rob55] we have for any integer $M \ge 1$:

$$\sqrt{2\pi M}\left(\frac{M}{e}\right)^M e^{\frac{1}{12(M+1)}} \le M! \le \sqrt{2\pi M}\left(\frac{M}{e}\right)^M e^{\frac{1}{12M}}$$

so in particular

$$M\log M - M + \log\sqrt{2\pi M} + \frac{1}{12(M+1)} \le \sum_{k=1}^{M}\log k,$$

and in (iii) we used the fact that $M \le \frac{d}{\sqrt{Ny}} \le M+1$ implies

$$1 \le \frac{d^2}{M^2Ny} \le \left(\frac{M+1}{M}\right)^2,$$

and the inequality (iv) holds because for any $M \ge 1$ we have

$$\frac{M}{M+1}\left[2 + 2\log\frac{M+1}{M} - \frac{1}{6M(M+1)} - \frac{\log(2\pi M)}{M}\right] \le 2,$$

which can be verified through direct computation.

Let us bound the second sum in the right hand side of inequality (5.15).

$$\begin{aligned}
&\sum_{\substack{d|N,\\ d\ge\sqrt{Ny}}} \sum_{k=1}^{M} k\varphi(r)\left(2\log\frac{d}{\sqrt{Ny}} - 2\log k\right)\\
&= \sum_{\substack{d|N,\\ d\ge\sqrt{Ny}}} \left(\varphi(r)M(M+1)\log\frac{d}{\sqrt{Ny}} - 2\varphi(r)\sum_{k=1}^{M} k\log k\right).
\end{aligned}$$

We bound $-\sum_{k=1}^{M} k\log k$ as follows. By Abel's summation formula,

$$\sum_{k=1}^{M} k\log k = \frac{M(M+1)}{2}\log M - \int_1^M \frac{\lfloor u\rfloor(\lfloor u\rfloor+1)}{2}\frac{1}{u}du,$$

hence, as $\frac{\lfloor u\rfloor}{u}\leq 1$,

$$\begin{aligned}-2\sum_{k=1}^{M} k\log k &= -M(M+1)\log M + \int_1^M \frac{\lfloor u\rfloor(\lfloor u\rfloor+1)}{u}du\\ &\leq -M(M+1)\log M + \frac{M(M+1)}{2} - 1.\end{aligned}$$

We set $\tilde{M} := \frac{d}{\sqrt{Ny}}$, so $M\leq\tilde{M}\leq M+1$. Therefore,

$$\begin{aligned}&\sum_{d|N,\ d\geq\sqrt{Ny}}\left(\varphi(r)M(M+1)\log\frac{d}{\sqrt{Ny}} - 2\varphi(r)\sum_{k=1}^{M}k\log k\right)\\
&\leq \sum_{d|N,\ d\geq\sqrt{Ny}}\varphi(r)\left(M(M+1)\log\left(\frac{M+1}{M}\right)+\frac{M(M+1)}{2}-1\right)\\
&\leq \sum_{d|N,\ d\geq\sqrt{Ny}}\varphi(r)\left(\tilde{M}(\tilde{M}+1)\log 2+\frac{\tilde{M}(\tilde{M}+1)}{2}\right)\\
&= \sum_{d|N,\ d\geq\sqrt{Ny}}\varphi(r)\left(\left(\log 2+\frac{1}{2}\right)\tilde{M}(\tilde{M}+1)\right) \qquad (5.17)\\
&\leq \left(\log 2+\frac{1}{2}\right)\left(\sum_{d|N}\varphi(r)\frac{d^2}{Ny}+\sum_{d|N,\ d\geq\sqrt{N}}\varphi(r)\frac{d}{\sqrt{Ny}}\right) \qquad (5.18)\\
&\leq \left(\log 2+\frac{1}{2}\right)\left(\frac{3}{2}+\frac{1}{2\sqrt{N}}\right)\psi(N). \qquad (5.19)\end{aligned}$$

For the first sum in (5.18), remark that $a=\frac{N}{d}$ also runs through the divisors of N and that $r=(d,a)$, hence $a\geq r$, and

$$\sum_{d|N}\varphi(r)\frac{d^2}{Ny} = \frac{1}{y}\sum_{a|N}\varphi(r)\frac{N}{a^2}\leq\frac{1}{y}\sum_{a|N}\frac{\varphi(r)}{r}\frac{N}{a} = \frac{1}{y}\sum_{d|N}\frac{\varphi(r)}{r}d = \frac{1}{y}\psi(N)\leq\psi(N).$$

For the second sum in (5.18), if we also set $\varphi(x)=0$ if $x\notin\mathbb{N}$, we get

$$\sum_{d|N,\ d\geq\sqrt{N}}\varphi(r)\frac{d}{\sqrt{Ny}} = \frac{1}{\sqrt{y}}\sum_{a|N,\ a\leq\sqrt{N}}\varphi(r)\frac{\sqrt{N}}{a}\leq\sqrt{N}\sum_{a\leq\sqrt{N}}\frac{\varphi(r)}{r}\leq\frac{\psi(N)+\varphi(\sqrt{N})}{2}. \qquad (5.20)$$

The last inequality in (5.20) comes from

$$\begin{aligned}\psi(N) &= \sum_{d>\sqrt{N}} \frac{\varphi(r)}{r} d + \sum_{d<\sqrt{N}} \frac{\varphi(r)}{r} d + \frac{\varphi(\sqrt{N})}{\sqrt{N}} \sqrt{N} \\ &= \sum_{d\le\sqrt{N}} \frac{\varphi(r)}{r} \left(d + \frac{N}{d}\right) - \varphi(\sqrt{N}) \ge 2\sqrt{N} \sum_{d\le\sqrt{N}} \frac{\varphi(r)}{r} - \varphi(\sqrt{N})\end{aligned}$$

as $(d + \frac{N}{d}) \ge 2\sqrt{N}$ for any $1 \le d \le N$. Notice also that $\frac{\varphi(\sqrt{N})}{\psi(N)} \le \frac{1}{\sqrt{N}}$, because $\varphi(\sqrt{N}) \le \sqrt{N}$ and $\psi(N) \ge N$. Equation (5.19) now follows.

This finishes the proof. □

Remark 5.12. *We remark that we can obtain the slightly worse bound* $(8 + 2\log 2)\psi(N)$ *in Lemma 5.11 with a simpler argument. With the notations in (5.14), it can be checked that the following inequality is true,*

$$\frac{2d}{(M+1)kr} \ge 1$$

for any $1 \le k \le M$*. This implies that the second sum in (5.15) is bounded by the first, so we could use the bound of Equation (5.16) for both of them.*

5.3.3 Small d

Now we consider the sum over $\gamma \in C_N$ with $d_\gamma < \sqrt{Ny}$.

Lemma 5.13. *Let* $\tau = iy$ *with* $y \ge 1$ *and* $N \ge 1$*. Then we have*

$$\sum_{\substack{\gamma\in C_N \\ d_\gamma<\sqrt{Ny}}} \log \operatorname{Im} \tilde{\tau}_\gamma \le \psi(N) \left(\frac{1}{e} + \log \operatorname{Im} \tau\right).$$

Proof. In this case $\operatorname{Im} \tau_\gamma > 1$, so $\operatorname{Im} \tilde{\tau}_\gamma = \operatorname{Im} \tau_\gamma$. We write $a = \frac{N}{d}$ and compute

$$\begin{aligned}\sum_{\substack{\gamma\in C_N \\ d_\gamma<\sqrt{Ny}}} \log \operatorname{Im} \tau_\gamma &= \sum_{\substack{d|N \\ d<\sqrt{Ny}}} \sum_{\substack{b<d \\ (b,r)=1}} \log \frac{ay}{d} \\ &= \sum_{\substack{d|N \\ d<\sqrt{Ny}}} \frac{d\varphi(r)}{r} \left[\log \frac{a}{d} + \log y\right] \\ &\le \psi(N) \log y + \sum_{\substack{d|N \\ d<\sqrt{Ny}}} \frac{d\varphi(r)}{r} \log \frac{a}{d}.\end{aligned}$$

The crude estimate $\log \frac{a}{d} \le \frac{1}{e} \frac{a}{d}$ (which holds as $\frac{\log x}{x}$ has a maximum at $x = e$ for $x > 0$) gives us

$$\sum_{\substack{d|N \\ d<\sqrt{Ny}}} \frac{d\varphi(r)}{r} \log \frac{a}{d} \le \frac{1}{e} \sum_{a|N} \frac{a\varphi(r)}{r} = \frac{1}{e}\psi(N). \tag{5.21}$$

As $y = \operatorname{Im} \tau$, this concludes the proof. □

5.3.4 Final steps of the proof

As shown in [BP24], computations by Andrew Sutherland confirm Theorem 5.1 for $N \leq 400$, so we may assume $N \geq 401$. In this case, the coefficient in Lemma 5.11 is

$$4.75 + 3.5 \log 2 + \frac{0.5 + \log 2}{2\sqrt{N}} < 7.2059.$$

Adding the bounds in Lemma 5.11 and Lemma 5.13, we obtain

$$\sum_{\gamma \in C_N} \log \operatorname{Im} \tilde{\tau}_\gamma \leq \psi(N)(7.5737 + \log \operatorname{Im} \tau).$$

We choose $\tau = iy$ such that $j(\tau) \in [1728, 3456]$, so $1 \leq y < 1.2536$, for which we compute (using SageMath [Sage])

$$-\log\left[|\Delta(\tau)|(\operatorname{Im}\tau)^6\right] \leq 6.5296,$$

and so in equation (5.9) we have

$$\begin{aligned} S_N(\tau) &\leq 6\psi(N)\big[\log N - 2\lambda_N + 0.1878\big] + 6\sum_{\gamma \in C_N} \log \operatorname{Im} \tilde{\tau}_\gamma - \psi(N) \log\left[|\Delta(\tau)|(\operatorname{Im}\tau)^6\right] \\ &\leq 6\psi(N)\big[\log N - 2\lambda_N + 9.0756\big]. \end{aligned}$$

We add a classical interpolation lemma.

Lemma 5.14. *Let $N \geq 1$. For any real $L > 1$,*

$$h(\Phi_N) \leq \max_{L \leq j(\tau) \leq 2L} S_N(\tau) + \psi(N)\left(\frac{1 + \log L}{L} + 4 \log 2\right).$$

Proof. This is obtained in [BP24] equation (19), and comes from Lemma 10 in [BS10]. □

We can finally use Lemma 5.14 with $L = 1728$ (corresponding to the smallest permissible value of y, which gives the best constants), and obtain

$$\begin{aligned} h(\Phi_N) &\leq \max_{1728 \leq j(\tau) \leq 3456} S_N(\tau) + \psi(N)\left(\frac{1 + \log 1728}{1728} + 4 \log 2\right) \\ &\leq 6\psi(N)\big[\log N - 2\lambda_N + 9.5387\big]. \end{aligned}$$

This concludes the proof of the upper bound in Theorem 5.1.

5.4 Proof of the lower bound in Theorem 5.1

We now turn to the lower bound in Theorem 5.1. For any τ in the complex upper half plane, recall that the logarithmic Mahler measure of $\Phi_N(X, j(\tau))$ is equal to

$$m\big(\Phi_N(X, j(\tau))\big) = S_N(\tau) = \sum_{\gamma \in C_N} \log \max\{1, |j(\tau_\gamma)|\}.$$

We start with an upper bound that will be used later in the proof.

Lemma 5.15. *For every $N \geq 1$ we have*

(a) $S_N(\tau) \le 2\log(\psi(N)+1) + \psi(N)\log\max\{1,|j(\tau)|\} + h(\Phi_N)$.

(b) $S_N(\rho) \le \log(\psi(N)+1) + h(\Phi_N)$, *where* $\rho = e^{\frac{i\pi}{3}}$.

Proof. Write $\Phi_N(X,Y) = \sum_{k=0}^{\psi(N)} P_k(Y)X^k$, where each $P_k(Y) \in \mathbb{Z}[Y]$ has degree $\le \psi(N)$. Denoting $H(P_k) = e^{h(P_k)}$ the maximum absolute value of the coefficients of P_k, we have

$$\left|P_k\big(j(\tau)\big)\right| \le (\psi(N)+1)\max\{1,|j(\tau)|\}^{\psi(N)}H(P_k) \le (\psi(N)+1)\max\{1,|j(\tau)|\}^{\psi(N)}H(\Phi_N).$$

Comparing the Mahler measure to the *length* of a polynomial [BZ20, Lemma 1.7], we get

$$\begin{aligned} S_N(\tau) = m\big(\Phi_N(X,j(\tau))\big) &\le \log\left[\sum_{k=0}^{\psi(N)}\left|P_k\big(j(\tau)\big)\right|\right] \\ &\le \log\left[(\psi(N)+1)^2\max\{1,|j(\tau)|\}^{\psi(N)}H(\Phi_N)\right]. \end{aligned} \tag{5.22}$$

This proves part (a).

Since $P_k(0)$ is the constant coefficient of $P_k(Y)$, we see that

$$\log|P_k(0)| \le h(P_k) \le h(\Phi_N).$$

As $j(\rho) = 0$, in this case (5.22) gives

$$S_N(\rho) \le \log\left[\sum_{k=0}^{\psi(N)}|P_k(0)|\right] \le \log(\psi(N)+1) + h(\Phi_N).$$

Part (b) follows. □

To obtain a lower bound on $h(\Phi_N)$, it is thus enough to bound $S_N(\rho)$ from below, which is the goal of the next lemma.

Lemma 5.16. *Let* $\rho = e^{\frac{i\pi}{3}}$. *Then for for any* $N \ge 1$,

$$S_N(\rho) \ge 6\psi(N)\left(\log N - 2\lambda_N - \frac{1}{6}\log\left|\frac{3^3}{(2\pi)^{24}}\Gamma\left(\frac{1}{3}\right)^{36}\right| - \frac{5.5335}{6}\right).$$

Proof. We bound the two sums in equation (5.8) for $S_N(\tau)$ from below, starting with Lemma 5.6 which gives us

$$\sum_{\gamma\in C_N}\log\max\{|\Delta(\tilde{\tau}_\gamma)|,|j(\tau_\gamma)\Delta(\tilde{\tau}_\gamma)|\} \ge -5.5335\psi(N). \tag{5.23}$$

Also, for any $\gamma \in C_N$, we have $\operatorname{Im}\tilde{\tau}_\gamma \ge \frac{\sqrt{3}}{2}$. We thus obtain

$$S_N(\tau) \ge -5.5335\psi(N) + 6\psi(N)(\log N - 2\lambda_N) + 6\psi(N)\log\frac{\sqrt{3}}{2} - \psi(N)\log|\Delta(\tau)(\operatorname{Im}\tau)^6|. \tag{5.24}$$

We will now specialize $\tau = \rho$. We obtain via Lemma 5.4 and equation (5.24):

$$S_N(\rho) \ge -5.5335\psi(N) + 6\psi(N)(\log N - 2\lambda_N) - \psi(N)\log\left|\frac{3^3}{(2\pi)^{24}}\Gamma\left(\frac{1}{3}\right)^{36}\right|, \tag{5.25}$$

which leads to the claim. □

By combining Lemma 5.15(b) and Lemma 5.16, we finally obtain

$$h(\Phi_N) \geq -5.5335\psi(N) + 6\psi(N)(\log N - 2\lambda_N) - \psi(N)\log\left|\frac{3^3}{(2\pi)^{24}}\Gamma\left(\frac{1}{3}\right)^{36}\right| - \log(\psi(N)+1),$$

and

$$\frac{1}{6}\left(\log\left|\frac{3^3}{(2\pi)^{24}}\Gamma\left(\frac{1}{3}\right)^{36}\right| + \frac{\log(\psi(N)+1)}{\psi(N)} + 5.5335\right) \leq 0.0351$$

when $N \geq 401$. This proves the lower bound from Theorem 5.1 in the case $N \geq 401$, whereas the numerical computations by Andrew Sutherland (see [BP24]) show that the Theorem also holds when $N \leq 400$.

5.5 Explicit Hecke points estimates

So far, we have only obtained bounds on $\sum_{\gamma\in C_N} \log \operatorname{Im}\tilde{\tau}_\gamma$ for the special values $\tau = iy$. One can deduce a general bound from Theorem 5.1, which we record in the following result.

Proposition 5.17. *Let $\tau \in \mathcal{F}$. Then*

(a) $$\max\left\{\log\frac{\sqrt{3}}{2}, \log\operatorname{Im}\tau - \log N + 2\lambda_N\right\} \leq \frac{1}{\psi(N)}\sum_{\gamma\in C_N}\log\operatorname{Im}\tilde{\tau}_\gamma \leq 10.832 + \log\operatorname{Im}\tau.$$

(b) If $\operatorname{Im}\tau \geq N$, *then* $$\frac{1}{\psi(N)}\sum_{\gamma\in C_N}\log\operatorname{Im}\tilde{\tau}_\gamma = \log\operatorname{Im}\tau - \log N + 2\lambda_N.$$

Proof. For each $\gamma \in C_N$ we have $\operatorname{Im}\tilde{\tau}_\gamma \geq \operatorname{Im}\tau_\gamma$, and also $\operatorname{Im}\tilde{\tau}_\gamma \geq \frac{\sqrt{3}}{2}$. Thus

$$\log\operatorname{Im}\tilde{\tau}_\gamma \geq \max\left\{\log\operatorname{Im}\tau_\gamma, \log\frac{\sqrt{3}}{2}\right\}.$$

Now (5.7) implies the lower bound in part (a).

Furthermore, if $\operatorname{Im}\tau \geq N$, then $\operatorname{Im}\tilde{\tau}_\gamma = \operatorname{Im}\tau_\gamma$ for all $\gamma \in C_N$, and so (5.7) implies part (b).

We now prove the upper bound in part (a). From Lemma 5.15(a) we obtain

$$S_N(\tau) \leq 2\log(\psi(N)+1) + \psi(N)\log\max\{1, |j(\tau)|\} + h(\Phi_N).$$

Next, we replace the left hand side by (5.8) and extract

$$\begin{aligned}
&\frac{1}{\psi(N)}\sum_{\gamma\in C_N}\log\operatorname{Im}\tilde{\tau}_\gamma \qquad (5.26)\\
&\quad\leq \left[\frac{1}{6\psi(N)}h(\Phi_N) - \log N + 2\lambda_N\right]\\
&\quad+ \frac{1}{6}\left[\log\max\left\{|\Delta(\tau)|, |j(\tau)\Delta(\tau)|\right\} - \frac{1}{\psi(N)}\sum_{\gamma\in C_N}\log\max\left\{|\Delta(\tilde{\tau}_\gamma)|, |j(\tilde{\tau}_\gamma)\Delta(\tilde{\tau}_\gamma)|\right\}\right]\\
&\quad+ \log\operatorname{Im}\tau + \frac{\log(\psi(N)+1)}{3\psi(N)}.
\end{aligned}$$

Now Theorem 5.1 and Lemma 5.6 give us

$$\frac{1}{\psi(N)}\sum_{\gamma\in C_N}\log \operatorname{Im}\tilde{\tau}_\gamma \le 9.5387+\frac{1}{6}[1.1266+5.5335]+\frac{\log 3}{6}+\log\operatorname{Im}\tau.$$

The result follows. □

We now prove Theorem 5.2, which is an explicit version of Silverman's Theorem 5.1 page 417 of [Sil90].

Proof. (of Theorem 5.2) Fix N and E, and let K be a sufficiently large number field that E and every E/C as well as the isogenies linking them are defined over K.

It follows from [Sil90, Prop. 2] that only the infinite places contribute to the difference, so

$$\begin{aligned} &h_\infty(j_E)-\frac{1}{\psi(N)}\sum_{\substack{C\ cyclic\\ \#C=N}} h_\infty(j_{E/C}) \qquad (5.27)\\ =&\frac{1}{[K:\mathbb{Q}]}\sum_{\sigma:K\hookrightarrow\mathbb{C}}\left[\log\max\left\{1,|\sigma(j_E)|\right\}-\frac{1}{\psi(N)}\sum_{\substack{C\ cyclic\\ \#C=N}}\log\max\left\{1,|\sigma(j_{E/C})|\right\}\right]. \end{aligned}$$

Notice that the Hecke sum in [Sil90] is over all subgroups $C\subset E$ of order N, not just the cyclic ones, but the argument in [Sil90, Prop. 2] gives the same result in our situation.

Let $\tau_\sigma\in\mathcal{F}$ be such that $\sigma(j_E)=j(\tau_\sigma)$, then

$$\sum_{\substack{C\ cyclic\\ \#C=N}}\log\max\left\{1,|\sigma(j_{E/C})|\right\}=S_N(\tau_\sigma)=m\big(\Phi_N(X,\sigma(j_E))\big)$$

is the Mahler measure of $\Phi_N(X,j(\tau_\sigma))$. Now Lemma 5.15(a) gives

$$m\big(\Phi_N(X,\sigma(j_E))\big)\le 2\log(\psi(N)+1)+\psi(N)\log\max\{1,|\sigma(j_E)|\}+h(\Phi_N).$$

Part (a) now follows from Theorem 5.1 and the estimate $\dfrac{\log(\psi(N)+1)}{\psi(N)}\le\dfrac{\log 3}{2}$.

To show part (b), we write $\tau=\tau_\sigma\in\mathcal{F}$ and combine (5.8), Lemma 5.6 and Proposition

5.17:

$$\begin{aligned}
&\log\max\{1,|\sigma(j_E)|\} - \frac{1}{\psi(N)}\sum_{\substack{C \text{ cyclic}\\ \#C=N}} \log\max\{1,|\sigma(j_{E/C})|\} \qquad (5.28)\\
&= \log\max\{1,|j(\tau)|\} - \frac{1}{\psi(N)} S_N(\tau)\\
&= \left[\log\max\{|\Delta(\tau)|,|j(\tau)\Delta(\tau)|\} - \frac{1}{\psi(N)}\sum_{\gamma\in C_N}\log\max\{|\Delta(\tilde\tau_\gamma)|,|j(\tilde\tau_\gamma)\Delta(\tilde\tau_\gamma)|\}\right]\\
&\quad - \frac{6}{\psi(N)}\sum_{\gamma\in C_N}\log\operatorname{Im}\tilde\tau_\gamma + 6\big[\log\operatorname{Im}\tau - \log N + 2\lambda_N\big]\\
&\le [1.1266 + 5.5335]\\
&\quad - 6\max\{\log\frac{\sqrt{3}}{2}, \log\operatorname{Im}\tau - \log N + 2\lambda_N\} + 6\big[\log\operatorname{Im}\tau - \log N + 2\lambda_N\big]\\
&\le 6.6601 + 6\min\{-\log\frac{\sqrt{3}}{2} + \log\operatorname{Im}\tau_\sigma - \log N + 2\lambda_N, 0\}.
\end{aligned}$$

We now insert this into (5.27) and invoke [Paz19a, Lemma 2.6], which gives us

$$\begin{aligned}
&h_\infty(j_E) - \frac{1}{\psi(N)}\sum_{\substack{C \text{ cyclic}\\ \#C=N}} h_\infty(j_{E/C})\\
&\le 6.6601 + 6\min\big\{0, -\log\frac{\sqrt{3}}{2} + \log\big(1+h_\infty(j_E)\big) + 1.94 - \log 2\pi - \log N + 2\lambda_N\big\}.
\end{aligned}$$

This proves part (b) of Theorem 5.2.

□

Remark 5.18. *The inequalities (5.28) and (5.22) imply the following lower bound on the height of the specialised polynomial $\Phi_N(X,j)$, which can be seen as a measure of non-cancellation:*

$$\begin{aligned}
h(\Phi_N(X,j)) &\ge S_N(\tau) - \log(\psi(N)+1)\\
&\ge \psi(N)\left[\log\max\{1,|j(\tau)|\} - 6.6601\right] - \log(\psi(N)+1)\\
&\ge \psi(N)\left[\log\max\{1,|j(\tau)|\} - 7.2095\right]. \qquad (5.29)
\end{aligned}$$

Remark 5.19. *If $N \le \operatorname{Im}\tau_\sigma$ for every $\sigma : K \hookrightarrow \mathbb{C}$, then the above proof, together with Proposition 5.17(b) gives*

$$\left| h_\infty(j_E) - \frac{1}{\psi(N)}\sum_{\substack{C \text{ cyclic}\\ \#C=N}} h_\infty(j_{E/C}) \right| \le 6.6601.$$

Remark 5.20. *Theorem 5.2 may be regarded as a "Hecke-averaged" version of [Paz19a, Thm 1.1], with improved bounds.*

If we replace the Weil height of the j-invariant with the stable Faltings height (see Definition 5.21) of elliptic curves in Theorem 5.2, Autissier [Aut03, Cor. 3.3] obtained the even neater result:

$$\frac{1}{\psi(N)} \sum_{\substack{C \text{ cyclic} \\ \#C=N}} h_{\mathrm{Falt}}(E/C) = h_{\mathrm{Falt}}(E) + \frac{1}{2}\log N - \lambda_N.$$

5.6 What is the size of $X_0(N)$?

In this final section we give a proof of Theorem 5.3. We start with the first item and recall the definition of the Faltings height of an abelian variety and of a curve.

5.6.1 Faltings height and modular polynomials

Let A be a semi-stable abelian variety defined over a number field k, of dimension $g \geq 1$. Let $\pi\colon \mathcal{A} \longrightarrow \mathrm{Spec}(\mathcal{O}_k)$ be the Néron model of A over $\mathrm{Spec}(\mathcal{O}_k)$, where $\mathcal{O}_k$ is the ring of integers of k. Let $\varepsilon\colon \mathrm{Spec}(\mathcal{O}_k) \longrightarrow \mathcal{A}$ be the zero section of π and let $\omega_{\mathcal{A}/\mathcal{O}_k}$ be the maximal exterior power of the sheaf of relative differentials

$$\omega_{\mathcal{A}/\mathcal{O}_k} := \varepsilon^\star \Omega^g_{\mathcal{A}/\mathcal{O}_k}.$$

For any archimedean place v of k, let σ be an embedding of k in $\mathbb{C}$ associated to v. The associated line bundle

$$\omega_{\mathcal{A}/\mathcal{O}_k,\sigma} = \omega_{\mathcal{A}/\mathcal{O}_k} \otimes_{\mathcal{O}_k,\sigma} \mathbb{C} \simeq H^0(\mathcal{A}_\sigma(\mathbb{C}), \Omega^g_{\mathcal{A}_\sigma}(\mathbb{C}))$$

is equipped with a natural L^2-metric $\|.\|_v$ given by

$$\|s\|_v^2 = \frac{i^{g^2}}{(2\pi)^g} \int_{\mathcal{A}_\sigma(\mathbb{C})} s \wedge \overline{s}\ .$$

The $\mathcal{O}_k$-module $\omega_{\mathcal{A}/\mathcal{O}_k}$ is of rank 1 and together with the hermitian norms $\|.\|_v$ at infinity it defines an hermitian line bundle $\overline{\omega}_{\mathcal{A}/\mathcal{O}_k} = (\omega_{\mathcal{A}/\mathcal{O}_k}, (\|.\|_v)_{v\in M_k^\infty})$ over $\mathcal{O}_k$.

Recall that for any hermitian line bundle $\overline{\mathcal{L}}$ over $\mathrm{Spec}(\mathcal{O}_k)$ the Arakelov degree of $\overline{\mathcal{L}}$ is defined as

$$\widehat{\deg}(\overline{\mathcal{L}}) = \log \#\left(\mathcal{L}/s\mathcal{O}_k\right) - \sum_{v\in M_k^\infty} d_v \log \|s\|_v\ ,$$

where s is any non zero section of $\mathcal{L}$. The resulting real number does not depend on the choice of s in view of the product formula on the number field k.

The natural idea is then to consider $\widehat{\deg}(\overline{\omega}_{\mathcal{A}/\mathcal{O}_k})$. This Arakelov degree of the metrized bundle $\overline{\omega}_{\mathcal{A}/\mathcal{O}_k}$ will give a translate (by a term of the form gc_0 with c_0 an absolute constant) of the classical Faltings height.

Definition 5.21. The stable height of A is defined as

$$h_{\mathrm{Falt}}(A) := \frac{1}{[k:\mathbb{Q}]} \widehat{\deg}(\overline{\omega}_{\mathcal{A}/\mathcal{O}_k})\ .$$

In the same spirit, we can also define the Faltings height of a stable curve.

Definition 5.22. Let k be a number field and C/k a smooth algebraic curve defined over k, with semi-stable reduction and genus $g \geq 1$. Let $p : C \to S$ be a semi-stable integral model of C on $S = \mathrm{Spec}(\mathcal{O}_k)$. The Faltings height of C/k is the quantity

$$h_{\mathrm{Falt}}(C) = \frac{1}{[k:\mathbb{Q}]} \widehat{\deg}(\det p_* \omega_{C/S}),$$

where the hermitian metrics are chosen as $\|\alpha\|_v^2 = \frac{i^{g^2}}{(2\pi)^g} \int \alpha \wedge \overline{\alpha}$.

This height is often referred to as the *stable* height, as it is stable by extension of the base field k. The following proposition is well known to experts.

Proposition 5.23. *Let k be a number field and C/k a smooth algebraic curve defined over k, with semi-stable reduction and genus $g \geq 1$. Let J_C denote the Jacobian of C. Then we have*

$$h_{\mathrm{Falt}}(J_C) = h_{\mathrm{Falt}}(C).$$

Proof. See for instance Proposition 6.5 in [Paz19b]. □

By specializing to $X_0(N)$, we get $h_{\mathrm{Falt}}(X_0(N)) = h_{\mathrm{Falt}}(J_0(N))$. We now recall a result of Jorgenson and Kramer on the asymptotic of the Faltings height of the modular Jacobian.

Theorem 5.24. *(Theorem 6.2 page 36 of [JK09]) Let N be square-free and coprime to 6. Let $g(N)$ be the dimension of the abelian variety $J_0(N)$. When N tends to infinity, one has*

$$h_{\mathrm{Falt}}(J_0(N)) = \frac{g(N)}{3} \log N + o(g(N) \log N).$$

We now need an estimate on the size of $g(N)$ as a function of N. The formula for $g(N)$ is classical and estimates abound in the literature (see for instance [Paz10]), with various conditions on N. The following variant is most relevant to this paper. We give a proof here for the sake of completeness.

Lemma 5.25. *Let N be square-free and coprime to 6. When N tends to infinity, we have for any $\varepsilon > 0$*

$$g(N) = \frac{N}{12} \prod_{p|N} \left(1 + \frac{1}{p}\right) + O(\sigma(N)) = \frac{\psi(N)}{12} + O_\varepsilon(N^\varepsilon),$$

where $\sigma(N) = \sum_{d|N} 1$. For a general N,

$$g(N) = \frac{\psi(N)}{12} + O(\sqrt{N} \log\log(2N)).$$

Proof. The dimension of $J_0(N)$ equals the genus of $X_0(N)$, which is given in Proposition 1.43 page 25 of [Shi94] by the formula, valid for N coprime to 6,

$$g(N) = 1 + \frac{N}{12} \prod_{p|N} \left(1 + \frac{1}{p}\right) - \frac{1}{4} \prod_{p|N} \left(1 + \left(\frac{-1}{p}\right)\right) - \frac{1}{3} \prod_{p|N} \left(1 + \left(\frac{-3}{p}\right)\right) - \frac{1}{2} \sum_{d|N} \varphi((d, \frac{N}{d})), \tag{5.30}$$

where φ is Euler's function and $\left(\frac{\cdot}{p}\right)$ is the quadratic residue symbol. In the general case, the formula has the same structure, with the products vanishing according to some divisibility conditions.

Let us solve first the square-free case. One can check that the second and third products in the above expression either vanish or coincide with $\sigma(N)$ up to the corresponding constant factor in front of the product. With respect to the sum, if N square-free then $(d, \frac{N}{d}) = 1$ for any $d|N$, and the sum equals $\sigma(N)$. The statement follows from the known growth rate of $\sigma(N) = O_\varepsilon(N^\varepsilon)$ (see Theorem 315 from [HW08]).

In the general case, the products are still bounded by $\sigma(N)$. Define now

$$\widetilde{\psi}(N) := \sum_{d|N} \varphi\left(\left(d, \frac{N}{d}\right)\right),$$

where (a, b) denotes the greatest common divisor of the integers a and b, and φ is Euler's totient function. Let us study this arithmetic function. Note that $\widetilde{\psi}$ is a multiplicative arithmetic function, i.e. $\widetilde{\psi}(ab) = \widetilde{\psi}(a)\widetilde{\psi}(b)$ if $(a, b) = 1$.

For p a prime number, $k \geq 1$ odd,

$$\begin{aligned}\widetilde{\psi}(p^k) &= \sum_{i=0}^{k} \varphi((p^i, p^{k-i})) = 2\sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \varphi(p^i) = 2(1 + \sum_{i=1}^{\lfloor \frac{k}{2} \rfloor}(p^i - p^{i-1})) = 2p^{\lfloor \frac{k}{2} \rfloor} \\ &= 2p^{\frac{k}{2}-\frac{1}{2}} = \frac{2}{\sqrt{p}}\sqrt{p^k} \leq \left(1 + \frac{1}{p}\right)\sqrt{p^k},\end{aligned}$$

since $\frac{2}{\sqrt{p}} < \left(1 + \frac{1}{p}\right)$ for any prime.

Likewise, if $k \geq 1$ is even,

$$\begin{aligned}\widetilde{\psi}(p^k) &= \sum_{i=0}^{k} \varphi((p^i, p^{k-i})) = 2\sum_{i=0}^{\frac{k}{2}-1} \varphi(p^i) + \varphi(p^{\frac{k}{2}}) = p^{\frac{k}{2}-1} + p^{\frac{k}{2}} = \\ &= \left(1 + \frac{1}{p}\right)\sqrt{p^k},\end{aligned}$$

Therefore, as $\widetilde{\psi}$ is multiplicative,

$$\widetilde{\psi}(N) \leq \sqrt{N}\prod_{p|N}\left(1 + \frac{1}{p}\right).$$

As $\psi(N) = N\prod_{p|N}(1 + \frac{1}{p})$,

$$\widetilde{\psi}(N) \leq \frac{\psi(N)}{\sqrt{N}}, \tag{5.31}$$

and it is known that $\psi(N) = O(N\log(\log(2N)))$ (see [BDGP96][Lemme 2 (i)]). This finishes the proof. □

Remark 5.26. *It follows further from the proof of Lemma 5.25 that for any $\varepsilon > 0$, $\widetilde{\psi}(N) = O_\varepsilon(N^{\frac{1}{2}+\varepsilon})$ with explicit constant*

$$C_\varepsilon = \prod_{1 > p^{\varepsilon - \frac{1}{p}}} p^{-\varepsilon}\left(1 + \frac{1}{p}\right). \tag{5.32}$$

We can therefore give an explicit (but worse) error term in the genus formula (5.30). In particular, from (5.31), $\psi(N) \geq N$ and $\sigma(N) \leq 2\sqrt{N}$ we can deduce:

$$\left|g(N) - \left(1 + \frac{\psi(N)}{12}\right)\right| \leq \frac{1}{2} C_\varepsilon N^{\frac{1}{2}+\varepsilon} + \frac{7}{12}\sigma(N) \leq \sqrt{N}\left(\frac{C_\varepsilon}{2} N^\varepsilon + \frac{7}{6}\right), \quad \text{and}$$
$$\frac{\left|g(N) - \left(1 + \frac{\psi(N)}{12}\right)\right|}{\psi(N)} \leq \frac{1}{2}\frac{\widetilde{\psi}(N)}{\psi(N)} + \frac{7}{12}\frac{\sigma(N)}{\psi(N)} \leq \frac{5}{3}\frac{1}{\sqrt{N}}.$$

It can also be shown, by inspecting how many primes verify the condition under the product in (5.32), that the constant C_ε verifies:

- $C_\varepsilon < 1$, *for* $\varepsilon > 0.585$ *(as the product is empty),*
- $C_\varepsilon < 1.2527$ *for* $\varepsilon > 0.26$ *(as the product only has the prime 2),*
- $C_\varepsilon < 1.5788$ *for* $\varepsilon > 0.132$ *(as the product only has the primes 2 and 3).*

We can now conclude on the first item of Theorem 5.3: by Proposition 5.23, $h_{\mathrm{Falt}}(X_0(N)) = h_{\mathrm{Falt}}(J_0(N))$. By Theorem 5.24, $h_{\mathrm{Falt}}(J_0(N)) \sim \frac{g(N)}{3}\log N$ when N tends to infinity and is square-free, coprime to 6. By Lemma 5.25, $g(N) \sim \frac{\psi(N)}{12}$. Use the Corollary page 390 of [Coh84] which gives $h(\Phi_N) \sim 6\psi(N)\log N$ to conclude that

$$h_{\mathrm{Falt}}(X_0(N)) \sim \frac{1}{6^3} h(\Phi_N). \tag{5.33}$$

5.6.2 Hecke correspondences and modular polynomials

Let us move to the second item of Theorem 5.3. In [Aut03], Autissier uses a morphism $i_N : X_0(N) \longrightarrow \mathbb{P}^1 \times \mathbb{P}^1$, which for two elliptic curves E_1, E_2 and a cyclic isogeny $\alpha : E_1 \to E_2$ is defined by $i_N((E_1, E_2, \alpha)) = (j(E_1), j(E_2))$. He denotes by T_N the image of $X_0(N)$ by i_N, and by $\hat{\mathcal{L}}$ a natural metrized lined bundle on $\mathbb{P}^1 \times \mathbb{P}^1$. Theorem 3.2 page 427 of [Aut03] gives

$$h_{\hat{\mathcal{L}}}(T_N) = 12\psi(N)(\log N - 2\lambda_N + 4\kappa_1),$$

where $\kappa_1 = 12\zeta'(-1) - \log\pi - \frac{1}{2}$. This implies that for any $N \geq 1$, $|h_{\hat{\mathcal{L}}}(T_N) - 2h(\Phi_N)|$ is bounded by a quantity linear in $\psi(N)$. This further implies, as the main term is of order of magnitude bigger than $\psi(N)$, the fact that when N tends to infinity

$$h_{\hat{\mathcal{L}}}(T_N) \sim 2h(\Phi_N).$$

5.6.3 Heegner points and modular polynomials

The third item in Theorem 5.3 comes from an asymptotic estimate computed in [Paz10] and heavily based on the Gross-Zagier computations [GZ86]. Corollaire 1 page 164 in [Paz10] provides us, when N tends to infinity and satisfies the Heegner conditions (there are infinitely many such N for each fixed discriminant D_k), with

$$\hat{h}_{J_0(N)}(c_{D_k}) \sim \frac{3h_k u_k}{g(N)} h_{\mathrm{Falt}}(J_0(N)),$$

where $\hat{h}_{J_0(N)}$ is the Néron-Tate height on the Jacobian $J_0(N)$ as defined in [GZ86]. Use $h_{\mathrm{Falt}}(J_0(N)) = h_{\mathrm{Falt}}(X_0(N))$ and (5.33) to obtain this third item.

5.6.4 Arakelov canonical sheaf of $X_0(N)$

The fourth item in Theorem 5.3 comes from the following asymptotic estimate, first computed in Théorème 1.1 page 646 of [MU98] in the case where N is coprime to 6 and square-free, and recently generalised to any N coprime to 6 in Theorem 1.1 of [DM24]:

$$\overline{\omega}^2 \sim 3g(N)\log N. \tag{5.34}$$

As we have $g(N) \sim \frac{\psi(N)}{12}$ by Lemma 5.25 and by Corollary page 390 of [Coh84] we have $h(\Phi_N) \sim 6\psi(N)\log N$, hence we get the result. This concludes the proof of Theorem 5.3.

Chapter 6

On the CM exception to a generalization of the Stéphanois theorem

This chapter is a reproduction of [Gij25] with minor modifications except for the following: [Gij25, Section 4.1] is now in Section 3.1, and we have included a second appendix.

The material referenced in [Gij25, Section 5.2 and Section 5.3] (here as Sections 6.5.2.1 and 6.5.2.2) was presented and extended in Sections 4.1.2 and 4.2, but here it has been kept as in [Gij25] for readability. Finally, we gave a different proof of Appendix 6.5.3 in Section 3.3.1.

Abstract

There are two classical theorems related to algebraic values of the j-invariant: Schneider's theorem and the Stéphanois theorem. Schneider's theorem for the j-invariant states that the transcendence degree $\operatorname{trdeg} \mathbb{Q}(\tau, j(\tau)) \geq 1$ with the sole exception of CM points. In contrast, CM points do not constitute an exception to the Stéphanois theorem, which states $\operatorname{trdeg} \mathbb{Q}(q, j(q)) \geq 1$ for the Fourier expansion (q-expansion) of the j-invariant, for any q. Schneider's theorem has been generalized to higher dimensions, and in particular holds for the Igusa invariants of a genus 2 curve. These functions have Fourier expansions, but a result of Stéphanois type is unknown. In this paper, we find that there are positive dimensional sources of exceptions to the generic behavior expected in genus 2, and we discuss their relation to CM points. We utilize Humbert singular relations, putting them into the transcendental framework. The computations of the transcendence degree for CM points are conditional to Schanuel's conjecture.

Keywords: transcendence theory, genus 2 curves, complex multiplication, Shimura varieties
Mathematics Subject Classification: 11J89, 11G10, 14K22, 14G35

6.1 Introduction

Consider the elliptic j-invariant $j : \mathbb{H} \to \mathbb{C}$, where $\mathbb{H} = \{\tau \in \mathbb{C} | \operatorname{Im} \tau > 0\}$, a $\mathrm{SL}_2(\mathbb{Z})$-invariant modular function, which classifies elliptic curves over $\overline{\mathbb{Q}}$. It admits a Fourier expansion:

$$j(q) = \frac{1}{q} + 744 + 196884q + \cdots \in \mathbb{Z}[[q]],\ q = e^{2\pi i \tau},$$

which defines a meromorphic function on $\mathbb{D} = \{q \in \mathbb{C} : |q| < 1\}$.

For a field $\mathbb{Q} \subset L$ we denote $\operatorname{trdeg} L$ for the transcendence degree of L over $\mathbb{Q}$. The following two results are classical.

Theorem 6.1 (Schneider's theorem [Sch37]). *For $\tau \in \mathbb{H}$, $\operatorname{trdeg} \mathbb{Q}(\tau, j(\tau)) \geq 1$ with the sole exception of τ quadratic imaginary.*

Theorem 6.2 (The Stéphanois theorem[1] [BDGP96]). *For $0 < |q| < 1$, $\operatorname{trdeg} \mathbb{Q}(q, j(q)) \geq 1$.*

The j-invariant admits a generalization for curves of genus two, which we call Igusa invariants j_1, j_2, j_3, see [Igu60]. Remark that, in the literature, Igusa invariants, or Igusa-Clebsch invariants, usually refer to weighted projective invariants, while we mean absolute ones. Also note that in [Igu60] there are five weighted invariants instead of four, so that the theory extends to fields of even characteristic.

A curve of genus two is necessarily hyperelliptic and admits a (singular) model $C : y^2 = f(x)$ for f a polynomial of degree 6. Set α_i for its six complex roots. We define the following:

$$I_2(f) = \sum (\alpha_1 - \alpha_2)^2 (\alpha_3 - \alpha_4)^2 (\alpha_5 - \alpha_6)^2,$$
$$I_4(f) = \sum (\alpha_1 - \alpha_2)^2 (\alpha_2 - \alpha_3)^2 (\alpha_3 - \alpha_1)^2 (\alpha_4 - \alpha_5)^2 (\alpha_5 - \alpha_6)^2 (\alpha_6 - \alpha_4)^2,$$
$$I_6(f) = \sum \prod_{l=1}^{3} \prod_{(i,j) \in C_l} (\alpha_i - \alpha_j)^2,$$

where $C_1 = \{(1,2),(2,3),(3,1)\}, C_2 = \{(4,5),(5,6),(6,4)\}, C_3 = \{(1,4),(2,5),(3,6)\}$,

$$I_{10}(f) = \prod (\alpha_i - \alpha_j)^2,$$

where the sums range among all permutations of six elements. Note that I_{10} is the discriminant of f. Therefore, as we are considering *smooth* curves of genus two, we assume in the following that $I_{10} \neq 0$.

We set the following normalization of the Igusa invariants

$$j_1(C) = \frac{I_2^5}{I_{10}},$$
$$j_2(C) = \frac{I_2^3 I_4}{I_{10}},$$
$$j_3(C) = \frac{I_2^2 I_6}{I_{10}},$$

[1] This is the standard name for this result in French, but to the best of our knowledge this theorem does not have a consistent name in English, save for the previous denomination as the Mahler-Manin conjecture. The French name comes from the result being proven in St-Étienne and "stéphanois" is its gentilic in French.

if $I_2 \neq 0$. There are other normalizations to choose when $I_2 = 0$, but we would not need them in this paper. As I_α are symmetric polynomial functions on the roots of f, they are functions on the coefficients of f. This mimics the situation with the j-invariant of an elliptic curve, which also admits a formula in terms of its Weierstrass model. They characterize completely the geometric isomorphism class of the curve ([Igu60, Section 6, Theorem 2]), and they detect algebraicity of genus two curves (see [Mes91]): a curve C is defined over $\overline{\mathbb{Q}}$ if and only if $j_l(C) \in \overline{\mathbb{Q}}$ for $l = 1, 2, 3$. In contrast to the case of elliptic curves, this does not hold over a fixed number field K: if $j_i(C) \in K$ for $i = 1, 2, 3$, then C may not be defined over K, but over a quadratic extension of K.

However, we need to consider the Igusa invariants as analytic functions on the moduli space of genus two curves $\mathcal{M}_2$. These analytic functions are not directly defined on $\mathcal{M}_2$, but on its image under the Torelli map in $\mathcal{A}_2$ (Torelli locus), the moduli space of principally polarized abelian surfaces. In other words, they can also be seen as invariants of the Jacobian of the curves (endowed the principal polarization). By the Torelli theorem (see [CS86, Theorem 12.1]), this map is injective, or equivalently, the Jacobian variety equipped with the principal polarization completely distinguishes the curve.

It is a classical result ([Wei57, Satz 2], or [BL04, Corollary 11.8.2 a)]) that the Torelli locus is precisely the indecomposable locus of $\mathcal{A}_2$, which we denote $\mathcal{A}_2^{ind} = \mathcal{A}_2 \setminus (\mathcal{A}_1 \times \mathcal{A}_1)$, where for $\mathcal{A}_1 \times \mathcal{A}_1$ we mean the identification with the locus of products of elliptic curves with the product principal polarization. We remark that an abelian surface in the indecomposable locus can be isogenous to a product of elliptic curves, or even isomorphic, if considering another principal polarization in the product of elliptic curves that is not induced as a product polarization.

The moduli space $\mathcal{A}_2$ is coarsely represented by a quotient of a symmetric space, analogously to $\mathcal{A}_1$ with respect to $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. The analytic space is the Siegel upper half-space of degree two:

$$\mathbb{H}_2 = \{\boldsymbol{\tau} \in \mathrm{Mat}_2(\mathbb{C}) |\ \boldsymbol{\tau} = \boldsymbol{\tau}^t,\ \mathrm{Im}\,\boldsymbol{\tau} \text{ is positive definite}\}.$$

In the following, we set $\boldsymbol{\tau} = \begin{pmatrix} \tau_1 & \tau_2 \\ \tau_2 & \tau_3 \end{pmatrix}$. Note that $\mathrm{Im}\,\boldsymbol{\tau}$ is positive definite if and only if:

$$\tau_1, \tau_3 \in \mathbb{H},\ \mathrm{Im}(\tau_2)^2 < \mathrm{Im}(\tau_1)\,\mathrm{Im}(\tau_3).$$

We denote the symplectic group

$$\mathrm{Sp}_4(\mathbb{Z}) = \left\{ M \in \mathrm{Mat}_4(\mathbb{Z}) |\ M^t J M = J,\ \text{where } J = \left(\begin{array}{c|c} 0 & I_2 \\ \hline -I_2 & 0 \end{array}\right) \right\},$$

which acts on $\mathbb{H}_2$ by linear fractional transformations: if $M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z})$ then

$$M\boldsymbol{\tau} = (\alpha\boldsymbol{\tau} + \beta)(\gamma\boldsymbol{\tau} + \delta)^{-1}.$$

The theory of modular forms generalizes to these functions, which are Siegel modular functions. In particular, they admit higher dimensional Fourier expansions (as they satisfy analogous relations to $j(\tau + 1) = j(\tau)$) of the form

$$\sum_{\substack{n \in \mathrm{GL}_2(\mathbb{Q}) \\ n \text{ half integral symmetric}}} a(n) e^{2\pi i \,\mathrm{tr}(n\boldsymbol{\tau})},$$

where $\mathrm{tr}(n\boldsymbol{\tau}) = n_1\tau_1 + n_3\tau_3 + 2n_2\tau_2$. Therefore, they admit series expansions in terms of $q_l = q_l(\boldsymbol{\tau}) = e^{2\pi i \tau_l}$, for $l = 1, 2, 3$.

As we are interested in geometric invariants of abelian surfaces, they are always considered as defined over an algebraically closed field, and all our related definitions (morphisms, isogenies, simplicity, etc.) are geometric as well.

The analog of Schneider's theorem is a known result for every genus, and in more general settings.

Theorem 6.3 (Cohen-Shiga-Wolfart[2] [SW95] and [Coh96])**.** *Consider a principally polarized abelian variety A with a (small) period matrix $\boldsymbol{\tau}$. Then the following are equivalent:*

- *the abelian variety A is defined over $\overline{\mathbb{Q}}$ and $\boldsymbol{\tau} \in \mathcal{M}_g(\overline{\mathbb{Q}})$,*
- *the abelian variety A has complex multiplication.*

In our set-up, it implies that the components of the matrix $\boldsymbol{\tau}$ and $j_l(\boldsymbol{\tau})$ for $l = 1, 2, 3$ are all simultaneously algebraic if and only if $\boldsymbol{\tau}$ parametrizes a CM abelian surface.

We recall the definition of a CM (complex multiplication) abelian surface; they are the natural generalization of CM elliptic curves. By Poincaré's irreducibility theorem, an abelian variety A admits an isogeny to a product

$$\prod_{k=1}^{s} A_k^{r_k}$$

for A_k simple abelian varieties. We use the notation $A \simeq B$ for isogenous abelian varieties. We denote $\mathrm{End}(A)$ for the (geometric) endomorphism ring and $\mathrm{End}_0(A) = \mathrm{End}(A) \otimes \mathbb{Q}$ for the endomorphism algebra.

Definition 6.4. *We say that a* simple *abelian variety A has CM if its endomorphism algebra $\mathrm{End}_0(A) = K$ for K a CM field with $[K : \mathbb{Q}] = 2\dim(A)$. An abelian variety is said to have CM if all the simple factors given by its isogeny decomposition have CM. We call $\boldsymbol{\tau} \in \mathbb{H}_g$ a CM* point *if it parametrizes a CM abelian variety.*

We also remark that this property does not involve the polarization of the abelian variety.

An abelian surface A with CM is necessarily one of the following. We follow the notation from [DO21, Table 1].

- Either A is simple, then $\mathrm{End}_0(A)$ is a quartic CM field *(CM simple)*;
- or $A \simeq E \times E'$ for two not isogenous CM elliptic curves (with necessarily different CM fields K, K'), then $\mathrm{End}_0(A) = K \times K'$ *(CM split)*;
- or $A \simeq E^2$, then $\mathrm{End}_0(A) = \mathcal{M}_2(K)$ *(CM isotypic)*.

We remark that in the literature, the CM split case is sometimes also called CM *non-simple* (but that does not necessarily mean non simple), for clarity we will use *split* in this document.

More generally, we will say an abelian surface A is *split* if it is isogenous to $E \times E'$, for E, E' *non isogenous* elliptic curves, and *isotypic* if it is isogenous to E^2, E an elliptic curve.

[2] The set-up of abelian *surfaces* with quaternionic multiplication was historically first study in [Mor72].

6.1.1 Questions and our main result

A higher dimensional analog of the Stéphanois theorem is still open, but the corresponding *functional* transcendence statement does hold: it follows from [BZ01, Theorem 2]. (On that direction of functional transcendence, see [Pil13, Theorem 2.5].) One could tentatively predict a "generic" behavior

$$\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q})) \geq 3, \tag{6.1}$$

where $q = (q_1, q_2, q_3)$, that would be satisfied outside a set of exceptions.

The purpose of this paper is to answer two questions, posed to us by Daniel Bertrand.

Question 1. *Can CM points be exceptions to Equation (6.1)?*

Question 2. *In case of an affirmative answer, are such exceptions exclusively due to being CM, or are they explained by belonging to a larger "exceptional subset"?*

This last question makes sense from the perspective of Shimura varieties, and the work behind the proof of the André-Oort conjecture. Likewise, the functional transcendence statement in [Pil13, Theorem 2.5] has weakly special subvarieties as unique sources of exceptions, as well as [CFN20, Theorem 1.1] and [PT16, Theorem 1.3]. Also note that, under the Grothendieck's period conjecture, in [And04, Proposition 23.2.4.1], for abelian varieties defined over $\overline{\mathbb{Q}}$, there is an equality between $\operatorname{trdeg} \mathbb{Q}(\boldsymbol{\tau})$ (for the field generated by the coefficients of the matrix $\boldsymbol{\tau}$) and the smallest dimension of Shimura subvariety containing it. More precisely, in [Fon23], special subvarieties also appear as obstructions to algebraic independence results (and conjecturally, the only ones), see [Fon23, Conjecture 13.1.2] for Hilbert modular surfaces and Hirzebruch-Zagier divisors.

In very broad terms, in the higher dimensional setting the CM points are understood as special points, or zero dimensional special subvarieties. Because $\mathcal{A}_2$ (and any $\mathcal{A}_g$ for $g \geq 2$) does have positive dimensional special subvarieties, one would be inclined to think that special subvarieties should play a role in this generalization.

The answer we give to Question 1 is positive, and for Question 2, we can associate to every CM point a special subvariety of $\mathcal{A}_2$ along which this type of exception for the transcendence degree
$\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q}))$ is achieved for other points, not necessarily CM, in a "generic" sense, cf. Theorem 6.6. This is a marked contrast with the statement of Theorem 6.3, which implies that for problems in terms of $\operatorname{trdeg} \mathbb{Q}(\tau_1, \tau_2, \tau_3, j_1(\boldsymbol{\tau}), j_2(\boldsymbol{\tau}), j_3(\boldsymbol{\tau}))$, the CM points are "isolated" exceptions. The computation of the transcendence degree $\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q}))$ for some CM points is conditional to Schanuel's conjecture 6.7, or more precisely to the Gelfond-Schneider conjecture 6.8.

It is noteworthy to highlight the role played by the Humbert singular relations (see Section 3.1), and that, *unconditionally*, knowledge about Humbert singular relations in moduli spaces of abelian surfaces gives insight on these questions stemming from the transcendence framework. We believe, however, that it might be difficult to generalize these results to higher genus curves: the antisymmetric relations in Remark 3.5 in the case of *surfaces* lead to the elegant (and one dimensional) Humbert relations (6.3) (below in Section 4), but this changes as soon as the abelian varieties have larger dimension.

Remark 6.5. *We state the result in terms of* $\min_{\mathrm{Sp}_4(\mathbb{Z})}$ *because, as we will see in the following section (Example 6.13),* $\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q}))$ *is not invariant under isomor-*

phism of principally polarized abelian surfaces. With this notation, we mean

$$\min_{M\in \mathrm{Sp}_4(\mathbb{Z})}\{\mathrm{trdeg}\,\mathbb{Q}(q_1(M\boldsymbol{\tau}),q_2(M\boldsymbol{\tau}),q_3(M\boldsymbol{\tau}),j_1(M\boldsymbol{\tau})),j_2(M\boldsymbol{\tau})),j_3(M\boldsymbol{\tau}))\}.$$

Remark that it is only the exponential functions $q_l = e^{2\pi i\tau_l}$ for $l = 1, 2, 3$ that are not invariant under $\mathrm{Sp}_4(\mathbb{Z})$.

Throughout this paper, for $\tau \in \mathbb{H}_2$, by abuse of notation, we write $A_{\boldsymbol{\tau}} \in \mathcal{A}_2$ instead of $[A_{\boldsymbol{\tau}}] \in \mathcal{A}_2$ for the corresponding isomorphism class of ppas (principally polarized abelian surface).

Theorem 6.6. *Let $\boldsymbol{\tau} \in \mathbb{H}_2$ a CM point with associated CM ppas $A_{\boldsymbol{\tau}} \in \mathcal{A}_2^{ind}$, and set $q_l = e^{2\pi i\tau_l}$, $q = (q_1, q_2, q_3)$ and $j_l(q) = j_l(\boldsymbol{\tau})$, for $l = 1, 2, 3$.*

1. *If $A_{\boldsymbol{\tau}}$ is simple, it belongs to a* unique *Humbert surface, where, for any ppas defined over $\overline{\mathbb{Q}}$ (which we associate with the notation $\tilde{q}$), it holds* $\min_{\mathrm{Sp}_4(\mathbb{Z})} \mathrm{trdeg}\,\mathbb{Q}(\tilde{q}_1,\tilde{q}_2,\tilde{q}_3,j_1(\tilde{q}),j_2(\tilde{q}),j_3(\tilde{q})) \leq 2$. *Moreover, under Schanuel's conjecture 6.7,* $\min_{\mathrm{Sp}_4(\mathbb{Z})} \mathrm{trdeg}\,\mathbb{Q}(q_1,q_2,q_3,j_1(\boldsymbol{q}),j_2(\boldsymbol{q}),j_3(\boldsymbol{q})) = 2$.

2. *If $A_{\boldsymbol{\tau}} \simeq E' \times E''$ with E', E'' not isogenous CM elliptic curves, then the same statement holds verbatim. In this case, the special subvariety can also be taken as a special curve of type $\mathbb{Q} \times CM$. Moreover, under Schanuel's conjecture 6.7,* $\min_{\mathrm{Sp}_4(\mathbb{Z})} \mathrm{trdeg}\,\mathbb{Q}(q_1,q_2,q_3,j_1(\boldsymbol{q}),j_2(\boldsymbol{q}),j_3(\boldsymbol{q})) = 2$.

3. *If $A_{\boldsymbol{\tau}} \simeq E^2$ with E a CM elliptic curve. Then $A_{\boldsymbol{\tau}}$ belongs to one of the modular curves in the collection specified in Section 6.5.2.2, and for any ppas in C defined over $\overline{\mathbb{Q}}$, it holds* $\min_{\mathrm{Sp}_4(\mathbb{Z})} \mathrm{trdeg}\,\mathbb{Q}(\tilde{q}_1,\tilde{q}_2,\tilde{q}_3,j_1(\tilde{q}),j_2(\tilde{q}),j_3(\tilde{q})) \leq 1$. Unconditionally, $\mathrm{trdeg}\,\mathbb{Q}(q_1,q_2,q_3,j_1(\boldsymbol{q}),j_2(\boldsymbol{q}),j_3(\boldsymbol{q})) = 1$.

6.1.2 Special subvarieties of $\mathcal{A}_2$

The definition of special subvarieties requires the set-up of Shimura varieties. In $\mathcal{A}_2$, they may be explicitly described as loci of principally polarized abelian surfaces with some conditions (in particular, they are all of PEL type). We simply list them, as they appear in [DO21, Table 1]. We will eventually focus on the special subvarieties completely contained in $\mathcal{A}_2^{ind}$. We do not impose that a special subvariety needs to be connected.

The special surfaces are of the following two types, where the second one can be considered as a "degenerate case" of the first. They are jointly called Humbert surfaces.

- Surfaces that parametrize ppas with real multiplication by a quadratic field.
- Surfaces that parametrize ppas isogenous to a product of elliptic curves.

For more details on these surfaces see Section 3.1. As an example, note that $\mathcal{A}_1 \times \mathcal{A}_1$ is a special surface.

The special curves come in three types that we denote as follows.

- Type $\mathbb{Q} \times CM$: they parametrize ppas isogenous to a product $E \times E'$ with E CM.

- Shimura curves: they parametrize ppas with quaternionic multiplication (QM) by an indefinite division quaternion algebra over $\mathbb{Q}$.

- Modular curves: they parametrize ppas isogenous to the square of an elliptic curve.

Modular curves can also be seen as "degeneration" of Shimura curves, with the quaternion algebra over $\mathbb{Q}$ taken as $\mathrm{Mat}_2(\mathbb{Q})$.

One can see, as an application of Theorem 6.28, that points belonging to modular and Shimura curves always lie at the intersection of special surfaces, whereas a curve of type $\mathbb{Q} \times CM$ cannot be realized as an intersection of two special surfaces. Conversely, the intersection of two Humbert surfaces (when it is one dimensional) consists of a (finite union of) Shimura or modular curves. This perspective has been used several times in the literature to study said curves, for example, see [HM95], [Run99], [Kan19] and [BG08].

Moreover, for each of the special subvarieties, not all the three types of special points can occur:

- For the special surfaces parametrizing real multiplication, a CM point cannot be split (by [Gor02, Corollary 2.7]), so we find CM points which are simple or isotypic.

- For the degenerate special surfaces, a CM point cannot be simple, by construction.

- For the special curve of type $\mathbb{Q} \times CM$, a CM point cannot be simple, by construction.

- For both modular and Shimura curves, a CM point can only be isotypic, via examination of the possible embeddings $B \hookrightarrow \mathrm{End}_0(A)$, for B the indefinite quaternion algebra over $\mathbb{Q}$.

Special subvariety \ CM point	simple	split	isotypic
$\mathcal{H}_\Delta$		NO	
$\mathcal{H}_{\delta^2}$	NO		
curve of type $\mathbb{Q} \times CM$	NO		
Shimura curve	NO	NO	
Modular curve	NO	NO	

Table 6.1: CM points on special subvarieties.

6.1.3 Organization

The organization of this paper is as follows. In Section 6.2 we reduce our main problem to study linear relations between the coefficients of a period matrix $\boldsymbol{\tau} \in \mathbb{H}_2$ of a ppas, and for CM points that is sufficient under the Gelfond-Schneider conjecture 6.8. In Section 6.3 we exemplify that Theorem 6.6 needs to be stated in terms of the $\mathrm{Sp}_4(\mathbb{Z})$-orbit, because linear relations between the coefficients of $\boldsymbol{\tau}$ are not invariant under the $\mathrm{Sp}_4(\mathbb{Z})$-action, and give the minimum value of $\mathrm{trdeg}\,\mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q}))$ for CM points. In Section 6.4 we present Humbert singular relations, or linear relations involving the coefficients of $\boldsymbol{\tau}$ *and* $\det \boldsymbol{\tau}$,

which behave better for this action. We introduce the lattice of singular relations and the lattice of singular linear relations $\mathcal{L}_{\boldsymbol{\tau}}$ and $\mathcal{L}_{\boldsymbol{\tau}}^{lin}$, which are positive definite lattices. We study $\mathcal{L}_{\boldsymbol{\tau}}$ and, in particular, compute $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}$, which is determined by $\mathrm{End}_0(A_{\boldsymbol{\tau}})$. Section 6.5 is devoted to study $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}^{lin}$ and prove the main result Theorem 6.6. The first two items of Theorem 6.6 are proved first, thanks to an application of Humbert's lemma Proposition 3.12. The last item requires distinguishing between Shimura and modular curves: both have $\operatorname{rank}\mathcal{L} = 2$ generically, hence we need more information to distinguish them. We study modular embeddings for both curves, searching for $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}^{lin} = 2$ generically in the image. Modular curves in the collection described in Section 6.5.2.2, the ones studied in [Kan16], admit embeddings with such property (and are the ones we use to finish the proof of Theorem 6.6), while the classical quaternionic embeddings in [Has95] for Shimura curves do not. We further conjecture that $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}^{lin} \neq 2$ for whole $\mathrm{Sp}_4(\mathbb{Z})$-orbit of any given point of these curves. In the Appendix, we give a partial answer to that, via Hirzebruch-Zagier divisors in Hilbert modular surfaces.

Acknowledgements: We thank Daniel Bertrand for numerous fruitful discussions during and after a stay of the author at IMJ-PRG during Spring 2024, and valuable comments on previous drafts. We also thank Damien Robert for the conversations during a visit to Université de Bordeaux; Aurel Page for answering our questions on quaternion algebras; Christopher Daw and Martin Orr for discussions about parametrization of quaternionic curves; and Yunqing Tang for pointing us to Hirzebruch-Zagier divisors.

We thank Fabien Pazuki for guidance and useful remarks throughout this whole project, and also Tim With Berland and Fadi Mezher for several office conversations.

6.2 On linear dependence relations

Let us first answer the analogous question for the j-invariant. If $\tau \in \mathbb{H}$ is quadratic imaginary, then $j(\tau)$ is algebraic and $\operatorname{trdeg}\mathbb{Q}(q, j(q)) = \operatorname{trdeg}\mathbb{Q}(q)$. But by the Gelfond-Schneider theorem [FN98, Chapter 3, Section 2, Theorem 3.1], it holds $q = e^{2\pi i\tau} = (-1)^{2\tau} \notin \overline{\mathbb{Q}}$, as $2\tau \in \overline{\mathbb{Q}} \setminus \mathbb{Q}$.

Suppose now that $\boldsymbol{\tau} \in \mathbb{H}_2$ is a CM point. By Theorem 6.3,

$$\{\boldsymbol{\tau} \in \mathbb{H}_2 |\ \boldsymbol{\tau} \text{ is a CM point}\} = \{\boldsymbol{\tau} \in \mathbb{H}_2 \cap \mathrm{Mat}_2(\overline{\mathbb{Q}}),\ j_1(\boldsymbol{\tau}), j_2(\boldsymbol{\tau}), j_3(\boldsymbol{\tau}) \in \overline{\mathbb{Q}}\}.$$

Therefore,

$$\operatorname{trdeg}\mathbb{Q}(q_1, q_2, q_3, j_1(q), j_2(q), j_3(q)) = \operatorname{trdeg}\mathbb{Q}(q_1, q_2, q_3) = \operatorname{trdeg}\mathbb{Q}(e^{2\pi i\tau_1}, e^{2\pi i\tau_2}, e^{2\pi i\tau_3}).$$

We now state Schanuel's conjecture.

Conjecture 6.7. (Schanuel's conjecture [FN98, Chapter 6, page 260]) *Suppose $x_1, \dots, x_n \in \mathbb{C}$ are $\mathbb{Q}$-linearly independent, then* $\operatorname{trdeg}\mathbb{Q}(x_1, \dots, x_n, e^{x_1}, \dots, e^{x_n}) \geq n$.

Conditionally to this conjecture, we have the following result, which can be seen as a Lindemann-Weierstrass type of statement for $e^{i\pi(\cdot)}$. Likewise, it is also a particular case of the Gelfond-Schneider conjecture (for $\alpha = -1 = e^{i\pi}$), independently attributed to Schneider [Sch57] and Gelfond [FN98, Chapter 6, Conjecture 1 page 259].

Conjecture 6.8 (The Gelfond-Schneider conjecture). *Consider $\alpha \in \overline{\mathbb{Q}}$, $\alpha \neq 0, 1$ and $x_1, \dots, x_n \in \overline{\mathbb{Q}}$ such that $1, x_1, \dots, x_n$ are $\mathbb{Q}$-linearly independent. Then $\alpha^{x_1}, \dots, \alpha^{x_n}$ are algebraically independent.*

We have included the classical deduction from Schanuel's conjecture (in the particular case $\alpha = e^{i\pi}$) for the reader's convenience, although historically it is an older conjecture.

Lemma 6.9 (Under Conjecture 6.8)**.** *Assume $1, x_1, \dots x_n \in \overline{\mathbb{Q}}$ are $\mathbb{Q}$-linearly independent. Then $e^{i2\pi x_1}, \dots, e^{i2\pi x_n}$ are algebraically independent.*

More generally, let $r = \dim \operatorname{span}_{\mathbb{Q}}(1, x_1, \dots x_n)$, then $\operatorname{trdeg} \mathbb{Q}(e^{i2\pi x_1}, \dots, e^{i2\pi x_n}) = r - 1$.

Remark 6.10. *The $\mathbb{Q}$-linearly independence with 1 cannot be omitted in the statement. If $x_1, \dots, x_n$ satisfy an equation $\sum_{i=1}^n a_i x_i = b$ with $a_i, b \in \mathbb{Z}$ then $q_l = e^{i2\pi x_l}$ for $l = 1, \dots, n$ solve a multiplicative dependence relation $q_1^{a_1} \cdots q_n^{a_n} = 1$.*

Proof. It follows that $i2\pi, i2\pi x_1, \dots, i2\pi x_n$ are $\mathbb{Q}$-linearly independent, so by Schanuel's conjecture,

$$\begin{aligned} n+1 &\leq \operatorname{trdeg} \mathbb{Q}(i2\pi, i2\pi x_1, \dots, i2\pi x_n, e^{i2\pi}, e^{i2\pi x_1}, \dots, e^{i2\pi x_n}) =^* \\ &= \operatorname{trdeg} \mathbb{Q}(i\pi, \underbrace{x_1, \dots, x_n, 1}_{\in \overline{\mathbb{Q}}}, e^{i2\pi x_1}, \dots, e^{i2\pi x_n}) = \operatorname{trdeg} \mathbb{Q}(i\pi, e^{i2\pi x_1}, \dots, e^{i2\pi x_n}) \end{aligned}$$

where the equality in (*) comes from $\overline{\mathbb{Q}}(i2\pi, i2\pi x_1, \dots, i2\pi x_n) = \overline{\mathbb{Q}}(i\pi, x_1, \dots x_n)$. Therefore, $i\pi, e^{i2\pi x_1}, \dots, e^{i2\pi x_n}$ are algebraically independent, which in particular implies our statement.

The second part has already been proven for $r = n + 1$. For general r, take r linearly independent elements among $1, \dots, x_n$. Assume first that we can take $1 = x_1$, then the claim follows as for the case $r = n + 1$. Otherwise, assume $x_1, \dots, x_r$ linearly independent, but with $\sum_{i=1}^r a_i x_i = b$ for not all zero $a_1, \dots, a_r, b \in \mathbb{Z}$. Applying Schanuel's conjecture as above for $i2\pi x_1, \dots, i2\pi x_r$ results in $\operatorname{trdeg} \mathbb{Q}(e^{i2\pi x_1}, \dots, e^{i2\pi x_r}) \geq r - 1$, but as in Remark 6.10, the transcendence degree cannot be maximal, as there is a multiplicative dependence among $e^{i2\pi x_1}, \dots, e^{i2\pi x_r}$. Therefore, $\operatorname{trdeg} \mathbb{Q}(e^{i2\pi x_1}, \dots, e^{i2\pi x_r}) = r - 1$. □

Corollary 6.11 (Under Conjecture 6.8)**.** *Let $\boldsymbol{\tau} \in \mathbb{H}_2$ a CM point. Then*

$$\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(q), j_2(q), j_3(q)) = \dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3) - 1.$$

Remark 6.12. *In the proof of Lemma 6.9, it was explicitly used that $x_i \in \overline{\mathbb{Q}}$. In the general case, $\dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3)$ will say nothing about $\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q}))$. But if we additionally assume that $j_l(q) \in \overline{\mathbb{Q}}$, them $\dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3)$ gives an upper bound.*

Therefore, Question 1 has naturally led us to study non-homogeneous linear dependence relations between the coefficients of $\boldsymbol{\tau}$, and for CM points this analysis is sufficient (conditionally to the Gelfond-Schneider conjecture, or Schanuel's conjecture). In any case, these non-homogeneous linear dependence relations produce multiplicative dependence relations between the q_l's, and give some partial information about the transcendence degree we are interested in.

Our strategy for Theorem 6.6 is to prove that for every CM point, there exists a special subvariety of $\mathcal{A}_2$ containing it and for which $\dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3)$ is generically constant through it, and equal to the value attained at the CM point. Therefore, we attach to every CM point a subvariety such that the points in this subvariety parametrizing abelian surfaces defined over $\overline{\mathbb{Q}}$ are at least "of the same type" of exception as the CM point.

6.3 First obstructions and the simplest case

We first show via an example that the answer to Question 1, as stated, unfortunately must be answered with "it depends".

Example 6.13. *One of the most "famous" genus 2 curves with CM is*

$$C : y^2 = x^6 - x.$$

In [BMM90, page 92] a period matrix is computed. For $\xi = e^{2\pi i/5}$, it is given by

$$\begin{pmatrix} \tau_1 & \tau_2 \\ \tau_2 & \tau_3 \end{pmatrix} := \begin{pmatrix} -\xi^4 & \xi^2+1 \\ \xi^2+1 & \xi^2-\xi^3. \end{pmatrix}$$

It is easy to check that the τ_l's are $\mathbb{Q}$-linearly independent, using that $-\xi^4 = 1+\xi+\xi^2+\xi^3$, and that $1, \xi, \xi^2, \xi^3$ are $\mathbb{Q}$-linearly independent. Therefore, by Corollary 6.11, $\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(q), j_2(q), j_3(q)) = 3$, and hence this is not an exception to (6.1).

On the other hand, an application of Humbert's lemma (Proposition 3.12, or more precisely of the algorithm described in [BW03, Proposition 4.5]), gives us the matrix

$$M = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ -3 & 0 & -1 & 0 \\ 0 & 2 & 0 & 3 \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z}).$$

Then $\boldsymbol{\tau}' = M\boldsymbol{\tau}$ belongs to the $\mathrm{Sp}_4(\mathbb{Z})$-orbit of τ, so the associated ppas are isomorphic. One can compute that

$$\boldsymbol{\tau}' = M\boldsymbol{\tau} = \begin{pmatrix} \frac{2}{55}\xi + \frac{9}{55}\xi + \frac{4}{55}\xi^2 + \frac{1}{11}\xi^3 & \frac{1}{11}\xi + \frac{1}{55}\xi^2 + \frac{4}{55}\xi^3 \\ \frac{1}{11}\xi + \frac{1}{55}\xi^2 + \frac{4}{55}\xi^3 & \frac{2}{5} + \frac{4}{55}\xi^2 + \frac{3}{55}\xi^2 + \frac{1}{55}\xi^3 \end{pmatrix},$$

which solves the linear equation

$$-\tau_1' + \tau_2' + \tau_3' = 0.$$

Here is the verification with SageMath [Sage]:

```
k = CyclotomicField(5)
xi = k.gen()
tau = Matrix([[-xi^4, xi^2 + 1],[xi^2 + 1,xi^2-xi^3]])
A = Matrix([[-1,0],[0,1]])
B = Matrix([[0,0],[0,1]])
C = Matrix([[-3,0],[0,2]])
D = Matrix([[-1,0],[0,3]])
Om = (A*tau + B)*((C*tau + D).inverse())
-Om[0][0] + Om[0][1] + Om[1][1]
```

Hence, by Corollary 6.11, for $\boldsymbol{\tau}'$ we do get an exception to (6.1). Question 1 should then be reformulated to a statement for the whole orbit under the action of $\mathrm{Sp}_4(\mathbb{Z})$ by fractional linear transformations.

On a different note, we will first consider the "most exceptional" type of exceptions we have found among the CM points (as in they minimize $\dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3)$, see Corollary 6.11). Remark that by definition of $\mathbb{H}_2$, it follows that $\tau_1, \tau_3 \notin \mathbb{Q}$, therefore

$$\dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3) \geq \dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1) = 2.$$

Definition 6.14. *For a complex abelian variety $A \cong \mathbb{C}^g/\Lambda$ of dimension g, by* big period matrix *we mean a matrix $\Pi \in \mathcal{M}_{g\times 2g}(\mathbb{C})$ constructed from a choice of basis vectors $\lambda_1, \dots, \lambda_{2g} \in \Lambda$ in terms of another choice of basis vectors $e_1, \dots, e_g \in \mathbb{C}^g$. If we set $\Pi = (\Omega_1, \Omega_2)$, then by a* (small) period matrix *we mean $\Omega_1 \in \mathcal{M}_g(\mathbb{C})$, after a choice of basis such that $\Omega_2 = I_g$. Given a big period matrix for A, we will say that $\Omega_2^{-1}\Omega_1$ is a period matrix.*

Alternatively, for ppav (by [BL04, paragraph before Proposition 8.1.1])) a period matrix comes from the choice of $\lambda_1, \dots, \lambda_{2g} \in \Lambda$ a *symplectic* basis with respect to the principal polarization in A, and setting the first g vectors $\lambda_1, \dots \lambda_g$ as a basis of $\mathbb{C}^g$.

Lemma 6.15. *An abelian surface A admits a big period matrix belonging to $\mathcal{M}_{2\times 4}(K)$ for K quadratic imaginary field if and only if A is isogenous to E^2, with E an elliptic curve with complex multiplication.*

We now prove Lemma 6.15, which we found as part of a longer exercise in [BL04, Exercise 10, section 5.6, page 142]. It is stated for abelian varieties of any dimension, and it gives a larger chain of equivalences:

1. The Picard number of A (*i.e.* the rank of its Néron-Severi group $\mathrm{NS}(A)$) is maximal, $\operatorname{rank} \mathrm{NS}(A) = g^2$.
2. A is isogenous to a E^g, E elliptic curve with complex multiplication.
3. A admits a big period matrix in $\mathcal{M}_{g\times 2g}(K)$, for K quadratic imaginary field.
4. A is isomorphic (as unpolarized abelian varieties) to a product of $E_1 \times \ldots \times E_g$, E_i pairwise isogenous elliptic curves with complex multiplication

Proof. We are inspired by the proof of [BL04, Corollary 10.6.3]. A big period matrix for $A \simeq E^2$ necessarily has the shape

$$\begin{pmatrix} \tau & 1 & 0 & 0 \\ 0 & 0 & \tau & 1 \end{pmatrix} R, \tag{6.2}$$

for $R \in \mathrm{Mat}_4(\mathbb{Q})$ and τ such that $E = \mathbb{C}/(\tau\ 1)\mathbb{Z}^2$. Hence $\begin{pmatrix} \tau & 1 & 0 & 0 \\ 0 & 0 & \tau & 1 \end{pmatrix} \in \mathcal{M}_{2\times 4}(\mathbb{Q}(\tau))$, with $\mathbb{Q}(\tau)$ quadratic imaginary, so the big period matrix (6.2) of A also belongs to $\mathbb{Q}(\tau)$.

On the other direction, if A admits a big period matrix in $\mathcal{M}_{2\times 4}(\mathbb{Q}(\alpha))$ with α quadratic imaginary, then can solve the matrix R as

$$\begin{pmatrix} r_0\alpha + r_4 & r_1\alpha + r_5 & r_2\alpha + r_6 & r_3\alpha + r_7 \\ r_8\alpha + r_{12} & r_9\alpha + r_{13} & r_{10}\alpha + r_{14} & r_{11}\alpha + r_{15} \end{pmatrix} = \begin{pmatrix} \alpha & 1 & 0 & 0 \\ 0 & 0 & \alpha & 1 \end{pmatrix} \begin{pmatrix} r_0 & r_1 & r_2 & r_3 \\ r_4 & r_5 & r_6 & r_7 \\ r_8 & r_9 & r_{10} & r_{11} \\ r_{12} & r_{13} & r_{14} & r_{15} \end{pmatrix}$$

We need to verify that R defines an isogeny between A and E^2, with $E = \mathbb{C}/(\alpha\ 1)\mathbb{Z}^2$ (and it is enough to show it as complex tori). First, there exists $m \in \mathbb{Z}$ such that $mR \in \mathrm{Mat}_4(\mathbb{Z})$. If we set Π' for the big period matrix of A, and $\Pi := \begin{pmatrix} \alpha & 1 & 0 & 0 \\ 0 & 0 & \alpha & 1 \end{pmatrix}$, then it holds

$$(mI_2)\Pi = \Pi'(mR),$$

and by [BL04, Equation (1.1) before Proposition 1.2.3], this is the compatibility condition sufficient to have a homomorphism of complex tori $E^2 \to A$. It is an isogeny because it is surjective (its analytic representation is multiplication by m) and between tori of the same dimension. □

Corollary 6.16. *Let $\boldsymbol{\tau} \in \mathbb{H}_2$ such that if $A_{\boldsymbol{\tau}} \simeq E^2$, with E a CM elliptic curve. Then* $\mathrm{trdeg}\, \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q})) = 1$.

Proof. This follows immediately from Lemma 6.15. If $A_{\boldsymbol{\tau}} \simeq E^2$ as in the statement, then it admits a big period matrix $(\Omega\ \Omega')$ lying in $\mathcal{M}_{2\times 4}(K)$ for K a quadratic imaginary field. Therefore, $\boldsymbol{\tau} \in \mathrm{Mat}_2(K)$, so $\dim \mathrm{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3) \leq \dim_{\mathbb{Q}} K = 2$. Hence $\mathrm{trdeg}\, \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q})) \leq 1$, and equality holds by the Gelfond-Schneider theorem [FN98, Chapter 3, Section 2, Theorem 3.1], as $q_1 = e^{2\pi i \tau_1}$ is transcendental.

We notice that $\boldsymbol{\tau}' \in \mathrm{Mat}_2(K)$ along the whole $\mathrm{Sp}_4(\mathbb{Z})$-orbit, because for $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z})$, it follows that $A\boldsymbol{\tau} + B \in \mathrm{Mat}_2(K)$ and $(C\boldsymbol{\tau} + D)^{-1} \in \mathrm{Mat}_2(K)$, as K is a number field. □

Remark 6.17. *The proof of Corollary 6.16 is thus unconditional, in particular is not relying on Schanuel's conjecture.*

6.4 Humbert singular relations

To go beyond in our study of linear relations between τ_1, τ_2, τ_3 we will know focus on studying linear relations between $1, \tau_1, \tau_2, \tau_3, \tau_2^2 - \tau_1\tau_3$, *i.e.* equations (over $\mathbb{Z}$) of the form

$$a\tau_1 + b\tau_2 + c\tau_3 + d(\tau_2^2 - \tau_1\tau_3) + e = 0, \tag{6.3}$$

which behave better under the action of $\mathrm{Sp}_4(\mathbb{Z})$. Returning to Example 6.13, the coefficients of both $\boldsymbol{\tau}$ and $\boldsymbol{\tau}'$ satisfy such a relation, and the one for $\boldsymbol{\tau}'$ happens to have $d = 0$. For $\boldsymbol{\tau}$, as $\boldsymbol{\tau} \in \mathrm{Mat}_2(\mathbb{Q}(\xi))$ with ξ a primitive 5-th root of unity, then $\dim \mathrm{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3, \tau_2^2 - \tau_1\tau_3) \leq \dim \mathbb{Q}(\xi) = 4$, so there must exist a relation as in (6.3).

For the basics on Humbert singular relations, we redirect to Section 3.1.

6.4.1 The lattice of singular relations

Definition 6.18. *For $\boldsymbol{\tau} \in \mathbb{H}_2$, we denote* the lattice of Humbert singular relations *the following $\mathbb{Z}$-module*

$$\mathcal{L}_{\boldsymbol{\tau}} := \{(a, b, c, d, e) \in \mathbb{Z}^5 : a\tau_1 + b\tau_2 + c\tau_3 + d(\tau_2^2 - \tau_1\tau_3) + e = 0\},$$

equipped with the positive definite (by Lemma 3.7) quadratic form induced by the discriminant Δ. It forms a positive definite integral lattice (as in finite rank free abelian group with a symmetric bilinear form).

Likewise, we consider the sublattice $\mathcal{L}_{\boldsymbol{\tau}}^{lin} := \mathcal{L}_{\boldsymbol{\tau}} \cap \{d = 0\}$*, or equivalently, the sublattice spanned by singular relations that do not involve* $\tau^2 - \tau_1\tau_3$*. Therefore, one can consider* $\mathcal{L}_{\boldsymbol{\tau}} \otimes \mathbb{Q} \subset \mathbb{Q}^5$ *and* $\mathcal{L}_{\boldsymbol{\tau}}^{lin} \otimes \mathbb{Q} \subset \mathbb{Q}^4$*.*

We are ultimately interested in the rank of $\mathcal{L}_{\boldsymbol{\tau}}^{lin}$. We will study $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}$ and then compare it with $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}^{lin}$.

Some of these results admit alternative proofs. First, by Lemma 3.2, Remark 3.3 and Lemma 3.4, there is a short exact sequence of abelian groups:

$$0 \to \mathbb{Z} \to \operatorname{End}^s(A_{\boldsymbol{\tau}}) \to \mathcal{L}_{\boldsymbol{\tau}} \to 0, \tag{6.4}$$

where the homomorphism $\operatorname{End}^s(A_{\boldsymbol{\tau}}) \to \mathcal{L}_{\boldsymbol{\tau}}$ is given by

$$f \mapsto \rho_{r,\boldsymbol{\tau}}(f) = \begin{pmatrix} \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} & \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & c \\ -c & 0 \end{pmatrix} & {}^t\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \end{pmatrix} \mapsto (a_2, (a_4 - a_1), -a_3, b, c).$$

If one combines (6.4) and Proposition 3.19, it follows that $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = \operatorname{rank} \mathrm{NS}(A_{\boldsymbol{\tau}}) - 1$. Hence, one can use results for $\operatorname{rank} \mathrm{NS}(A_{\boldsymbol{\tau}})$ instead. However, we believe that a proof via explicit manipulation of the lattices is insightful in our setting, for that we are also able to understand the effect on the quadratic form Δ.

In that direction, there is a more intrinsic definition of $(\mathcal{L}_{\boldsymbol{\tau}}, \Delta)$ in $\mathrm{NS}(A_{\boldsymbol{\tau}})/\mathbb{Z}H_0$, with the quadratic form defined in terms of the intersection pairing, see [Kan16, Section 2].

Remark 6.19. *The following is true:*

- *We have the following short exact sequences of* $\mathbb{Q}$*-vector spaces*

$$0 \mapsto \mathcal{L}_{\boldsymbol{\tau}} \otimes \mathbb{Q} \mapsto \mathbb{Q}^5 \mapsto \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3, \tau_2^2 - \tau_1\tau_3) \mapsto 0,$$

  and

$$0 \mapsto \mathcal{L}_{\boldsymbol{\tau}}^{lin} \otimes \mathbb{Q} \mapsto \mathbb{Q}^4 \mapsto \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3) \mapsto 0.$$

  Because $\tau_1 \notin \mathbb{Q}$*, the dimension of both spans is at least* 2*. Consequently* $0 \leq \operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} \leq 5 - 2 = 3$*, and* $0 \leq \operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}^{lin} \leq 4 - 2 = 2$*.*

The extremal bounds for $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}$ are realized.

Lemma 6.20. *Let* $\boldsymbol{\tau} \in \mathbb{H}_2$*. Then*

- $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 0$ *if and only if* $\operatorname{End}_0(A_{\boldsymbol{\tau}}) = \mathbb{Q}$*.*
- $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 3$ *if and only if* $A_{\boldsymbol{\tau}} \simeq E^2$ *for* E *a CM elliptic curve.*

Proof. The case $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 0$ corresponds to the absence of any Humbert singular relation whatsoever, hence, $\operatorname{End}^s(A_{\boldsymbol{\tau}}) = \mathbb{Z}$ by Lemma 3.2 and Lemma 3.4. By [BL04, Theorem 5.3.2], any abelian subvariety is associated to a symmetric idempotent, so $A_{\boldsymbol{\tau}}$ is necessarily simple if $\operatorname{End}_0^s(A_{\boldsymbol{\tau}}) = \mathbb{Q}$. By Albert's classification of endomorphisms algebra of simple abelian varieties [BL04, Proposition 5.5.7], and the fact that the case III does not occur for abelian surfaces, it follows that $\operatorname{End}_0(A_{\boldsymbol{\tau}}) = \mathbb{Q}$.

For the second part, note that $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}} = 3$ is equivalent to $\dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3, \tau_2^2 - \tau_1\tau_3) = 2$ and hence,

$$\operatorname{span}_{\mathbb{Q}}(1, \tau_1) = \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3) = \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3, \tau_2^2 - \tau_1\tau_3).$$

Hence, there exists $a, b, a', b', a'', b'' \in \mathbb{Q}$ such that:

$$\begin{aligned} \tau_2 &= a + b\tau_1, \\ \tau_3 &= a' + b'\tau_1, \\ \tau_2^2 - \tau_1\tau_3 &= a'' + b''\tau_1. \end{aligned}$$

On the other hand,

$$a'' + b''\tau_1 = \tau_2^2 - \tau_1\tau_3 = (b^2 - b')\tau_1^2 + (2ab - a')\tau_1 + a^2.$$

Then either τ_1 solves a quadratic equation (remark that $\tau_1, \tau_3 \in \mathbb{H}$, so in that case τ_1 is quadratic imaginary), and as $\boldsymbol{\tau} \in \operatorname{Mat}_2(\mathbb{Q}(\tau_1))$ we are finished by Lemma 6.15, or $b^2 - b' = 0$. The latter is impossible, as $\operatorname{Im}\boldsymbol{\tau}$ is a positive definite symmetric matrix:

$$\operatorname{Im}\boldsymbol{\tau} = \begin{pmatrix} \operatorname{Im}(\tau_1) & b\operatorname{Im}(\tau_1) \\ b\operatorname{Im}(\tau_1) & b'\operatorname{Im}(\tau_1) \end{pmatrix},$$

so in particular it has positive determinant, hence $(b' - b^2)\operatorname{Im}(\tau_1) > 0$, so $b^2 \neq b'$. □

This result already suggests that $\mathcal{L}_{\boldsymbol{\tau}}$ is a more suitable object for the $\mathrm{Sp}_4(\mathbb{Z})$-action. More is true:

Lemma 6.21. *If $\boldsymbol{\tau}, \boldsymbol{\tau}' \in \mathbb{H}_2$ parametrize isomorphic ppas, then $(\mathcal{L}_{\boldsymbol{\tau}}, \Delta)$ and $(\mathcal{L}_{\boldsymbol{\tau}'}, \Delta)$ are isomorphic as positive definite lattices. More generally, if they are isogenous, there is $\mathbb{Z}$-linear map $(\mathcal{L}_{\boldsymbol{\tau}}, \Delta) \to (\mathcal{L}_{\boldsymbol{\tau}'}, \Delta')$ such that $\Delta'(\cdot) = \kappa\Delta(\cdot)$ for some constant $\kappa \geq 0$, which extends to a $\mathbb{Q}$-linear isomorphism $\mathcal{L}_{\boldsymbol{\tau}} \otimes \mathbb{Q} \to \mathcal{L}_{\boldsymbol{\tau}'} \otimes \mathbb{Q}$. In particular, $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}$ is invariant under isogenies.*

The proof of Lemma 6.21 will be split into two parts. Let us rewrite the statements in terms of actions on $\mathbb{H}_2$. Consider $\boldsymbol{\tau} \in \mathbb{H}_2$ and $(A_{\boldsymbol{\tau}}, H)$ the corresponding principally polarized abelian surface, and assume that we have an isogeny $\phi : B \to A_{\boldsymbol{\tau}}$ with exponent $e(\phi)$, *i.e.* there exists $\psi : A_{\boldsymbol{\tau}} \to B$ with $\psi\phi = e(\phi)_B$ and $\phi\psi = e(\phi)_A$, by [BL04, Proposition 1.2.6]. Hence, ϕ induces a polarization of degree $e(\phi)^4$.

The following comes from the proof of [BL04, Proposition 8.1.2]. By the theory of elementary divisors, the induced polarization on B has a type $D = \operatorname{diag}(d_1, d_2)$ with $d_1 | d_2$ and $\det(D) = e(\phi)^4$. We can choose a symplectic basis for $A_{\boldsymbol{\tau}}$ and B such that have matrices $R \in \operatorname{Mat}_4(\mathbb{Z})$ and $L \in \operatorname{Mat}_2(\mathbb{C})$ (the rational and analytic representation of ϕ, respectively) and period matrices

$$L(\boldsymbol{\tau}'\ D) = (\boldsymbol{\tau}\ I_2)R.$$

If we set

$${}^tM = R\begin{pmatrix} I_2 & 0 \\ 0 & D^{-1} \end{pmatrix},$$

then

$$L(\boldsymbol{\tau}'\ I_2) = (\boldsymbol{\tau}\ I_2){}^tM,$$

and by the same arguments to prove that isomorphisms of principally polarized abelian varieties correspond to matrices in $\mathrm{Sp}_4(\mathbb{Z})$ acting on $\mathbb{H}_2$, one proves that $M \in \mathrm{Sp}_4(\mathbb{Q})$ and $\boldsymbol{\tau}' = M\boldsymbol{\tau}$, via a linear fractional transformation.

We claim that $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}$ is invariant by the action of matrices $M \in \mathrm{Sp}_4(\mathbb{Q})$ such that there exists $D = \operatorname{diag}(d_1, d_2)$, $d_1 | d_2$ and $d_1 d_2 \in \mathbb{Z}_{>0}$ a perfect square with

$$ {}^tM \begin{pmatrix} I_2 & 0 \\ 0 & D \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Z}). \tag{6.5} $$

Let us first prove the invariance under the $\mathrm{Sp}_4(\mathbb{Z})$-action separately.

Lemma 6.22. *For $\boldsymbol{\tau} \in \mathbb{H}_2$ and $\boldsymbol{\tau}' = M\boldsymbol{\tau}$ with $M \in \mathrm{Sp}_4(\mathbb{Z})$, $(\mathcal{L}_{\boldsymbol{\tau}}, \Delta)$ and $(\mathcal{L}_{\boldsymbol{\tau}'}, \Delta)$ are isomorphic as positive definite lattices.*

Proof. By (3.3) and (3.4), the action of M in both the rational and analytic representation corresponds to conjugation by suitable matrices. On the other hand, one can check $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ if and only if $f_M := \phi^{-1} f \phi \in \mathrm{End}^s(A_{\boldsymbol{\tau}'})$ for $\phi : A_{\boldsymbol{\tau}'} \to A_{\boldsymbol{\tau}}$ the isomorphism corresponding to M, by the effect on ϕ on the principal polarization and on the Rosati involution (equivalently, one can prove it directly for the rational representations $\rho_{r,\boldsymbol{\tau}}(f)$ and $\rho_{r,\boldsymbol{\tau}'}(f)$, for that the corresponding matrix equation (3.5) is invariant under conjugation by matrices in $\mathrm{Sp}_4(\mathbb{Z})$).

For $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$, by (the proof of) Lemma 3.2, we read the Humbert singular relation as the non-trivial equation in

$$ \boldsymbol{\tau} B \boldsymbol{\tau} - C = \boldsymbol{\tau} A - {}^t(\boldsymbol{\tau} A), $$

where $\rho_{r,\boldsymbol{\tau}}(f) = \begin{pmatrix} A & B \\ C & {}^tA \end{pmatrix}$.

Therefore, as $f_M \in \mathrm{End}^s(A_{\boldsymbol{\tau}'})$ with rational representation given by (3.3), we have a Humbert singular relation for $\boldsymbol{\tau}'$ given by

$$ \boldsymbol{\tau}' B' \boldsymbol{\tau}' - C' = \boldsymbol{\tau}' A' - {}^t(\boldsymbol{\tau}' A'), $$

with coefficients given by ${}^tM^{-1} \rho_{r,\boldsymbol{\tau}}(f) {}^tM$. Remark that $\rho_{r,\boldsymbol{\tau}}(f) = nI_4$ for some $n \in \mathbb{Z}$ if and only if $\rho_{r,M\boldsymbol{\tau}}(f) = nI_4$, so $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 0$ if and only if $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}'} = 0$. Furthermore, $\Delta(f) = \Delta(f_M)$, because $\Delta(f) = \operatorname{tr}(\rho_{a,\boldsymbol{\tau}}(f))^2 - 4 \det \rho_{a,\boldsymbol{\tau}}(f)$ is invariant under conjugation.

Finally, suppose that we have two endomorphisms $f, g \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$, and consider the corresponding Humbert singular relations l_f, l_g. They are linearly dependent if and only if there exists $n, m, l \in \mathbb{Z}$ such that

$$ n\rho_{r,\boldsymbol{\tau}}(f) = m\rho_{r,\boldsymbol{\tau}}(g) + lI_4. \tag{6.6} $$

This follows from the fact that a Humbert singular relation characterizes the matrix of the rational representation up to addition by a multiple of I_4, and from $\rho_{r,\boldsymbol{\tau}}(sf) = s\rho_{r,\boldsymbol{\tau}}(f)$ for all $s \in \mathbb{Z}$. As (6.6) is invariant under conjugation, $f, g \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ produce linearly independent Humbert singular relations if and only if the same is true for $f_M, g_M \in \mathrm{End}^s(A_{\boldsymbol{\tau}'})$. It then follows $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = \operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}'}$. □

Proof of Lemma 6.21. Consider M as in (6.5) and the set-up described after Lemma 6.21. For an endomorphism $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ with rational representation $\rho_{r,\boldsymbol{\tau}}(f)$, could consider $\rho_{r,\boldsymbol{\tau}'}(f)$, which gives the rational representation of f in the basis $(\boldsymbol{\tau}' I_2)$, *i.e.* $\rho_{r,\boldsymbol{\tau}'}(f) = {}^tM^{-1} \rho_{r,\boldsymbol{\tau}}(f) M$. As $M \in \mathrm{Sp}_4(\mathbb{Q})$, it is symmetric with respect to the (extension of) the Rosati involution to

$\mathrm{End}^0(A_{\boldsymbol{\tau}'})$, but it is only defined on $\mathrm{End}^0(A_{\boldsymbol{\tau}'})$, as we have formally inverted the isogeny ϕ. However, we just need to "clear denominators" and consider the correct multiple of f so that we get an honest endomorphism in $\mathrm{End}^s(A_{\boldsymbol{\tau}'})$.

By setting

$$R = {}^tM \begin{pmatrix} I_2 & 0 \\ 0 & D \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Z}),$$

we recover the rational representation of the isogeny (with integer coefficients), hence

$$\rho_{r,\boldsymbol{\tau}'}(f) = \begin{pmatrix} I_2 & 0 \\ 0 & D \end{pmatrix} R^{-1} \rho_{r,\boldsymbol{\tau}}(f) R \begin{pmatrix} I_2 & 0 \\ 0 & D \end{pmatrix}^{-1}.$$

It is easy to check that although $R^{-1} \notin \mathrm{Mat}_4(\mathbb{Z})$, we have $\begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} R^{-1} \in \mathrm{Mat}_4(\mathbb{Z})$.

In short, multiplying by a power of $e(\phi)$ will make our matrix integral. From another point of view, in our chosen basis $\rho_{r,\boldsymbol{\tau}'}(f)$ is the rational representation of $\psi^{-1} f \psi$, formally inverting ϕ, so we need to instead consider $\phi f \psi$ (from this description we can also see that $\phi f \psi$ is symmetric). Because $\psi^{-1} = \frac{1}{e(\phi)}\phi$, this means that

$$e(\phi)\rho_{r,\boldsymbol{\tau}'}(f) \in \mathrm{Mat}_4(\mathbb{Z}).$$

Now that we have defined a map $\mathcal{L}_{\boldsymbol{\tau}} \to \mathcal{L}_{\boldsymbol{\tau}'}$, we can argue, as in the proof of Lemma 6.22, that $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}'} = \operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}$, and $\Delta' = e(\phi)^2 \Delta$.

As an alternative proof of the invariance of the rank under isogenies, note that by Proposition 3.19, there exists an isomorphism of abelian group between $\mathrm{NS}(A_{\boldsymbol{\tau}})$ and $\mathrm{End}^s(A_{\boldsymbol{\tau}})$. In addition, by (6.4), $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = \operatorname{rank} \mathrm{NS}(A_{\boldsymbol{\tau}}) - 1$. Finally, it is known that the rank of the Néron-Severi group of an abelian variety is invariant under isogenies, see [BL99, Chapter 1, Prop 3.2]. □

We devote the rest of this section to the computation of $\operatorname{rank}(\mathcal{L}_{\boldsymbol{\tau}})$.

Proposition 6.23. *If $\boldsymbol{\tau} \in \mathbb{H}_2$ corresponds to a ppas $A_{\boldsymbol{\tau}}$ such that $\mathrm{End}_0(A_{\boldsymbol{\tau}})$ is commutative but distinct from $\mathbb{Q}$, which means that it is either*

1. *a real quadratic field,*
2. *a CM quartic field,*
3. *$\mathbb{Q} \times \mathbb{Q}$,*
4. *$\mathbb{Q} \times K$ for K an imaginary quadratic,*
5. *or $K_1 \times K_2$ for $K_1 \neq K_2$ imaginary quadratic fields,*

then $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 1$.

Proof. We remark that $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 1$ is equivalent to $A_{\boldsymbol{\tau}}$ solving a unique (up to sign) primitive HSR, hence $A_{\boldsymbol{\tau}}$ belonging to a unique Humbert surface of minimal discriminant.

The cases (1) and (2) correspond to $\mathrm{End}(A)$ without zero divisors (equivalently $A_{\boldsymbol{\tau}}$ simple), then there cannot exist $f \in \mathrm{End}_0(A_{\boldsymbol{\tau}})$ such that $\Delta(f)$ is a perfect square, by Proposition 3.10. In addition, there is a *unique* quadratic real field embedded in $\mathrm{End}_0(A_{\boldsymbol{\tau}})$: in the case (1)

$\mathrm{End}_0(A_{\boldsymbol{\tau}})$ is already a quadratic field, and in (2) it corresponds to the (unique) maximal real subfield of the CM field. By Albert's classification of endomorphisms algebras of simple abelian varieties [BL04, Proposition 5.5.7], in (1) the Rosati involution is trivial, and in (2) the Rosati involution corresponds to complex conjugation, so in both cases the embeddings restrict to $\mathrm{End}_0^s(A_{\boldsymbol{\tau}})$, and are surjective. Hence $\mathrm{End}_0^s(A_{\boldsymbol{\tau}})$ is a real quadratic field, and $\mathrm{End}^s(A_{\boldsymbol{\tau}})$ is an order in said field. Taking $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ with minimal non-zero $\Delta(f)$ (the discriminant of the order by which $A_{\boldsymbol{\tau}}$ has real multiplication), we necessarily have $\{n + mf\} = \mathrm{End}^s(A)$. Therefore, we have one (up to sign) primitive HSR, and any other is necessarily a multiple of it.

The other cases correspond to $A_{\boldsymbol{\tau}} \simeq E \times E'$ with E and E' nonisogenous elliptic curves. First, $\mathrm{End}_0^s(A_{\boldsymbol{\tau}})$ has (non trivial) symmetric idempotents by [BL04, Theorem 5.3.2], in particular there exists $f \in \mathrm{End}^s(A_{\boldsymbol{\tau}})$ with $f \notin \mathbb{Z}$, so that $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} \geq 1$. We note that $A_{\boldsymbol{\tau}}$ cannot have real multiplication by any real quadratic field, as by [Gor02, Corollary 2.7], the abelian surfaces with real multiplication by a quadratic field are either simple or isotypic (isogenous to the square of an elliptic curve).

Therefore, Δ induces a positive definite quadratic form on $\mathcal{L}_{\boldsymbol{\tau}}$ that only represents squares. By contradiction, assume that there is a sublattice of $\mathcal{L}_{\boldsymbol{\tau}}$ of rank two, then Δ will induce a positive definite *binary* quadratic form $rf(x, y)$ with $f = ax^2 + bxy + cy^2$ primitive (*i.e.* $\gcd(a, b, c) = 1$) and $r \in \mathbb{Z}_{>0}$. It is known that primitive positive definite quadratic forms represent infinitely many prime numbers (see [Cox22, Theorem 9.12]), which is a contradiction to rf only representing squares. Therefore, $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 1$.

□

The lattice of singular relations is a more interesting object in the quaternionic multiplication case. Note that if we assume A ppas admitting an embedding $B \hookrightarrow \mathrm{End}_0(A_{\boldsymbol{\tau}})$, then either $B \cong \mathrm{End}_0(A)$ or $\mathrm{End}_0(A)$ is strictly larger (and in particular, $\operatorname{rank}_{\mathbb{Q}}(\mathrm{End}_0(A)) \geq 4$). In the latter case, one can check by inspection with all the other possibilities for $\mathrm{End}_0(A)$ with $\operatorname{rank}_{\mathbb{Q}}(\mathrm{End}_0(A)) \geq 4$, that necessarily $\mathrm{End}_0(A) \cong M_2(K)$ for K a imaginary quadratic field. Therefore $A_{\boldsymbol{\tau}} \simeq E^2$ for E an elliptic curve with CM by K, and by Lemma 6.20, it is completely characterized by $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 3$.

Proposition 6.24. *Let B be an indefinite quaternion* division *algebra over $\mathbb{Q}$. Suppose that $A_{\boldsymbol{\tau}}$ admits quaternionic multiplication by an order in B.*

- *If $A_{\boldsymbol{\tau}}$ is simple (equivalently* $\mathrm{End}_0(A) \cong B$*) then* $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 2$.
- *Otherwise, $A_{\boldsymbol{\tau}}$ is CM and* $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 3$.

Proof. If $A_{\boldsymbol{\tau}}$ is simple, by [Rém17, Lemma 7.9], then we can fix some Eichler order $\mathcal{O}$ (intersection of maximal orders) of square-free level in B, and there exists an isogeny from $A_{\boldsymbol{\tau}}$ to an abelian surface A' with $\mathrm{End}(B) = \mathcal{O}$. This abelian variety A' further admits a principal polarization, as we will see in Section 6.5.2.1. By Lemma 6.22, it is enough then to compute $\operatorname{rank} \mathcal{L}_{\tilde{\boldsymbol{\tau}}}$, with $\tilde{\boldsymbol{\tau}}$ a period matrix for A'. In Section 6.5.2.1, we will state more information about ppas with quaternionic multiplication by Eichler orders; in particular, $\tilde{\boldsymbol{\tau}}$ is in the $\mathrm{Sp}_4(\mathbb{Z})$-orbit of a matrix in the image of the morphism in Proposition 6.29, from which it follows that $\operatorname{rank} \mathcal{L}_{\tilde{\boldsymbol{\tau}}} = 2$. By Lemma 6.21, $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 2$.

Alternatively, for a more intrinsic argument, we could have also fixed $\mathcal{O}$ as a maximal order in B. In the case of maximal orders, by [LY20, Lemma 16 and Lemma 17], the lattice of singular relations is isomorphic to a lattice of rank two contained in B.

The second statement was proven in Lemma 6.20. □

To finish the study of $\mathcal{L}_{\boldsymbol{\tau}}$, we are only missing the isotypic case $A_{\boldsymbol{\tau}} \simeq E^2$, where E is *without* complex multiplication. By Lemma 6.21, assume that $A_{\boldsymbol{\tau}}$ is *isomorphic* to the product of elliptic curves with product polarization.

Lemma 6.25. *Consider an elliptic curve E and the abelian surface E^2 (with the canonical principal polarization). There exists an embedding*

$$\mathcal{O}_K \hookrightarrow \mathrm{End}(E^2),$$

for all real quadratic fields K.

With respect to the canonical principal polarization, there exist infinitely many quadratic fields such that

$$\mathcal{O}_K \hookrightarrow \mathrm{End}^s(E^2).$$

Remark 6.26. *The endomorphism algebra of an abelian variety is invariant under isogenies, but that is not true for the endomorphism ring. What we can say is that if $A_{\boldsymbol{\tau}} \simeq E^2$, then $A_{\boldsymbol{\tau}}$ admits real multiplication by an* order *in K for any real quadratic field K, but not necessarily by the ring of integers $\mathcal{O}_K$.*

Proof. This is a classical observation from the theory of real multiplication, for example in [Gor02, Example 2.2 (2)], if d is square free, we can choose integers a, b, c with $d = a^2 + bc$.

If $d \equiv 2, 3(4)$ then the embedding from $\mathcal{O}_{\mathbb{Q}(\sqrt{D})} \to \mathrm{Mat}_2(\mathbb{Z})$ is given by

$$\sqrt{d} \mapsto \begin{pmatrix} a & b \\ c & -a \end{pmatrix}.$$

If $d \equiv 1(4)$ we can choose a odd and b, c even (take $a = 1$, $b = 2$, $c = (D-1)/2$), and the embedding corresponds to

$$\frac{1+\sqrt{d}}{2} \mapsto \begin{pmatrix} \frac{1+a}{2} & \frac{b}{2} \\ \frac{c}{2} & \frac{1-a}{2} \end{pmatrix}.$$

The canonical polarization induces the Rosati involution $\mathrm{Mat}_2(\mathbb{Q}) \to \mathrm{Mat}_2(\mathbb{Q})$ given by $Z \mapsto {}^tZ$. Hence, the previous embeddings map to $\mathrm{End}^s(A_{\boldsymbol{\tau}})$ when can take $b = c$ in $d = a^2 + bc$. By Jacobi's two square theorem, it follows that $\mathcal{O}_K \hookrightarrow \mathrm{End}^s(A_{\boldsymbol{\tau}})$ for square free integers $d = 2^j p_i \cdots p_r$ with $j = 0, 1$ and all $p_i \equiv 1 \pmod 4$.

□

Proposition 6.27. *If $A_{\boldsymbol{\tau}} \simeq E^2$, where E does* not *have CM, then* $\mathrm{rank}\, \mathcal{L}_{\boldsymbol{\tau}} = 2$.

Proof. As we have said previously, it is enough to show it for E^2 with the canonical product polarization. By Lemma 6.20, it cannot be 0 or 3. Assume that $\mathrm{rank}\, \mathcal{L}_{\boldsymbol{\tau}} = 1$, hence there exists a primitive singular relation l such that $\mathcal{L}_{\boldsymbol{\tau}} = ml$ for $m \in \mathbb{Z}$, and Δ restricts to the quadratic form $f(m) = m^2\Delta(l)$, where $\Delta(l)$ is a fixed constant. On the other hand, by Lemma 6.25, $m^2\Delta(l)$ must represent infinitely many fundamental discriminants, which is a contradiction. Therefore, $\mathrm{rank}\, \mathcal{L}_{\boldsymbol{\tau}} = 2$. □

We collect all the past results (Lemma 6.20, Proposition 6.23, Proposition 6.24 and Proposition 6.27) of this section in the following theorem. Remark that the proofs of 2) and 3) are for one implication only, but as we have exhausted all the possible cases for $\mathrm{End}_0(A_{\boldsymbol{\tau}})$, they are also equivalences.

Theorem 6.28. *Let $\boldsymbol{\tau} \in \mathbb{H}_2$ and consider the corresponding ppas $A_{\boldsymbol{\tau}} \in \mathcal{A}_2$.*

- *It holds $\mathrm{End}_0(A_{\boldsymbol{\tau}}) = \mathbb{Q}$ if and only if $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 0$.*
- *It holds $\mathrm{End}_0(A_{\boldsymbol{\tau}})$ is commutative but not $\mathbb{Q}$ if and only if $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 1$.*
- *It holds $\mathrm{End}_0(A_{\boldsymbol{\tau}})$ is an indefinite quaternion algebra over $\mathbb{Q}$ if and only if $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 2$.*
- *It holds $\mathrm{End}_0(A_{\boldsymbol{\tau}}) = \mathrm{Mat}_2(K)$, for K a quadratic imaginary field, if and only if $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 3$.*

Alternatively, this is also proven in [BL99, Chapter 2, Proposition 7.1], by computing $\operatorname{rank} \mathrm{NS}(A)$. We underline the case of $\mathrm{End}_0(A_{\boldsymbol{\tau}})$ with zero divisors, which follows from a computation of $\operatorname{rank} \mathrm{NS}(E \times E')$ with E, E' elliptic curves. By [Kan16, Proposition 23] there is a group isomorphism:

$$\mathrm{NS}(E \times E') \cong \mathbb{Z} \oplus \mathbb{Z} \oplus \mathrm{Hom}(E, E').$$

Therefore, for $A_{\boldsymbol{\tau}} \simeq E \times E'$ (with E and E' not necessarily isogenous) depending on $\operatorname{rank} \mathrm{Hom}(E, E') = 0, 1, 2$, we have

$$\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = \begin{cases} 1, & \text{if } E \not\simeq E', \\ 2, & \text{if } E \simeq E', \text{ without CM}, \\ 3, & \text{if } E \simeq E', \text{ with CM}. \end{cases}$$

6.5 From Humbert singular relations to linear relations

We now have a good understanding of $\mathcal{L}_{\boldsymbol{\tau}}$ from Theorem 6.28. However, we were initially interested in $\mathcal{L}^{lin}_{\boldsymbol{\tau}}$. This section is devoted to infer the relevant information about it to prove Theorem 6.6.

Let us first notice that $\operatorname{rank} \mathcal{L}^{lin}_{\boldsymbol{\tau}} \leq \min(\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}, 2)$, and if $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 0$, then $\operatorname{rank} \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 0$ too, and since $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}}$ is invariant under $\mathrm{Sp}_4(\mathbb{Z})$-orbit, then $\operatorname{rank} \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 0$ holds for the whole orbit of $\boldsymbol{\tau}$.

6.5.1 The case $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 1$, and proof of Theorem 6.6 1), 2)

If $\boldsymbol{\tau} \in \mathbb{H}_2$ is a CM point and $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 1$ then there exists (up to sign) only one primitive Humbert singular relation l satisfied by $\boldsymbol{\tau}$. It can involve $\tau_2^2 - \tau_1\tau_3$ or not. In the latter case, $\dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3, \tau_2^2 - \tau_1\tau_3) = 3$ and by Corollary 6.11, $\operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q})) = 2$.

If l involves the determinant, then by Humbert's lemma Proposition 3.12, there exists $\boldsymbol{\tau}'$ in the $\mathrm{Sp}_4(\mathbb{Z})$-orbit with $\dim \operatorname{span}_{\mathbb{Q}}(1, \tau_1', \tau_2', \tau_3') = 3$, and can apply the previous argument now to $\boldsymbol{\tau}'$. It then follows that $\min_{\mathrm{Sp}_4(\mathbb{Z})} \operatorname{trdeg} \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q})) = 2$.

Set $\Delta = \Delta(l)$ and consider the Humbert surface $\mathcal{H}_\Delta$. Consider now other $\tilde{\boldsymbol{\tau}} \in \mathcal{H}_\Delta$ (non-CM) such that $j_1(\tilde{\boldsymbol{\tau}}), j_2(\tilde{\boldsymbol{\tau}}), j_3(\tilde{\boldsymbol{\tau}}) \in \overline{\mathbb{Q}}$. Then $\operatorname{trdeg} \mathbb{Q}(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3, j_1(\tilde{q}), j_2(\tilde{q}), j_3(\tilde{q})) = \operatorname{trdeg} \mathbb{Q}(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3)$. On the other hand, they also satisfy the Humbert singular relation l, so we can argue as above and have $\min_{\mathrm{Sp}_4(\mathbb{Z})} \operatorname{trdeg} \mathbb{Q}(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3, j_1(\tilde{q}), j_2(\tilde{q}), j_3(\tilde{q})) \leq 2$.

We have therefore proven the first two statements in Theorem 6.6. In the second statement, note that if $A_{\boldsymbol{\tau}}$ belongs to a special curve of type $\mathbb{Q} \times CM$, this special curve can also be taken as the special subvariety, for that $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 1$ generically in that curve by Proposition 6.23 (4).

6.5.2 The case $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 2, 3$

If $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 3$, then necessarily $\operatorname{rank} \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 2$. Also, by Theorem 6.28, this case can only happen for $A_{\boldsymbol{\tau}}$ isotypic CM.

If $\operatorname{rank} \mathcal{L}_{\boldsymbol{\tau}} = 2$ then $\operatorname{rank} \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 2$ or 1, and we will show that both cases occur. By Proposition 6.23, they have to be either Shimura curves or modular curves. For the proof of Theorem 6.6, we want a special curve such that $\operatorname{rank} \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 2$, so our discussion in Section 6.5.2.1 is not relevant for the proof of Theorem 6.6, though it is still insightful to complete the study for $\mathcal{L}^{lin}_{\boldsymbol{\tau}}$.

6.5.2.1 Shimura curves by hereditary orders

For the rest of the section, we need to introduce the quaternionic modular embeddings for an hereditary order in an indefinite quaternion algebra over $\mathbb{Q}$. This construction is standard in the literature for Shimura curves, see, for example, [GY19, Section 2.1] or [Has95, Section 3], and comes from a classical paper by Shimura [Shi63]. We have also included it in Section 4.1 and Subsection 4.1.2.

Let B an indefinite quaternion algebra over $\mathbb{Q}$ of discriminant D (with $D > 1$, equivalently, D is a division algebra). Let $\mathcal{O}$ be an hereditary order, equivalently, an Eichler order of level N square-free (by definition, an intersection of two maximal orders).

Set for $\tau \in \mathbb{H}$ the vector $v_\tau = (\tau\ 1)^t \in \mathbb{C}^2$, and fix an embedding $\phi : B \to \mathrm{Mat}_2(\mathbb{R})$. We remark that such an embedding exists as B is indefinite, *i.e.* $B \otimes_{\mathbb{Q}} \mathbb{R} \cong \mathrm{Mat}_2(\mathbb{R})$, and any two embeddings differ by conjugation by an invertible matrix in $\mathrm{Mat}_2(\mathbb{R})$, by the Skolem-Noether theorem.

We want to assign to every $\tau \in \mathbb{H}$ a complex torus with a principal polarization. We set:

- the lattice $\Lambda_\tau := \phi(\mathcal{O})v_\tau$, with $\phi(\mathcal{O}) \subset \mathrm{Mat}_2(\mathbb{R})$ acting as 2×2 matrices, and the torus $\mathbb{C}^2/\Lambda_\tau$;
- for the principal polarization, let $\mu \in \mathcal{O}$ satisfying the following conditions
  - it holds $\operatorname{tr}(\mu) = 0$, $\operatorname{nrd}(\mu) > 0$, and $\mu^2 + DN = 0$, and
  - if we set $\phi(\mu) = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ then $c > 0$,

  then there is a symplectic form E_μ on $\Lambda_\tau \times \Lambda_\tau \to \mathbb{Z}$ (and extended to $\mathbb{C}^2$ by $\mathbb{R}$-linearity) defined by
$$E_\mu(\phi(\alpha)v_\tau, \phi(\beta)v_\tau) := \operatorname{tr}(\mu^{-1}\alpha\bar{\beta}).$$

The conditions imposed to μ make E_μ a Riemann form, and $A_\tau = (\mathbb{C}^2/\Lambda_\tau, E_\mu)$ a principally polarized abelian surface. Likewise, we have an embedding $\mathcal{O} \hookrightarrow \mathrm{End}(A_\tau)$ induced by ϕ, compatible with the Rosati involution induced by E_μ. The restriction of the Rosati involution to $\mathcal{O}$ is given by $\alpha \mapsto \mu^{-1}\bar{\alpha}\mu$.

This assignment induces a well-defined map $X_0^D(N) \to \mathcal{A}_2$, for $X_0^D(N)$ the Shimura curve associated to $\mathcal{O}$. Set $\mathfrak{X}_\mu$ for its image in $\mathcal{A}_2$.

We can further define a map $\mathbb{H} \to \mathbb{H}_2$ by choosing a symplectic basis $\alpha_1, \dots, \alpha_4$ in $\mathcal{O}$ with respect to E_μ. Then the big period matrix is given by $\Pi(\tau) = (\alpha_1 v_\tau, \alpha_2 v_\tau, \alpha_3 v_\tau, \alpha_4 v_\tau) = (\Omega_1(\tau), \Omega_2(\tau))$, and the period matrix by $\Omega_2(\tau)^{-1}\Omega_1(\tau)$. This basis also induces an embedding via ϕ of $\mathcal{O}^1$ into $\mathrm{Aut}_{\mathbb{Z}}(\mathcal{O}, E_\mu) \cong \mathrm{Sp}_4(\mathbb{Z})$, see [Has95, Proposition 2.5].

Those maps can be made explicit, see [Has95, Theorem 3.5 and Theorem 5.1]. It then follows that $\mathrm{rank}\, \mathcal{L}_{\boldsymbol{\tau}} = 2$ and $\mathrm{rank}\, \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 1$ for $\boldsymbol{\tau}$ in the image of these embeddings. There is a choice of an odd prime p and an integer a, see [Has95, page 535 above Equation (1)] for more details. Likewise, see [GY19, Lemma 38] and [LY20, Lemma 5], where these choices are reinterpreted in terms of $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of positive definite binary quadratic forms of discriminant $-16DN$.

Proposition 6.29 ([Has95])**.** *Let $\mathcal{O}$ and B be as above, and set $\varepsilon = \frac{1+\sqrt{p}}{2}$ and $\bar{\varepsilon} = \frac{1-\sqrt{p}}{2}$. Then the following map Ω gives a modular embedding of $\mathbb{H}$ into $\mathbb{H}_2$ with respect to $\phi(\mathcal{O}^1)$ and $\mathrm{Sp}_4(\mathbb{Z})$:*

$$\Omega(z) = \frac{1}{pz}\begin{pmatrix} -\bar{\varepsilon}^2 + \frac{(p-1)aDN}{2}z + DN\varepsilon^2 z^2, & \bar{\varepsilon} - (p-1)aDNz - DN\varepsilon z^2 \\ \bar{\varepsilon} - (p-1)aDNz - DN\varepsilon z^2, & -1 - 2aDNz + DNz^2 \end{pmatrix}.$$

Furthermore, $\boldsymbol{\tau} = \Omega(z)$ satisfy simultaneously the following singular relations parametrized by two independent integers $x, y \in \mathbb{Z}$:

$$x\tau_1 + (x + 2aDNy)\tau_2 - \frac{p-1}{4}x\tau_3 + y\left(\tau_2^2 - \tau_1\tau_3\right) + \left(a^2DN - b\right)DNy = 0,$$

where we put $a^2DN + 1 = pb$. Moreover, if $z \in \mathbb{H}$ is not a CM point, then it does not have another singular relation.

One can then see that $\mathrm{rank}\, \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 1$ for $\boldsymbol{\tau}$ in $\mathfrak{X}_\mu$.

Let us understand the subset covered by $\cup_\mu \mathfrak{X}_\mu$. The following is a consequence of [Shi63, Section 2, Theorem 1 and Theorem 2], that we state as in [GY19, page 6, paragraph before section 2.2]. Set $\mathcal{Q}_{D,N} \subset \mathcal{A}_2$ the set of ppas with QM by an Eichler order $\mathcal{O}$ of level N in a quaternion algebra of discriminant D, such that the Rosati involution restricted to $\mathcal{O}$ coincides with $\alpha \mapsto \mu^{-1}\bar{\alpha}\mu$, for $\mu \in \mathcal{O}$ allowed as above.[3] Then

$$\mathcal{Q}_{D,N} = \bigcup_{\mu \in \mathcal{O}} \mathfrak{X}_\mu, \tag{6.7}$$

where μ runs through the allowed elements in $\mathcal{O}$.

We can have two distinct allowed elements $\mu_1 \neq \mu_2 \in \mathcal{O}$ such that $\mathfrak{X}_{\mu_1} = \mathfrak{X}_{\mu_2}$. Actually, there are only *finitely many* connected components in (6.7), by [Rot04b, Proposition 4.3].

[3] As by [Voi21, §43.6.2], when $\mathrm{End}(A) = \mathcal{O}$, or equivalently for the simple abelian surfaces in $\mathcal{Q}_{D,N}$, the last condition is always satisfied.

Remark 6.30. *As* $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}^{lin}$ *is not invariant under the* $\mathrm{Sp}_4(\mathbb{Z})$*-orbit, we can only say that for this explicit period matrix,* $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}^{lin}=1$*, but it could be possible that for another element of the orbit* $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}^{lin}=2$*.*

We conjecture that this is not the case, and that $\operatorname{rank}\mathcal{L}^{lin}\neq\operatorname{rank}\mathcal{L}$ *throughout the* $\mathrm{Sp}_4(\mathbb{Z})$*-orbit. We partially solve this for some Shimura curves in the Appendix.*

6.5.2.2 A collection of modular curves

The modular embedding construction for Shimura curves $X_0^D(N)$ can be applied to $\mathrm{Mat}_2(\mathbb{Q})$ and $\mathcal{O}_N:=\begin{pmatrix}\mathbb{Z} & \mathbb{Z}\\ N\mathbb{Z} & \mathbb{Z}\end{pmatrix}\subset\mathrm{Mat}_2(\mathbb{Q})$, N square free. We follow here [LY20, Section 6] for the case N square-free, and [Kan16] for general N. See also Section 4.2.

We take $\mu\in\mathcal{O}_N:=\begin{pmatrix}\mathbb{Z} & \mathbb{Z}\\ N\mathbb{Z} & \mathbb{Z}\end{pmatrix}\subset\mathrm{Mat}_2(\mathbb{Q})$ with trace 0, determinant N and positive $(2,1)$ entry, *i.e.*

$$\mu=\begin{pmatrix}a & b\\ cN & -a\end{pmatrix},$$

with $c>0$ and $-a^2-bcN=N$. This implies $b<0$ and, as N is square free, we have $N\mid a$. Let us rewrite μ as

$$\mu=\begin{pmatrix}aN & -b\\ cN & -aN\end{pmatrix},$$

with $b,c>0$ and $a\in\mathbb{Z}$ such that $bc-Na^2=1$.

Lemma 6.31. *For* $\mu\in\mathcal{O}_N\subset\mathrm{Mat}_2(\mathbb{Q})$ *as above, we have a map* $Y_0(N)\to\mathcal{A}_2$ *induced from*

$$\begin{aligned}\mathbb{H}&\to\mathbb{H}_2\\ \tau&\mapsto\begin{pmatrix}b\tau & aN\tau\\ aN\tau & cN\tau\end{pmatrix}.\end{aligned}$$

For $\boldsymbol{\tau}$ *of this shape, we have* $\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}=\operatorname{rank}\mathcal{L}_{\boldsymbol{\tau}}^{lin}=2$.

Remark 6.32. *Taking* $\mu=\begin{pmatrix}0 & -1\\ N & 0\end{pmatrix}$ *we recover the diagonal embedding*

$$\tau\mapsto\begin{pmatrix}\tau & 0\\ 0 & N\tau\end{pmatrix}$$

of $Y_0(N)\to\mathcal{A}_1\times\mathcal{A}_1$*, with the product of elliptic curves and product polarization. However, different choices of* μ *allow us to consider modular curves contained in the indecomposable locus, so we can evaluate the Igusa invariants on them.*

If N *is not square-free, Lemma 6.31 still constructs some explicit maps from* $Y_0(N)$ *to* $\mathcal{A}_2$*, although not necessarily all of them.*

Finally, notice that the explicit map in Lemma 6.31 gives an alternative proof of Corollary 6.16, for the images of quadratic imaginary $\tau\in\mathbb{H}$*.*

Proof. Set the embedding $\phi:\mathrm{Mat}_2(\mathbb{Q})\hookrightarrow\mathrm{Mat}_2(\mathbb{R})$ as simply the natural inclusion. We remark that we are considering complex tori $\mathbb{C}^2/\Lambda\tau$ where $\Lambda\tau=\mathcal{O}_N v_\tau$, and the Riemann form

$E_\mu(\alpha v_\tau, \beta v_\tau) = \mathrm{tr}(\mu^{-1}\alpha\bar{\beta})$. Note that the quaternionic conjugation in $\mathrm{Mat}_2(\mathbb{Q})$ is adjugation of matrices.

One can check by computation that the following basis for $\mathcal{O}_N$

$$\alpha_1 = \begin{pmatrix} b & 0 \\ aN & 0 \end{pmatrix}, \ \alpha_2 = \begin{pmatrix} aN & 0 \\ cN & 0 \end{pmatrix}, \ \alpha_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ \alpha_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

is a symplectic basis with respect to E_μ. It is a basis because from $bc - Na^2 = 1$ it follows that $\gcd(b, aN) = \gcd(a, c) = 1$, and we have verified that this basis is symplectic with SageMath [Sage]. Here is the code

```
var('x,y,a,b,c,N')
mu = matrix([[a*N,-b], [c*N,-a*N]])
al1 = matrix([[b, 0],[a*N,0]])
al2 = matrix([[a*N,0],[c*N,0]])
al3 = matrix([[0,1],[0,0]])
al4 = matrix([[0,0],[0,1]])
al = [al1, al2, al3, al4]
def E(x,y):
    return ((mu.inverse())*x*(y.adjugate())).trace()
List = [] #computes the values of E(x,y) to check they are either 0 or 1
List2 = [] #returns the pairs of indices with E(x,y) = 1
for i in range(4):
    List += [E(al[i], al[i]).full_simplify()]
    for j in range(i,4):
        List += [E(al[i], al[j]).full_simplify()]
        if E(al[i], al[j]) != 0:
            List2 += [(i,j)]
List
```

The big period matrix is then $\Pi(\tau) = (\alpha_1 v_\tau, \alpha_2 v_\tau, \alpha_3 v_\tau, \alpha_4 v_\tau)$ and the embedding into $\mathbb{H}_2$ is given by the statement. This choice of symplectic basis also induces a compatible embedding from $\mathcal{O}_N^1 = \Gamma_0(N)$ to $\mathrm{Sp}_4(\mathbb{Z})$, so it induces a map $Y_0(N) \to \mathcal{A}_2$. It follows that $\det(\boldsymbol{\tau}) = N\tau^2$, and there cannot be a generic linear dependence relation between $\det(\boldsymbol{\tau})$ and τ_1, τ_2, τ_3. □

We can also explicitly determine the quadratic form associated to $\mathcal{L}_{\boldsymbol{\tau}}$. Solving for the singular relations in (p, q, r, s, t)

$$(bp + aNq + cNr)\,\tau + sN\tau^2 + t = 0$$

forces $s = t = 0$ and

$$bp = -N\,(aq + cr)\,.$$

As $bc - Na^2 = 1$, in particular $(b, N) = 1$, therefore, necessarily $N|p$. One can check that

$$(0, -c, a, 0, 0)$$

and

$$(-N, -Nab, b^2, 0, 0)$$

are generators of the lattice (every possible value of p is attained, and for a fixed value of p, (q,r) are the solutions to a Bézout type equation and one can check that all such solutions are generated). Therefore,

$$\mathcal{L}_{\boldsymbol{\tau}} = \{(-Ny, -cx - Naby, ax + b^2y, 0, 0),\ x, y \in \mathbb{Z}\},$$

and

$$\Delta(x,y) = c^2x^2 + 2aN(bc+2)xy + b^2N(bc+3)y^2. \tag{6.8}$$

We have recovered the same positive definite quadratic form as in [Kan16, Equation (5)], with parameters $(c,b,a) \in P(N)$ in their notation. By [Kan16, Theorem 13], this quadratic form is characterized by the following properties ("quadratic forms of type N"):

- its discriminant as a binary quadratic form is $-16N$,
- for any $x, y \in \mathbb{Z}$, $\Delta(x,y) \equiv 0, 1 \pmod 4$,
- Δ primitively represents a square prime to N (note that $\Delta(1,0) = c^2$, $\Delta(b^2, -a) = b^2$).

In addition, every such quadratic form of type N is $\mathrm{SL}_2(\mathbb{Z})$-equivalent to one as in (6.8) for some parameters (c,b,a) such that $bc - Na^2 = 1$. We call the quadratic forms in (6.8) *standard*.

Consider now a general N. As we mentioned in Remark 6.32, not all admissible μ are covered by the explicit formula in Lemma 6.31. However, it follows from the treatment in [Kan16] (where N is not required to be square-free), that it is enough to construct the standard ones, as they serve as a full set of representatives of the finitely many distinct images of these maps, in bijection with the $\mathrm{GL}_2(\mathbb{Z})$-equivalence class of the quadratic forms of type N.

Analogously to the Shimura curves, let us study the loci $\cup_\mu \mathfrak{X}_\mu$. In [Kan16], these loci are considered in $\mathcal{M}_2 \cong \mathcal{A}_2^{ind}$, which is sufficient to us, as we eventually want to evaluate the Igusa invariants. The problem studied is the following: it is classical ([Wei57, Satz 2]) that for $C \in \mathcal{M}_2$, the ppas $(\mathrm{Jac}(C), \theta)$ belongs to the indecomposable locus of $\mathcal{A}_2$, in other words, for every E, E' pair of elliptic curves:

$$(\mathrm{Jac}(C), \theta) \ncong (E \times E', L_E \oplus L_{E'}),$$

as ppas, where $L_E \oplus L'_E$ is the canonical induced principal polarization on $E \times E'$. However, one could ask if there exist another principal polarization $\tilde{L}$ on $E \times E'$ such that

$$(\mathrm{Jac}(C), \theta) \cong (E \times E', \tilde{L}), \tag{6.9}$$

or equivalently, if $\mathrm{Jac}(C)$ is isomorphic to $E \times E'$ as *unpolarized* abelian varieties.

It follows that if (6.9) holds for some $C \in \mathcal{M}_2$, then necessarily $\mathrm{Hom}(E, E') \neq 0$ by [Kan16, Proposition 26], or also [Lan06, Lemma 2.2] or [DO21, Lemma 2.1]. Furthermore, if $\mathrm{Jac}(C)$ is not CM, then the elliptic curves are not CM and there exists a *unique* cyclic isogeny of minimal degree N (the degree of an isogeny generating $\mathrm{Hom}(E, E') \cong \mathbb{Z}$), see [Kan16, Corollary 27].

Definition 6.33 ([Kan16])**.** *Let $N \geq 1$ an integer, a curve $C \in \mathcal{M}_2$* has type N *if there exists E, E' elliptic curves, a cyclic isogeny $\alpha : E \to E'$ with $N = \deg(\alpha)$ and an isomorphism of abelian varieties $\mathrm{Jac}(C) \cong E \times E'$. Equivalently, there exists a principal polarization $\tilde{L} = \tilde{L}(\alpha)$ on $E \times E'$ such that (6.9) holds for $(\mathrm{Jac}(C), \theta)$. The polarization $\tilde{L}$ depends on α in an explicit way.*

Consider the locus of $\mathcal{M}_2$, for $N \in \mathbb{Z}_{>0}$ defined by

$$K(N) = \{C \in \mathcal{M}_2 |\ C \text{ has type } N\}.$$

By [Kan16, Theorem 12], these loci $K(N)$ are individually covered by modular curves associated to $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of quadratic forms of type N, and any such quadratic form is $\mathrm{SL}_2(\mathbb{Z})$-equivalent to one of the form of (6.8).

One could additionally ask if *all* modular curves are of this shape. We do not need to ponder about that, because it is true that every possible CM point belongs to one of $K(N)$, as a consequence of the following result.

Theorem 6.34 (Shioda and Mitani)**.** *If an abelian surface is isogenous to E^2 with E elliptic curve with complex multiplication, then it is* isomorphic *(as unpolarized abelian surfaces) to a product of elliptic curves.*

Proof. See [BL04, Corollary 10.6.3]. □

We finish this section with a slightly more general result via an alternative approach. Consider $\boldsymbol{\tau} \in \mathbb{H}_2$ and the lattice $\Lambda = (I_2\ \boldsymbol{\tau})\mathbb{Z}^4 \in \mathbb{C}^2$. Remark that $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$ admits a natural left $\mathrm{End}_0(A_{\boldsymbol{\tau}})$-structure. If $A_{\boldsymbol{\tau}}$ is isotypic, then $\mathrm{Mat}_2(\mathbb{Q}) \subset \mathrm{End}_0(A_{\boldsymbol{\tau}})$, hence $\Lambda \otimes_{\mathbb{Z}} \mathbb{Q}$ admits a $\mathrm{Mat}_2(\mathbb{Q})$-structure.

In the previous examples, this $\mathrm{Mat}_2(\mathbb{Q})$-structure was given simply by matrix multiplication, as the embedding $\mathrm{Mat}_2(\mathbb{Q}) \hookrightarrow \mathrm{Mat}_2(\mathbb{R})$ was the natural inclusion.

Lemma 6.35. *Let $\boldsymbol{\tau} \in \mathbb{H}_2$ with $A_{\boldsymbol{\tau}} \simeq E^2$. Suppose that the $\mathrm{Mat}_2(\mathbb{Q})$-module structure of $\Lambda_{\boldsymbol{\tau}} \otimes \mathbb{Q}$ is given by matrix multiplication. Then $\dim \mathrm{span}_{\mathbb{Q}}(1, \tau_1, \tau_2, \tau_3) = 2$.*

Proof. By hypothesis, for any $M \in \mathrm{Mat}_2(\mathbb{Q})$, we have $Mv \in (I_2\ \boldsymbol{\tau})\mathbb{Q}^4$ for v any of the basis vectors. In particular, taking $M = E_{ij}$ elementary matrices, we derive the following relations for $\boldsymbol{\tau}$: there exists $a, b, c, d \in \mathbb{Q}$ such that:

$$\begin{pmatrix}\tau_1\\0\end{pmatrix} = \begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}\tau_1\\\tau_2\end{pmatrix} = a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix} + c\begin{pmatrix}\tau_1\\\tau_2\end{pmatrix} + d\begin{pmatrix}\tau_2\\\tau_3\end{pmatrix}$$

In other words,

$$\begin{aligned}\tau_1 &= a + c\tau_1 + d\tau_2,\\ 0 &= b + c\tau_2 + d\tau_3.\end{aligned}$$

From the second equation, we have $c \neq 0$, because $\tau_3 \notin \mathbb{Q}$ (and $b = d = c = 0$ implies $\tau_1 = a \in \mathbb{Q}$, which cannot happen). Then $\tau_2 \in \mathrm{span}_{\mathbb{Q}}(1, \tau_3)$. If $d = 0$, then $\tau_2 \in \mathbb{Q}$, otherwise τ_2 depends non trivially on τ_3.

A similar argument with $\begin{pmatrix}0&0\\1&0\end{pmatrix}$ yields equations

$$\begin{aligned}0 &= a' + c'\tau_1 + d'\tau_2,\\ \tau_1 &= b' + c'\tau_2 + d'\tau_3,\end{aligned}$$

where from the first equation $d' \neq 0$ and $\tau_2 \in \mathrm{span}_{\mathbb{Q}}(1, \tau_1)$. If $c' = 0$, it follows that $\tau_2 \in \mathbb{Q}$ and the second equation gives a non trivial linear relation between τ_1 and τ_3 (as $d' \neq 0$), and

we have proven our result. If $\tau_2 \in \mathbb{Q}$, as the first case above, this argument still aplies, and we arrive at the desired conclusion.

Finally, in the more general case we have $\tau_2 \in \mathrm{span}_{\mathbb{Q}}(1,\tau_1) \cap \mathrm{span}_{\mathbb{Q}}(1,\tau_3)$ with non trivial dependence on τ_1 and τ_3. Solving for τ_2 gives a non zero linear equation that involves $1, \tau_1, \tau_3$. The conclusion follows. □

6.5.3 End of proof of Theorem 6.6

We finish the proof of Theorem 6.6. If $A_{\boldsymbol{\tau}} \simeq E^2$, with E a CM elliptic curve, as we discuss at the beginning of Section 6.5.2, we have

$$\mathrm{trdeg}\, \mathbb{Q}(q_1, q_2, q_3, j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q})) = 1.$$

By Theorem 6.34, $A_{\boldsymbol{\tau}} \in K(N)$ for some N, and hence by [Kan16, Theorem 12], it belongs to one of the curves in the collection described in Lemma 6.31, then generically in this curve $\mathrm{rank}\, \mathcal{L}_{\boldsymbol{\tau}} = \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 2$, and hence, if $A_{\tilde{\boldsymbol{\tau}}} \in C$ is defined over $\overline{\mathbb{Q}}$,

$$\min_{\mathrm{Sp}_4(\mathbb{Z})} \mathrm{trdeg}\, \mathbb{Q}(\tilde{q}_1, \tilde{q}_2, \tilde{q}_3, j_1(\tilde{q}), j_2(\tilde{q}), j_3(\tilde{q})) \leq 1.$$

Appendix 1: On a "simultaneous Humbert's lemma" for QM abelian surfaces

Recall the collection of modular curves from Section 6.5.2.2. For $A_{\boldsymbol{\tau}}$ belonging to one of those, we have explicitly given an element in the $\mathrm{Sp}_4(\mathbb{Z})$-orbit such that $\mathrm{rank}\, \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 2$. We can rephrase that as an answer to a different question. Humbert's lemma Proposition 3.12 states that if $\boldsymbol{\tau} \in \mathbb{H}_2$ solves a HSR, there is always an element in the orbit solving a linear HSR. For the elements in the locus in $\mathcal{A}_2$ covered by the collection in Section 6.5.2.2, they solve *two* linear ones. Then the question that we ask is: given $\boldsymbol{\tau} \in \mathbb{H}_2$ solving two Humbert singular relations, can we always find an element $M \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\boldsymbol{\tau}' = M\boldsymbol{\tau}$ solves two linear HSR *simultaneously?* We made the conjecture in Remark 6.30 that it was not possible for abelian surfaces with $\mathrm{End}_0(A)$ a division quaternion algebra.

The explicit embedding from Proposition 6.29 satisfies only $\mathrm{rank}\, \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 1$. We claim that, under some extra conditions, $\mathrm{rank}\, \mathcal{L}^{lin}_{\boldsymbol{\tau}} = 1$ holds for the rest of the $\mathrm{Sp}_4(\mathbb{Z})$-orbit. In other words, there is no "simultaneous Humbert lemma" for Shimura curves that allow us to linearize two HSR. We will now give a proof.

Our strategy is to translate the set-up to a Hilbert modular surface, as both Shimura and modular curves are compatible with real multiplication by the ring of integers of some quadratic field. We prove that both types of curves analytically correspond to Hirzebruch-Zagier divisors in $\mathbb{H}^2_1$. Furthermore, linear HSR will correspond to linear Hirzebruch-Zagier divisors. In this context, we have at our disposition a criterion that allows us to distinguish both types of curves by the coefficients or the divisors, and in particular implies that for compact curves (Shimura curves) the divisor cannot be linear.

We set up the notation for Hilbert modular surfaces from [MR20], and the convention for Hirzebruch-Zagier divisors from [Gee88, Chapter V].

Set a fundamental discriminant $\Delta > 0$ with $\Delta \equiv 1 (\bmod 4)$ and consider the real quadratic field $K = \mathbb{Q}(\sqrt{\Delta})$. There is an explicit embedding from the (symmetric) Hilbert modular space to the Humbert surface $\mathcal{H}_\Delta$.

Consider the following notations.

- $\mathcal{O}_K$ for the ring of integers of K;
- w for the generator of $\mathcal{O}_K$, *i.e.* $w = \frac{1+\sqrt{\Delta}}{2}$ and $\{1, w\}$ is a $\mathbb{Z}$-basis of $\mathcal{O}_K$;
- $\partial_K = \sqrt{\Delta}\mathcal{O}_K$ and $\partial_K^{-1} = \frac{1}{\sqrt{\Delta}}\mathcal{O}_K$ for the different and codifferent of K. Remark that the norm of the different ideal $\mathrm{N}(\partial_K) = \Delta$;
- for $\alpha \in K$, write $\bar{\alpha}$ for the non-trivial Galois conjugate;
- for $\mathfrak{a}$ a fractional ideal in K,

$$\Gamma(\mathfrak{a}) := \mathrm{SL}_2(\mathcal{O}_K \oplus \mathfrak{a}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(K)\ a, d \in \mathcal{O}_K,\ b \in \mathfrak{a}^{-1}, c \in \mathfrak{a} \right\} =$$
$$= \begin{pmatrix} \mathcal{O}_K & \mathfrak{a}^{-1} \\ \mathfrak{a} & \mathcal{O}_K \end{pmatrix} \cap \mathrm{SL}_2(K),$$

  and we fix $\mathfrak{a} = \partial_K$;
- the variables of the Hilbert space are denoted $\boldsymbol{z} = (z_1, z_2) \in \mathbb{H}_1^2$.

Consider the matrix $R = \begin{pmatrix} 1 & w \\ 1 & \bar{w} \end{pmatrix}$ and the map

$$\begin{aligned} \phi : \mathbb{H}_1^2 &\to \mathbb{H}_2 \\ \boldsymbol{z} &\mapsto {}^tR \begin{pmatrix} z_1 & 0 \\ 0 & z_2 \end{pmatrix} R = \begin{pmatrix} z_1 + z_2 & z_1 w + z_2 \bar{w} \\ z_1 w + z_2 \bar{w} & z_1 w^2 + z_2 \bar{w}^2 \end{pmatrix}. \end{aligned}$$

Then by [MR20, Proposition 2.11]), it defines an embedding between the symmetric Hilbert modular surface associated to K and $\mathcal{H}_\Delta$ (after properly setting an embedding $\Gamma = \Gamma(\partial_K) \to \mathrm{Sp}_4(\mathbb{Z})$) and it is independent of the choice of $\mathbb{Z}$-basis for $\mathcal{O}_K$). Setting $\boldsymbol{\tau} = \phi(\boldsymbol{z})$, it satisfies the normalized HSR (note the change of convention in [MR20])

$$\frac{\Delta - 1}{4}\tau_1 + \tau_2 - \tau_3 = 0.$$

We invert ϕ in $\mathcal{H}_\Delta$:

$$\begin{aligned} z_1 &= \frac{-\bar{w}}{\sqrt{\Delta}}\tau_1 + \frac{1}{\sqrt{\Delta}}\tau_2 = \frac{-\bar{w}\tau_1 + \tau_2}{\sqrt{\Delta}}, \\ z_2 &= \frac{w}{\sqrt{\Delta}}\tau_1 - \frac{1}{\sqrt{\Delta}}\tau_2 = \frac{w\tau_1 + \tau_2}{\sqrt{\Delta}}. \end{aligned}$$

Let us consider now $\boldsymbol{\tau}$ with $A_{\boldsymbol{\tau}}$ belonging to a Shimura curve in $\mathcal{H}_\Delta$. After applying Humbert's lemma to the relation of discriminant Δ, can consider an intersection

$$\begin{cases} \frac{\Delta-1}{4}\tau_1 + \tau_2 - \tau_3 = 0 \\ a'\tau_1 + b'\tau_2 + c\tau_3 + d(\tau_2 - \tau_1\tau_3) + e = 0 \end{cases} \implies \begin{cases} \frac{\Delta-1}{4}\tau_1 + \tau_2 = \tau_3 \\ a\tau_1 + b\tau_2 - d\det(\boldsymbol{\tau}) + e = 0, \end{cases}$$

where $a = a' + c(\Delta - 1)/4$, $b = b' + c$.

Lemma 6.36. *Under the map ϕ, there is a bijection:*

$$\left\{ \begin{array}{c} (a,b,d,e) \in \mathbb{Z}^4 : \\ a\tau_1 + b\tau_2 + d\det \boldsymbol{\tau} + e = 0 \end{array} \right\} \to \left\{ \begin{array}{c} (p,q,\gamma) \in \mathbb{Z}^2 \times \partial_K^{-1} : \\ p\sqrt{\Delta} z_1 z_2 - \bar{\gamma} z_1 + \gamma z_2 + \frac{q}{\sqrt{\Delta}} = 0 \end{array} \right\}$$
$$(a,b,d,e) \mapsto \left(d, e, \frac{1}{\sqrt{\Delta}} (a + b\bar{w}) \right).$$

In particular, linear equations in $\boldsymbol{\tau}$ correspond to linear equations in $\boldsymbol{z}$.

Proof. First, by definition of ϕ we have the following relation between determinants:

$$\det \boldsymbol{\tau} = z_1 z_2 (w - \bar{w})^2 = \Delta z_1 z_2.$$

In one direction, setting $\tau_1 = z_1 + z_2$, $\tau_2 = wz_1 + \bar{w} z_2$:

$$\begin{aligned} a\tau_1 + b\tau_2 + d\det \boldsymbol{\tau} + e &= 0 \\ a(z_1 + z_2) + b(wz_1 + \bar{w} z_2) + d\Delta z_1 z_2 + e &= 0 \\ d\Delta z_1 z_2 + (a + bw) z_1 + (a + b\bar{w}) z_2 + e &= 0 \\ d\sqrt{\Delta} z_1 z_2 + \left(\frac{a + bw}{\sqrt{\Delta}} \right) z_1 + \underbrace{\left(\frac{a + b\bar{w}}{\sqrt{\Delta}} \right)}_{=:\gamma} z_2 + \frac{e}{\sqrt{\Delta}} &= 0 \end{aligned}$$

Clearly $\gamma \in \partial_K^{-1}$, and $-\bar{\gamma} = -\frac{a+bw}{-\sqrt{\Delta}} = \frac{a+bw}{\sqrt{\Delta}}$.

In the other direction,

$$\begin{aligned} p\sqrt{\Delta} z_1 z_2 - \bar{\gamma} z_1 + \gamma z_2 + \frac{q}{\sqrt{\Delta}} &= 0 \\ p\sqrt{\Delta} \frac{\det \boldsymbol{\tau}}{\Delta} - \bar{\gamma} \left(\frac{-\bar{w}\tau_1 + \tau_2}{\sqrt{\Delta}} \right) + \gamma \left(\frac{w\tau_1 + \tau_2}{\sqrt{\Delta}} \right) + \frac{q}{\sqrt{\Delta}} &= 0 \\ (\gamma w + \overline{\gamma w})\, \tau_1 - (\bar{\gamma} + \gamma)\, \tau_2 + p \det(\boldsymbol{\tau}) + q &= 0, \end{aligned}$$

and we just need to check that the coefficients are integers (remark that neither γ nor $w\gamma$ are algebraic integers). Write $\gamma = (m + nw)/\sqrt{\Delta}$, then:

$$\gamma + \bar{\gamma} = \frac{m + nw - m - n\bar{w}}{\sqrt{\Delta}} = \frac{n(w - \bar{w})}{\sqrt{\Delta}} = n \in \mathbb{Z},$$

and the same argument applies to $w\gamma$, as it also belongs to ∂_K^{-1}. □

Let us arrange the coefficients of the equations in $\boldsymbol{z}$ in a matrix

$$B = \begin{pmatrix} p\sqrt{\Delta} & \gamma \\ -\bar{\gamma} & \frac{q}{\mathrm{N}(\partial_K)} \sqrt{\Delta} \end{pmatrix},$$

with $p, q \in \mathbb{Z}, \gamma \in \partial_K^{-1}$. Then this matrix is what in [Gee88, Chapter V, Definition(1.2)] is called a *skew-hermitian matrix integral with respect to ∂_K*. Its determinant satisfies:

$$pq + \mathrm{N}(\gamma) \in \frac{1}{\mathrm{N}(\partial_K)} \mathbb{Z}.$$

One can think of $\Delta \det(B)$ as an analog to the discriminant of the HSR, in the sense that the action of Γ changes the equation but respects the determinant. Following [Gee88], we define the subset of $\mathbb{H}_1^2$:

$$T(M) = \bigcup_{\substack{B=\begin{pmatrix} p\sqrt{\Delta} & \gamma \\ -\bar{\gamma} & \frac{q}{\mathrm{N}(\partial_K)}\sqrt{\Delta} \end{pmatrix} \\ \det(B)=\frac{M}{\mathrm{N}(\partial_K)}}} \{(z_1, z_2) \in \mathbb{H}_1^2 : \; p\sqrt{\Delta} z_1 z_2 - \bar{\gamma} z_1 + \gamma z_2 + \frac{q}{\mathrm{N}(\partial_K)}\sqrt{\Delta} = 0\},$$

which is Γ-invariant and defines a divisor on $\mathbb{H}_1^2$, the *Hirzebruch-Zagier divisor of discriminant* M. We remark that the standard definition of these divisors requires $\det B$ positive, and the ones considered in Lemma 6.36 do not necessarily satisfy that. However, the relevant results to us (namely the discussion between [Gee88, Chapter V, Lemma 1.4 and Proposition 1.5]) do not require the determinant to be positive.

As we have said, this divisor can define a Shimura or modular curve, and one can tell them apart because the former are naturally compact (equivalently, they do not meet the resolution of the cusps of the compactification of the Hilbert surface). More precisely, it follows from [Gee88, Chapter V, Proposition 1.5] (and from the prior discussion, which does not require $\det(B) > 0$), that the indefinite quaternion algebra over $\mathbb{Q}$ corresponding to $T(M)$ is

$$\left(\frac{\Delta, \frac{-M}{\mathrm{N}(\partial_K)\Delta}}{\mathbb{Q}}\right).$$

We are finally able to prove that Shimura curves do not admit a simultaneous Humbert lemma. Arguing by contradiction, assume that it admits a description by intersection of two linear HSR.

Assumption 1. *Assume that one of them is a normalized HSR with respect to a fundamental discriminant* $\Delta \equiv 1 \pmod 4$.

By Lemma 6.36, the linear HSR gives a linear equation for the Hirzebruch-Zagier divisor corresponding to the Shimura curve in the Hilbert modular surface of discriminant Δ. The corresponding matrix is of the shape

$$B = \begin{pmatrix} 0 & \gamma \\ -\bar{\gamma} & \frac{q}{\sqrt{\Delta}} \end{pmatrix},$$

with $\det(B) = \gamma\bar{\gamma} \neq 0$. Consider the quaternion algebra

$$\left(\frac{\Delta, \frac{-\det(B)}{\Delta}}{\mathbb{Q}}\right) = \left(\frac{\Delta, \frac{-\gamma\bar{\gamma}}{\Delta}}{\mathbb{Q}}\right) = \left(\frac{\Delta, \mathrm{N}(\alpha)}{\mathbb{Q}}\right)$$

where $\alpha := \gamma/\sqrt{\Delta}$ (remark the change of sign, because $\mathrm{N}(\sqrt{\Delta}) = -\Delta$). It is finally enough to see that the quaternion algebra splits over $\mathbb{Q}$ to arrive to a contradiction.

Lemma 6.37. *Consider* $\Delta > 0$ *a fundamental discriminant,* $K = \mathbb{Q}(\sqrt{\Delta})$ *and* $\alpha \in K^\times$. *It follows*

$$\left(\frac{\Delta, \mathrm{N}(\alpha)}{\mathbb{Q}}\right) \cong \mathrm{Mat}_2(\mathbb{Q})$$

Proof. Consider the standard $\mathbb{Q}$-basis $\{1, I, J, IJ\}$ with $I^2 = \Delta$, $J^2 = \mathrm{N}(\alpha)$ and $IJ = -JI$. We can then identify $I = \sqrt{\Delta}$. By classical theory of quaternion algebras over $\mathbb{Q}$, the statement is equivalent to the quaternion algebra having a non-trivial element of norm zero. We can explicitly find one: write $\alpha = a + b\sqrt{\Delta}$ with $a, b \in \mathbb{Q}$. Then

$$\mathrm{N}(\alpha) = a^2 - \Delta b^2.$$

Consider the element $\mu := a + bI + J$. Then

$$(a + bI + J)(a - bI - J) = a^2 - (bI)^2 - J^2 = a^2 - \Delta b^2 - \mathrm{N}(\alpha) = 0.$$

□

The condition $\Delta \equiv 1 \pmod 4$ in Assumption 1 could be removed by repeating this analysis for the quadratic fields of discriminant divisible by four. The true impositions are that one of the linear HSR is *normalized* and with respect to a *fundamental* discriminant. Hence this proof does not rule out that the linear HSR could be not normalized, or normalized with respect to an order in a quadratic field that it is not the ring of integers.

Appendix 2: Examples of lattices of Humbert singular relations

In Subsection 6.4.1 and Section 6.5 we considered the positive definite lattice of Humbert singular relations and worked explicitly on this object. In this appendix, we recompile specific examples of the lattices on the quadratic forms. All the computations have been made with [MAGMA], and the examples come from [LMFDB].

Definition 6.38. *For $\boldsymbol{\tau} \in \mathbb{H}_2$, we denote* the lattice of Humbert singular relations *the following $\mathbb{Z}$-module*

$$\mathcal{L}_{\boldsymbol{\tau}} := \{(a, b, c, d, e) \in \mathbb{Z}^5 : a\tau_1 + b\tau_2 + c\tau_3 + d(\tau_2^2 - \tau_1\tau_3) + e = 0\},$$

equipped with the positive definite quadratic form induced by the discriminant Δ. It is a positive definite integral lattice (as in finite rank free abelian group with a symmetric bilinear form).

Here we simply include some examples of computations of period matrices of abelian surfaces.

Example 6.39. *The Jacobian of the genus 2 curve given by*

$$y^2 = x^6 + 4x^3 + 8,$$

has endomorphism algebra isomorphic to $\mathrm{Mat}_2(\mathbb{Q}(\sqrt{-6}))$, it admits an isogeny to E^2, for E is a CM elliptic curve for the order $\mathbb{Z}[\sqrt{-6}]$ (`https://www.lmfdb.org/Genus2Curve/Q/5184/a/46656/1` in LMFDB).

Computing numerically its period matrix with Magma, where e−20 is for scientific notation,

$$\begin{pmatrix} -5.4210108624275221700\text{e}{-20} + i1.6329931618554520654 & -2.7105054312137610850\text{e}{-20} + i0.81649658092772603267 \\ 2.0328790734103208138\text{e}{-20} + i0.81649658092772603273 & 2.0328790734103208138\text{e}{-20} + i1.6329931618554520654 \end{pmatrix}.$$

One can guess that all the entries are purely imaginary numbers (and they belong to $\overline{\mathbb{Q}}$ as this is a CM point), and after squaring them, they coincide up to high precision with rational numbers ($-0.666..$ in the case of τ_2). More precisely, it coincides numerically up to high precision with the matrix:

$$\boldsymbol{\tau} = \begin{pmatrix} i2\sqrt{\frac{2}{3}} & i\sqrt{\frac{2}{3}} \\ i\sqrt{\frac{2}{3}} & i2\sqrt{\frac{2}{3}} \end{pmatrix}$$

```
C :=ComplexField(200);
P<x> := PolynomialRing(C);
f := x^6 + 4*x^3 + 8;
A := AnalyticJacobian(f);
tau := SmallPeriodMatrix(A);
a := Sqrt(-2/3);
Abs(tau[1][2] - a);
Abs(tau[1][1] - 2*a);
Abs(tau[2][2] - 2*a);
```

Solving for the lattice of singular relations in (a, b, c, d, e) yields the system

$$\begin{aligned} 2d - e &= 0 \\ 2a + b + 2c &= 0 \end{aligned}$$

hence the HSR are parametriced by $(a, -2a - 2c, c, d, 2d)$, with discriminant $b^2 - 4ac - 4de = (-2a - 2c)^2 - 4ac - 8d = 4(a^2 + c^2 + ac - 2d)$, so

$$(\mathcal{L}_{\boldsymbol{\tau}}, \Delta) = (\{(x, y, z) \in \mathbb{Z}^3 : (x, -2x - 2y, y, z, 2z)\}, 4(x^2 + y^2 + xy - 2z)$$

There are three (up to isomorphism over $\overline{\mathbb{Q}}$) genus two curves in the LMFDB database such that their jacobian has endomorphism algebra $\mathrm{Mat}_2(CM)$ for a quadratic imaginary field. Those are given by:

$$\begin{aligned} C_1 &: y^2 = x^6 + 4x^3 + 8 \\ C_2 &: y^2 = 4x^6 + 1 \\ C_3 &: y^2 = x^5 - x \end{aligned}$$

with $\mathrm{End}_0(\mathrm{Jac}(C_i))$ isomorphic to $\mathrm{Mat}_2(\mathbb{Q}(\sqrt{-6}))$, $\mathrm{Mat}_2(\mathbb{Q}(\sqrt{-3}))$ and $\mathrm{Mat}_2(\mathbb{Q}(\sqrt{-2}))$, respectively. Analogously, it can be checked with Magma that their period matrices coincide numerically with:

$$\begin{aligned} \boldsymbol{\tau}_1 &= \begin{pmatrix} \frac{2}{3}\sqrt{-6} & \frac{1}{3}\sqrt{-6} \\ \frac{1}{3}\sqrt{-6} & \frac{2}{3}\sqrt{-6} \end{pmatrix} \\ \boldsymbol{\tau}_2 &= \begin{pmatrix} -1 - \frac{2}{3}\sqrt{-3} & \frac{1}{3}\sqrt{-3} \\ \frac{1}{3}\sqrt{-3} & 2 - \frac{2}{3}\sqrt{-3} \end{pmatrix} \\ \boldsymbol{\tau}_3 &= \begin{pmatrix} \frac{1}{3} + \frac{2}{3}\sqrt{-2} & \frac{-1}{3} + \frac{1}{3}\sqrt{-2} \\ \frac{-1}{3} + \frac{1}{3}\sqrt{-2} & \frac{-1}{3} + \frac{2}{3}\sqrt{-2} \end{pmatrix} \end{aligned}$$

Example 6.40. *The Jacobian of the curve $y^2 = x^6 + 3x^4 + x^2 - 1$ admits a $(2,2)$-isogeny (over $\mathbb{Q}$) to the product $E_1 \times E_2$, E_i with CM by $\mathbb{Q}(\sqrt{-2})$ and $\mathbb{Q}(i)$, respectively ([FKRS12, page 1422, first paragraph after item viii)]).*

Using Magma one can check that its normalized period matrix has the form:

$$\boldsymbol{\tau} = \begin{pmatrix} 2i & i \\ i & \frac{-1}{2} + \frac{\sqrt{-2}+i}{2} \end{pmatrix}$$

with determinant $-(\sqrt{2}+i)$.

As $\mathrm{Jac}(C)$ is isogenous to $E_1 \times E_2$, $\mathrm{End}_0(\mathrm{Jac}(C)) \cong \mathbb{Q}(\sqrt{-2}) \times \mathbb{Q}(i)$ as *product of rings*. It is true that the minimal field of definition of $\mathrm{Jac}(C)$ and all its endomorphisms can be taken as the compositum $\mathbb{Q}(\sqrt{-2})\mathbb{Q}(i) = \mathbb{Q}(\sqrt{-2}, i) = \mathbb{Q}(\xi_8)$, for ξ_8 a primitive 8-th root of unity [FKRS12, Prop 4.5])

For computing all singular relations, let's rewrite in terms of the $\mathbb{Q}$-base of $\{1, \xi, \xi^2, \xi^3\}$ for $\xi = \xi_8$. Hence

$$\boldsymbol{\tau} = \begin{pmatrix} 2\xi^2 & \xi^2 \\ \xi^2 & \frac{-1+\xi+\xi^2+\xi^3}{2} \end{pmatrix}$$

with $\tau_2^2 - \tau_1\tau_3 = \xi + \xi^2 + \xi^3$. Finding all singular relations

$$a\tau_1 + b\tau_2 + c\tau_3 + d(\tau_2^2 - \tau_1\tau_3) + e = 0$$

for $(a,b,c,d,e) \in \mathbb{Z}$ it's then equivalent to solving the system :

$$\begin{aligned} \frac{-c}{2} + e &= 0 \\ \frac{c}{2} + d &= 0 \\ 2a + b + \frac{c}{2} + d &= 0 \\ \frac{c}{2} - d &= 0 \end{aligned}$$

which has solutions given by $(a', -2a', 0, 0, 0)$, so the only primitive one is $(1, -2, 0, 0, 0)$, of discriminant 4.

Example 6.41. *Consider for $N \in \mathbb{Z}_{>0}$ fixed and $\tau \in \mathbb{H}$ the period matrix*

$$\boldsymbol{\tau} = \begin{pmatrix} \tau & 0 \\ 0 & N\tau \end{pmatrix},$$

which induces one of the modular embeddings of $Y_0(N) \to \mathcal{A}_2$, that in this case is contained in $\mathcal{A}_1 \times \mathcal{A}_1$.

Solving for

$$\begin{aligned} a\tau + b0 + cN\tau - d(N\tau^2) + e &= 0 \\ -dN\tau^2 + (a + cN)\tau + e &= 0 \end{aligned}$$

which implies $e = 0$ and $d = 0$ if τ is not quadratic, and hence

$$(\mathcal{L}_{\boldsymbol{\tau}}, \Delta_{\boldsymbol{\tau}}) = (\{(x,y) \in \mathbb{Z}^2 : (-Nx, y, x, 0, 0)\},\ y^2 + 4Nx^2),$$

and any other restriction can only come from τ being quadratic imaginary.

Example 6.42. *In [LY20, Example 34], they give another embedding for $Y(N)$ into $\mathcal{A}_2$ that doesn't lie in the indecomposable locus ([LY20, Example 37], χ_{10} doesn't vanish identically). This is given for N odd and by*

$$\boldsymbol{\tau} = \begin{pmatrix} \frac{N+1}{2}\tau & N\tau \\ N\tau & 2N\tau \end{pmatrix}.$$

One checks that $\det \boldsymbol{\tau} = -N\tau^2$, so the equation for singular relations:

$$\left(a\frac{N+1}{2} + bN + 2cN\right)\tau + dN\tau^2 + e = 0.$$

which implies $e = 0$ and $d = 0$ (unless τ is quadratic), so we just need to solve for (a, b, c) in $a(N+1)/2 + bN + 2cN = 0$, or $-a\frac{N+1}{2} = N(b + 2c)$. Remark that $\gcd(N, (N+1)/2) = 1$, so $N|a$. Writing $a = -fN$,

$$b + 2c = f\frac{N+1}{2},$$

so (b, c) are given by the solutions to Bézout's identity. As $\gcd(1, 2) = 1$, by the extended Euclidean algorithm, those are given by one initial solution $(g_0, h_0) = (f\frac{N+1}{2}, 0)$ and then for $t \in \mathbb{Z}$ by $(g_0 - 2t, h_0 + t)$. In conclusion,

$$\mathcal{L}_{\boldsymbol{\tau}} = \{(x, y) \in \mathbb{Z}^2\ (-xN, x\frac{N+1}{2} - 2y, y, 0, 0)\},$$

so $\Delta = (x(N+1)/2 - 2y)^2 + 4xyN = x^2(N+1)^2/4 + 2(N-1)xy + 4y^2$.

Remark that by the general formula in the proof of Lemma 6.31, our initial solution would have been $(g_0, h_0) = (f\frac{N+1}{2}, f\left(\frac{N+1}{2}\right)^2)$, so the lattice and the quadratic form have more complicated formulas, that are however $\mathrm{GL}_2(\mathbb{Z})$-equivalent. This one agrees with the formula from [LY20, Example 39], and they claim that the quadratic form is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to

$$\begin{cases} 4x^2 + Ny^2, & N \equiv 1(4), \\ 4x^2 + 4xy + (N+1)y^2, & N \equiv 3(4). \end{cases}$$

Example 6.43. *The curve $y^2 = x^6 + 4x^5 + 6x^4 + 2x^3 + x^2 + 2x + 1$ is isogenous to the square of an (non-CM) elliptic curve (https://www.lmfdb.org/Genus2Curve/Q/169/a/169/1).*

With Magma one checks that its period matrix coincides numerically with

$$\begin{pmatrix} \alpha i & \frac{-1+\alpha i}{2} \\ \frac{-1+\alpha i}{2} & \alpha i \end{pmatrix}$$

for $\alpha \in \mathbb{R} \setminus \overline{\mathbb{Q}}$ (that numerically is approximately 1.173688920770601387588)
using the code

```
C<i> :=ComplexField(200);
P<x> := PolynomialRing(C);
f := x^6 + 4*x^5 + 6*x^4 + 2*x^3 + x^2 + 2*x + 1;
A := AnalyticJacobian(f);
tau := SmallPeriodMatrix(A);
al := -(tau[1][1])*i;
tau[1][1] - al*i;
tau[1][2] - (-0.5 + 0.5*al*i);
tau[2][2] - al*i;
```

One checks further that $\tau_2^2 - \tau_1\tau_3 = (3\alpha^2 + 1 - 2\alpha i)/4$. Hence the singular relations for this curve come from solving:

$$a\alpha i + \frac{b}{2}(-1 + \alpha i) + c\alpha i + \frac{d}{4}\left(3\alpha^2 + 1 - 2\alpha i\right) + e = 0.$$

As the term α^2 only appears with d (and α is transcendental), that forces $d = 0$. Then

$$(\mathcal{L}_{\boldsymbol{\tau}}, \Delta_{\boldsymbol{\tau}}) = (\{(x, 2y, y - x, 0, y)\}, 4(y^2 - xy + x^2).$$

Hence its rank is 2, but the singular relations do not involve the determinant, so we have two linear relations and hence $\mathrm{trdeg}(q_1, q_2, q_3) \leq 1$ for this curve.

The following examples are of type $Q \times CM$.

Example 6.44. *The curve $y^2 = x^6 - 6x^4 + x^2 + 28$ is of type $Q \times CM$ (https://www.lmfdb.org/Genus2Curve/Q/448/a/448/2).*

We calculate its period matrix

$$\begin{pmatrix} -\frac{3}{2} + \beta i & \frac{-1}{2} + (\beta - 2)i \\ \frac{-1}{2} + (\beta - 2)i & -\frac{3}{2} + \beta i \end{pmatrix}$$

and

$$\tau_2^2 - \tau_1\tau_3 = -6 + 4\beta + (2\alpha + 2)i,$$

for $\beta \in \mathbb{R} \setminus \overline{\mathbb{Q}}$ (that numerically is approximately 2.121098670864189300993018782 58) using the code

```
C<i> :=ComplexField(200);
P<x> := PolynomialRing(C);
f := x^6 - 6*x^4 + x^2 +28;
A := AnalyticJacobian(f);
tau := SmallPeriodMatrix(A);
al := -(tau[1][1] + 1.5)*i;
tau[1][1] - (-1.5 + al*i);
tau[1][2] - (-0.5 + (al - 2)*i);
tau[2][2] - (-1.5 + al*i);
tau[1][2]^2 - tau[1][1]*tau[2][2] - (-6 + 4*al + (2*al + 2)*i);
```

Taking real parts, it follows that $d = -3a - b - 3c + 2e = 0$. Taking imaginary parts,

$$\beta\,(a + b + c) - 2b = 0.$$

As $\beta \notin \mathbb{Q}$, it follows $2b = 0$ and $a + b + c = 0$. Therefore, $\mathcal{L}_{\boldsymbol{\tau}} = \{(a, 0, -a, 0, 0)\}$.

Example 6.45. *For another example with $CMxQ$, we have the curve $y^2 = 4x^5 + 5x^4 + 10x^3 + 5x^2 + 4x$ with endomorphism algebra $\mathbb{Q} \times \mathbb{Q}(\sqrt{-7})$.*

Its periods matrix coincides numerically with

$$\begin{pmatrix} 2 + \gamma i & \frac{1}{2} - \left(\frac{\sqrt{7}}{2} - \gamma\right) i \\ \frac{1}{2} - \left(\frac{\sqrt{7}}{2} - \gamma\right) i & 1 + \gamma i \end{pmatrix}$$

with $\gamma = 1.22198716319824229812191332969730705\ldots \in \mathbb{R} \setminus \overline{\mathbb{Q}}$.

```
C<i> :=ComplexField(200);
P<x> := PolynomialRing(C);
f := 4*x^5 + 5*x^4 + 10*x^3 + 5*x^2 + 4*x;
A := AnalyticJacobian(f);
tau := SmallPeriodMatrix(A);
al := -(tau[1][1] - 2)*i;
tau[1][1] - (2 + al*i);
tau[1][2] - (0.5 + (Sqrt(7)/2 -al)*i);
tau[2][2] - (1 + al*i);
tau[1][2]^2 -tau[1][1]*tau[2][2] - (-7/2 + Sqrt(7)*al + (Sqrt(7)/2 - 4*al)*i);
```

Can check too that $\tau_2^2 - \tau_1\tau_3 = \frac{-7}{2} + \sqrt{7}\gamma + \left(\frac{\sqrt{7}}{2} - 4\gamma\right) i$. Solving now for singular relations, taking real part:

$$2a + \frac{b}{2} + c + d(\frac{-7}{2} + \sqrt{7}\gamma) + e = 0.$$

So $d = 0$ and $2a + c + e = 0$. From the imaginary part,

$$a\gamma - b\left(\frac{\sqrt{7}}{2} - \gamma\right) + c\gamma = 0,$$

and so $b = 0$, and $a + c = 0$. Finally, $\mathcal{L}_{\boldsymbol{\tau}} = \{(a, 0, -a, 0, -a)\}$. Then $q_1 = -q_3$ and $q_2 = iq_1e^{\pi\frac{\sqrt{7}}{2}}$. Hence $\operatorname{trdeg}\mathbb{Q}(q_1, q_2, q_3) = \operatorname{trdeg}\mathbb{Q}(q_1, e^{\pi\frac{\sqrt{7}}{2}})$. By Gelfond-Schneider, $e^{\pi\frac{\sqrt{7}}{2}} \notin \overline{\mathbb{Q}}$, hence have $1 \leq \operatorname{trdeg}(\mathbb{Q}(q_1, q_2, q_3) \leq 2$.

Chapter 7

Modular proof of the one dimensional Stéphanois theorem for prime levels

Here is a generalization of the proof of the Stéphanois theorem [BDGP96]. We do it for the modular version in [Wal96, Théorème 1], but it also works for the original proof. The generalization relies on being able to only use prime powers q^p, instead of q^n for every $n \in \mathbb{Z}_{>0}$. This observation is useful, as in genus two the modular polynomials in the literature are only considered for prime levels.

Furthermore, one can even restrict to primes in fixed arithmetic sequence.

On the proof itself, the only crucial difference appears in [Wal96, Cinquième pas, page 5] proof, and on the final contradiction in [Wal96, Septième pas, page 7], but for readability we present a complete proof with this modification.

7.1 Introduction

Consider the elliptic j-invariant $j : \mathbb{H} \to \mathbb{C}$, a $\mathrm{SL}_2(\mathbb{Z})$-invariant modular function, which classifies elliptic curves over $\overline{\mathbb{Q}}$. It admits a Fourier expansion with integer coefficients:

$$J(q) = \frac{1}{q} + 744 + 196884q + \cdots \; q = e^{2\pi i \tau},$$

which defines a meromorphic function on $\mathbb{D} = \{q \in \mathbb{C} : |q| < 1\}$, the only pole being at $q = 0$ (and simple). We are going to present the classical proof of the following theorem.

Theorem 7.1. *Let* $q \in \mathbb{D} \setminus \{0\}$, *then* $\operatorname{trdeg} \mathbb{Q}(q, J(q)) \geq 1$.

We are going to present here the “modular" proof of the Stéphanois theorem ([BDGP96]) in [Wal96]. The “modular" approach comes from Daniel Bertrand [Ber97]. The main change in both approaches rely on how to “cancel out" the pole at infinity of the j-function. The original was simply to consider $\tilde{J}(q) := qJ(q)$. The modular approach uses the modular discriminant form Δ, as it has a simple zero at infinity, and consider $\Delta(q)J(q)$. Both of them follow the same strategy, the main difference steam from:

- The auxiliary function $F(q)$ requires bound on the (integral) Fourier coefficients of either $\tilde{J}^l(q)$ or $\Delta^k J(q)^l$ for $k \geq l$. In the first case, they require Mahler's bound on the

coefficients of powers of the j-invariant. In the modular proof, one observes that, *independently of* l, $\Delta^k J(q)^l$ is a *cusp* for of weight $12k$, hence one can invoke Hecke's bound on the Fourier coefficients.

- There is a step requiring lower bounds on the auxiliary function F on q^S for a fixed $S \in \mathbb{Z}_{>0}$, where q is an (eventually impossible) point in $\mathbb{D}$ such that $(q, J(q)) \in \overline{\mathbb{Q}}^2$. On the original proof, by construction $F(q) = A(q^S, J(q^S))$ is a polynomial with integer coefficients evaluated in the pair of algebraic numbers $(q^S, J(q^S))$ (remark that $J(q^S) \in \overline{\mathbb{Q}}$ because $J(q) \in \overline{\mathbb{Q}}$ and $\Phi_S(J(q), J(q^S)) = 0$ with $\Phi_S \in \mathbb{Z}[X, Y]$ the modular polynomial of level S). Then the proof of a lower bound relies on Liouville's inequality for the algebraic number $F(q^S, J(q^S))$. In the modular proof, the auxiliary function is now $F(q) = \Delta^{2N} A(q, J(q))$ (where $\deg_X A < N$ and $\deg_Y A < N$). Then $\Delta^{2N}(q^S)$ is not algebraic anymore, but one combines the previous lower bound on $\Delta^{-2N}(q^S)F(q^S) = A(q^S, J(q^S)) \in \overline{\mathbb{Q}}$, and a lower bound on $\Delta^{2N}(q^S)$ that is negligible in comparison with the first one.

We focus on the modular proof with a goal of generalize it to the Igusa invariants for genus two curves (or for Gundlach invariants for genus two curves with real multiplication by $\mathbb{Q}(\sqrt{5})$). The corresponding invariants are also defined as quotients of (Siegel or Hilbert) modular forms, with the denominator being (powers of) a cusp form, and it zero locus is the only subset where the invariants are not defined. Hence one could use it to "get rid of poles", as it happens with Δ. More importantly, Hecke's bound on Fourier coefficients on cusp forms generalizes to both Siegel and Hilbert cusp forms.

Finally, in this presentation on the modular proof of the Stéphanois theorem, we change a key point in the type of isogenies considered. The main strategy goes back to the Mahler method: one starts with a pair $(q, J(q)) \in \overline{\mathbb{Q}}^2$. For any $n \in \mathbb{Z}_{>0}$, then $(q^n, J(q^n)) \in \overline{\mathbb{Q}}^2$, because as we said before $J(q^n)$ is algebraic over $\mathbb{Q}(J(q))$ via the modular polynomial of level n. For the proof, one requires more information on the modular polynomial (namely degree and height).

What we present here, is that one could restrict oneself to only work with prime powers q^p for the proof. It is a small change in the strategy (following [Wal96, Théorème 1], the only modifications come at Cinquième pas, on the determination on the fixed power P, and Septième pas, on the final contradiction in the transcendence proof. In general terms, restricting oneself to working only with prime powers gives a worse bound on P, but it is still enough to conclude in the contradiction.

Morally, one should expect this modification to still work for the proof because every isogeny between elliptic curves factors as composition of prime isogenies. In a way, all the information encoded in the modular polynomials in every level is already encoded in the modular polynomials in prime level. Again, with a goal in the generalizations, it is good news to us, because in higher dimensions, only the modular polynomials for prime levels have been described in the literature.

Furthermore, one can restrict even more to work with primes in fixed arithmetic sequences.

7.1.1 List of parameters

Let us present now the proof. First, as it happens with a standard transcendence proof, we work with integral parameters, dependent on each other, such that we have an impossible

inequality in the end. Alongside, there will be a sequence of constants (some absolute, other dependent only on q) appearing on our relations, for which we reserve the notation C_i for $i = 0, 1, \dots$. Likewise, we start the proof with a fixed $q \in \mathbb{D}$ such that $(q, J(q)) \in \overline{\mathbb{Q}}$. For the complex variable in $\mathbb{D}$, we set z. We construct an auxiliary polynomial (depending on $N \in \mathbb{Z}_{>0}$) $A \in \mathbb{Z}[X, Y]$ via Siegel's lemma, with $\deg_X A < N$ and $\deg_Y A < N$, and consider an auxiliary function $F(z) := \Delta^{2N}(z)A(z, J(A)$. We present the parameters first on this list.

$N \in \mathbb{Z}_{>0}$ will appear first as $\deg_X A < N$ and $\deg_Y A < N$, for $A \in \mathbb{Z}[X, Y]$ the auxiliary polynomial. All other functions and parameters depend on it. Its sole purpose is being large enough for the contradiction to happen in the end. It depends on $|q|$ only in (7.3), and it is a lower bound for N.

$L \in \mathbb{Z}_{>0}$ is set in the application of the Siegel's lemma for $F(z) = \Delta^{2N} A(z, J(z))$. The polynomial A is constructed so that F vanishes at $z = 0$ with order at least L. In particular one should have $N^2 \geq 2L$ for the bounds of the Siegel's lemma, and eventually it will be set $L := [N^2/2]$, for $[\cdot]$ the integer part, therefore

$$2L \leq N^2 < 2(L + 1)$$

$M \in \mathbb{Z}_{>0}$ is defined as the actual order of vanishing of the auxiliary function F at 0, hence

$$M \geq L.$$

$P \in \mathbb{Z}_{>0}$ will be the prime power q^P that we consider eventually. Its the only parameter depending on q. Surprisingly, consider the strategy of the proof is build on Mahler's method, it cannot be set to be arbitrarily large. Hence it will be defined as the minimum P possible, and one requires upper bounds of it in terms of the other parameters, more precisely, in (7.7), it will be set

$$\frac{P^2}{\log P} = O_{|q|}\left(N \log M\right)$$

7.2 Preliminary lemmas

Here we put the preliminary necessary results on cusp modular forms and modular polynomials.

Proposition 7.2. *The meromorphic function $J : \mathbb{D} \to \mathbb{C}$ induced by the Fourier expansion of the j-invariant is transcendental over $\mathbb{Q}(z)$.*

Proof. See [BDGP96, Lemme 4].

□

7.2.1 Modular cusp forms

Lemma 7.3 (Hecke bound on cusp forms)**.** *There exists an absolute constant $C_1 > 0$ such that the following holds. For any $N \in \mathbb{Z}_{>0}$ and any $1 \leq l \leq N$, the functions $\Delta^{2N} J^l$ are modular cusp forms of weight $24N$, and if we consider their Fourier expansions at infinity:*

$$\Delta^{2N} J^l := \sum_{k \geq 1} c_{N,l}(k) q^k,$$

then $c_{N,l}(k) \in \mathbb{Z}$ *for all* N, k, k *and*

$$|c_{N,l}(k)| \leq C_1^N k^{12N}$$

Proof. First $\Delta^{2N} J^l$ is holomorphic for that Δ has a simple zero at infinity that cancels out with the pole of J for our choices of parameters $l \leq N$, and have that $\Delta^{2N} J^i$ vanishes at infinity. Its obvious that is a modular form of weight $24N$.

Finally, one carries out the classical Hecke proof, keeping track of the corresponding explicit constants, see [Ser73, Chapter VII, Section 4, Theorem 5]. □

7.2.2 Height lemmas

We also state Liouville's inequality[1] for algebraic numbers, and another result of height nature. Remark the following notation for $\alpha \in \overline{\mathbb{Q}}$:

$\deg(\alpha)$ for its degree, i.e. $\deg(\alpha) = [\mathbb{Q}(\alpha) : \mathbb{Q}]$,

$h(\alpha)$ for its logarithmic (Weil) height.

$M(\alpha)$, for $\alpha \in \overline{\mathbb{Q}}$ for the Mahler measure of the minimal polynomial of α in $\mathbb{Z}[X]$.

$m(\alpha) = \log M(\alpha)$ for its logarithmic Mahler measure, remark then that $m(\alpha) = \deg(\alpha) h(\alpha)$.

for any polynomial $P \in \mathbb{Z}[X_1, \dots X_r]$ we denote its *height* and its *length* as $H(P) := \max |c|$ and $L(P) := \sum |c|$ with c ranging over the coefficients of P. Likewise, we have lowercase notation for its logarithms: $h(P) := \log h(P)$, and $l(P) = \log L(P)$.

Lemma 7.4 (Liouville's inequality)**.** *Let* $0 \neq \alpha \in \overline{\mathbb{Q}}$*, then it holds that*

$$\log |\alpha| \geq -\deg(\alpha) h(\alpha).$$

Proof. See [BG06, Theorem 11.5.21]. □

The second lemma we use of height theory allows us, for $\alpha, \beta \in \overline{\mathbb{Q}}$ linked via $P(\alpha, \beta) = 0$ for $P \in \mathbb{Z}[X, Y]$, to control arithmetic information of β in terms of arithmetic information of α and P. It will eventually be used for the modular polynomials.

Lemma 7.5. *Consider* $\alpha, \beta \in \overline{\mathbb{Q}}$ *and a polynomial* $P \in \mathbb{Z}[X, Y]$ *such that* $P(\alpha, \beta) = 0$*. Assume that* $P(\alpha, Y)$ *is not a constant polynomial. Then*

$$m(\beta) \leq \deg(\alpha) \left(\log L(P) + \deg_x(P) h(\alpha) \right).$$

Proof. See [BDGP96, Lemme 5]. □

[1] There seems to be a variety of results in Diophantine approximation with this name.

Finally, we are going to require bounds of the height of the evaluation of an integral polynomial in algebraic values. Remark that this is a generalization of the standard properties of the height for $\alpha, \beta \in \overline{\mathbb{Q}}$:

$$h(\alpha\beta) \leq h(\alpha) + h(\beta), \text{ and } h(\alpha \pm \beta) \leq \log 2 + h(\alpha) + h(\beta).$$

Lemma 7.6. *Consider $P \in \mathbb{Z}[X_1, \dots X_n]$ a non zero polynomial, and $\alpha_1, \dots \alpha_n \in \overline{\mathbb{Q}}$. Then the following holds*

$$h(P(\alpha_1, \dots, \alpha_n)) \leq \log L(P) + \sum_{i=1}^{n} \deg_{X_i}(P) h(\alpha_i).$$

Proof. See [Wal00, Lemma 3.7]. □

We state the properties of the modular polynomials in the following subsection for that the original proof in [BDGP96] requires then for comparing $h(J(q^p))$ with $h(J(q))$, but one can alternatively use a more geometric approach in [Ber97, Lemma 2], by passing through the Faltings height:

Proposition 7.7. *There exists $C_2 > 0$ such that for any $q \in \mathbb{D}$ such that $J(q) \in \overline{\mathbb{Q}}$, and any $n \in \mathbb{Z}$*

$$h(J(q^n)) \leq 2h(J(q)) + 6\log(1+n) + C_2.$$

Proof. See [Ber97, Lemma 2]. □

7.2.3 Modular polynomials

Another main object in this argument is the modular polynomials (that we restrict to prime levels) $\Phi_p(X, Y) \in \mathbb{Z}[X, Y]$. They are characterized for being the (irreducible and monic) polynomial in two variables such that $\phi(j_1, j_2) = 0$ if and only if j_1 and j_2 are the j-invariants of elliptic curves which are p-isogenous. We have the following results on its degree and height from Chapter 5

Lemma 7.8. *Consider the modular polynomial $\Phi_N(X, Y) \in \mathbb{Z}[X, Y]$ for a* general *level N.*

- *Its degree* $\deg(\Phi_N) := \deg_X(\Phi_N) = \deg_Y(\Phi_N) = \psi(N)$ *where ψ is the Dedekind psi function given by*
$$\psi(N) = N \prod_{p|N \text{ prime}} \left(1 + \frac{1}{p}\right).$$
- (Cohen's bound) *It follows that*
$$h(\phi_N) = 6\psi(N)(\log N - 2\kappa_N + O(1)),$$
where $\kappa_N = \sum_{p|N \text{ prime}} \frac{\log p}{p}$.

- (Explicit bound) *If one replaces* κ_N *with* $\lambda_N = \sum_{p^n || N} \frac{p^n - 1}{p^{n-1}(p^2-1)} \log p$, *then one can take the explicit constants:*

$$6\psi(N)\left(\log N - 2\lambda_N - 0.0351\right) \leq h(\Phi_N) \leq 6\psi(N)\left(\log N - 2\lambda_N + 9.5387\right).$$

We state the best results known so far, but for the sake of the proof, Mahler's bound $h(\phi_N) = O(N^{3/2})$ would have been enough. We specialize for prime level, for future use. The explicit bound in this case is from [BS10].

Lemma 7.9. *Consider* $\Phi_p(X, Y) \in \mathbb{Z}[X, Y]$ *for* p *a prime number.*

- *Its degree* $\deg(\phi_p) = p + 1$.
- *We have that*

$$h(\phi_p) = 6p \log p + O(p).$$

- *(Explicit bound, Bröker-Sutherland) We have that*

$$h(\phi_p) \leq 6p \log p + 16p + 14\sqrt{p} \log p.$$

Therefore, we have the bound on $L(\phi_p) \leq (\psi(p) + 1)^2 e^{h(\phi_p)}$, so

$$\log L(\phi_p) \leq 2 \log(p + 2) + 6p \log p + 16p + 14\sqrt{p} \log p \leq C' p \log p$$

for some absolute constant C'.

Finally, we state the following result about the degree of the specialization of the modular polynomials, from [Ber97, lemma 2 ii)]

Proposition 7.10. *Let* $q \in \mathbb{D}$ *such that* $J(q) \in \overline{\mathbb{Q}}$, *and let* E *the elliptic curve with* $J(E) = J(q)$. *Then*

- *If* E *is CM, consider* $\tau \in \mathbb{H}$ *a period for* E *and equation* $a\tau^2 + b\tau + c = 0$. *Then if* p *does not divide* a*:*

$$\frac{p - 1}{3} \leq [\mathbb{Q}(J(q), J(q^p)) : \mathbb{Q}(J(q))].$$

- *If* E *is* not *CM, consider* p_E *the prime number given by Serre's theorem in [Ser72, Théorème 2]. Then if* $p > p_E$,

$$p + 1 = [\mathbb{Q}(J(q), J(q^p)) : \mathbb{Q}(J(q))].$$

Equality to $p + 1$ *also holds in the first case for* p *being inert in the CM field of* E.

7.2.4 Bounds on sums of primes

This last subsection is for stating some results of classical analytic number theory, as consequences of the prime number theorem.

Proposition 7.11. *Let $\pi(x) := \#\{p \text{ prime}, p \leq x\}$ the prime counting function.*

- *For $x \geq 2$,*
$$\pi(x) = \frac{x}{\log x} + \frac{x}{(\log x)^2} + O\left(\frac{x}{(\log x)^3}\right).$$
But we may use the following weaker version:[2] *there exists an absolute constant $C_{12} > 0$ such that*
$$\frac{x}{\log x} \leq \pi(x) \leq C_{12}\frac{x}{\log x}.$$
- *For x big enough,*
$$\frac{x^2}{2\log x} \leq \sum_{p \text{ prime}, p<x} p \leq \frac{x^2}{\log x}.$$

Let $\delta \in \mathbb{Z}_{>0}$ and $a \in \mathbb{Z}_{>0}$ such that $\gcd(a, \delta) = 1$, and consider
$$\pi(x, a, \delta) := \#\{p \text{ prime}, p \equiv a \pmod{\delta}\}.$$

- *For x large enough only in terms of δ,*
$$\pi(x) = \frac{x}{\phi(\delta)\log x} + \frac{x}{\phi(\delta)(\log x)^2} + O_\delta\left(\frac{x}{(\log x)^3}\right),$$
for ϕ Euler's totient function.
- *For x large enough only in terms of δ.*
$$\frac{x^2}{\phi(\delta)2\log x} \leq \sum_{p \text{ prime}, p\equiv a(\delta), p<x} p \leq \frac{x^2}{\phi(\delta)\log x}.$$

Proof. The first is the the quantitative version of the prime number theorem in [MV07, Chapter 6, equation under (6.15)]. For the second part,[3] set $\Sigma(x) = \sum_{p \text{ prime}, p<x} p$. We require the prime number theorem until the second order term to check the sign of the second order term in $\Sigma(x)$, so that our bounds hive nicer constants. It follows from Abel summation formula and the prime number theorem above
$$\Sigma(x) = \pi(x)x - \int_{t=2}^{x} \pi(t)dt = \frac{x^2}{\log x} + \frac{x^2}{\log x^2} + O\left(\frac{x}{(\log x)^3}\right) - \int_{t=2}^{x} \pi(t)dt.$$
For any $m \geq 1$, by integration by parts,
$$\int_2^x \frac{t}{(\log t)^m}dt = \int_2^x \frac{1}{(\log t)^m}tdt = \frac{t^2}{2(\log t)^m}\Big|_2^x + \int_2^x \frac{t}{m(\log t)^{m+1}}dt = \frac{x^2}{2(\log x)^m} + O\left(\frac{x^2}{m(\log x)^{m+1}}\right),$$

[2] sometimes called Tchebychev's theorem.
[3] we saw this here.

hence

$$\int_{t=2}^{x}\left(\frac{t}{\log t}+\frac{t}{(\log t)^2}\right)dt=\frac{x^2}{2\log x}+\frac{3}{2}\int_{t=2}^{x}\frac{t}{(\log t)^2}dt+O(1)=\frac{x^2}{2\log x}+\frac{3}{4}\frac{x^2}{(\log x)^2}+O\left(\frac{x^2}{(\log x)^3}\right).$$

so

$$\int_{t=2}^{x}\pi(t)dt=\frac{x^2}{2\log x}+\frac{3}{4}\frac{x^2}{(\log x)^2}+O\left(\frac{x^2}{\log x^3}\right),$$

and it follows

$$\Sigma(x)=\frac{x^2}{(\log x)}+\frac{x^2}{4(\log x)^2}+O\left(\frac{x^2}{(\log x)^3}\right).$$

And one can check computationally with SageMath [Sage] that already for $x\geq 11$,

$$\frac{x^2}{2\log x}\leq\sum_{p\text{ prime},\ p<x}p\leq\frac{x^2}{\log x}.$$

For primes in arithmetic progression, the relevant statement is [MV07, Corollary 11.21]. Also, for more explicit versions see [BMOR18, Theorem 1.1] □

7.3 Construction of the auxiliary function

Our auxiliary function is a polynomial $A\in\mathbb{Z}[X,Y]$ with $\deg_X A<N$ and $\deg_Y A<N$ such that the auxiliary function $F(z)=\Delta(z)^{2N}A(z,J(z))$ is holomorphic and with high vanishing order at $z=0$. One then writes $A(x,y)=\sum_{0\leq i,l<N}a_{i,l}x^iy^l$, considers the Fourier expansion of $\Delta(z)^{2N}A(z,J(z))$ and sets the first L coefficients to be zero. More precisely, if we set $\Delta^{2N}J^l=\sum_{k\geq 1}c_{N,l}(k)z^k$ as in Lemma 7.3, then we write

$$\begin{aligned}\Delta(z)^{2N}A(z,J(z))&=\sum_{0\leq i,l<N}a_{i,l}z^i\left(\Delta(z)^{2N}J(z)^l\right)=\sum_{0\leq i,l<N}a_{i,l}z^i\sum_{k\geq 1}c_{N,l}(k)z^k=\\&=\sum_{0\leq i,l<N}\sum_{k\geq 1}a_{i,l}c_{N,l}(k)z^{i+k}=\sum_{\nu\geq 0}\left(\sum_{\substack{0\leq i\leq\min\nu,N-1\\0\leq L<N}}a_{i,l}c_{N,l}(\nu-i)\right)z^\nu\end{aligned}\tag{7.1}$$

So we set the system in N^2 variables and L equations:

$$\sum_{0=i}^{\min\nu,N-1}\sum_{l=0}^{N-1}c_{N,l}(\nu-i)a_{i,l},\ \text{for } 0\leq\nu<L,$$

and we apply Siegel's lemma:

Proposition 7.12 (Siegel's lemma)**.** *Consider the linear system with coefficients in $\mathbb{Z}$ in X variables and Y equations with $X>Y$:*

$$\sum_{j=1}^{X}m_{i,j}x_{i,j}=0,\quad 1\leq j\leq Y,$$

Set $B=\max|m_{i,j}|$, then there exists a non zero solution in $\mathbb{Z}^X$ such that

$$|x_{i,j}|\leq(XB)^{\frac{Y}{X-Y}},$$

A standard extra assumption for Siegel's lemma is $X \geq 2Y$, so that $Y/(X-Y) \leq 1$. The upper bound on the system comes from Lemma 7.3, so $|c_{N,l}(\nu - i)| \leq C_1^N(\nu - i)^{12N}$, so we take $B = C_1^N L^{12N}$, and hence

$$\begin{aligned} |a_{i,l}| &\leq N^2 C_1^N L^{12N} \\ L(A) := \sum_{0 \leq i,l < N} |a_{i,l}| &\leq N^4 C_1^N L^{12N} \leq C_3^N L^{12N}. \end{aligned} \tag{7.2}$$

Finally, F is not the constantly zero function by Proposition 7.2.

7.4 Upper bound on $|F(z)|$

Set $M = ord_{z=0}F(z)$, for F our auxiliary function. Then by construction $M \geq L$. The following is an upper bound on F on closed disks for $0 < r < 1$, $\mathbb{D}_r \subset \mathbb{D}_1$ of the form $|F(z)| = O_{M,N}(z^M)$. As $z^{-M}F(z)$ is holomorphic on $\mathbb{D}_1$, by definition of M, one would apply the maximum principle to $G(z) := z^{-M}F(z)$ on closed disks[4] $\mathbb{D}_r$, and furthermore one would need the dependence of $\max_{D_r} G$ in terms of N, M. So instead of an application of the maximum principle, one uses that the expansion of $G(z) = \sum d_k z^k$ are of polynomial growth, meaning the terms $d_k = O(k^{rN})$ for $r \in \mathbb{Z}_{>0}$, so they are comparable to derivatives and powers of $\frac{1}{1-z}$.

With more details now, by (7.1),

$$G(z) := z^{-M}F(z) = \sum_{k \geq 0} d_k z^k$$

with

$$d_k = \sum_{0 \leq i,l < N} a_{i,l} c_{N,l}(M + k - i),$$

where the sum has N^2 terms,[5] $|c_{N,l}(M + k - i)| \leq C_1^N(M + k - i)^{12N}$ by Lemma 7.3 and $|a_{i,l}| \leq N^2 C_1^N L^{12N}$ by 7.2, therefore exists $C_4 > 0$ with

$$|d_k| \leq N^4 C_1^N L^{12N} C_1^N (M + k)^{12N} \leq C_4^N L^{12N} (M + k)^{12N}.$$

The following is a pretty general argument for power expansion with coefficient of this type of growth: bound a function $\sum_{t=1}^{\infty}(M + t)^{12N}$ via comparing it with the expansion of $1/(1 - x)$, powers of it, or derivatives.

Lemma 7.13 (Explicit Schwarz lemma for polynomial Taylor series)**.** *Assume $h(z) = \sum_{t \geq 0} h_t z^n \in \mathbb{Z}[[z]]$ is a holomorphic function on $\mathbb{D}$ such that $|h_t| \leq (M + t)^K$ for fixed $K \in \mathbb{Z}_{>0}$. Then*

$$|h(z)| \leq (M + 1)^K K! \frac{1}{(1 - |z|)^{K+1}}.$$

[4] The boundary of $\mathbb{D}_1$ is the natural boundary of modular forms, so this arguments need to be carried out on $\mathbb{D}_r$.

[5] Remark that the restriction $i < k$ from (7.1) does not apply anymore.

Proof. Note that the Taylor expansion for any $k \geq 1$ and any $0 < x < 1$,

$$\frac{k!}{(1-x)^{k+1}} = \sum_{n=0}^{\infty} (n+1)\cdots(n+k)x^k,$$

hence

$$1 + \sum_{n\geq 1} n^k x^k \leq \sum_{n\geq 0} (n+1)\cdots(n+k)x^k = \frac{k!}{(1-x)^{k+1}},$$

so we have the bounds (remark $(M+t)^K \leq (M+1)^K t^K$):

$$\sum_{t\geq 0} (M+t)^K |z|^t \leq (M+1)^K \left(1 + \sum_{t\geq 1} t^K |z|^t\right) \leq (M+1)^K \frac{K!}{(1-|z|)^{K+1}}.$$

□

Then by Lemma 7.13 above, it follows that

$$|F(z)| \leq |z|^M C_5^N L^{12N} M^{12N} N^{12N} \frac{1}{(1-|z|)^{12N+1}},$$

therefore, we set the only dependence of N on q, setting N big enough so that for[6] $r = \frac{1+|q|}{2}$,

$$\left(\frac{1}{(1-r)}\right)^{12N+1} \leq \left(\frac{N^2}{2}\right)^N, \tag{7.3}$$

that in turn it is upper bounded by M^N. in addition $L \leq M$ and $2L \leq N^2 < 2(L+1) \leq 2(M+1) \leq 3M$

$$(LMN)^{12N} \leq \left(M^2\sqrt{3N}\right)^{12N} = 3^{6N} M^{\frac{5}{2}12N} \leq M^{31N}$$

can gather all powers in terms of M, so for all z with $0 < |z| < r = \frac{1+|q|}{2}$,

$$|F(z)| \leq |z|^M M^{31N}. \tag{7.4}$$

7.5 Lower bound on prime powers $|F(q^P)|$

Consider a prime number P such that $F(q^P) \neq 0$. Remark that such a P exists, for that arguing over closed discs on $\mathbb{D}_{|q|} \subset \mathbb{D}$, F is a holomorphic function non constantly zero, so it can only have finitely many zeros. Set now $\alpha = A(q^P, J(q^P))$, then $F(q^P) = \Delta(q^P)^{2N}\alpha$. First observe that as $z^{-1}\Delta(z)$ is holomorphic and nonvanishing on $\mathbb{D}$, by the maximum principle we have an absolute constant $C_6 > 0$ such that $|\Delta(q^P)| \geq C_6 |q|^P$, hence

$$|\Delta(q^P)|^{2N} \geq C_6^{2N} |q|^{P2N} \geq \exp(-(C_7 \log(|q|^{-1})NP).$$

[6] This is simply so one argues on a disk containing q in its interior and this is a simple enough choice of radius.

Let us focus now on $\alpha = A(q^P, J(q^P))$. As $q, J(q) \in \overline{\mathbb{Q}}$, $q^P \in \overline{\mathbb{Q}}$ and $J(q^P) \in \overline{\mathbb{Q}}$ as $\Phi_P(J(q), J(q^P)) = 0$, therefore $\alpha \in \overline{\mathbb{Q}}$ and we will use Liouville's inequality Lemma 7.4. We then need upper bounds on $\deg(\alpha)$ and $h(\alpha)$. As $\alpha \in \mathbb{Q}(q, J(q^P))$, we have and $J(q^P)$ belongs to an algebraic extension of $\mathbb{Q}(J(q))$ of degree bounded by $S+1$, we have

$$\begin{aligned}\deg(\alpha) &\leq [\mathbb{Q}(q, J(q^P)) : \mathbb{Q}] \leq [\mathbb{Q}(q) : \mathbb{Q}][\mathbb{Q}(J(q^P)) : \mathbb{Q}] = \deg(q)\deg(J(q^P)) \\ &\leq (P+1)\deg(q)\deg(J(q)).\end{aligned}$$

For the height, by Lemma 7.6 and properties of the height,

$$h(A(q^P, J(q^P)) \leq \log L(A) + Nh(q^P) + Nh(J(q^P)) = \log L(A) + NPh(q) + Nh(J(q^P)).$$

For $h(J(q^P))$, the original proof of [BDGP96] compared the $h(J(q^P))$ with $h(J(q))$ via Lemma 7.5, with the modular polynomial Φ_P and $J(q)$,

$$\begin{aligned}\deg(J(q^P))h(J(q^P)) &\leq \deg(J(q))\left(\log L(\Phi_P) + \deg_x(\Phi_P)h(J(q))\right) \\ &\leq \deg(J(q))\left(C'(P+1)\log P + (P+1)h(J(q))\right).\end{aligned}$$

Alternatively, for a more geometric version, we use Proposition 7.7:

$$h(J(q^P)) \leq 2h(J(q)) + 6\log(1+P) + C_2.$$

Therefore,

$$\begin{aligned}\deg(\alpha)h(\alpha) &\leq \deg(\alpha)\left(\log L(A) + NPh(q) + Nh(J(q^P))\right), \\ &\leq (P+1)\deg(J(q))\deg(q)\left(\log L(A) + NPh(q) + N\left(2h(J(q)) + 6\log(1+P) + C_2\right)\right), \\ &\leq (P+1)\deg(J(q))\deg(q)\left(\log L(A) + NPh(q) + NC_7\log(P) + 2Nh(J(q))\right)\end{aligned}$$

and as $\log L(A) \leq \log(C_3^N(N^2/2)^{12N}) \leq 25N\log N$, if we set
$C_8(q) = \deg(q)\deg(J(q))\max\{1, 25, C_7, h(q), h(J(q))\}$, then

$$\deg(\alpha)h(\alpha) \leq C_8(q)N(P+1)\left(P + \log N + \log P + 1\right) \leq C_9(q)NP\left(P + \log N\right).$$

Then Liouville's inequality $\log|\alpha| \geq -\deg(\alpha)h(\alpha)$ and our previous bound on $\Delta(q^P)^{2N}$ allow us to conclude that exists $C_9(q)$ such that

$$|F(q^P)| \geq \exp\left(-C_{10}(q)NP\left(P + \log N\right)\right). \tag{7.5}$$

7.6 Definition of P and upper bound

Combining (7.4) and (7.5):

$$\begin{aligned}\exp\left(-C_{10}(q)NP\left(P+\log N\right)\right) \leq |q|^{PM}M^{31N} &= \exp\left(-\log(|q|^{-1})PM + 31N\log M\right), \\ \log(|q|^{-1})PM &\leq C_9(q)NP\left(P+\log N\right) + 31N\log M, \\ M &\leq_{\exists C_{10}(q)} C_{10}(q)N\left(P + \log P + \log N + \frac{31\log M}{P}\right),\end{aligned}$$

and as $N^2 \leq 2(L+1) \leq 2(M+1) \leq 3M$, it follows that $\log N \leq \log M$ and $N \leq \sqrt{3M}$ and there exists $C_{11}(q)$ such that

$$M \leq C_{11}(q)\sqrt{M}\left(P + \log M\right). \tag{7.6}$$

Observe that at this point we are aiming for a contradiction of the sort $M \leq \sqrt{M}\log M$, and we would have that as long as we may control P well enough in terms of N, M. This is also the step where the strategy differs from the original Mahler method.

The only condition that we impose on P is that $F(q^P) \neq 0$. We said before that such an P exists, so we may consider the smallest of it, i.e. P such that $F(q^p) = 0$ for all primes $p < P$.

Set $r = \frac{1+|q|}{2}$. We can think of considering the following holomorphic function on $\mathbb{D}_r$

$$\frac{F(z)}{z^M} \prod_{p \text{ prime},\, p<P} \frac{1}{z - q^p},$$

we simply going to change the factors $\frac{1}{z-q^p}$ to the following (inverse) Blaschke-type product

$$H(z) := \frac{F(z)}{z^M} \prod_{p \text{ prime},\, p<P} \frac{r^2 - z\bar{q}^p}{r(z - q^p)}.$$

We consider this modified factors because at the boundary $\{|z| = r\}$, it holds (this argument comes from [Nes96, Under Equation (6)])

$$|r^2 - \bar{q}^p z| = |r^2 - \bar{q}^p z|\frac{|\bar{z}|}{r} = \frac{1}{r}|r^2\bar{z} - \bar{q}^p r^2| = r|z - \bar{q}^p| = |r(z - q^p)|,$$

so $\max_{|z|=r} |H(z)| = r^{-M}|F(z)|$. By the maximum principle, for $z \in \mathbb{D}_r$:

$$|H(z)| \leq r^{-M}|F(z)| \leq M^{31N},$$

by (7.4). On the other hand, $H(0) = d_0 \prod_{p \text{ prime},\, p<P} \frac{-r}{q^p}$, with d_0 the first coefficients of the expansion of $G(z) = z^{-M}F(z)$. By construction, $d_0 \in \mathbb{Z}$ and it is non-zero, so $|d_0| \geq 1$. Together with $|H(0)| \leq M^{31N}$, we have our bound involving P, M, N. We require to bound $\pi(P) := \sum_{p \text{ prime},\, p<P} 1$ and $\sum_{p \text{ prime},\, p<P} p$, and we use Proposition 7.10. Therefore, we have:

$$\begin{aligned} r^{\pi(P)}|q|^{-\sum_{p \text{ prime},\, p<P} p} &\leq M^{31N}, \\ \left(\frac{1}{r}\right)^{-C_{12}P/\log P} \left(\frac{1}{|q|}\right)^{P^2/2\log P} &\leq M^{31N} \\ \frac{P^2}{2\log P}\log(|q|^{-1}) - C_{12}\log(r^{-1})\frac{P}{\log P} &\leq 31N\log M, \end{aligned}$$

and as $r \geq |q|$ it follow that $-\log r^{-1} \geq -\log|q|^{-1}$, so there exists $C_{13} > 0$ such that $(P/2 - C_{12}) \geq (C_{13})^{-1}P$ and

$$\frac{P^2}{2\log P}\log(|q|^{-1}) - C_{12}\log(r^{-1})\frac{P}{\log P} \geq \frac{P}{\log P}\log(|q|^{-1})\left(\frac{P}{2} - C_{12}\right) \geq \frac{P^2}{\log P}\log(|q|^{-1})\frac{1}{C_{13}},$$

hence

$$\frac{P^2}{\log P} \leq C_{13}\frac{1}{\log(|q|^{-1})}31N\log M. \tag{7.7}$$

7.7 Final Contradiction

In (7.7), we can take $\frac{P^2}{\log P} \geq P\sqrt{P}$, so the following arguments are easier. Using $N \leq \sqrt{3M}$, it follows

$$P^{3/2} \leq \frac{C_{13}31}{\log(|q|^{-1})} N \log M \leq C_{14}(q)\sqrt{M}\log M.$$

Therefore, together with (7.6),

$$M \leq C_{16}(q)\sqrt{M}\left(\log M + \left(\sqrt{M}\log M\right)^{2/3}\right) \leq C_{17}(q)\sqrt{M}\left(\sqrt{M}\log M\right)^{2/3},$$
$$M \leq C_{17}(q)M^{5/6}(\log M)^{2/3},$$

which is a contradiction for M big enough. As $M \geq [N^2/2]$, and N can be taken arbitrarily large,[7] we have our final contradiction.

7.8 Other bounds on P that do not yield a contradiction

This section is no part of the proof, but it is included with sights to possible generalizations for genus 2.

The following bounds are weaker, of $P = O(N)$ and $P = O(N \log M)$, and one can see that in the previous section that they do not yield a contradiction. Actually, any bound of the form $P^{1+\varepsilon} = O(N \log M)$ for any $\varepsilon > 0$ gives the contradiction, but not when only $\varepsilon = 0$.

We spend some time in these arguments for that they are the only ones that can be generalized to genus two, even though they are also not strong enough for a contradiction.

7.8.1 Jensen's inequality

We have a holomorphic function $G(z) = z^{-M}F(z)$ such that on D_r it follows $|G(z)| \leq M^{31N}$, and our question boils down to control the number of zeros using the size of the function, and Jensen's inequality tells you precisely that: [MV07, Lemma 6.1], using $r' = |q| < r = \frac{1+|q|}{2}$, the number of zeros of G in $\{|z| < |q|\}$ does not exceed

$$\frac{1}{\log(r/r')} \underbrace{\frac{1}{|G(0)|}}_{\geq 1} \log(M^{31N}) \leq O_q(N \log M),$$

as $G(0) = d_0 \in \mathbb{Z}$. In particular, $P = O_q(N \log M)$. From this perspective, one sees the strategy in the previous section is an strengthening of the bound given by Jensen's inequality, by using that the zeros one is counting are in a geometric sequence $\{q^p\}$ for p prime.

7.8.2 A purely algebraic argument

This argument shows well-definedness of P without using complex analysis. Suppose $(A(q^p, J(q^p)) = 0$ for p prime (with $(q, J(q)) \in \overline{\mathbb{Q}}^2$ so $(q^p, J(q^p)) \in \overline{\mathbb{Q}}^2$), as has always been in this chapter). Then $J(q^p)$ is a root of a polynomial in $\mathbb{Q}(q^p)[X] \subset \mathbb{Q}(q)[X]$ of degree N, so

$$\deg(J(q^p)) = [\mathbb{Q}(J(q^p)) : \mathbb{Q}] \leq [\mathbb{Q}(J(q^p) : \mathbb{Q}(q)][\mathbb{Q}(q) : \mathbb{Q}] \leq \deg(q)N,$$

[7] recall that the only dependence of N on q is in Equation (7.3).

but this bound is independent on p, which is absurd: generically $\mathbb{Q}[J(q^p) : J(q)] = p + 1$, and actually by [Ber97, Lemma 2 ii)], for p large enough in terms of $J(q)$,

$$\frac{p-1}{3} \leq [\mathbb{Q}(J(q), J(q^p)) : \mathbb{Q}(J(q)].$$

This lower bound is only for CM elliptic curves, for the non-CM for p large enough,[8] the degree is generically $p + 1$.

Hence, as $[\mathbb{Q}(J(q^p)) : \mathbb{Q}] \geq [\mathbb{Q}(J(q)(J(q^p) : \mathbb{Q}(J(q)] = \deg(J(q))[\mathbb{Q}(J(q), J(q^p)) : \mathbb{Q}(J(q)]$, it follows that

$$\frac{p-1}{3} \leq \frac{\deg(q)}{\deg(J(q))} N,$$

so taking P as the smallest p such that this bound does not apply, we have $A(q^p, J(q^p)) \neq 0$ and a bound $P = O_q(N)$.

[8] larger than the prime in Serre's uniformity theorem.

Chapter 8

On the steps to the Stéphanois theorem for genus two

Our goal is to generalized the strategy from the modular proof of the Stéphanois theorem to the Igusa invariants of curves of genus two. Following the guideline from [BDGP96] (and [Wal96]), that we have presented in Chapter 7, we have the following sequence of steps:

- construction of the auxiliary function(s),
- upper bounds,
- lower bounds,
- definition of the prime power and upper bound

What we present here is a generalization of all the steps, except the last one, and only for values $\boldsymbol{q} = (q_1, q_2, q_3)$ belonging to a subdomain of of the domain of definition of the Fourier expansion of the Igusa invariants. We can then state the subsequent result.

Theorem 8.1. *Let $\boldsymbol{\varrho} = (\varrho_1, \varrho_2, \varrho_3)$ belonging to the domain of Equation (8.2), and furthermore assume it belongs to the subset*[1] *satisfying the property*

the prime number $P = P(\boldsymbol{\varrho}, N)$ defined in Section 8.5 satisfies $P = O_{\boldsymbol{\varrho}}(N^{1/3-\varepsilon})$ for some $\varepsilon > 0$.

Then we have

$$\operatorname{trdeg}_{\mathbb{Q}} \mathbb{Q}(\varrho_1, \varrho_2, \varrho_3, j_1(\boldsymbol{\varrho}), j_2(\boldsymbol{\varrho}), j_3(\boldsymbol{\varrho})) \geq 1.$$

8.1 Preliminary results

We gather here the preliminaries results needed for our proof, together with the results from Section 7.2.

Proposition 8.2. *Let $K = \mathbb{C}(\mathcal{A}_2)$, the field of functions of $\mathcal{A}_2$, or equivalently the field generate by Siegel modular functions with respect to $\mathrm{Sp}_4(\mathbb{Z})$. As functions on $\boldsymbol{\tau}$, the exponentials $q_k = e^{i2\pi\tau_k}$ are algebraically independent over $K(\boldsymbol{\tau})$.*

Proof. This follows from [BZ01, Theorem 1], where it holds for every genus. □

[1] conjecturally, this subset should be the whole domain.

8.1.1 Fourier expansion of Siegel modular forms

Consider the Siegel upper half space of degree two

$$\mathbb{H}_2 = \{\boldsymbol{\tau} = \begin{pmatrix} \tau_1 & \tau_2 \\ \tau_2 & \tau_3 \end{pmatrix}, \operatorname{Im}\boldsymbol{\tau} > 0\},$$

and a holomorphic Siegel modular forms $f : \mathbb{H}_2 \to \mathbb{C}$ with respect to $\mathrm{Sp}_4(\mathbb{Z})$. They admit a Fourier expansion

$$\sum_{M \in \mathrm{Sym}_2(\mathbb{Z})^\vee,+} a(M) e^{i2\pi \operatorname{Tr}(M\boldsymbol{\tau})}, \tag{8.1}$$

where $\mathrm{Sym}_2(\mathbb{Z})^\vee = \left\{ \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}, \ a,b,c \in \mathbb{Z} \right\}$. We call $M \in \mathrm{Sym}_2(\mathbb{Z})^\vee$ *half integral symmetric*, and the superscript $+$ means positive semidefinite matrix. We will also use the notations $M \geq 0$ and $M > 0$ for positive semidefinite and positive definite matrix, respectively. That means that M is of the form $\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$, with $a, c \geq 0$ and $b^2 \leq 4ac$. In particular $\operatorname{Im}(\operatorname{Tr}(N\boldsymbol{\tau})) = \operatorname{Tr}(N \operatorname{Im}\boldsymbol{\tau}) \geq 0$, because both N and $\operatorname{Im}\boldsymbol{\tau}$ are positive semi-definite,[2] it follows that $|\exp(i2\pi \operatorname{Tr}(M\boldsymbol{\tau}))| = \exp(i2\pi \operatorname{Im}(\operatorname{Tr}(N\boldsymbol{\tau}))) \leq 1$.

Notice that $\operatorname{Tr} N\boldsymbol{\tau} = a\tau_1 + b\tau_2 + c\tau_3$, so if we set

$$q_l := e^{i2\pi\tau_l}, \ l = 1,2,3,$$

we can see Equation (8.1) as a power series in the variables $\boldsymbol{q} := (q_1, q_2, q_3)$. Note that

- the exponents of q_1, q_3 are always in $\mathbb{Z}_{\geq 0}$,
- and the exponents of q_2 can be negative, but from $b^2 \leq 4ac$, they are bounded in terms of the exponents of q_1 and q_3. Hence, as a power series on q_1 and q_3, the coefficients are Laurent polynomials in q_2.

We can see the Fourier expansion as naturally defined in $(R[q_2, q_2^{-1}])[[q_1, q_3]]$, for R an appropriate ring of the Fourier coefficients, and $R[q_2, q_2^{-1}]$ denoting the ring of Laurent polynomials.

From a different perspective, we may want to consider the complex domain that the assignment $\boldsymbol{q} : \mathbb{H}_2 \to \mathbb{C}^3$ defines, and see a Fourier expansions as an holomorphic function on a complex submanifold of $\mathbb{C}^3$, with possibly singular boundary. If $\boldsymbol{\tau} \in \mathbb{H}_2$ then $\tau_1, \tau_3 \in \mathbb{H}$ and $\operatorname{Im}(\tau_2)^2 < \operatorname{Im}(\tau_1)\operatorname{Im}(\tau_3)$.

Lemma 8.3. *The image of the assignment $\boldsymbol{q} : \mathbb{H}^2 \to \mathbb{C}^3$ is given by*

$$\boldsymbol{q}(\mathbb{H}^2) := \{(q_1, q_3) \in (\mathbb{D}_1)^2 : \ |q_2| < e^{\sqrt{\log|q_1^{-1}| \log|q_3^{-1}|}}, \ \text{or } |q_2| > e^{-\sqrt{\log|q_1^{-1}| \log|q_3^{-1}|}}\},$$

i.e. on fixed (q_1, q_3), it is the complement on two discs of the complex plane centered around 0 and ∞, respectively.[3] *Furthermore, $\mathbb{D}_1^3$ acts on it via $(q_1, q_2, q_3) \mapsto (\xi_1 q_1, \xi_2 q_2, \xi_3 q_3)$ for $|\xi_j| = 1$, $j = 1,2,3$, i.e. $\boldsymbol{q}(\mathbb{H}^2)$ is a* Reinhardt *domain.*

[2] see this link.

[3] Observe that this complement always contains the unit circle.

The Fourier expansion of a holomorphic Siegel modular form is uniformly convergent on compact subsets of $\mathbb{H}_2$. Considering the induced series in q_1, q_2, q_3, for any compact subset away from $q_1q_2q_3 = 0$, we can invert the q-variables (they are locally biholomorphic mappings), and prove uniform convergence of the series. Therefore, it induces holomorphic functions on this complex submanifold of $\mathbb{C}^3$.[4] The boundary at infinity of $\mathbb{H}_2$ is singular, so we expect our manifold to be singular too at $q_1q_2q_3 = 0$.

Remark 8.4. *We will see below in Item 1 that the symmetries of the Fourier expansion of a Siegel modular form include* $(q_1, q_2, q_3) \to (q_1, q_2^{-1}, q_3)$ *and* $(q_1, q_2, q_3) \to (q_3, q_2, q_1)$*, hence one could always restrict to* $|q_2| \leq 1$ *and* $|q_1| \geq |q_3|$*.*

$$\{q_1, q_3 \in (\mathbb{D}_1)^2 : |q_1| \geq |q_3|,\ 1 \geq |q_2| > e^{-\sqrt{\log|q_1|\log|q_3|}}\} \subset \mathbb{D}_1 \times \overline{\mathbb{D}_1} \times \mathbb{D}_1$$

8.1.2 Siegel's fundamental domain and Minkowski's domain

$\mathrm{Sp}_4(\mathbb{Z})$ contains the following subgroup isomorphic to $\mathrm{GL}_2(\mathbb{Z})$,

$$\left\{ \begin{pmatrix} U & 0 \\ 0 & {}^tU^{-1} \end{pmatrix},\ \text{for } U \in \mathrm{GL}_2(\mathbb{Z}) \right\},$$

which acts on $\mathbb{H}_2$ by $\boldsymbol{\tau} \mapsto U\boldsymbol{\tau}{}^tU$. We call them *unimodular transformations*. We can see $\mathrm{GL}_2(\mathbb{Z})$ acting on positive define real matrices (on $\mathrm{Im}\,\boldsymbol{\tau}$), and consider a fundamental domain for this action. By the theory of Minkowski reduction [Kli90, section I.2], there exists such a domain composed on Minkowski reduced matrices.

Definition 8.5. *Let* $n \in \mathbb{Z}_{>0}$ *and* $Y = (y_{ij}) \in \mathrm{Sym}_n^+(\mathbb{R})$ *a positive definite matrix. We say that* Y *is* Minkowski reduced *if*

- *it holds* $y_{k,k+1} \geq 0$ *for* $k = 1, \ldots, n-1$*,*
- *and the successive minima of the associated quadratic form to* Y *on* $\mathbb{Z}^n$ *are the diagonal elements, meaning that* ${}^t[g]Yg \geq y_{k,k}$ *for all* $g \in \mathbb{Z}^n$ *with* $\gcd(g_k, \ldots g_n) = 1$*, for* $k = 1, \ldots n$*.*

We call the subdomain of $\mathrm{Sym}_n^+(\mathbb{R})$ *composed of such matrices* Minkowski's reduced domain *of* $\mathrm{Sym}_n^+(\mathbb{R})$*, or simply* Minkowski's domain*.*

It follows that Minkowski reduced matrices have a easily described appearance.

Proposition 8.6. *Let* $Y = (y_{ij}) \in \mathrm{Sym}_n^+(\mathbb{R})$ *be Minkowski reduced. Then it satisfies the following:*

- $y_{k,k} \leq y_{k+1,k+1}$ *for* $1 \leq k \leq n-1$*,*
- *it holds* $|2y_{k,l}| \leq y_{k,k}$ *for* $l \neq k$*, and any* $1 \leq k \leq n$*.*
- *there exists a constant* $c(n) > 0$ *only depending on* n*, such that it holds*

$$\det(y) \leq \prod_{i=1}^{n} y_{i,i} \leq c(n)\det(y)$$

[4] see also [Kli90, paragraph under Equation II.4.1, page 44].

Proof. This is [Kli90, Proposition 1 in Section I.2]. □

Remark 8.7. *A way to understand this result is to notice that the analogous minimization problem of Y* over $\mathbb{R}$ *simply produces its eigenvalues (real and positive) in increasing order, and the "reduced" matrix in that sense will be its diagonalization. However, the minimizing problem we consider is* over $\mathbb{Z}$*. In particular, its successive minima are not necessarily its eigenvalues, and one cannot diagonalize in general a symmetric matrix over* $\mathrm{GL}_2(\mathbb{Z})$*. Hence the above result would say that the best one can do over* $\mathbb{Z}$ *is to make the off-diagonal elements small.*[5]

In dimension two, more can be proven:

Lemma 8.8. *In dimension two, a matrix* $Y = \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix} \in \mathrm{Sym}_2(\mathbb{R})$ *is Minkowski reduced if and only if* $0 \leq 2y_2 \leq y_1 \leq y_3$ *and* $y_1, y_3 \neq 0$.

Proof. If Y is Minkowski reduced it follows for the general case in any dimension in Proposition 8.6. For the other direction, first notice that Y is positive definite, because

$$\det Y = y_1y_3 - y_2^2 \geq 4y_2^2 - y_2^2 = 3y_2^2,$$

we need to verify that $y_1 \leq y_3$ are the successive minima for $\mathbb{Z}^2$ for the quadratic form induced by Y, $f_Y(m,n) := y_1m^2 + 2y_2mn + y_3n^2$.

For y_1, that means verifying $y_1 = \min_{(0,0)\neq(m,n)\in\mathbb{Z}^2} f_Y(m,n)$ If $mn \geq 0$,

$$y_1m^2 + \underbrace{2y_2mn}_{\geq 0} + y_3n^2 \geq y_1m^2 + y_3n^2 \geq \min\{y_1, y_3\} = y_1 = f_Y(1,0),$$

and if $mn < 0$:

$$y_1m^2 + \underbrace{2y_2mn}_{\geq y_1mn} + \underbrace{y_3n^2}_{\geq y_1n^2} \geq y_1(m^2 + mn + n^2) \geq y_1,$$

which follows from $\min_{(0,0)\neq(m,n)\in\mathbb{Z}^2} m^2 + mn + n^2 = 1$, which can be checked by elementary arguments. For y_3 being the second minima, the minimizing problem is under the extra condition that $n = 1$, so minimizing $y_1m^2 + 2y_2m + y_3 = m(y_1m + 2y_2) + y_3$. It now follows from $\min_{m\in\mathbb{Z}} m(y_1m + 2y_2) = 0$, which can be checked analogously. □

Definition 8.9. *We call the following domain the Minkowski's domain of* $\mathbb{H}^2$*:*

$$\{\boldsymbol{\tau} \in \mathbb{H}_2 : \ \mathrm{Im}\,\boldsymbol{\tau} \text{ is Minkowski reduced, } |\mathrm{Re}\,\tau_i| \leq \frac{1}{2}\}.$$

Remark 8.10. *The Minkowski's domain maps via* $\boldsymbol{q}$ *to:*

$$\{\boldsymbol{q} : \ 1 \geq |q_2|^2 \geq |q_1| \geq |q_3| > 0\} \subset \mathbb{D}_1 \times \overline{\mathbb{D}_1} \times \mathbb{D}_1, \tag{8.2}$$

which is easier to describe that $\boldsymbol{q}(\mathbb{H}^2)$.

We also consider the following domain, that also contains the Siegel's fundamental domain:

$$\mathcal{K} = \{|\,\mathrm{Re}(\tau_l)| \leq \frac{1}{2}, 0 \leq 2\,\mathrm{Im}\,\tau_2 \leq \mathrm{Im}\,\tau_1 \leq \mathrm{Im}\,\tau_3, \mathrm{Im}\,\tau_1 \geq \frac{\sqrt{3}}{2}\}.$$

[5] for $n = 2, 3$ they are diagonally dominant matrices, so one can relate its eigenvalues with the succesive minima by Gershgorin's circle method.

8.1.2.1 Symmetries of the Fourier coefficients

Any holomorphic Siegel modular form g with Fourier expansion

$$\sum_{M\in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+}} a(M)e^{i2\pi\,\mathrm{Tr}(M\boldsymbol{\tau})}$$

satisfies that, for any $U \in \mathrm{GL}_2(\mathbb{Z})$ and $M \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+}$, $a({}^tUMU) = a(M)$. In particular, if we write $M = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$, for the positive semidefinite binary quadratic form $Q_M = ax^2 + bxy + cy^2$ with discriminant $b^2 - 4ac \leq 0$, then $a(M)$ depends on the $\mathrm{GL}_2(\mathbb{Z})$-equivalence class of Q_M.

Definition 8.11. *We say that $\delta \in \mathbb{Z}$ is a* discriminant *if $\delta \equiv 0, 1 \pmod 4$. We say that it is a* fundamental discriminant *if it is the discriminant of a quadratic field.*

Remark 8.12. *Every discriminant is of the form $f^2\Delta$ for $f \in \mathbb{Z}_{>0}$ and Δ a fundamental discriminant.*

1. Taking $U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, it follows that

$$a\left(\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}\right) = a\left(\begin{pmatrix} a & -b/2 \\ -b/2 & c \end{pmatrix}\right),$$

   and taking $U = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, then

$$a\left(\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}\right) = a\left(\begin{pmatrix} c & b/2 \\ b/2 & a \end{pmatrix}\right),$$

2. If $\det(M) = 0$, then every $\mathrm{GL}_2(\mathbb{Z})$-equivalence class has a representative:

$$\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix},$$

   for $a \in \mathbb{Z}$. In particular, there are infinitely many classes.

3. If $\det(M) < 0$, then there are only finitely many classes, bounded by $h(-4\det(M))$, for the class number of the discriminant $-4\det(M)$. Two quadratic forms f, g are $\mathrm{GL}_2(Z)$-equivalent if either f and g or f and g^{-1} are in the same $\mathrm{SL}_2(\mathbb{Z})$-equivalence class, where g^{-1} means the inverse under Dirichlet composition (see [Cox22, Theorem 3.9]). In particular, the exact number of equivalence classes lies between $h(-4\det(M))/2$ and $h(-4\det(M))$.

For any given weight k, we have a subspace, *the Maass Spezialschar*, of Siegel forms that are lifts of Jacobi forms of weight k and index 1. For modular forms in this subspace, we have extra symmetries: a given Fourier coefficient $a(M)$ only depends on $\det(M)$ and $\gcd(a, b, c)$. More precisely, if $-4\det(M)$ is a fundamental discriminant, then necessarily $\gcd(a, b, c) = 1$, so there is only *one* Fourier coefficient. if $-4\det(M) = f^2D'$ with D' fundamental discriminant, then we have one Fourier coefficient for every divisor of f.

8.1.2.2 Ordering by determinant and ordering by trace

Given a *cusp* Siegel modular form, $a(M) = 0$ for all M with $\det(M) = 0$. Write the Fourier expansion,

$$\sum_{M \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+}, M>0} a(M)\boldsymbol{q}^M,$$

On first instance, the symmetries of the Fourier coefficients would make us thing on ordering *by determinant*. However, on a fixed determinant there are infinitely many h.i.s matrices. More precisely, if we write $M = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$, we consider the corresponding positive definite binary quadratic form $ax^2+bxy+cy^2$ of discriminant $-4\det(M) = b^2-4ac$. Fixing $\delta = 4\det(M) > 0$ (hence $-\delta$ is a negative discriminant), then $b^2 + \delta = 4ac$. It follows that $b \equiv \delta \pmod 2$ and that (a, c) are pairs of divisors of $(b^2 + \delta)/4$. One can write

$$\sum_{0<M \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+}} a(M)\boldsymbol{q}^M = \sum_{\substack{\delta>0 \\ -\delta \text{ discriminant}}} \sum_{b \equiv \delta\ (2)} \sum_{a \mid \frac{b^2+\delta}{4}} a(M)\boldsymbol{q}^M$$

where we are already accounting for the symmetry $b \to -b$ and $(a, c) \to (c, a)$. On fixed δ and b, we have exactly $\sigma_0((b^2 + \delta)/4)$ matrices.

This is *not* the ordering that we are going to use, as for the construction of the auxiliary functions in the transcendence proof we do not have a way to preserve all the symmetries of the Fourier coefficients (except for $b \to -b$ and $(a, c) \to (c, a)$). We are going to order the terms of the Fourier expansion *by trace*.

Lemma 8.13. *There are finitely many h.i.s. positive definite matrices with fixed trace. More precisely, there are* 3 *matrices for* $t = 2$ *and* 10 *for* $t = 3$*, and for an integer* $t \geq 4$*:*

$$t^2 + (t-1) \leq \{M \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+},\ M > 0,\ \mathrm{tr}(M) = t\} \leq \frac{5}{3}t^2 + (t-1),$$

and

$$\frac{t^2}{4} + \left\lfloor \frac{t+1}{2} \right\rfloor \leq \left\{ M = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+},\ M > 0\ \ \mathrm{tr}\, M = t, a \leq c,\ b \geq 0 \right\} \leq \frac{5}{12}t^2 \left\lfloor + \frac{t+1}{2} \right\rfloor.$$

Proof. Set a matrix $M = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+}$, M positive definite with $\mathrm{tr}\, M = t$. First notice that there are $t-1$ pairs $a, c \geq 1$ such that $a + c = t$, given explicitly by $(a, t-a)$ with $a = 1, \ldots t-1$. For each pair $(a, t-a)$, there are only finitely many b such that $b^2 < 4a(t-a)$. Observe that for any a, $4a(t-a) \leq 4\frac{t}{2}\frac{t}{2} = t^2$, so $|b| < t$. Hence it follows a first upper bound of $(t-1)(2(t-1)+1) = (2t-1)(t-1)$. We are going to improve it alongside the calculations of a lower bound.

For any $a = 1, \ldots, t-1$, denote $b_t(a)$ the number of possible values of $|b| > 0$ with fixed t, a. Our total number of matrices of fixed t, a is $2b(t, a) + 1$. For $b^2 < 4a(t-a)$, then:

$$b(t,a) = \begin{cases} \lfloor 2\sqrt{a(t-a)} \rfloor & \text{if } 2\sqrt{a(t-a)} \notin \mathbb{Z} \\ 2\sqrt{a(t-a)} - 1 & \text{otherwise} \end{cases},$$

from where it follows

$$2\sqrt{a(t-a)} - 1 \leq b_t(a) \leq 2\sqrt{a(t-a)}.$$

By construction, we have the symmetry $b_t(a) = b_t(t-a)$. Assume first that t is odd, so $\sum_{a=1}^{t-1} b_t(a) = 2\sum_{a=1}^{\lfloor t/2 \rfloor} b_t(a)$, and $\lfloor t/2 \rfloor = (t-1)/2$. As $\sqrt{a(t-a)}$ is an increasing function on $0 \leq a \leq t/2$, we can compare to the integrals:

$$2\int_{u=1}^{(t+1)/2} (2\sqrt{u(t-u)} - 1)du \leq \sum_{a=1}^{t-1} b_t(a) \leq 4\int_{u=1}^{(t+1)/2} \sqrt{u(t-u)}du. \tag{8.3}$$

Remark the integration until $\lfloor t/2 \rfloor + 1 = (t+1)/2$. One can check that the following function is a primitive of the right hand side, and we have the upper bounds

$$\begin{aligned}\frac{1}{2}t^2\left(\arcsin\left(\frac{t-2}{t}\right) + \arcsin\left(\frac{1}{t}\right)\right) + \sqrt{t-1}\,(t-2) + \frac{1}{2}\sqrt{t^2-1} \\ \leq \frac{t^2}{2}\frac{\pi}{2} + \sqrt{t-1}\,(t-2) + \frac{t}{2}\left(\leq \frac{5}{6}t^2\right);\end{aligned}$$

and for the lower bound

$$\begin{aligned}\frac{1}{2}t^2\left(\arcsin\left(\frac{t-2}{t}\right) + \arcsin\left(\frac{1}{t}\right)\right) + \sqrt{t-1}\,(t-2) + \frac{1}{2}\sqrt{t^2-1} - (t-1) \\ \geq \frac{t^2}{2}2(\arcsin(1/3) + \sqrt{t-1}\,(t-2) - \frac{1}{2}(t-1)\left(\geq_{\text{ if } t\geq 4} \frac{t^2}{2}\right),\end{aligned}$$

where for the bounds we have used that $\arcsin\left(\frac{t-2}{t}\right) + \arcsin\left(\frac{1}{t}\right)$ is an increasing function on t and that $t \geq \sqrt{t^2-1} \geq (t-1)$. The last inequalities have been checked by software.

Assume now that t is even. We prove that $b_t(t/2-1) = b_t(t/2) = t/2-1$ if $t \geq 3$. By our definition of b_t, as $2\sqrt{t^2/4} = t$ is an integer, we have $b_t(t/2) = t-1$. For $b_t(t/2-1)$, we have

$$2\sqrt{\left(\frac{t}{2}-1\right)\left(\frac{t}{2}+1\right)} = \sqrt{t^2-4},$$

implying

$$t > \sqrt{t^2-4} \geq (t-1),$$

where the last inequality comes from $t^2 - 4 \geq t^2 - (2t-1) = (t-1)^2$. We have therefore $\sum_{a=1}^{t-1} b_t(a) = 2\sum_{a=1}^{t/2-1} b_t(a) + (t-1)$, so we bound,

$$2\int_{u=1}^{t/2} (2\sqrt{u(t-u)} - 1)du + (t-1) \leq \sum_{a=1}^{t-1} b_t(a)) \leq 4\int_{u=1}^{t/2} \sqrt{u(t-u)}du + (t-1).$$

Notice again the endpoints of the integrals. Similarly as Equation (8.3), we get the upper bound

$$\frac{1}{2}t^2 \arcsin\left(\frac{t-2}{t}\right) + \sqrt{t-1}\,(t-2) + (t-1) \leq \frac{t^2}{2} + \sqrt{t-1}\,(t-2) + (t-1)\left(\leq \frac{3}{4}t^2\right);$$

and the lower bound

$$\frac{1}{2}t^2 \arcsin\left(\frac{t-2}{t}\right) + \sqrt{t-1}\,(t-2) - (t-2) + (t-1) \geq \frac{t^2}{2}\arcsin(1/3) + \sqrt{t-1}\,(t-2) + 1\left(\geq \frac{t^2}{2}\right).$$

Finally, we count $\sum_{a=1}^{t-1}(2b(t,a)+1)$, (for $t \geq 4$):

$$t^2 + (t-1) \leq \{M \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+},\ M > 0,\ \mathrm{tr}\, M = t\} \geq \frac{5}{3}t^2 + (t-1),$$

and one can compute that there are 3 matrices for $t = 2$ and 10 for $t = 3$.

The second statement follows in the same way. □

8.1.2.3 Bound on Fourier coefficients of cusp forms

We can prove the following bound for Siegel cusp forms as a direct generalization of the Hecke estimate for elliptic cusp forms, that furthermore scales properly after taking powers. It is true that for Siegel cusp forms there are better bounds available, but this is enough for our purposes.

Proposition 8.14 (Hecke estimate for Siegel cusp forms)**.** *Consider a Siegel cusp form g of weight k with Fourier expansions*

$$g = \sum_{T>0} a_g(T)\boldsymbol{q}^T,$$

then the function on $\mathbb{H}_2$ given by

$$G := \det(\mathrm{Im}\,\boldsymbol{\tau})^{k/2} g(\boldsymbol{\tau})$$

is bounded in $\mathbb{H}_2$. Setting $C(g) := \sup_{\mathbb{H}_2} G$ and $C^*(g) = e^{4\pi}C(g)$, we have the following bound for all $T > 0$:

$$|a_g(T)| \leq C^*(g)\det(T)^{k/2}.$$

In particular for any $l > 0$ for any power g^l it holds:

$$|a_{g^l}(T)| \leq (C^*(g))^l \det(T)^{lk/2}.$$

Proof. The proof is a straightforward generalization of the known Hecke bound for elliptic cusp forms. By construction it follows that G is invariant under the $\mathrm{Sp}_4(\mathbb{Z})$-action, and it is furthermore bounded at infinity because g is a cusp form: this follows from the fact that for T h.i.s and positive *definite*, we have the bound:

$$|\det(\boldsymbol{y})^{k/2} e^{i2\pi\,\mathrm{tr}(T\boldsymbol{\tau})}| = \det(\boldsymbol{y})^{k/2} e^{-2\pi\,\mathrm{tr}(T\boldsymbol{y})} \leq \mathrm{tr}(\boldsymbol{y}))^k e^{-2\pi\delta(T)\,\mathrm{tr}(\boldsymbol{y})},$$

where $\delta(T) > 0$ is the minimum of the eigenvalues of T. The inequality

$$\det(\boldsymbol{y}) \leq \frac{\mathrm{tr}(\boldsymbol{y})^2}{4}$$

follows from the fact that $\boldsymbol{y}$ is positive definite, as in dimension two implies $\mathrm{tr}(\boldsymbol{y})^2 - 4\det(\boldsymbol{y}) > 0$. The inequality in the exponential comes from Lemma 8.36 (applied to T instead of $\mathrm{Im}\,\boldsymbol{\tau}$). See also [Igu72, Lemma V.5.21] for its generalization to arbitrary genus.

Set $C(g)$ as above. Recall the definition of the Fourier coefficients, setting $\boldsymbol{\tau} = \boldsymbol{x} + i\boldsymbol{y}$

$$a_g(T) = \int_{\boldsymbol{x} \bmod 1} g(\boldsymbol{x} + i\boldsymbol{y}) e^{-2\pi i\,\mathrm{tr}(T\boldsymbol{\tau})} d\boldsymbol{x},$$

so

$$|a_g(T)| \leq \int_{\boldsymbol{x} \bmod 1} C(g)\det(\boldsymbol{y})^{-k/2} e^{2\pi \operatorname{tr}(T\boldsymbol{y})} d\boldsymbol{x} = C(g)\det(\boldsymbol{y})^{-k/2} e^{2\pi \operatorname{tr}(T\boldsymbol{y})},$$

and this bound holds for every $\boldsymbol{y} = \operatorname{Im}\boldsymbol{\tau}$ with $\boldsymbol{\tau} \in \mathbb{H}_2$. Setting $\boldsymbol{y} = T^{-1}$ (remark that because T is positive definite, its inverse is too), we have

$$|a_g(T)| \leq C(g) e^{4\pi} \det(T)^{k/2}.$$

□

8.1.3 Igusa invariants and Igusa-Streng invariants

We want to generalize the"modular" proof of the Stéphanois theorem, because considering the Fourier expansion of the Igusa invariants themselves may not be as convenient, as they are not well defined on $\chi_{10} = 0$ and we do not currently have in the literature known bounds for this coefficients.

The modular proof is more suitable for our purposes, as the Igusa invariants have common denominators by powers of χ_{10}, and we can use bounds for Fourier coefficients of cusp Siegel modular forms by Proposition 8.14.

Consider the Siegel Eisenstein series E_4, E_6, E_{10}, E_{12}, and define the normalized cusp forms of Igusa [Igu62, page 195]

$$\chi_{10} = \frac{-43867}{2^{12}\cdot 3^5 \cdot 5^2 \cdot 7 \cdot 53}(E_4E_6 - E_{10})$$

and

$$\chi_{12} = \frac{131 \cdot 593}{2^{13}\cdot 3^7 \cdot 5^3 \cdot 7^2 \cdot 337}\left(3^2 \cdot 7^2 E_4^3 + 2 \cdot 5^3 E_6^2 - 691 E_{12}\right).$$

The Igusa invariants are given by

$$j_1 = 2 \cdot 3^5 \frac{\chi_{12}^5}{\chi_{10}^6},\ j_2 = \frac{3^3}{2^3}\frac{E_4\chi_{12}^3}{\chi_{10}^4},\ j_3 = \frac{3}{2^5}\frac{E_6\chi_{12}^2}{\chi_{10}^3} + \frac{3^2}{2^3}\frac{E_4\chi_{12}^3}{\chi_{10}^4},$$

Remark that χ_{10}^6 is a common denominator for j_l, and $\chi_{10}^6 j_l$ is a cusp form of weight 60.

These are the more standard Igusa invariants, but we want to work instead with the Igusa-Streng's invariants of [Str10]. We presented them already in Section 1.1 and Subsection 1.3.1, and they admit the expressions:

$$\tilde{j_1} = \frac{1}{2^{10}}\frac{E_4E_6}{\chi_{10}},\ \tilde{j_2} = 2^7 \cdot 3\frac{\chi_{12}E_4^6}{\chi_{10}},\ \tilde{j_3} = \frac{1}{2^{18}}\frac{E_4^5}{\chi_{10}^2}.$$

One normalizes instead by $\chi_{10}^2\tilde{j_1}$, $\chi_{10}\tilde{j_2}$, $\chi_{10}^3\tilde{j_3}$.,

In addition, from [Str10, Appendix 1],the Fourier coefficients of $E_4, E_6, 4\chi_{10}, 12\chi_{12}$ are all integers, and each one of them has one coefficients equal to 1.

Finally, as $\mathbb{C}(A_2) = \mathbb{C}(j_1, j_2, j_3)$, together with Proposition 8.2 we have the following result of algebraic independence, necessary for the construction of the auxiliary functions.

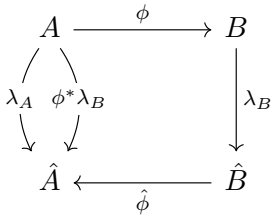
Theorem 8.15. *The functions j_1, j_2, j_3 are algebraically independent over $\mathbb{C}(q_1, q_2, q_3)$.*

8.1.4 Modular polynomials

An important element in the proofs are the modular polynomials, both Siegel and Hilbert. Here we recall them briefly with its most important properties. We based this presentation on [BCCK24, Section 2.1], [KPR25, Section 2.5], [MR20, Section 4.1] and [BLS13, Section 2].

One first key difference with the one dimensional analog is that we encounter different types of isogenies in the two dimensional version. Consider an isogeny between complex abelian surfaces $\phi : A \to B$, which over $\mathbb{C}$ only means an isogeny of complex tori. Once we consider polarizations, one could ask what natural extra condition to impose on the isogeny to consider it a "polarised isogeny" of sorts.[6]

For this section, we are going to consider polarizations as (special type of) isogenies $\lambda_A : A \to \hat{A}$. If (B, λ_B) is principally polarized, ϕ induces an polarization on A given by $\phi^*\lambda_B = \hat{\phi} \circ \lambda_B \circ \phi$, for $\hat{\phi} : \hat{B} \to \hat{A}$ the dual isogeny of B.



One could then think that a natural condition for ϕ is to impose $\lambda_A = \phi^*\lambda_B$, but this is too restrictive (see [Orr17, end of p. 674]) as by degree computations $\deg(\lambda_A) = \deg(\phi)^2 \deg(\lambda_B)$. Hence, unless ϕ is an isomorphism, it cannot induced a *principal* polarization.

Therefore, we do not restrict our notion of isogeny once we consider polarized abelian varieties, but as we will see below in Definition 8.16, the effect of the isogeny on the polarizations will be used to classify them.

Hence our set up is, given one principally polarized abelian surface A, consider all usual isogenies $\phi : A \to B$, and ask which B admits a principal polarization.[7] Equivalently, as $B \cong A/\ker\phi$ as unpolarized abelian varieties, we ask for which finite subgroups $\ker\phi$ the quotient $A/\ker\phi$ admits a principal polarization.

Let us resume $\phi : A \to B$ an isogeny, and principal polarizations λ_A and λ_B. Consider the induced polarization on A, $\phi^*\lambda_B$, and set

$$f := \lambda_A^{-1}\phi^*\lambda_B \in \mathrm{End}(A).$$

It extends to a group homomorphism $\mathrm{NS}(A) \to \mathrm{End}(A)$, and restricts to $\mathrm{NS}(A) \to \mathrm{End}^s(A, \lambda_A)$, which is a group isomorphism, as we stated in Proposition 3.19. Recall that the Rosati involution with respect to λ_A is given as $' : \mathrm{End}(A) \to \mathrm{End}(A)$, $g \mapsto \lambda_A^{-1} \circ \hat{g} \circ \lambda_A$. Furthermore, the polarizations of A under this morphism are identified with the *totally positive* elements of $\mathrm{End}^s(A, \lambda_A)$, see [BL04, Theorem 5.2.4]. We denote the subset of totally positive elements of $\mathrm{End}^s(A, \lambda_A)$ by $\mathrm{End}^{s,++}(A, \lambda_A)$.

Reciprocally, if we start with $f \in \mathrm{End}^{s,++}(A, \lambda_A)$, and the polarization of A given by $\lambda_A \circ f$, to recover the isogeny $\phi : A \to B$ such that $\phi^*\lambda_B = \lambda_A \circ f$, we just need $V = \ker\phi$: in this case $B \cong A/V$. This quotient admits a principal polarization by Proposition 8.20 below.

[6] We have found different definitions of this concept, so for the sake of clarity we are not going to use this denomination.

[7] a polarization on B always exist, but it is not necessarily principal.

Let us focus now on the $V = \ker \phi$, and how it depends on f. Here, we restrict to the generic case that A is a *simple* abelian surface. Then, by our analysis of Humbert singular relations in Section 3.1, in particular f solves a quadratic equation over $\mathbb{Z}$ with positive discriminant in $\mathbb{Z}_{>0}$ (that is not a perfect square), and so f is either a positive integer $n \in \mathbb{Z}_{>0}$, or β is a real quadratic integer.

On the latter case, set $\mathcal{O} = \mathbb{Z}[f] \subset \mathrm{End}^s(A, \lambda_A)$: then (A, λ_A) has real multiplication by the order $\mathcal{O} \subset \mathbb{Q}(f)$. We will see below in Remark 8.23 that (B, λ_B) also has real multiplication by $\mathcal{O}$. More generally, it is natural to consider from the beginning a real multiplication structure $\mathcal{O} \hookrightarrow \mathrm{End}^s(A, \lambda_A)$, and ask whether the isogeny respects it.

Consequently, we define the following type of isogenies.

Definition 8.16. *Consider (A, λ_A) and (B, λ_B) principally polarised abelian surfaces, and consider $\phi : A \to B$ an isogeny.*

- *We say that ϕ is an n-isogeny for $n \in \mathbb{Z}_{>0}$ if*
$$\hat{\phi} \circ \lambda_B \circ \phi = \lambda_A \circ [n_A],$$
for $[n_A] \in \mathrm{End}(A)$ the multiplication-by-n endomorphism.[8] *If furthermore $\ker \phi \cong \left(\mathbb{Z}/n\mathbb{Z}\right)^2$, we will say that ϕ is a (n, n)-isogeny.*
- *Let K a real quadratic field and consider an order $\mathcal{O} \subset K$. Assume furthermore that A and B have real multiplication by $\mathcal{O}$ with given endomorphism $\iota_A : \mathcal{O} \hookrightarrow \mathrm{End}^s(A, \lambda_A)$ and $\iota_B : \mathcal{O} \hookrightarrow \mathrm{End}^s(B, \lambda_B)$. Consider $\beta \in \mathcal{O}$ a totally positive number, then we say that ϕ is a β-isogeny if*
$$\hat{\phi} \circ \lambda_B \circ \phi = \lambda_A \circ \iota_A(\beta).$$

Remark 8.17. *We will see in Proposition 8.24 below that a n-isogeny satisfies $\deg(f) = n^2$. It follows then that for $n = p$ prime, a p-isogeny is always a (p, p)-isogeny.*

The first type of isogenies, when $n = p$ is a prime number, are parametrized by the Siegel modular polynomials. The second type, for β a (totally positive) prime number in $\mathcal{O}_K$, the ring of integers of K, are parametrized by the Hilbert modular polynomials.

Remark 8.18. *These two types of isogeny have a non-trivial (non empty, and not a subset) intersection. They are not exclusive, as one is allowed to take $\beta \in \mathbb{Z}_{>0}$ in the definition of a β-isogeny. In that case, it is a n-isogeny that preserves the real multiplication, which is a non trivial extra condition.*

Let us return to $\ker \phi$. From $f = \lambda_A^{-1} \circ \hat{\phi} \circ \lambda_B \circ \phi$, necessarily $\ker \phi \subset \ker f$. If $f = [n_A]$, then $\ker f = A[n]$, so for $f = \iota_A(\beta)$, for $\iota_A : \mathcal{O}_K \to \mathrm{End}^s(A, \lambda_A)$, we extend the notation $A[\beta] := \ker \iota_A(\beta)$.

The polarization $\lambda_A \circ [n_A]$ induces a symplectic pairing[9] on $A[n] \times A[n] \to \mu_n$, given by the Weil pairing, which can be written in terms of the associated Riemann form E_A as

$$\begin{aligned} e_n : A[n] \times A[n] &\to \mu_n \\ (u, v) &\mapsto \exp(-i2\pi n E_A(u, v)). \end{aligned}$$

[8] In some reference this is what is called polarised isogeny, without making reference to n.

[9] in multiplicative notation.

More generally, for a polarization of the form $\lambda_A \circ f$ with $f \in \mathrm{End}^{s,++}(A, \lambda_A)$, one can define a symplectic pairing as the Weil paring, see [Mum08, Definition p. 210] (and [Mum08, p. 228] for its functorial properties), or also see [Mil86a, Section 16]. We write it $e_{\lambda_A \circ f}$, or if the context is clear, simply e_f. From the specialization for complex abelian varieties in [Mum08, Theorem 1, p.237] it follows that $e_f(u, v) = \exp(-i2\pi E_f(u, v))$, for E_f the associated Riemann form to the polarization $\lambda_A \circ f$. From [BL04, Remark 5.2.2, Proposition 5.1.1], if $\rho_r(f) \in \mathrm{Mat}_4(\mathbb{Z})$ the rational representation of f, then $E_f(u, v) = E_A(\rho_r(f)u, v)$. Alternatively, as $\phi^*\lambda_B = \lambda_A \circ f$, it can be recovered as $E_f(u, v) = (\phi^* E_A)(u, v) = E_B(\rho_r(\phi)u, \rho_r(\phi)v)$. Then we have two properties for $\ker \phi$ in terms of the Weil pairing:

- $\ker \phi \subset \ker f$ is isotropic[10] for the Weil pairing, meaning that $e_f|(V \times V) \equiv 1$,
- and it is maximal among the isotropic subgroups of $(\ker f, e_f)$.

Remark 8.19. *Observe that as e_f is symplectic, for any isotropic subgroup $(\#V) \leq \sqrt{\#\ker f}$, and by [Mum08, Theorem 4, p. 234], a subgroup is maximal isotropic if and only if equality holds.*

The first property follows from $\phi^*\lambda_B = \lambda_A \circ f$ and by [Mum08, Property (1) p. 228], $e_{\phi^*\lambda_B}(u, v) = e_{\lambda_B}(\phi u, \phi v)$, so e_f is trivial on $\ker \phi \times \ker \phi$.

In general, we have the following more general criteria for descend of polarizations to quotients:

Proposition 8.20. *Given (A, λ_A) a principally polarized abelian variety, $f \in \mathrm{End}^{s,++}(A, \lambda_A)$ with associated polarization $\lambda_A \circ f$, and Weil pairing e_f. Consider $V \subset \ker f$, then the following are equivalent:*

- *$V \subset \ker f$ is isotropic for e_f,*
- *There exists a polarization μ in A/V such that $P^*\mu = \lambda_A \circ f \in \mathrm{NS}(A)$, for the projection $P : A \to A/V$*

In such a case, $\deg(f) = \deg(\mu)(\#V)^2$

Proof. This is [Mil86a, Proposition 16.8], or more generaly [Mum08, Corollary p.232], for line bundles. The second part follows from $\lambda_A \circ f = \hat{p} \circ \mu \circ p$ and computing degrees. □

Maximality of $\ker \phi$ among isotropic subgroups follows from the fact that B is principally polarized: if $\ker \phi \subset V$ with V isotropic, then there is a polarization λ'_B on $B' = A/V$ inducing $\phi^*\lambda_B$ on A via the projection $A \to A/V$. If we consider instead

$$A \to A/\ker \phi \to A/V,$$

it then follows that $\lambda_B = (P')^*\lambda_{B'}$ for $P' : A/\ker \phi \to A/V$, so $\deg(\lambda_B) = 1$ implies in particular $\deg P' = 1$ and $V = \ker \phi$.

Proposition 8.21. *([BCCK24, Lemma 2.1]) Consider (A, λ_A) a ppas. There is a bijection between:*

[10] sometimes this is called *totally* isotropic.

- *isomorphisms classes of isogenies*[11] $\phi : A \to B$ *with* B *admitting a principal polarization* λ_B,
- *pairs* (f, V), *where* $f \in \mathrm{End}^{s,++}(A, \lambda_A)$ *and* $V \subset \ker f$ *is maximal isotropic with respect to the Weil pairing* e_f.

The assignment is given by $\phi \mapsto (\lambda_A^{-1}\phi^*\lambda_B, \ker\phi)$, *where* $\phi^*\lambda_B = \hat{\phi}\lambda_B\phi$ *is the polarization induced by* ϕ.

Proof. One implication has already been argued. For the other one, starting with a pair (f, V) as above by Proposition 8.20, the polarization $\lambda_A \circ f$ descends to a polarization μ in the quotient A/V, which is principal.

This is an inverse to the other assignment, because for $(f, \ker\phi)$ coming from $\phi : A \to B$, the natural isomorphism $A/\ker\phi \to B$ we have the following commutative diagram

$$\begin{array}{ccc} (A, \lambda_A) & \xrightarrow{\phi} & (B, \lambda_B) \\ & \searrow^{P} & \cong\uparrow \\ & & (A/\ker\phi, \mu) \end{array}$$

which also proves that the vertical map is an isomorphism of ppas. □

Observe that for β-isogenies, it is natural to consider how it affects the real multiplication structure. Under an isogeny $\phi : A \to B$, the endomorphism algebra is invariant but that is not true for the endomorphism *ring*. Assuming $\mathcal{O} \subset K$ a real multiplication order, and an embedding $\iota_A : \mathcal{O} \hookrightarrow \mathrm{End}^s(A, \lambda_A)$, then ι_A induces a map $\iota_B : \mathcal{O} \hookrightarrow \mathrm{End}\left(A/\ker\phi\right)$ if and only if $\ker\phi$ is invariant under $\mathcal{O}$ (via ι_A).

Note then that the real multiplication structure *does* restrict the isogenies we consider, without the polarization.

Proposition 8.22. *Consider* $\mathcal{O} \subset K$ *a real quadratic order. Consider* (A, λ_A) *a ppas with RM by* $\mathcal{O}$*: meaning there exists an embedding* $\iota_A : \mathcal{O} \hookrightarrow \mathrm{End}^s(A, \lambda_A)$. *There is a bijection between:*

- *isomorphisms classes of isogenies* $\phi : A \to B$ such that B has RM by $\mathcal{O}$ compatible with ϕ, *with* B *admitting a principal polarization* λ_B,
- *pairs* (f, V), *where* $f \in \mathrm{End}^{s,++}(A, \lambda_A)$ *and* $V \subset \ker f$ *is* invariant under the action of $\iota_A(\mathcal{O})$ *and is maximal isotropic with respect to the Weil pairing* e_f.

Remark 8.23. *Consider the set-up at the beginning* $(A, \lambda_A), (B, \lambda_B) \in \mathcal{A}_2$, *an isogeny* $\phi : A \to B$ *and* $f = \lambda_A^{-1} \circ \phi^*\lambda_B \in \mathrm{End}^{s,++}(A, \lambda_A)$. *Then* $\mathbb{Z}[f]$, *that we identify with* $[m]_A + [n]_A \circ f$ *acts on* $\ker(f)$ *by* $[n]_A$, *so* $\ker(f)$ *is stable under* $\mathbb{Z}[f]$. *Hence, when* $\mathbb{Z}[f]$ *is not trivial, the isogeny* ϕ *induces RM by* $\mathbb{Z}[f]$ *on* (B, λ_B).

[11] meaning up to $(B, \lambda_B) \cong (B', \lambda_{B'})$ isomorphims of ppas, and $\phi' : A \to B'$ compatible with said isomorphism, which in particular implies $\ker\phi = \ker\phi'$ and that $\phi^*\lambda_B = (\phi')^*\lambda_{B'}$.

Proposition 8.24. *For ϕ a n-isogeny, $\deg\phi = n^2$, and they are never cyclic isogenies. For $n = p$ a prime number, one furthermore has $\ker\phi \cong \left(\mathbb{Z}/p\mathbb{Z}\right)^2$.*

For K a real quadratic field and ϕ a β-isogeny, for $\beta \in \mathcal{O}_K^{++}$ a prime element, then $\deg\phi = N_{K|\mathbb{Q}}(\beta)$. Hence, if p is a prime that splits in K as $p = \beta\widetilde{\beta}$, ϕ is a cyclic isogeny.

Proof. By [BL04, Proposition 1.2.5] that $A[n] \cong \left(\mathbb{Z}/n\mathbb{Z}\right)^4$, so by Remark 8.19, $\deg\phi = n^2$, and $\ker\phi \ncong \left(\mathbb{Z}/n^2\mathbb{Z}\right)$, as there are no elements of order n^2.

In the case of real multiplication, analogously as above, $A[\beta] \cong \left(\frac{1}{\beta}\Lambda\right)/\Lambda \cong \Lambda_{\boldsymbol{z}}/\beta\Lambda_{\boldsymbol{z}}$ with $\Lambda = \mathcal{O}_K\boldsymbol{z} \oplus (\partial_K)^{-1}$, for $\boldsymbol{z} = (z_1, z_2) \in \mathbb{H}^2$ parameterizing A, so

$$\begin{aligned}\deg\beta &= \left[\mathcal{O}_K \oplus (\partial_K)^{-1} : \beta(\mathcal{O}_K \oplus \partial_K^{-1})\right] = [\mathcal{O}_K : \beta\mathcal{O}_K]\left[\partial_K^{-1} : \beta\partial_K^{-1}\right] \\ &= [\mathcal{O}_K : \beta\mathcal{O}_K]^2 = N_{K|\mathbb{Q}}(\beta)^2\end{aligned}$$

so as $\deg\phi = \sqrt{\deg\beta}$, the conclusion follows. □

Our distinguished isogeny. We are considering the isogeny given by $\boldsymbol{\tau} \mapsto p\boldsymbol{\tau}$, for p prime. For $\boldsymbol{\tau} \in \mathbb{H}_2$ the kernel of the induced isogeny $\phi : A_{\boldsymbol{\tau}} \to A_{p\boldsymbol{\tau}}$ is isomorphic (over $\mathbb{C}$) as an abelian group to $\mathbb{Z}^2 \oplus \boldsymbol{\tau}\mathbb{Z}^2/\mathbb{Z}^2 \oplus p\boldsymbol{\tau}\mathbb{Z}^2 \cong \left(\mathbb{Z}/p\mathbb{Z}\right)^2$. On the other hand, the induced Riemann form is $p\begin{pmatrix} 0 & I \\ -I & 0\end{pmatrix}$, hence the isogeny ϕ is a (p,p)-isogeny.

Hilbert and Siegel isogenies. By Remark 8.18, it is not always the case that a n-isogeny respects the real multiplication. We are going to show that via Humbert singular relations for our distinguished isogeny.

Lemma 8.25. *Consider p a prime number and $\mathcal{O}$ a real quadratic order with discriminant Δ, and $\boldsymbol{\tau} \in \mathbb{H}_2$ such that $A_{\boldsymbol{\tau}} \in \mathcal{H}_\Delta$, and consider a primitive Humbert singular relation (a,b,c,d,e) for $\boldsymbol{\tau}$. Then $A_{\boldsymbol{\tau}} \in \mathcal{G}_{p^2\Delta}$. If $d = 0$ then $A_{p\boldsymbol{\tau}} \in \mathcal{G}_\Delta$, and if $d = e = 0$, then $A_{p\boldsymbol{\tau}} \in \mathcal{H}_\Delta$.*

Proof. There exists $(a,b,c,d,e) \in \mathbb{Z}$ coprime with $b^2 - 4(ac+de) = \Delta$ with

$$a\tau_1 + b\tau_2 + c\tau_3 + d(\tau_2^2 - \tau_1\tau_3) + e = 0.$$

It follows that $p\boldsymbol{\tau}$ solves

$$(ap)\tau_1' + (bp)\tau_2' + (cp)\tau_3' + d(\tau'^2 - \tau_1'\tau_3') + (p^2e) = 0,$$

of discriminant $p^2\Delta$. A priori, this relation could not be primitive, so we can only guarantee $\mathcal{G}_{p^2\Delta}$. In particular, if $d = 0$, then it simplifies to

$$a\tau_1' + b\tau_2' + c\tau_3' + (pe) = 0,$$

which has the same discriminant. □

Proposition 8.26. *Consider K a real quadratic field and the Hilbert modular surface of maximal real multiplication by K. Consider p an inert prime number in K. Then the multiplication-by-p isogeny, given by $\boldsymbol{z} \to (pz_1, pz_2)$ in the Hilbert variables, is a (Hilbert) p-isogeny.*

Proof. This follows from the above lemma, together with the Hilbert modular embeddings from Theorem 2.23. □

8.1.4.1 Construction of modular polynomials

Considering the classical modular polynomials $\phi_N(X, Y) \in \mathbb{Z}[X, Y]$, the are two equivalent approaches to its construction:

- *Prescribing the isogenies.* The set

$$C_N := \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \ : \ a, b, d \in \mathbb{Z},\ ad = N,\ a \geq 1,\ 0 \leq b \leq d-1,\ \gcd(a, b, d) = 1 \right\},$$

  is a set of representatives for $\{\gamma \in \mathrm{Mat}_N(Z) \ : \ \det \gamma = N,\ \gcd((\gamma_{i,j})) = 1\}$ under the action of $\mathrm{SL}_2(\mathbb{Z})$, and it parametrizes all cyclic isogenies of degree N between elliptic curves up to isomorphism. Then one can prescribe a polynomial such that the roots are precisely the j-invariants of the cyclic N isogenies to a given $\tau \in \mathbb{H}$. In terms of functions, one sets

$$\Phi_N(X, j(\tau)) = \prod_{\gamma \in C_N} (X - j(\gamma\tau)) \in \mathbb{C}(j)[X]$$

  as a polynomial in the field of functions $\mathbb{C}(j)$. One can then prove that $\Phi_N(X, j) \in \mathbb{Z}[j][X]$, and considering as polynomials in two variables $\Phi_N(X, Y)$.

- *Modular curves as covers of $X(1)$.* Considering the modular function $j_N(\tau) := j(N\tau)$, it is a modular function not for the full modular group but for the Hecke congruence subgroup $\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),\ c \equiv 0\ (N) \right\}$. Considering the modular curve $X_0(N)$, that satisfies $X_0(N)(\mathbb{C}) = \Gamma_0(N)\backslash\mathbb{H}^*$, the inclusion $\Gamma_0(N) \hookrightarrow \mathrm{SL}_2(\mathbb{Z})$ induces a cover $X_0(N) \to X(1)$. In turn it induces a field extension $\mathbb{C}(X(1)) \subset \mathbb{C}(X_0(N))$ (actually, it is also defined over $\mathbb{Q}$: $\mathbb{Q}(X(1)) \subset \mathbb{Q}(X_0(N))$). This extension is finite, the degree being the degree of the cover of $X_0(N) \to X(1)$. Furthermore, as $\mathbb{C}(X(1)) = \mathbb{C}(j)$ and $\mathbb{C}(X_0(N)) = \mathbb{C}(j, j_N)$, Φ_N is then defined as the minimal polynomial of j_N over $\mathbb{Q}(j)$. One can likewise prove that $\Phi_N(X, Y) \in \mathbb{Z}[X, Y]$ as a polynomial in two variables.

  The cover is Galois, and fixing $\{\alpha\}$ representatives for the cosets $\Gamma_0(N)\backslash \mathrm{SL}_2(\mathbb{Z})$, one see that $\{\tau \mapsto j(N\alpha\tau)\}$ are the distinct conjugates to j_N in $\mathbb{C}(X_0(N))$ and hence

$$\Phi_N(X, j) = \prod_{\alpha \in \Gamma_0(N)\backslash \mathrm{SL}_2(\mathbb{Z})} (X - j(N\alpha\tau)$$

Both approaches are equivalent and define the exact same polynomials, but it is the second one the more suitable for generalization, as $\Gamma_0(N)$ generalizes more directly to suitable subgroups of $Sp_4(\mathbb{Z})$ and $\mathrm{SL}_2(\mathcal{O}_K)$. The abstract construction of the modular polynomials follows from considering the corresponding finite extension at the level of fields of functions.

We sketch it for the Siegel modular polynomials, following [BL09, Section 3 and Section 4], and [Mil15, Section 1]. Fix p a prime number. One considers

$$\Gamma_0^{(2)}(p) := \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z}) : \ \gamma \equiv 0_2 \ (p) \right\} \subset \mathrm{Sp}_4(\mathbb{Z}),$$
$$\Gamma^{(2)}(p) := \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z}) : \ \beta \equiv \gamma \equiv 0_2 \ (p), \ \alpha \equiv \delta \equiv I_2 \ (p) \right\} \subset \Gamma_0^{(2)}(p),$$

and note that $\Gamma^{(2)}(p) = \ker\left(\mathrm{Sp}_4(\mathbb{Z}) \to \mathrm{Sp}_4\left(\mathbb{Z}/p\mathbb{Z}\right)\right)$

Lemma 8.27. *We have that, for p prime, the index $[\mathrm{Sp}_4(\mathbb{Z}) : \Gamma_0^{(2)}(p)]$ equals:*

- *the number of 2-dimensional isotropic subspaces of the symplectic $\mathbb{Z}/p\mathbb{Z}$-vector space $\left(\mathbb{Z}/p\mathbb{Z}\right)^4$,*
- *the number of lines in $\left(\mathbb{Z}/p\mathbb{Z}\right)^4$,*

and hence $[\mathrm{Sp}_4(\mathbb{Z}) : \Gamma_0^{(2)}(p)] = \frac{p^4-1}{p-1} = 1 + p + p^2 + p^3$.

Then $\Gamma_0^{(2)}(p)\backslash\mathbb{H}_2$ solves the folloing moduli problem.

Theorem 8.28. *We have the following*

- *There is a bijection*

$$\Gamma_0^{(2)}(p)\backslash\mathbb{H}_2 \to \left\{ \begin{smallmatrix} \text{pairs } (A,V),\ A\in\mathcal{A}_2,\ V\subset A[p] \\ \text{maximal isotropic subgroup for } e_p \end{smallmatrix} \right\} \Big/ \phi : (A,V) \to (A',V') \text{ isom of ppas}$$
$$\boldsymbol{\tau} \mapsto \left(A_{\boldsymbol{\tau}}, \left\langle \left(\frac{1}{p},0,0,0\right), \left(0,\frac{1}{p},0,0\right) \right\rangle \right).$$

- *Define*

$$Y_0^{(2)}(p) := \Gamma_0^{(2)}(p)\backslash\mathbb{H}_2,$$

the $Y_0^{(2)}(p)$ is not a compact space, but admits the following compactification:[12]

$$X_0^{(2)}(p) := Y_0^{(2)}(p) \cup Y_0(p) \cup \mathbb{P}^1(\mathbb{Q}),$$

which makes $Y_0^2(p)$ a quasi projective variety.

Proof. This is [BL09, Theorem 3.2]. □

As in the one dimensional case, the inclusion $\Gamma_0^{(2)}(p) \subset \mathrm{Sp}_4(\mathbb{Z})$ induces a cover $Y_0^2(p) \to \mathcal{A}_2$ (that extents to the compactifications), and hence we have a finite degree extension of the field of functions, with degree given by the index $[\mathrm{Sp}_4(\mathbb{Z}) : \Gamma_0^2(p)]$. The Siegel modular polynomials will then describe the field extension $\mathbb{C}(Y_0^2(p)) \supset \mathbb{C}(\mathcal{A}_2)$.

[12] Recall the Satake compactification of $\mathcal{A}_2^{Sat} = \mathcal{A}_2 \sqcup \mathcal{A}_1 \sqcup \mathcal{A}_0$.

We set $F = \mathbb{C}(j_1, j_2, j_3)$ for the field of functions of $\mathcal{A}_2$. Consider the functions for $i = 1, 2, 3$:

$$j_{i,p}(\boldsymbol{\tau}) := j_i(p\boldsymbol{\tau}), \tag{8.4}$$

meaning multiplication by p as a scalar on the matrix $\boldsymbol{\tau}$. Written as a fractional linear transformation, $\begin{pmatrix} pI_2 & 0 \\ 0 & I_2 \end{pmatrix}$.

They can be proven to be $\Gamma_0^{(2)}(p)$-invariant[13] ([BL09, Lemma 4.1]), and furthermore, each one of them it is a generator of the field of functions of $Y_0^{(2)}(p)$.

Theorem 8.29. *The field of functions $\mathbb{C}(Y_0^{(2)}(p))$ equals*

$$F(j_{1,p}) = F(j_{2,p}) = F(j_{3,p}),$$

for $F = \mathbb{C}(j_1, j_2, j_3)$, and the extension has degree $[\mathrm{Sp}_4(\mathbb{Z}) : \Gamma_0^2(p)] = \delta(p)$.

That implies that there exists a polynomial $\Phi_{1,p} \in \mathbb{Q}(j_1, j_2, j_3)[Y]$ and rational functions $Q_{i,p}, Q_{2,p} \in \mathbb{C}(j_1, j_2, j_3)(Y)$ such that:

- *$\Phi_{1,p}$ is the minimal polynomial of $j_{1,p}$ over $\mathbb{C}(j_1, j_2, j_3)$, with degree $\deg_Y \Phi_{1,p} = \delta(p)$*
- *$j_{i,p} = Q_{i,p}(j_{1,p})$ for $i = 2, 3$.*

Outside of a subvariety of $\mathcal{A}_2$, the rational functions $Q_{i,p}$ can be taken of the form $\frac{\Phi_{i,p}}{\partial_Y \Phi_{1,p}}$ for $i = 2, 3$, for $\Phi_{i,p} \in \mathbb{Q}(j_1, j_2, j_3)[Y]$ a polynomial *of $\deg \Phi_{i,p} = \delta(p) - 1$, and $\partial_Y \Phi_{1,p}$ the derivative of $\Phi_{1,p}$ with respect to Y.*

Proof. This is [BL09, Lemma 4.2, Theorem 5.2, paragraph between Remark 6.2 and Lemma 6.3]. □

Definition 8.30. *Let p a prime number. We define the p-th* Siegel modular polynomials *as $\Phi_{i,p} \in \mathbb{Q}(j_1, j_2, j_3)[Y]$ as the polynomials given by the Theorem 8.29.*

Explicit formulas. It follows for similar reasons as in the classical modular polynomial that the first modular polynomial verifies, for any $\boldsymbol{\tau} \in \mathbb{H}_2$:

$$\Phi_{1,p}(Y) = \prod_{M \in \Gamma_0^2(p) \backslash \mathrm{Sp}_4(\mathbb{Z})} (Y - j_{1,p}(M\boldsymbol{\tau})) \in \mathbb{Q}(j_1, j_2, j_3)[Y], \tag{8.5}$$

one could then consider analogously the minimal polynomials of $j_{2,p}$ and $j_{3,p}$ over $\mathbb{C}(j_1, j_2, j_3)$, which have an analogous formula, but in practice that would imply solving three different polynomials in terms of the given invariants $j_1(\boldsymbol{\tau}), j_2(\boldsymbol{\tau}), j_3(\boldsymbol{\tau})$, and getting $\delta(d)^3$ triples of invariants, instead of $\delta(d)$, that we know is the correct generic number of surfaces ([GHKRW06, Section 3]). What Theorem 8.29 says instead, is that one only needs to solve for $\Phi_{1,p}$ and the other invariants are completely determined by that choice of root. Of course, that requires for $\Phi_{1,p}$ to have distinct roots when evaluated at $j_1(\boldsymbol{\tau}), j_2(\boldsymbol{\tau}), j_3(\boldsymbol{\tau})$, or equivalently to not belong to the zero locus of $\partial_Y \Phi_{1,p}$. If that were the case, one can swap $j_{1,p}$ for one of the other invariants. That misses precisely the locus of $\mathcal{A}_2$ of abelian surfaces admitting an (p, p)-endomorphism to itself, that was studied in [BLS13].

[13] It uses that pI_2 commutes with any other matrix.

There are explicit formulas too for $\Phi_{i,p}$, $i = 2, 3$, given by Lagrange interpolation, that come originally from [GHKRW06, Section 3], for Igusa class polynomials.[14] Set C_p a set of representatives of $\Gamma_0^2(p)\backslash \mathrm{Sp}_4(\mathbb{Z})$. If $\partial_Y \Phi_{1,p}(\boldsymbol{\tau}) \neq 0$, then by Lagrange interpolation, for $i = 2, 3$,

$$Y \mapsto \sum_{M \in C_p} \left(\prod_{M' \neq M \in C_p} \frac{Y - j_{1,p}(M'\boldsymbol{\tau})}{j_{1,p}(M\boldsymbol{\tau}) - j_{1,p}(M'\boldsymbol{\tau})} \right) j_{i,p}(M\boldsymbol{\tau}),$$

is a rational function such that $j_{1,p}(\boldsymbol{\tau}) \mapsto j_{i,p}(\boldsymbol{\tau})$ for $i = 2, 3$, for any $\boldsymbol{\tau}$ such that $\partial_Y \Phi_{1,p}(\boldsymbol{\tau}) \neq 0$. Hence, one can set

$$\Phi_{i,p}(Y) := \sum_{M \in C_p} \left(\prod_{M' \neq M \in C_p} \big(Y - j_{1,p}(M'\boldsymbol{\tau})\big) \right) j_{i,p}(M\boldsymbol{\tau}), \in \mathbb{Q}(j_1, j_2, j_3)[Y],\ i = 2, 3. \quad (8.6)$$

Remark 8.31. *The coefficients of the Siegel modular polynomials are in $\mathbb{Q}(j_1, j_2, j_3)$ instead of $\mathbb{C}(j_1, j_2, j_3)$ by the same q-expansion argument principle as in the elliptic case. But they are honest rational functions with true denominators. This comes from the fact that the Igusa invariants are not defined on $\mathcal{A}_1 \times \mathcal{A}_1$: it can happen that starting with $A_{\boldsymbol{\tau}}$ on $\mathcal{A}_2 \setminus (\mathcal{A}_1 \times \mathcal{A}_1)$, one of the given (p, p)-isogenies ends up on $\mathcal{A}_1 \times \mathcal{A}_1$. That happens precisely when $\mathcal{A}_{\boldsymbol{\tau}}$ belongs to the Humbert surfaces of discriminant p^2.*

Hilbert modular polynomials The construction of the Hilbert modular polynomials works very similarly to the Siegel one, as it is done in [MR20]. We simply point out two interesting phenomena that are special to this case.

- *Splitting of the primes.* The Hilbert modular polynomials are indexed in the primes $\beta \in \mathcal{O}_K^{++}$. Depending on the splitting behavior of the integer prime they lie over, the modular polynomials have different degrees.
- *Symmetric and non-symmetric invariants.* One may choose to consider the covers as symmetric (or non-symmetric) Hilbert surfaces. And, in turn, one may choose corresponding generators of the field of functions to construct the polynomials (called correspondingly symmetric or non-symmetric invariants). The pull-back of the Igusa invariants via the modular embedding from Proposition 2.26 will always be symmetric invariants by [MR20, Lemma 2.9].

With respect to the differentiated behavior regarding splitting, from the start the number of isogenies is different.

Proposition 8.32. *Let $\beta \in \mathcal{O}_K^{++}$ lying over a prime $p \in \mathbb{Z}_{>0}$.*

- *If p is inert in K, then*

$$\#\{\textit{isom. classes of (Hilbert) } \beta\textit{-isogenies}\} =$$
$$\#\{\textit{isom. clases of (Siegel) } p\textit{-isogenies stable under } \mathcal{O}_K\} = p^2 + 1$$

- *If p is split or ramified, then all β-isogenies are cyclic and*

$$\#\{\textit{isom. classes of } \beta\textit{-isogenies}\} = p + 1$$

[14] A generalization of Hilbert class polynomials.

Proof. This is [MR20, Proposition 4.3]. Remark that we are counting as in Proposition 8.22. □

However, the degree of the isogenies do not quite match the expected degree of the modular polynomials in the non-inert case. If one works with a symmetric covers and symmetric invariants, then β-isogenies and $\widetilde{\beta}$-isogenies appear simultaneously an the polynomial is counting both. See [MR20, Example 4.7].

For the inert primes, this behavior does not show up, and in addition we recover our distinguished isogeny $\boldsymbol{z} \mapsto p\boldsymbol{z}$. Hence, for the restriction in Section 8.6, we will consider the Hilbert modular polynomials for inert primes.

8.2 Step 1: Auxiliary functions

We start the proof assuming that $\boldsymbol{q} := (q_1, q_2, q_3) \in \overline{\mathbb{Q}}^3$ and $j_1(\boldsymbol{q}), j_2(\boldsymbol{q}), j_3(\boldsymbol{q}) \in \overline{\mathbb{Q}}$ (either for the Igusa or the Streng-Igusa invariants, but we will work with the latter). By the integrality of the Fourier series of $E_4, E_6, 4\chi_{10}$ and $12\chi_{12}$, we use instead $(4\chi_{10})^{6N}$ and

$$J_1 := 2^{10}\tilde{\jmath}_1,\ J_2 = \frac{1}{2^5}\tilde{\jmath}_2,\ J_3 = 2^{18}\tilde{\jmath}_3. \tag{8.7}$$

That way, once we multiply the denominators with $(4\chi_{10})^{3N}$, we have series with integer coefficients

Remark 8.33. *We do not claim this is the best normalization to preserve integrality.*

We construct *three* auxiliary polynomials $A_k \in \mathbb{Z}[X_1, X_2, X_2^{-1}, X_3, Y]$, for $k = 1, 2, 3$ (where the exponents of $\boldsymbol{X} = (X_1, X_2, X_3)$ range over matrices in $\mathrm{Sym}_2(\mathbb{Z})^{\vee,+}$, positive definite, with $\mathrm{tr} < N$ and the matrix 0_2 for the constant term), one for each invariant, and consequently *three* auxiliary functions $F_k = (4\chi_{10})^{3N} A_k(q_1, q_2, q_3, J_k(\boldsymbol{q}))$. The polynomials depend on N too, but we drop it from the notation. The variable i_k goes with the powers $J_k^{i_k}$, and for the variables $\boldsymbol{q} = q_1, q_2, q_3$, we consider $\boldsymbol{q}^M$ for $0 < M \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+}$ such that $\mathrm{tr}\, M \leq N - 1$, or $M = 0_2$.

$$A_k(q_1, q_2, q_3, J_k) = \sum_{i_k=0}^{N-1} \sum_{M, \mathrm{tr}\, M \leq N-1} b_{i_k,M}^{(k)} \boldsymbol{q}^M J_k^{i_k}.$$

Our auxiliary functions are therefore

$$F_k(\boldsymbol{q}) = (4\chi_{10})^{3N} A_k(q_1, q_2, q_3, J_k(\boldsymbol{q})) = \sum_{i_k=0}^{N-1} \sum_{M, \mathrm{tr}\, M \leq N-1} b_{i_k,M}^{(k)} \boldsymbol{q}^M \left((4\chi_{10})^{3N} J_k^{i_k} \right).$$

Consider the Siegel cusp forms of $(4\chi_{10})^{3N} J_k^{i_k}$. They have weight $30N$, and we have the following bound for the Fourier coefficients. If we set

$$(4\chi_{10})^{3N} J_k^{i_k} = \sum_{0 < T \in \mathrm{Sym}_2(\mathbb{Z})^{\vee,+}} a_{i_k,N}(T) \boldsymbol{q}^T,$$

then by Proposition 8.14 there exists an absolute constant[15] $C_1 > 0$ such that

$$|a_{i_k,N}(T)| \leq C_1^N \det(T)^{15N}.$$

More precisely, if C^* means the constant from Proposition 8.14, then we set $C_1 = \max_{k=1,2,3}((4\chi_{10})^3, C^*((4\chi_{10})^3 J_k$, then for every $0 \leq i_k < N$,

$$C^*((4\chi_{10})^{3N} J_k^{i_k}) = C^*((4\chi_{10})^{3(N-i_k)}((4\chi_{10})^3 J_k)^{i_k}) \leq C_1^{(N-i_k)} C_1^{i_k} = C_1^N.$$

As we are going to order the coefficients by trace, notice that *on fixed trace* $\operatorname{tr}\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} = a + c = t$, the determinant is maximized when $b = 0$ and then $\det = ac$. Then, recall the GM-AM inequality

$$\frac{a+c}{2} \leq ac \leq \left(\frac{a+c}{2}\right)^2, \tag{8.8}$$

hence,

$$|a_{i_k,N}(T)| \leq C_1^N \operatorname{tr}(T)^{30N}.$$

Our auxiliary functions, for $k = 1, 2, 3$, $F_k(\boldsymbol{q}) = (4\chi_{10})^{3N} A_k(q_1, q_2, q_3, J_k(\boldsymbol{q}))$, have then aFourier expansion,

$$F_k(q_1,q_2,q_3) = (4\chi_{10})^{3N} \sum_{i_k=0}^{N-1} \sum_{M,\operatorname{tr} M \leq N-1} b_{i_k,M}^{(k)} \boldsymbol{q}^M J_k^{i_k} = \sum_{i_k=0}^{N-1} \sum_{M,\operatorname{tr} M \leq N-1} b_{i_k,M}^{(k)} \boldsymbol{q}^M (4\chi_{10})^{3N} J_k^{i_k}$$

$$= \sum_{i_k=0}^{N-1} \sum_{M,\operatorname{tr} M \leq N-1} b_{i_k,M}^{(k)} \boldsymbol{q}^M \sum_{0<T\in \operatorname{Sym}_2(\mathbb{Z})^{\vee,+}} a_{i_k,N}(T) \boldsymbol{q}^T$$

$$= \sum_{M,\operatorname{tr} M \leq N-1} \sum_{i_k=0}^{N-1} \sum_{0<T\in \operatorname{Sym}_2(\mathbb{Z})^{\vee,+}} b_{i_k,M}^{(k)} a_{i_k,N}(T) \boldsymbol{q}^{T+M}$$

$$= \sum_{0<T'\in \operatorname{Sym}_2(\mathbb{Z})^{\vee,+}} \sum_{i_k=0}^{N-1} \sum_{\substack{M,\operatorname{tr} M \leq N-1, \\ T'-M>0}} b_{i_k,M}^{(k)} a_{i_k,N}(T'-M) \boldsymbol{q}^{T'},$$

where in the last equation we rewrite the infinite sum in terms of $T' = T + M$. Notice that when $M \neq 0$, $T = T' - M$ being positive definite impose in particular that $\operatorname{tr} M < \operatorname{tr} T'$. Ordering the last expansion by trace, we set equations

$$\sum_{M,\operatorname{tr} M \leq \min(N,\operatorname{tr} T')} \sum_{i_k=0}^{N-1} b_{i_k,M}^{(k)} a_{i_k,N}(T'-M) = 0, \text{ for } 0 \leq \operatorname{tr} T' < L, \text{ for } k = 1,2,3$$

with L a high order of vanishing that we will determine later. Remark that by our construction, $T' = T + M$ is always definite, so the above equations are only for $T' > 0$.

We are going to apply Siegel's lemma Proposition 7.12, hence we set X for the number of variables and Y for the number of equations.

[15] depends of course on the functions, what we mean is that it does not depend on $\boldsymbol{q}$, or any of the parameters of the proof.

Lemma 8.34. *Setting* $L := \lfloor \sqrt[3]{N^4/4} \rfloor$*, it follows that* $X \geq 2Y$*.*

Proof. We need upper bounds on Y and lower bounds on X. Using Lemma 8.13, we have the following upper bound of the number of equations Y:

$$\sum_{l=2}^{L-1} \left(\frac{5}{3} l^2 + (l-1) \right) \leq \frac{5}{3} \sum_{l=1}^{L-1} l^2 + \sum_{l=1}^{L-1} l = \frac{5}{3} \frac{(L-1)L(2L-1)}{6} + \frac{L(L-1)}{2}$$
$$= \frac{L(L-1)}{6} \left(\frac{5}{3}(2L-1) + 3 \right) = \frac{(L-1)L(5L+2)}{9}.$$

For the number of variables X, we have $0 \leq i_k \leq N-1$, and for $\operatorname{tr} M \leq N-1$ (and $M = 0_2$), we have the lower bound:

$$1 + \left(3 + 10 + \sum_{t=4}^{N-1} \left(t^2 + t - 1 \right) \right)$$
$$= 13 + \frac{(N-1)N(2N-1)}{6} - (1+4+9) + \frac{(N-2)(N-1)}{2} - (1+2)$$
$$= -4 + \frac{(N-1)(2N-1)}{6} (N - (N-2)) + \frac{(N-2)(N-1)}{6} (2N - 1 + 3)$$
$$\geq \left(-4 + \frac{(N-1)(2N-1)}{3} \right)_{\geq 0} + \frac{(N-2)(N-1)(2N+2)}{6}$$
$$\geq \frac{(N-2)(N-1)(N+1)}{3}$$

and the upper bound

$$\sum_{t=1}^{N-1} \left(\frac{5}{3} t^2 + (t-1) \right) \leq \frac{(N-1)N(5N+2)}{9}$$

We want to guarantee the relations $X \geq 2Y$, so we can impose

$$\frac{(N-2)(N-1)N(N+1)}{3} \geq \frac{2(L-1)L(5L+2)}{9}.$$

Furthermore, imposing the middle inequality in

$$(N-2)(N-1)N(N+1) \geq_{\text{if } N \geq 3} N^4 \geq 4L^3 \geq_{\text{if } L \geq 3} \frac{2(L-1)L(5L+2)}{3},$$

hence we can set in the end $L := \lfloor \sqrt[3]{N^4/4} \rfloor$, and the lemma is proven. □

The variable k is not taken into consideration here, as we see it as three different applications of Siegel's lemma Proposition 7.12, one for each J_k. Siegel's lemma therefore gives a solution for the coefficients $b_{i_k,M}^{(k)} \in \mathbb{Z}$ of each $A_k(q_1, q_2, q_3, J_k)$ (note that $Y/(Y-X) \leq 1$)

$$L(A_k) = \sum_{i_k, M} |b_{i_k,M}^{(k)}| \leq X (X C_1^N L^{30N})^{\frac{Y}{X-Y}} \leq \frac{(N-1)^2 N^4 (5N+2)^2}{81} L^{30N} C_1^N \leq_{\exists C_2} C_2^N L^{30N}. \tag{8.9}$$

Finally, set $M_k = \operatorname{ord}_0(F_k)$. It is well defined because by Theorem 8.15, each J_k is transcendental over $\mathbb{C}(\boldsymbol{q})$, so F_k cannot be the zero function. By construction, $M_k \geq L$.

8.3 Step 2: Upper bounds

Our strategy involves simplifying to a one variable complex function defined in the unit disc by a power series with polynomial growth, with high vanishing order at zero.

We restricted at the beginning to the Minkowski's domain for $\boldsymbol{\tau}$ of $0 \leq 2\operatorname{Im}(\tau_2) \leq \operatorname{Im}(\tau_1) \leq \operatorname{Im}(\tau_3)$, which for the $\boldsymbol{q}$-variables entails $1 \geq |q_2|^2 \geq |q_1| \geq |q_3| > 0$. We distinguish according to if $|q_2| < 1$ or $|q_2| = 1$. Consider our auxiliary functions $A_k(\boldsymbol{q}) = \sum_{t \geq L} \sum_{\operatorname{tr} T = t} D^{(k)}(T) \boldsymbol{q}^T$, where

$$D^{(k)}(T) = \sum_{l=0}^{\min(N, \operatorname{tr} T)} \sum_{M, \operatorname{tr} M = l} \sum_{i_k=0}^{N-1} b^{(k)}_{i_k, M} a_{i_k, N}(T - M).$$

Case $|q_2| < 1$. Note that (we write $T = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$)

$$|F_k(\boldsymbol{q})| \leq \sum_{t \geq L} \sum_{\operatorname{tr} T = t} |D^{(k)}(T)||q_1|^a |q_2|^b |q_3|^c \leq \sum_{t \geq L} \sum_{\operatorname{tr} T = t} |D^{(k)}(T)||q_2|^{2(a+c)+b}$$

Hence consider the one variable function:

$$G_k(z) = \sum_m \left(\sum_{T,\, 2(a+c)+b=m} |D^{(k)}(T)| \right) z^m.$$

First note that as A_k does not vanish trivially, G_k does not either. Then, note that the set of matrices $0 < T = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \in \operatorname{Sym}_2(\mathbb{Z})^{\vee,+}$ such that $2(a+c)+b = m$ (remark that b can be negative) is finite, by a similar arguments as in Lemma 8.13 that we know outline, so G_k is well-defined.

Indeed, by the proof of Lemma 8.13, in particular for such a matrix T it follows that $|b| < \operatorname{tr} T = a + c$, hence $-(a+c) < b < (a+c)$, or

$$\underbrace{(a+c)}_{\geq 2} + \underbrace{((a+c)+b)}_{\geq 1} = m.$$

Hence for every possible value $l = 2, \ldots, m-1$ for $(a+c)$, can solve $a + c = l$ (therefore $b = m - 2l$). All in all, have an upper bound on the number of matrices T of

$$\sum_{l=2}^{m-1} (l-1) = (m-2)(m-1)/2 \leq m^2.$$

Furthermore, by construction $D^{(k)}(T) = 0$ for all $\operatorname{tr} T < M_k$, which forces $m \geq (a+c)+1 > M_k + 1$, so our function $G_k(z)$ vanishes with order at least $M_k + 1$ at 0. If $T_0 = [a_0, b_0, c_0]$ with $\operatorname{tr} T_0 = a_0 + b_0 = M_k$ is such that $D^{(k)}(T_0) \neq 0$, then we have a non-zero term in the power expansion G_k, and from $-(a_0 + c_0) < b_0 < (a_0 + c_0)$, it follows that

$$M_k + 1 \leq \operatorname{ord}_0 G_k \leq 2M_k - 1.$$

Set $\tilde{M}_k := \operatorname{ord}_0 G_k$. Let us furthermore bound the growth of the coefficients of G_k. By construction $G_k \in \mathbb{Z}[[z]]$.

Remark that $D^{(k)}(T) = \sum_{l=0}^{\min(N,\operatorname{tr} T)} \sum_{M,\operatorname{tr} M=l} \sum_{i_k=0}^{N-1} b^{(k)}_{i_k,M} a_{i_k,N}(T-M)$, and that the coefficient in m of G_k is given by $\sum_{T,\, 2(a+c)+b=m} |D^{(k)}(T)|$. We have the bound

$$\begin{aligned}
|D^{(k)}(T)| \leq & \sum_{l=2}^{\min(N,\operatorname{tr} T)} \sum_{M,\operatorname{tr} M=l} \sum_{i_k=0}^{N-1} \frac{(N-1)N^2(5N+2)}{9} L^{30N} C_1^N C_1^N \operatorname{tr}(T)^{30N} \leq \\
\leq & \sum_{l=2}^{\min(N,\operatorname{tr} T)} (2l-1)(l-1) \operatorname{tr}(T-M)^{30N} \frac{(N-1)N^2(5N+2)}{9} L^{30N} C_1^{2N} \qquad (8.10) \\
= & \frac{(N-1)N^2(5N+2)}{9} L^{30N} C_1^{2N} \sum_{l=2}^{\min(N,\operatorname{tr} T)} (2l-1)(l-1)(\operatorname{tr} T - l)^{30N} \\
\leq_{\exists C_3} & C_3^N L^{30N} \operatorname{tr} T^{30N+2}.
\end{aligned}$$

so for the coefficients of G_k we have a bound

$$\sum_{T,\, 2(a+c)+b=m} |D^{(k)}(T)| \leq m^2 C_3^N L^{30N} \operatorname{tr} T^{30N+2} \leq m^2 C_3^N L^{30N} m^{30N+2} \leq_{\exists C_4} C_4^N L^{30N} m^{30N+4}.$$

Hence, if we consider $G_k(z) z^{-\tilde{M}_k}$, it is holomorphic in $\mathbb{D}_1$ with a power expansion $\sum_{m\geq 0} g^{(k)}(\tilde{M}_k + m) z^m \in \mathbb{Z}[[z]]$ and coefficients bounded by

$$C_4^N L^{30N} (\tilde{M}_k + m)^{30N+4}.$$

By the explicit Schwarz lemma Lemma 7.13 applied to $G_k(z)$ with $K = 30N+4$, we can further more choose N big enough so that $K! \leq N^{30N+4}$, so conclude

$$|G_k(z)| \leq |z|^{\tilde{M}_k} C_4^N L^{30N} N^{30N+4} (\tilde{M}_k + 1)^{30N+4} \frac{1}{(1-|z|)^{30N+5}}$$

using $M_k + 1 \leq \tilde{M}_k \leq 2M_k - 1$, can wrap everything up on

$$|F_k(\boldsymbol{q})| \leq |q_2|^{M_k} \underbrace{C_4^N (L 2 M_k N)^{31N}}_{\leq C_5^N (L M_k N)^{31N}} \frac{1}{(1-|q_2|)^{30N+5}}. \qquad (8.11)$$

This takes care of the case that $|q_2| < 1$ (for a specific choice of $\boldsymbol{\varrho}$, for any r such that $0 < |\varrho_2| < r < 1$, one will choose N big enough for the fraction $1/(1-|\varrho_2|)^*$ gets bounded in terms of M^N). Also, the bounds $L \leq M_k$ and $N^4 < 4(L+1)^3$ are used to bound $(LMN)^{31N}$ in M^{*N}.

Case $|q_2| = 1$.We construct a function in terms of q_1, so use $|q_2| = 1$ and $|q_3| \leq |q_1|$

$$\begin{aligned}
|F_k(\boldsymbol{q})| \leq \sum_{t\geq M_k} \sum_{\operatorname{tr} T = t} |D^{(k)}(T)| |q_1|^a |q_2|^b |q_3|^c \leq \sum_{t\geq M_k} \sum_{\operatorname{tr} T = t} |D^{(k)}(T)| |q_1|^{a+c} \\
= \sum_{t\geq M_k} \left(\sum_{\operatorname{tr} T = t} |D^{(k)}(T)| \right) |q_1|^t,
\end{aligned}$$

so considering the auxiliary function

$$\tilde{G}_k(z) = \sum_{t\geq M_k} g_t^{(k)} z^t = \sum_{t\geq M_k} \left(\sum_{\operatorname{tr} T = t} |D^{(k)}(T)| \right) z^t,$$

we have the bound for the coefficients (using the previous bounds for $|D^k(T)|$ from Equation (8.10),

$$|g_t^{(k)}| = \sum_{\operatorname{tr} T=t} |D^{(k)}(T)| \leq (2t-1)(t-1)C_3^N L^{30N} t^{30N+2} \leq_{\exists C_6} C_6^N L^{30N} t^{30N+4},$$

and the similar analysis allow us to conclude that for $|q_2| = 1$,

$$|F_k(\boldsymbol{q})| \leq |q_1|^{M_k} C_7^N (LM_K N)^{31N} \frac{1}{(1-|q_1|)^{30N+5}}. \tag{8.12}$$

Hence we will control with either $|\varrho_1|$ or $|\varrho_2|$, depending on if $|\varrho_2| < 1$ or not.

Remark 8.35. *This upper bounds admit a generalization to the whole $\mathbb{H}_2$. We use the same ideas as in [BL14, Section 6], to relate our multivariable function with a one dimensional one, as above*

Given any $\boldsymbol{\tau} \in \mathbb{H}_2$, $\operatorname{Im} \boldsymbol{\tau}$ is positive definite, so we can consider:

$$\sigma(\boldsymbol{\tau}) := \sup\{\sigma \in \mathbb{R}_{>0} : \operatorname{Im} \boldsymbol{\tau} - \sigma I_2 \text{ is positive definite}\},$$

alternatively $\sigma(\boldsymbol{\tau})$ is the smallest eigenvalue of $\operatorname{Im} \boldsymbol{\tau}$.

For $\boldsymbol{q}$ as we are considering them, we analogously define $\sigma(\boldsymbol{q}) = \exp(i2\pi\sigma(\boldsymbol{\tau}))$ for any $\boldsymbol{\tau}$ such that $q_k = e^{i2\pi\tau_k}$ for $k = 1, 2, 3$. Remark that it is well-defined, as it only depends on $\operatorname{Im} \boldsymbol{\tau}$ and this one is uniquely recovered as $\operatorname{Im} \tau_k = -\log |q_k|/2\pi$.

Lemma 8.36. *[BL14, Lemma 6.1] For any $\boldsymbol{\tau} \in \mathbb{H}_2$ and any $T \in \operatorname{Sym}_2(\mathbb{Z})^{\vee,+}$ matrix:*

$$|\exp(i2\pi \operatorname{tr}(T\boldsymbol{\tau}))| \leq \exp(i2\pi \operatorname{tr}(T)\sigma(\boldsymbol{\tau})).$$

Hence, for a given power expansion as in our set-up $\sum_{T\in\operatorname{Sym}_2(\mathbb{Z})^{\vee,+}} c(T)\boldsymbol{q}^T$, it follows

$$\left| \sum_{T\in\operatorname{Sym}_2(\mathbb{Z})^{\vee,+}} c(T)\boldsymbol{q}^T \right| \leq \sum_{t>0} \left(\sum_{\operatorname{tr} T=t} |c(T)| \right) \sigma(\boldsymbol{q})^t.$$

Proof. Firs remark that $\operatorname{Im}(\operatorname{tr}(T\boldsymbol{\tau})) = \operatorname{tr}(T \operatorname{Im} \boldsymbol{\tau})$, so $|\exp(i2\pi \operatorname{tr}(T\boldsymbol{\tau})| = \exp(i2\pi \operatorname{tr}(T \operatorname{Im} \boldsymbol{\tau})|$. Now, as T is positive semidefinite, $T(\operatorname{Im}(\boldsymbol{\tau}) - \sigma(\boldsymbol{\tau})I_2)$ has non-negative trace

$$0 \leq \operatorname{tr}(T(\operatorname{Im}(\boldsymbol{\tau}) - \sigma(\boldsymbol{\tau})I_2)) = \operatorname{tr}(T \operatorname{Im}(\boldsymbol{\tau})) - \sigma(\boldsymbol{\tau}) \operatorname{tr}(T)$$

so $\sigma(\boldsymbol{\tau}) \operatorname{tr}(T) \leq \operatorname{tr}(T \operatorname{Im}(\boldsymbol{\tau}))$, and taking exponentials the bound follows. Finally, remark that if $T = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$, then $\exp(i2\pi \operatorname{tr}(T\boldsymbol{\tau})) = q_1^a q_2^b q_3^c = \boldsymbol{q}^T$, and for upper bounds on the absolute value, we again only need $\operatorname{Im}(\boldsymbol{\tau})$, which can be uniquely recovered from $\boldsymbol{q}$. □

Tracing back the same strategy as we did in the case of $|q_2| = 1$, one has a bound of the form

$$|F_k(q_1, q_2, q_3)| \leq |\sigma(\boldsymbol{q})|^{M_k} C_7^N (LM_K N)^{31N} \frac{1}{(1-\sigma(\boldsymbol{q}))^{30N+5}},$$

(and as in the previous case, for a specific choice of $\boldsymbol{\varrho}$, there will be a choice of N such that $1/(1-\sigma(\boldsymbol{\varrho}))$ gets bounded in terms of M^N.

8.4 Step 3: Lower bounds

Our contradiction boils from the assumption of a given $\boldsymbol{\varrho} = (\varrho_1, \varrho_2, \varrho_3) \in \overline{\mathbb{Q}}^3$ with $J_k(\boldsymbol{\varrho}) \in \overline{\mathbb{Q}}$ for $k = 1, 2, 3$. Assume there exists a given prime P and $\boldsymbol{\varrho}^P = (\varrho_1^P, \varrho_2^P, \varrho_3^P)$ such that $F_k(\boldsymbol{\varrho}^P) \neq 0$, for some $k = 1, 2, 3$. In the definition of F_k we have a factor of $\chi_{10}(\boldsymbol{\varrho}^P)^{3N}$, and it is $\chi_{10}(\boldsymbol{\varrho}^P)^{-3N} A(\boldsymbol{\varrho}^P)$ the value that we know it is algebraic. But the first thing that should happen is that on the lower bound on $|A(\boldsymbol{\varrho}^P)|$, $\chi_{10}(\boldsymbol{\varrho}^P)^{3N}$ *should not interfere.*

8.4.1 Lower bounds on χ_{10}

We have the following lower bound on $\chi_{10}(\boldsymbol{\tau})$ for $\boldsymbol{\tau}$ in the Siegel fundamental domain [HP17, Proposition 5.6].

$$|\chi_{10}(\boldsymbol{\tau})| \geqslant c_0 \min\{1, \pi|\tau_2|\}^2 e^{-2\pi(\operatorname{Tr}(\operatorname{Im}\boldsymbol{\tau}) - \operatorname{Im}\tau_2)} \geqslant c_0 \min\{1, \pi|\tau_2|\}^2 e^{-2\pi \operatorname{Tr}\operatorname{Im}\boldsymbol{\tau}} \tag{8.13}$$

with $c_0 = 8 \times 10^{-5}$. It is stated only for the Siegel domain, but one can check that the proof extends to our domain $\mathcal{K}$. Observe that if $\boldsymbol{\tau}$ belongs to the Minkowski's domain, then for p prime such that $p \operatorname{Im}(\tau_1) \geq \frac{\sqrt{3}}{2}$, $p\boldsymbol{\tau} \in \mathcal{K}$.

In the variables $\varrho_1, \varrho_2, \varrho_3$, as $\operatorname{Tr}(\operatorname{Im}\boldsymbol{\tau}) - \operatorname{Im}\tau_2 = \operatorname{Im}\tau_1 + \operatorname{Im}\tau_3 - \operatorname{Im}\tau_2$, the exponential is $|\varrho_1||\varrho_3||\varrho_2|^{-1}$. Remark that $-2\pi \operatorname{Im}\tau_i = \log|\varrho_i|$, for $i = 1, 2, 3$, and it is only $\operatorname{Re}\tau_i$ that it is sensitive to the determination of the complex logarithm. We are also going to distinguish if $|\varrho_2| < 1$ or $|\varrho_2| = 1$.

Case $|\varrho_2| \neq 1$. Set τ_2 the determination of $\log \varrho_2$ such that $|\operatorname{Re}(\tau_2)| \leq 1/2$

$$|\tau_2| \geq \operatorname{Im}\tau_2 = \frac{-\log(|\varrho_2|)}{2\pi},$$

Note that because $\boldsymbol{\varrho}^P$ stays in the fundamental domain (except for translations on $\operatorname{Re}\boldsymbol{\tau}$), if $|\varrho_2| \neq 1$ then

$$\chi_{10}(\boldsymbol{\varrho}^P) \geq \min\left\{1, \frac{(P\log(|\varrho_2|^{-1})}{2}\right\}^2 |\varrho_2|^{-P}|\varrho_1|^P|\varrho_3|^P \geq |\varrho_2|^{-P}|\varrho_1|^P|\varrho_3|^P$$

for $P \geq \frac{2}{\log(|\varrho_2|^{-1})}$, and therefore

$$(\chi_{10}(\boldsymbol{\varrho}^P))^{3N} \geq |\varrho_2|^{-3NP}|\varrho_1|^{3NP}|\varrho_3|^{3NP} \geq |\varrho_1|^{3NP}|\varrho_3|^{3NP} \tag{8.14}$$

Case $|\varrho_2| = 1$. Then $\tau_2 \in \mathbb{R}$ and it is its determination of the angle of $\log \varrho_2$ that we need to pay attention to. As we have assumed $\varrho_2 \in \overline{\mathbb{Q}}$, then $i2\pi\tau_2 = \log \varrho_2$ is a logarithm of an algebraic number.

As we have said before, $P\boldsymbol{\tau}$ belongs to $\mathcal{K}$ except for translations by an integral matrix in $\operatorname{Sym}_2(\mathbb{Z})$. In particular,

$$\chi_{10}(\boldsymbol{\varrho}^P) \geq \min\{1, \pi||P\tau_2||_{\mathbb{Z}}\}^2 |\varrho_1|^P|\varrho_3|^P$$

where $||\cdot||_{\mathbb{Z}}$ means the distance to the nearest integer. Denote $n = n_{\tau_2,S} \in \mathbb{Z}$ said integer, and set $\omega \in \mu_\infty$ such that $\pi i = \log(\omega)$ for our determination of the logarithm. It then follows that $|n| \leq P|\tau_2| + 1 \leq P/2 + 1$ as $|\tau_2| = |\operatorname{Re}(\tau_2)| \leq 1/2$, so:

$$\pi||P\tau_2||_{\mathbb{Z}} = \pi\left|P\frac{\log(\varrho_2)}{2\pi i} - \frac{n2\pi i}{2\pi i}\right| = \frac{1}{2}\left|P\log(\varrho_2) - n\log(\omega)\right|,$$

so we can apply lower bound for linear forms in logarithms of algebraic numbers. By [Bug23, Theorem 1.1], there exists a constant $C_8(\boldsymbol{\varrho}) > 0$ only depending on $[\mathbb{Q}(\varrho_2, \omega) : \mathbb{Q}]$, such that for $B = \max\{3, P, |n|\} \leq \max\{3, P, P/2 + 1\} \leq P$:

$$|P \log(\varrho_2) - n \log(\omega)| \geq B^{-C_8(\boldsymbol{\varrho})^2 \bar{h}(\varrho_2)\bar{h}(\omega)} \geq P^{-C_8(\boldsymbol{\varrho})^2 \bar{h}(\varrho_2)}$$

where $\bar{h}(\cdot) = \max\{1, h(\cdot)\}$ for the logarithmic Weil height. Hence, if $|\varrho_2| = 1$,

$$\chi_{10}(\boldsymbol{\varrho}^P)^{3N} \geq C_9(\boldsymbol{\varrho})^{-6N \log P} |\varrho_1|^{3NP} |\varrho_3|^{3NP}, \tag{8.15}$$

Therefore, combining (8.14) and (8.15), there exists $C_{10}(\boldsymbol{\varrho})$ such that

$$\chi_{10}(\boldsymbol{\varrho}^P)^{3N} \geq \exp(-C_{10}(\boldsymbol{\varrho})NP). \tag{8.16}$$

8.4.1.1 Partial extension to the whole Siegel upper half space

To extend our result to $\boldsymbol{\tau} \in \mathbb{H}_2$ instead of $\boldsymbol{\tau}$ in the Minkowski's domain, the lower bound on $|\chi_{10}|$ need to be extended. First observed that, in general, for $\boldsymbol{\tau} \in \mathbb{H}_2$, let $\widetilde{\boldsymbol{\tau}}$ the reduced element of the orbit $\mathrm{Sp}_4(\mathbb{Z}) \cdot \boldsymbol{\tau}$ that is *Siegel*-reduced. Let also $\gamma_{\boldsymbol{\tau}} \in \mathrm{Sp}_4(\mathbb{Z})$ such that $\widetilde{\boldsymbol{\tau}} = \gamma_{\boldsymbol{\tau}} \boldsymbol{\tau}$. Then, as $\widetilde{\boldsymbol{\tau}}$ is Siegel reduced, for any $\begin{pmatrix} * & * \\ C & D \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z})$, it follows that $|\det(C\widetilde{\boldsymbol{\tau}} + D)| \geq 1$. On the other hand, as χ_{10} is a modular form of weight 10,

$$\chi_{10}(\boldsymbol{\tau}) = \chi_{10}(\gamma_{\boldsymbol{\tau}}^{-1} \widetilde{\boldsymbol{\tau}}) = \det(C\widetilde{\boldsymbol{\tau}} + D)^{10} \chi_{10}(\widetilde{\boldsymbol{\tau}}),$$

for where it follows

$$|\chi_{10}(\boldsymbol{\tau})| \geq |\chi_{10}(\widetilde{\boldsymbol{\tau}})|.$$

Alternatively, it holds for any $\gamma \in \mathrm{Sp}_4(\mathbb{Z})$ and $\boldsymbol{\tau} \in \mathbb{H}_2$ that [Kli90, Proposition 1]

$$\mathrm{Im}(\gamma \boldsymbol{\tau}) = (C\boldsymbol{\tau} + D)^{-1} \,\mathrm{Im}(\boldsymbol{\tau}) (C\overline{\boldsymbol{\tau}} + D)^{-1}$$

where $\bar{\cdot}$ means complex conjugation. Therefore,

$$\det \mathrm{Im}(\gamma \boldsymbol{\tau}) = \frac{1}{|\det(C\boldsymbol{\tau} + D)|^2} \det \mathrm{Im}\, \boldsymbol{\tau},$$

and the automorphic factor is recovered as $\det(\mathrm{Im}\, \boldsymbol{\tau}) / \det(\mathrm{Im}(\gamma \boldsymbol{\tau}))$.

For our bounds to extend, we would not want to work with $\widetilde{P\boldsymbol{\tau}}$, as we do not control the dependence on P, but instead with $P\widetilde{\boldsymbol{\tau}}$, for that we have the bounds that we derived previously. We would need to compare $\chi_{10}(\widetilde{P\boldsymbol{\tau}})$ with $\chi_{10}(P\widetilde{\boldsymbol{\tau}})$, which in general are different elements in the Siegel domain.

As it happens with the one dimensional case, we have the following matrix identity (remark that P is a scalar):

$$\begin{pmatrix} PI_2 & 0 \\ 0 & I_2 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & PB \\ \frac{1}{P}C & D \end{pmatrix} \begin{pmatrix} PI_2 & 0 \\ 0 & I_2 \end{pmatrix}$$

from where it follows that $\widetilde{P\boldsymbol{\tau}} = P\widetilde{\boldsymbol{\tau}}$ if and only if $\gamma_{\boldsymbol{\tau}} \in \Gamma_0^{(2)}(P)$. Hence, in that case we can extend out lower bounds on χ_{10}.

8.4.2 Lower bounds on $A_k(q^P)$

We set $\alpha_k = \alpha_k(P,N) = A_k(\varrho_1^P, \varrho_2^P, \varrho_3^P, J_k(\boldsymbol{\varrho}^P)) \in \overline{\mathbb{Q}}$. We use Liouville inequality to give lower bounds for $|\alpha_k|$

$$\log|\alpha_k| \geq -\deg(\alpha_k)h(\alpha_k),$$

so we have to upper bound $\deg(\alpha_k)$ and $h(\alpha_k)$.

By Theorem 8.29, the three Siegel modular polynomials, for prime level P, for $i = 1,2,3$

$$\Phi_{i,S}(j_1,j_2,j_3,Y_1) \in \mathbb{Q}(j_1,j_2,j_3)[Y_1]$$

satisfy that $\Phi_{1,P}$ has degree $\delta(P) = P^3 + P^2 + P + 1$ on Y_i, by Lemma 8.27, and for given j_1, j_2, j_3, $\Phi_{1,P}$ has as roots the $\tilde{j}_1$ Streng-Igusa invariant of the (P,P)-isogenies. Then for any given choice of root $\tilde{j}_{1,P}$, the other invariants are determined by $\tilde{j}_{i,P} = \Phi_{i,P}(\tilde{j}_{1,P})$ for $i = 2,3$. Remark that our J_k are multiples of $\tilde{j}_k$ as by Equation 8.7.

Assume first now (for the degree computations) that the prime P is chosen so that $\Phi_{i,P}(j_1(\boldsymbol{\varrho}), j_2(\boldsymbol{\varrho}), j_3(\boldsymbol{\varrho}), Y_i)$ are well-defined polynomials (equivalently, $A_{\boldsymbol{\varrho}}$ is not (P,P)-isogenous to a product of elliptic curves with product principal polarization).

By construction $\alpha_k \in \mathbb{Q}(\varrho_1^P, \varrho_2^P, \varrho_3^P, J_k(\boldsymbol{\varrho}^P)) \subset \mathbb{Q}(\varrho_1, \varrho_2, \varrho_3, J_k(\boldsymbol{\varrho}^P))$. Now, by the Siegel modular polynomials give us that $J_1(\boldsymbol{\varrho}^P)$ belongs on a finite extension of $\mathbb{Q}(\tilde{j}_1(\boldsymbol{\varrho}), \tilde{j}_2(\boldsymbol{\varrho}), \tilde{j}_3(\boldsymbol{\varrho})) = \mathbb{Q}(J_1(\boldsymbol{\varrho}), J_2(\boldsymbol{\varrho}), J_3(\boldsymbol{\varrho}))$ of degree bounded by $\delta(P)$. For $k = 2,3$,
$J_k(\boldsymbol{\varrho}^P) \in \mathbb{Q}(J_1(\boldsymbol{\varrho}), J_2(\boldsymbol{\varrho}), J_3(\boldsymbol{\varrho}), J_1(\boldsymbol{\varrho}^P))$, by Theorem 8.29. Hence,

$$\deg(\alpha_k) \leq \underbrace{\mathbb{Q}(\varrho_1,\varrho_2,\varrho_3) : \mathbb{Q}]}_{:=\deg(\boldsymbol{\varrho})}[\mathbb{Q}(J_k(\boldsymbol{\varrho}^P) : \mathbb{Q}] \leq \deg(\boldsymbol{\varrho})\delta(P)\underbrace{[\mathbb{Q}(j_1(\boldsymbol{\varrho}), j_2(\boldsymbol{\varrho}), j_3(\boldsymbol{\varrho})) : \mathbb{Q}]}_{=:\deg(j(\boldsymbol{\varrho}))}. \quad (8.17)$$

Remark 8.37. *We thank Jean Kieffer for the following argument that allow us to extend the bound on the degree of $J_k(\boldsymbol{\varrho}^P)$ even when the specialization of the modular polynomials is not defined.*

Lemma 8.38. *Consider $(A,\lambda_A) \to (B,\lambda_B)$ and (l,l)-isogeny between principally polarised abelian surfaces. Assume A is defined over a field K. Then B is defined over a finite extension of K' satisfying $[K' : K] \leq \delta(l)$.*

Proof. By Proposition 8.21 the kernel of the isogeny is a subgroup $G \subset A[l]$, maximal isotropic with respect to the Weil pairing, and G becomes isomorphic (over $\overline{K}$) to $\left(\mathbb{Z}/l\mathbb{Z}\right)^2$. The total number of such subgroups is $\delta(l)$, by Lemma 8.27.

There is an action of $\mathrm{Gal}(\bar{K}/K)$ on $A[l]$, and this action restrict to the set of maximal isotropic subgroups. Hence G is defined on $K' \supset K$ with $[K' : K]$ equal to the index of $\mathrm{Stab}(G)$ in $\mathrm{Gal}(\bar{K}/K)$, which equals the size of the orbit of $\mathrm{Gal}(\bar{K}/K) \cdot G$. We can therefore upper bound it by the total number of such subgroups, $\delta(l)$. □

Therefore, we can bound by $2\delta(P)$ the degree of the finite extension of $\mathbb{Q}(J_1(\boldsymbol{\varrho}), J_2(\boldsymbol{\varrho}), J_3(\boldsymbol{\varrho}))$ where $J_k(\boldsymbol{\varrho}^P)$ belongs to, and carry out the rest of the argument. The factor of two comes from Mestre's obstruction, as the number field where $A_{\boldsymbol{\varrho}}$ is defined could be a quadratic extensions of $\mathbb{Q}(J_1(\boldsymbol{\varrho}), J_2(\boldsymbol{\varrho}), J_3(\boldsymbol{\varrho}))$.

For the height, recovering our bounds for $L(A_k)$, note that $\boldsymbol{\varrho}_2^{NP} A_k(\boldsymbol{\varrho}^P, J_k(\boldsymbol{\varrho}^P))$ is a polynomial in the four variables, so by standard height results.

$$h(\alpha) = h(\varrho_2^{-NP}\varrho_2^{NP}\alpha) = NPh(\varrho_2) + h(\varrho_2^{NP}\alpha) \leq \log L(A_k) + NS(h(\varrho_1) + 2h(\varrho_2) + h(\varrho_3)) + Nh(J_k(\boldsymbol{\varrho}^P).$$

For $h(J_k(\boldsymbol{\varrho}^P))$, we use the following result of Kieffer [Kie22] of comparison of heights: in [Kie22, Section 5.3] it is given the definition of the *j-height* for ppas defined over $\overline{\mathbb{Q}}$ (so that the Igusa invariants, or the Streng invariants, are algebraic) as

$$h_j(A) = h(j_1(A), j_2(A), j_3(A))$$

as the logarithmic Weil height of the affine truple. Set also $\bar{h_j}(A) = \max\{1, h_j(A)\}$. Then the following is proven, passing via Theta heights and Faltings heights.

Proposition 8.39. *[Kie22, Proposition 5.18] Let A, A' be ppas defined over $\overline{\mathbb{Q}}$, where j_1, j_2, j_3 are well defined and $j_3(A)j_3(A') \neq 0$. Let $d \geq 1$ an integer, if A and A' are linked via an isogeny of degree d, then*

$$\bar{h_j}(A') \leq 8000\bar{h_j}(A) + 1.08 \cdot 10^{11} \log(\bar{h_j}(A)) + 20 \log(d) 1.67 \cdot 10^{12}$$

Hence for our situation: (here $j_k = \tilde{j}_k$ and remark that the degree of the isogeny is P^2)

$$\bar{h}(j_1(\boldsymbol{\varrho}^P), j_2(\boldsymbol{\varrho}^P), j_3(\boldsymbol{\varrho}^P)) \leq C_{11} \left(\bar{h}(j_1(\boldsymbol{\varrho}), j_2(\boldsymbol{\varrho}), j_3(\boldsymbol{\varrho})) + 2 \log P + 1\right).$$

Now, we do have

$$h(j_k(\boldsymbol{\varrho}^P)) \leq h(j_1(\boldsymbol{\varrho}^P), j_2(\boldsymbol{\varrho}^P), j_3(\boldsymbol{\varrho}^P)) \leq \sum_{k=1}^{3} h(j_k(\boldsymbol{\varrho}^P))$$

and

$$\max\{1, h(j_1(\boldsymbol{\varrho}), j_2(\boldsymbol{\varrho}), j_3(\boldsymbol{\varrho}))\} \leq \sum_{k=1}^{3} h(j_k(\boldsymbol{\varrho})) + 1$$

so we have:

$$h(j_k(\boldsymbol{\varrho}^P)) \leq C_{11} \left(\sum_{k=1}^{3} h(j_k(\boldsymbol{\varrho})) + 2 \log P + 2 \right).$$

Finally, $h(J_k(\boldsymbol{\varrho}^P)) \leq C_{12} h(j_k(\boldsymbol{\varrho}^P))$, for our fixed choice of normalization of the Igusa-Streng invariants, so altogether:

$$\begin{aligned} h(\alpha) &\leq \log L(A_k) + NP(h(\varrho_1) + 2h(\varrho_2) + h(\varrho_3)) + h(j_k(\boldsymbol{\varrho}^P)) \\ &\leq \log L(A_k) + 2NP(\sum_{i=1}^{3} h(\varrho_i)) + NC_{13} \left(\sum_{i=1}^{3} h(j_i(q)) + 2 \log P + 2 \right) \end{aligned}$$

hence, setting $C_{14}(\boldsymbol{\varrho}) = \max\{2\sum_{i=1}^{3} h(\varrho_i), C_{11} \sum_{i=1}^{3} h(j_i(\boldsymbol{\varrho})) + 2, 2C_{11}\}$,

$$h(\alpha) \leq \log L(A_k) + C_{14}(\boldsymbol{\varrho})(NP + N \log P)), \tag{8.18}$$

and by our first step in Equation (8.9), $\log L(A_k) \leq 41N \log N$ (for N big enough), hence, combining it with Equations (8.17) and (8.18),

$$\begin{aligned} |\alpha_k| &\geq \exp(-C_{15} P^3 \deg(\boldsymbol{\varrho}) \deg(j(\boldsymbol{\varrho}))N \left(41 \log N + C_{14}(\boldsymbol{\varrho}) \left(P + \log P\right)\right) \\ &\geq \exp\left(-C_{13}(\boldsymbol{\varrho}) P^3 N \left(\log N + P + \log P\right)\right). \end{aligned}$$

The lower bounds on $\chi_{10}(\boldsymbol{\varrho}^P)^{3N}$ in (8.16) are negligible compared to this lower bound, hence

$$|F_k(\boldsymbol{\varrho}^P)| \geq \exp\left(-C_{14}(\boldsymbol{\varrho})P^3N\left(\log N + P + \log P\right)\right). \tag{8.19}$$

We compare with the upper bound on $|F_k(\boldsymbol{q}^P)|$ in (8.11) or (8.12), for N big enough in terms of $\boldsymbol{\varrho}$

$$\begin{aligned} -C_{14}(\boldsymbol{\varrho})P^3N\left(\log N + P + \log P\right) &\leq -PM\log(1/|\varrho_2|) + 32N\log(M), \\ M &\leq C_{15}(\boldsymbol{\varrho})\left(P^2N\left(\log N + P\right)\right) \end{aligned}$$

At this stage, observe that, very roughly, $M \geq L \cong N^{4/3}$, or $N = O(M^{3/4})$, hence we are aiming to a contradiction of the sort $M = O_{\boldsymbol{\varrho}}(M^{3/4}\log M)$.

Hence, the last ingredient are suitable upper bounds of P in terms on N and M.

8.5 Definition of P

Apart from the choices already taken so far for the size of P (that simply impose a lower bound in terms of $\boldsymbol{\varrho}$), we requre

1. $\chi_{10}(\boldsymbol{\varrho}^P) \neq 0$,
2. $A_k(\boldsymbol{\varrho}^P) \neq 0$ for some $k = 1, 2, 3$.

We recall that if $\boldsymbol{\tau} = \begin{pmatrix} \tau_1 & \tau_2 \\ \tau_2 & \tau_3 \end{pmatrix}$ is taken in Minkowski's domain, then χ_{10} vanishes if and only if the off-diagonal element is 0. Taking the multiples $P\boldsymbol{\tau}$, they stay in the Minkowski's domain, up to translation by a matrix in $\mathrm{Sym}^2(\mathbb{Z})$. Therefore, if $\mathrm{Im}(\tau_2) \neq 0$, $\mathrm{Im}(P\tau_2) \neq 0$ and this will be the imaginary part of the reduced coordinate, so $\chi_{10}(\boldsymbol{\varrho}^P) \neq 0$.

If $\mathrm{Im}(\tau_2) = 0$, then as we have assumed $\varrho_2 \in \overline{\mathbb{Q}}$, then either $\tau_2 \in \mathbb{Q}$ or $\tau_2 \in \mathbb{R} \setminus \overline{\mathbb{Q}}$. In the first case, it can happen that $P\tau_2 \in \mathbb{Z}$ and then $\chi_{10}(\boldsymbol{\varrho}^P) = 0$, but necessarily for a unique prime P_0. On the second case $P\tau_2$ is never an integer, and $\chi_{10}(\boldsymbol{\varrho}^P) \neq 0$.

The second condition is that which imposes upper bounds in terms of N. We will see that the argument from Subsection 7.8.2 admits an extension, and marks a limit case of the upper bound required for the transcendence proof to finish. Hence, we define P as the first prime (large enough only in terms of $\boldsymbol{\varrho}$) such that $A_k(\boldsymbol{\varrho}^P) \neq 0$ for some $k = 1, 2, 3$.

8.5.1 Vanishing of $A_k(q^P)$ and first bound on P

We see there are only *finitely many powers* p such that $A_k(\boldsymbol{\varrho}^p, J_k(\boldsymbol{\varrho}^p)) = 0$ for $k = 1, 2, 3$. This argument generalizes the one from Subsection 7.8.2, but it does not give a bound good enough to prove Stéphanois, and neither it is strong enough in our two dimensional setting. In any case, it does prove finiteness of the powers p.

We see A_k as polynomials in one variable with coefficients in $\mathbb{Q}(\varrho_1, \varrho_2, \varrho_3)$. Set $B_{k,p} := A_k(\varrho_1^p, \varrho_2^p, \varrho_3^p, Y_k) \in \mathbb{Q}(\varrho_1, \varrho_2, \varrho_3)[Y_k]$ of degree at most N, and $\alpha_{k,p} = J_k(\boldsymbol{\varrho}^p)$. Roughly speaking, increasing the power p does not change the field $\mathbb{Q}(\varrho_1, \varrho_2, \varrho_3)$, so one deduces a bound on the degree of $\alpha_{k,p}$ *independent* on p; but that is absurd. More precisely, we have $B_{k,p}(\alpha_{k,p}) = 0$ for all $k = 1, 2, 3$, hence

$$[\mathbb{Q}(\boldsymbol{\varrho}, \alpha_{k,p}) : \mathbb{Q}] = [\mathbb{Q}(\alpha_{k,p}) : \mathbb{Q}(\boldsymbol{\varrho})][\mathbb{Q}(\boldsymbol{\varrho}) : \mathbb{Q}] \leq N[\mathbb{Q}(\boldsymbol{\varrho}) : \mathbb{Q}].$$

For convenience, set $\beta_k = \alpha_{k,1}$, it follows then;

$$[\mathbb{Q}(\boldsymbol{\varrho}, \alpha_{k,p}, \beta_1, \beta_2, \beta_3) : \mathbb{Q}] \leq N[\mathbb{Q}(\boldsymbol{\varrho}) : \mathbb{Q}][\mathbb{Q}(\beta_1, \beta_2, \beta_3) : \mathbb{Q}].$$

On the other hand, $\alpha_{k,p}$ are solutions to the modular polynomials $\Phi_{k,p}(\beta_1, \beta_2, \beta_3)[X_k] \in \mathbb{Q}(\beta_1, \beta_2, \beta_3)[X_k]$. Assume first that that the denominators of the modular polynomials of level p do not vanish (so p does not verify the congruence of the previous sections) and that the degree of the specialization on this specific values has the maximal degree $\delta(p) = p^3+p^2+p+1$. More precisely,[16] we assume that for at least one $k = 1, 2, 3$,

$$[\mathbb{Q}(\alpha_{k,p}) : \mathbb{Q}(\beta_1, \beta_2, \beta_3)] = \delta(p),$$

therefore, we bound:

$$\begin{aligned} p^3 \leq \delta(p) = [\mathbb{Q}(\alpha_{k,p}) : \mathbb{Q}(\beta_1, \beta_2, \beta_3)] &\leq [\mathbb{Q}(\boldsymbol{\varrho}, \alpha_{k,p}, \beta_1, \beta_2, \beta_3) : \mathbb{Q}] \\ &\leq N[\mathbb{Q}(\boldsymbol{\varrho}) : \mathbb{Q}][\mathbb{Q}(\beta_1, \beta_2, \beta_3) : \mathbb{Q}] \leq C_{16}(\boldsymbol{\varrho})N. \end{aligned}$$

Remark 8.40. *A bound like this may actually only hold when* $\mathrm{End}(A_{\boldsymbol{\tau}}) = \mathbb{Z}$*, but we have included as a partial generalization of Subsection 7.8.2. As it happens in the Stéphanois proof, we conjecture that this results should not rely on extra information on geometric properties of* $A_{\boldsymbol{\varrho}}$*, but only follow on complex analytic properties of the auxiliary functions* A_k*.*

8.5.2 Extension to the whole fundamental domain

To eventually extend to the whole fundamental domain, we stumble upon another problem: if $\boldsymbol{\tau} \in \mathbb{H}_2$ is not in the Minkowski's domain, it can happen that $\chi_{10}(p\boldsymbol{\tau}) = 0$. Thankfully, that one ends up having a more forgiving solution if we impose the more restrictive

- For any B ppas (p, p)-isogenous to $A_{\boldsymbol{\tau}}$, $\chi_{10}(B) \neq 0$,

and this condition gets rephrased into conditions between the prime p and the Humbert invariant of $A_{\boldsymbol{\tau}}$. This subsection treats how to conjecturaly extend the bounds on P to this situation.

For $\boldsymbol{\varrho} \in \mathbb{H}_2$, we can consider an associated quadratic form $\Delta = \Delta(\boldsymbol{\varrho})$ to $A_{\boldsymbol{\varrho}}$, as in Subsection 6.4.1. Note that $\operatorname{rank} \mathcal{L}_{\boldsymbol{\varrho}} = 3$ if and only if $A_{\boldsymbol{\varrho}}$ is isotypical CM, by Theorem 6.28, and we can discard that case, hence $\operatorname{rank} \Delta \leq 2$. Then, we impose that p^2 is not primitively represented by the quadratic form Δ.

In the cases of rank $0, 1$, either there is no restrictions on the prime, or there is only one. In the case rank two, then write $f = \Delta = ax^2 + bxy + cy^2$, of (negative) discriminant $disc(f) = b^2 - 4ac$. If f is a quadratic form associated to a Shimura curve, then in particular it cannot represent any squares, by Corollary 4.17. Then, we focus on f associated to a modular curve. But as we start with $A_{\boldsymbol{\varrho}} \notin \mathbb{H}_1$, f cannot represent 1.

Suppose there exists (x, y) with $f(x, y) = p^2$. If (x, y) are not coprime, then by dividing by the gcd we get $l^2 f(x', y') = p^2$. But as $f(x', y') \neq 1$, then $l = 1$ and (x, y) were already coprime. That implies that f represents p^2 *primitively.*

[16] The analogous statement for the one dimensional modular polynomial is true as long as we do not evaluate on a CM elliptic curve, but this dos not generalize to higher dimensions, see [BLS13].

That gives general restrictions on p^2. Because f represents p^2 primitively, by [Cox22, Lemma 2.3], there exists g in the same $\mathrm{SL}_2(\mathbb{Z})$-class as f given by

$$g(x,y) = p^2x^2 + b'xy + c'y^2.$$

Hence, the discriminant $\mathrm{disc}(f) = (b')^2 - 4p^2c \equiv (b')^2 \bmod p^2$. Assume now that $p \nmid \mathrm{disc}(f)$, then by Hensel's lemma ([Neu99, Chapter II, Lemma 4.6]), it is equivalent to $\mathrm{disc}(f)$ being a (non-zero) quadratic residue mod p, i.e.

$$\left(\frac{\mathrm{disc}(f)}{p}\right) = 1.$$

using multiplicative properties of the Legendre symbol, and quadratic reciprocity, one ends up with congruence conditions for p, depending only on $\mathrm{disc}(f)$.

Therefore, one ensures non-vanishing of χ_{10} along $\boldsymbol{\varrho}^p$ by congruence conditions on p (that only depend on $\boldsymbol{\varrho}$). The argument from Section 7.7 admits a generalization along primes on arithmetic sequences and arrives to the same contradiction. That leads us to conjecture that, if one would generalize the argument from Section 7.7, it should also admit a generalization to primes under fixed congruences conditions.

8.6 Restriction to the Hilbert surface for $\mathbb{Q}(\sqrt{5})$

These results can be formulated analogously for Hilbert modular forms. We will write it down for $K = \mathbb{Q}(\sqrt{5})$ for simplicity of the exposition. It is a quadratic field with a unit of negative norm given by $\frac{1+\sqrt{5}}{2}$, which is also the standard generator of the ring of integers. That makes the modular embeddings of the Hilbert surface into the Humbert space and the expansion of the Fourier series simplify greatly.

The proof itself is a direct translation to what we have done for the Igusa invariants, so we simply list the necessary results on Hilbert modular forms, but do not repeat the proof itself. We remark that our translation of the proof makes the result intrinsic to Hilbert modular forms. This is part of an ongoing project to extend the last step of the proof in the Hilbert set-up, in the $\mathbb{Q}(\sqrt{5})$ case.

In addition, it is one of the finitely many Humbert surfaces that are rational, see Theorem 2.14, so its field of functions is generated by two algebraically independent Hilbert modular functions, customary called in the literature *Gundlach invariants*.

We are going to change our conventions slightly. The normalized HSR for $\Delta = 5$ by our convention in Chapter 3 is $-\tau_1 + \tau_2 + \tau_3 = 0$. By Proposition 3.16, applying the linear fractional transformation $J = \begin{pmatrix} 0 & -I_2 \\ I_2 & 0 \end{pmatrix} \in \mathrm{Sp}_4(\mathbb{Z})$, then $J\boldsymbol{\tau} = -\boldsymbol{\tau}^{-1}$ solves the Humbert singular relation

$$\tau_1 + \tau_2 - \tau_3 = 0,$$

and we consider this relation from now on.

Likewise, we explicit out the Hilbert modular embedding with respect to this new basis. Set $\varepsilon = \frac{1+\sqrt{5}}{2}$ and $\widetilde{\varepsilon}$ for its Galois conjugate.

Proposition 8.41. *The following maps*

$$\begin{aligned}\mathbb{H}^2 &\to \mathbb{H}_2\\ (z_1,z_2) &\mapsto {}^t\begin{pmatrix}1 & \widetilde{\varepsilon}\\ 1 & \varepsilon\end{pmatrix}\begin{pmatrix}\frac{\varepsilon}{\sqrt{5}}z_1 & 0\\ 0 & \frac{-\widetilde{\varepsilon}}{\sqrt{5}}z_2\end{pmatrix}\begin{pmatrix}1 & \bar{\varepsilon}\\ 1 & \varepsilon\end{pmatrix} = \begin{pmatrix}\frac{\varepsilon z_1-\widetilde{\varepsilon}z_2}{\sqrt{5}} & \frac{z_2-z_1}{\sqrt{5}}\\ \frac{z_2-z_1}{\sqrt{5}} & \frac{-\widetilde{\varepsilon}z_1+\varepsilon z_2}{\sqrt{5}}\end{pmatrix} =: \begin{pmatrix}\tau_1(\boldsymbol{z}) & \tau_2(\boldsymbol{z})\\ \tau_2(\boldsymbol{z}) & \tau_3(\boldsymbol{z})\end{pmatrix}\end{aligned}$$

and the embedding

$$\begin{aligned}\mathrm{SL}_2(\mathcal{O}_K) &\to \mathrm{Sp}_4(\mathbb{Z})\\ \begin{pmatrix}a_1+a_2\widetilde{\varepsilon} & b_1+b_2\widetilde{\varepsilon}\\ c_1+c_2\widetilde{\varepsilon} & d_1+d_2\widetilde{\varepsilon}\end{pmatrix} &\mapsto \begin{pmatrix}a_1 & a_2 & b_2 & (b_1+b_2)\\ a_2 & (a_1+a_2) & (b_1+b_2) & (b_1+2b_2)\\ (c_2-c_1) & c_1 & d_1 & d_2\\ c_1 & c_2 & d_2 & (d_1+d_2)\end{pmatrix}\end{aligned}$$

define a holomorphic map $\mathrm{SL}_2(\mathcal{O}_K)\backslash\mathbb{H}^2 \to \mathrm{Sp}_4(\mathbb{Z})\backslash\mathbb{H}_2$*, that it is finite degree map generically of degree 2.*

Proof. This follows from Theorem 2.23, in the Z-basis $\{1,\widetilde{\varepsilon}\}$ instead of $\{1,\varepsilon\}$. Also by [Run99, Lemma 4], the embedding is determined by the minimal equation of $\widetilde{\varepsilon}$, which is the same as for ε. □

8.6.1 Fourier expansion of a Hilbert modular form

We present the Fourier expansion of a Hilbert modular form. The material comes from [Bru08, Section 1.3]. Our convention was to consider the Hilbert modular group as $\mathrm{SL}_2(\mathcal{O}_K)$. For any $\mu\in\mathcal{O}_K$, and $\boldsymbol{z}\in\mathbb{H}^2$ $f(\boldsymbol{z}+\mu)=f(\boldsymbol{z})$: more precisely, the stabilizer at ∞ of $\mathrm{SL}_2(\mathcal{O}_K)$ is given by

$$\left\{\begin{pmatrix}\varepsilon & \mu\\ 0 & \varepsilon^{-1}\end{pmatrix},\ \mu\in\mathcal{O}_K \text{ and } \epsilon\in\mathcal{O}_K^*\right\}$$

Hence f admits a Fourier expansion of the form:

$$\sum_{\omega\in(\mathcal{O}_K^\vee)^+} a_f(\omega)e^{i2\pi(\omega z_1+\widetilde{\omega}z_2)}, \tag{8.20}$$

where $\mathcal{O}_K^\vee$ means dual with respect to the trace (hence $\mathcal{O}_K^\vee=\partial_K^{-1}$), the superscript $+$ means totally positive elements (or 0), and $\widetilde{\cdot}$ means the non-trivial Galois automorphism of K. We are going to rewrite the Fourier expansion so that it is indexed[17] over $\mathcal{O}_K^+$.

Lemma 8.42. *Let f a holomorphic Hilbert modular form for $\mathbb{Q}(\sqrt{5}$ with Fourier expansion as in Equation 8.20. Then writing $t=\frac{\sqrt{5}}{\varepsilon}\omega$, we can index the sum as*

$$f=\sum_{t=r+s\bar{\varepsilon}\in\mathcal{O}_K^+} a_f(t)e^{i2\pi(tz_1+\widetilde{t}z_2)}$$

[17] Under the convention of the Hilbert modular group $\mathrm{SL}_2(\mathcal{O}_K\oplus\partial_K^{-1})$ the Fourier expansion would be automatically in the form that we want, but we are consistent with the convention that we set.

Proof. For $\omega = \frac{a+b\varepsilon}{\sqrt{5}} \in \partial_K^{-1,+}$:

$$\omega z_1 + \overline{\omega} z_2 = \frac{a+b\varepsilon}{\sqrt{5}} z_1 + \frac{-a-b\overline{\varepsilon}}{\sqrt{5}} = b\left(\frac{\varepsilon z_1 - \overline{\varepsilon} z_2}{\sqrt{5}}\right) - a\left(\frac{z_2 - z_1}{\sqrt{5}}\right)$$

Recall that $\varepsilon = \frac{1+\sqrt{5}}{2}$, is both a generator of $\mathcal{O}_K$ and fundamental unit of norm -1, one can rewrite the Fourier expansions via

$$t = \frac{\sqrt{5}}{\varepsilon}\omega = -\sqrt{5}\bar{\varepsilon}\omega$$

Notice that $\frac{\sqrt{5}}{\varepsilon}$ is totally positive, and for $\omega = \frac{a+b\varepsilon}{\sqrt{5}} \in \partial_K^{-1,+}$ then

$$t = (-\bar{\varepsilon})(a+b\varepsilon) = b - a\bar{\varepsilon}.$$

Hence we write:

$$f = \sum_{t=r+s\bar{\varepsilon} \in \mathcal{O}_K^+} a_f(t) e^{i2\pi(tz_1 + \widetilde{t} z_2)}$$

□

Finally, we want to see then as power series with coefficients in $\mathbb{Z}$:

$$e^{i2\pi(tz_1+\widetilde{t}z_2)} = \left(e^{i2\pi\left(\frac{\varepsilon z_1 - \widetilde{\varepsilon} z_2}{\sqrt{5}}\right)}\right)^r \left(e^{i2\pi\left(\frac{z_2-z_1}{\sqrt{5}}\right)}\right)^s =: q_1^r q_2^s. \tag{8.21}$$

The notation q_1, q_2 is not merely artificial: for $\tau_1(\boldsymbol{z})$ and $\tau_2(\boldsymbol{z})$ as in Proposition 8.41, they correspond to $q_i = e^{i2\pi\tau_i}$ for $i = 1, 2$, as our q-variables for Siegel modular forms.

We have the following restriction for restriction of Siegel modular forms to Hilbert ones, and compatibility of the Fourier expansions.

Proposition 8.43. *Consider f a holomorphic Siegel modular form with the following Fourier expansion*

$$f = a_f(0_2) + \sum_{T \in \mathrm{Sym}_2(\mathbb{Z})^\vee, T \geq 0} a_f(T) \boldsymbol{q}^T,$$

then the pull back to the Hilbert space via the map in Proposition 8.41 is a symmetric Hilbert modular form for $\mathrm{SL}_2(K)$ with Fourier expansion:

$$g = a_g(0) + \sum_{t = p + q\widetilde{\varepsilon} \in \mathcal{O}_K^+} a_g(t) q_1^p q_2^q,$$

where $a_g(0) = a_f(0_2)$, q_1 and q_2 are as in Equation 8.21 and

$$a_g(t) := \sum_{T \in \mathrm{Sym}_2(\mathbb{Z})^\vee,\, T \geq 0,\, T(1,\widetilde{\varepsilon}) = t} a_f(T),$$

where the sum runs over the matrices $T = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$ with $t = a + b\widetilde{\varepsilon} + c\widetilde{\varepsilon}^2$.

Proof. This is [LY11, Proposition 3.2]. □

In this result As $\widetilde{\varepsilon}^2 = 1 + \widetilde{\varepsilon}$, if we write $t = r + s\widetilde{\varepsilon}$ for t as above, for $T = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$, $T(1, \widetilde{\varepsilon}) = r + s\widetilde{\varepsilon}$ is equivalent to

$$\begin{aligned} r &= a + c, \\ s &= b + c. \end{aligned}$$

From this, it follows that the Fourier expansion of g alternatively comes from simple $q_3 = q_1 q_2$ in the Fourier expansion of f. Therefore, in the Fourier expansion of the Hilbert modular form in the "correct" variables q_1, q_2, the overall effect is a restriction of the corresponding function in $\boldsymbol{q}(\mathbb{H}_2)$ to $q_3 = q_1 q_2$.

Remark 8.44. *Notice that the (non-zero) singular terms T in the Fourier expansion of the Siegel modular form have corresponding quadratic forms $Q_T = k(ax + by)^2$ with $k \in \mathbb{Z}_{>0}$ and $a, b \in \mathbb{Z}$, so they represent values of $t = k(a + b\bar{\varepsilon})^2$. However, for T non-singular can also represent multiples of perfect squares. In conclusion, the singular Fourier coefficients are not singled-out as in the Siegel case, except for the zero matrix. That is coherent with the fact that for Hilbert modular forms, being a cusp form means simply vanishing at the zero term.*

For the Fourier expansion of a Hilbert modular form, we can also order them in terms of $r + s\widetilde{\varepsilon}$. One could order them using the field trace $r + 2s$, but it is also natural to order them in terms of r.

Lemma 8.45. *For $t = r + s\widetilde{\varepsilon} \in \mathcal{O}_K$, it follows that $t \gg 0$ if and only if*

$$r \in \mathbb{Z}_{>0}, \text{ and } \widetilde{\varepsilon} r < s < \varepsilon r.$$

Hence, for r fixed, there are $\sqrt{5}r - 1 \leq \lfloor \varepsilon r \rfloor - \lfloor \widetilde{\varepsilon} r \rfloor \leq \sqrt{5}r + 1$ such elements $t \in \mathcal{O}_K^+$.

Furthermore, $N_{K|\mathbb{Q}}(r + s\bar{\varepsilon}) = r^2 - s^2 + rs \leq N_{K|\mathbb{Q}}(r + (r/2)\bar{\varepsilon}) = 5r^2/4$.

Proof. Observe that $t \gg 0$ if $r + s\varepsilon > 0$ and $r + s\widetilde{\varepsilon} > 0$. It follows that $r > 0$: it is clear if $s = 0$, if $s > 0$ then $\widetilde{\varepsilon} s < 0$ so $r > r + s\widetilde{\varepsilon} > 0$, and if $s < 0$ then $\varepsilon s < 0$ and $r > r + s\varepsilon > 0$. The second condition follows from

$$\begin{aligned} r + s\widetilde{\varepsilon} > 0 &\implies r > s\underbrace{(-\widetilde{\varepsilon})}_{>0} \implies s < \frac{-1}{\widetilde{\varepsilon}} r = \varepsilon r, \\ r + s\varepsilon > 0 &\implies r > s\underbrace{(-\varepsilon)}_{<0} \implies s > \frac{-1}{\varepsilon} r = \widetilde{\varepsilon} r. \end{aligned}$$

Hence $s \in \mathbb{Z} \cap (\widetilde{\varepsilon} r, \varepsilon r)$ and there are exactly $\lfloor \varepsilon r \rfloor - \lfloor \widetilde{\varepsilon} r \rfloor$ possibilities for s. The lower and upper bounds follows from the standard inequalities $\lfloor x \rfloor \leq x \leq \lfloor x \rfloor + 1$. For the bounds in the norm, we see that $(r + s\widetilde{\varepsilon})(r + s\varepsilon) = r^2 - s^2 + rs$. As a function of s on the interval $(\widetilde{\varepsilon} r, \varepsilon r)$ it is a concave parabola, hence it is maximum is at its vertex $s = r/2$. □

Finally, by Theorem 2.21, we have the suitable analogous bounds for the Fourier coefficients as in Proposition 8.14 in terms of the filed norm $N_{K/\mathbb{Q}}$.

8.6.2 Gundlach invariants

This comes from the presentation [LY11, section 4], or [MR20, section 2.2], from which we use the conventions for the definition of the Gundlach invariants, of the original results from [Gun63], and details can be found in [Nag83].

We have the Hilbert Eisenstein series of even weight $k \geq 2$ with Fourier expansion:

$$G_k(z) = 1 + \sum_{t=a+b\widetilde{\varepsilon}\in\mathcal{O}_K^+} b_k(t) q_1^a q_2^b$$

where

$$b_k(t) = \kappa_k \sum_{(\mu)\supset(t)} |\mathcal{O}_K/\mu\mathcal{O}_K|^{k-1}, \quad \kappa_k = \frac{(2\pi)^{2k}\sqrt{5}}{(k-1)!^2 5^k \zeta_K(k)},$$

and ζ_K is the Dirichlet zeta function for the field K. This coefficient is a rational number, and here are their values for low values of k, from [LY11, End of page 944, after Equation (4.2)]

$$\kappa_k = \begin{cases} 2^3 \cdot 3 \cdot 5 & \text{if } k = 2, \\ 2^4 \cdot 3 \cdot 5 & \text{if } k = 4, \\ \frac{1}{67} \cdot 2^3 \cdot 3^2 \cdot 5 \cdot 7 & \text{if } k = 6, \\ \frac{1}{412751} \cdot 2^3 \cdot 3 \cdot 5^2 \cdot 11 & \text{if } k = 10. \end{cases}$$

Furthermore, setting

$$\begin{aligned} \theta_6 &= -\frac{67}{2^5 3^3 5^2}\left(G_6 - G_2^3\right) \\ \theta_{10} &= 2^{-10}3^{-5}5^{-5}7^{-1}\left(412751 G_{10} - 5\cdot 67 \cdot 2293 G_2^2 G_6 + 2^2\cdot 3\cdot 7\cdot 4231 G_2^5\right) \\ \theta_{12} &= 2^{-2}\left(\theta_6^2 - G_2\theta_{10}\right), \end{aligned}$$

which are primitive integral symmetric Hilbert modular forms, we define the Gundlach invariants as:

$$J_1^H = \frac{G_2^5}{\theta_{10}} \text{ and } J_2^H = \frac{G_2^2 G_6}{\theta_{10}}.$$

Remark 8.46. *Our strategy as designed translates directly to the invariants from above, as we have them as quotients of modular forms for which we have bounds for the (integral) Fourier coefficients, and a common denominator given by a cusp form. We can also relate the Gundlach invariants to the pull-back of the Igusa invariants.*

The functions θ_i can be expressed in term of pullbacks (from Proposition 2.26) of Siegel modular forms and Eisenstein series. From [LY11, Theorem 4.4],

$$\begin{aligned} \phi^* E_4 &= G_2^2, \\ \phi^* E_6 &= -\frac{42}{25}G_2^3 + \frac{67}{25}G_6 = G_2^3 - 2^5 3^3 \theta_6, \\ -4\phi^*\chi_{10} &= \theta_{10}, \\ 12\phi^*\chi_{12} &= 3\theta_6^2 - 2G_2\theta_{10}. \end{aligned}$$

and by [LY11, Proposition 4.5] in terms of [MR20, Corollary A.13],

$$\phi^* j_1 = 8 J_1^H \left(\frac{3(J_2^H)^2}{J_1^H - 2}\right)^5, \ \phi^* j_2 = \frac{1}{2} J_1^H \left(\frac{3(J_2^H)^2}{J_1^H - 2}\right)^3,$$
$$\phi^* j_3 = \frac{1}{2^3} J_1^H \left(\frac{3(J_2^H)^2}{J_1^H - 2}\right)^2 \left(4\frac{(J_2^H)^2}{J_1^H} + 2^5 3^2 \frac{J_2^H}{J_1^H - 3}\right).$$

8.6.3 Isogenies that respect real multiplication and pullback of $\mathcal{K}$

We would want the Hilbert modular polynomials to parametrize, among others, isogenies given by $(z_1, z_2) \mapsto (pz_1, pz_2)$. If p is inert in K (i.e. for $p = 2$ and $\left(\frac{p}{5}\right) = -1$, equivalently $p \equiv 2, 3 \pmod 5$), then we can consider p-isogenies as in [MR20] and their modular polynomials. Then, if p is inert in K, their Hilbert p-isogenies in bijection with the Siegel (p, p)-isogenies that respect the endomorphism ring $\mathcal{O}_K$, by Proposition 8.32.

We notice that the normalized Humbert singular relation $\tau_1 + \tau_2 = \tau_3$ is invariant under multiplication by p (and that multiplying by p in $\boldsymbol{\tau}$ is multiplying by p in $\boldsymbol{z}$), by Proposition 8.26. That means that $p\boldsymbol{\tau}$ still belongs to $\mathcal{H}_5$, so this isogeny respect the endomorphism ring. In conclusion, one of the p-isogenies considered in their Hilbert modular polynomials gets written as $(z_1, z_2) \to (pz_1, pz_2)$.

Our choice of invariants for $\mathcal{H}_5$ have as denominators the pull-back of χ_{10}, hence we apply the same considerations as to restricting to a subdomain of $\mathbb{H}^2$ such that θ_{10} only vanishes when $z_2 = z_1$. A very straightforward way is to simply pull back the intersection of $\mathcal{K}$ with $\tau_3 = \tau_1 + \tau_2$.

If we impose $\operatorname{Im}(z_2) \geq \operatorname{Im}(z_1)$, then $\operatorname{Im}(\tau_2) \geq 0$ and as $\tau_1 + \tau_2 = \tau_3$, already have $\operatorname{Im}(\tau_1) \leq \operatorname{Im}(\tau_3)$. Hence one simply requires additionally $2\operatorname{Im}(\tau_2) \leq \operatorname{Im}(\tau_1)$, and

$$\begin{aligned}
2\frac{\operatorname{Im} z_2 - \operatorname{Im} z_1}{\sqrt{5}} &\leq \frac{\varepsilon \operatorname{Im} z_1 - \widetilde{\varepsilon} \operatorname{Im} z_2}{\sqrt{5}} \\
2\left(\operatorname{Im} z_2 - \operatorname{Im} z_1\right) &\leq -\frac{1}{2}\left(\operatorname{Im} z_2 - \operatorname{Im} z_1\right) + \frac{\sqrt{5}}{2}\left(\operatorname{Im} z_2 + \operatorname{Im} z_1\right) \\
5\left(\operatorname{Im} z_2 - \operatorname{Im} z_1\right) &\leq \sqrt{5}\left(\operatorname{Im} z_2 + \operatorname{Im} z_1\right) \\
\left(\sqrt{5} - 1\right)\operatorname{Im} z_2 &\leq \left(\sqrt{5} + 1\right)\operatorname{Im}(z_1) \\
\operatorname{Im} z_2 &\leq \frac{\sqrt{5} + 1}{\sqrt{5} - 1} \operatorname{Im} z_1 = (1 + \varepsilon) \operatorname{Im} z_1
\end{aligned}$$

One can invert the embedding and get

$$z_1 = \tau_1 + \widetilde{\varepsilon}\tau_2, \ z_2 = t_1 - \varepsilon\tau_2$$

and one deduces linear conditions $\operatorname{Re}(z_1), \operatorname{Re}(z_2)$ from $|\operatorname{Re}\tau_i| \leq 1/2$. Therefore, on the subdomain of $\mathbb{H}^2$ defined by the above conditions on $\operatorname{Re}(z_i)$,[18] together with

$$\operatorname{Im} z_1 \leq \operatorname{Im} z_2 \leq (1 + \varepsilon) \operatorname{Im}(z_1),$$

θ_{10} only vanishes if $z_1 = z_2$.

[18] probably a better solution would be to adapt the proof from [Kli90, Proposition 2] to the Hilbert theta series.

Appendix: On deducing lower bounds on χ_{10} in the whole $\mathbb{H}_2$

This is a continuation of Subsection 8.4.1.1. As a partial result to our goal of comparing $\chi_{10}(P\widetilde{\boldsymbol{\tau}})$ and $\chi_{10}(\widetilde{P\boldsymbol{\tau}})$, we have derived a comparison between $\det(\operatorname{Im}\widetilde{P\boldsymbol{\tau}})$ and $\det(P\operatorname{Im}\widetilde{\boldsymbol{\tau}})$, that may be of independent interest. It comes as a generalization of the following result from [Paz19a, Lemma 2.4], which states that for two isogenous elliptic curves $\phi : E_1 \to E_2$ with $E_i(\mathbb{C}) = \mathbb{C}/(\mathbb{Z} + \tau_i\mathbb{Z})$ for $i = 1, 2$, with τ_i in the fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$, it follows that

$$\frac{1}{\deg\phi} \leq \frac{\operatorname{Im}\tau_2}{\operatorname{Im}\tau_1} \leq \deg\phi.$$

We can generalize it to principally polarized abelian surfaces.

Proposition 8.47. *Let $A_1, A_2 \in \mathcal{A}_2$, with period matrices $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in \mathbb{H}_2$ that are Siegel-reduced. Assume there exists a (P, P)-isogeny between them*

Assume furthermore than $\det\operatorname{Im}\boldsymbol{\tau}_i \geq 1$ and set $y^{(i)} = (\operatorname{Im}\boldsymbol{\tau}_i)_{1,1}$. Then

$$\frac{1}{P^2} \leq \frac{y^{(1)}(\det\operatorname{Im}\boldsymbol{\tau}_1)^{-1}}{y^{(2)}(\det\operatorname{Im}\boldsymbol{\tau}_2)^{-1}} \leq P^2,$$

which we prove below that implies

$$\frac{3}{4}\frac{1}{P^2}\frac{1}{\sqrt{\det\operatorname{Im}\boldsymbol{\tau}_1}} \leq \frac{\det\operatorname{Im}\boldsymbol{\tau}_2}{\det\operatorname{Im}\boldsymbol{\tau}_1} \leq \frac{4}{3}P^2\sqrt{\det\operatorname{Im}\boldsymbol{\tau}_2}$$

Proof. We are going to relate $\det\operatorname{Im}\boldsymbol{\tau}_i$ with the *injectivity diameter* of A_i (see [GR14b, Section 2.4], [Aut13, Section 2] and [Paz19a, section 2.4]). Consider the positive definite hermitian form $H_i = z^t(\operatorname{Im}\boldsymbol{\tau})^{-1}\bar{w}$ for $z, w \in \mathbb{C}^2$ (identified with $T_{A_i}(\mathbb{C})$). Consider H_i restricted to the period lattice $\Lambda_{\boldsymbol{\tau}_i} = \mathbb{Z}^2 + \boldsymbol{\tau}_i\mathbb{Z}^2$. Then we define the injectivity diameter $\rho(A_i, H_i)$ as its first minimum, i.e.

$$\rho(A, H) = \min_{z\in\Lambda\setminus\{0\}}\{\sqrt{H(z,z)}\}$$

for Λ the period lattice of A, in our case given by $\mathbb{Z}^2+\boldsymbol{\tau}\mathbb{Z}^2$. Equivalently, $\rho(A, H)$ is the radius of the largest ball centered at zero where the exponential map $\exp : T_0A \to A$ is injective.

Alternatively, given a polarization μ on an abelian variety, we write $\rho(A, \mu) = \rho(A, H_\mu)$ for H_μ the induced hermitian form.

Let us set $\phi : (A_1, \mu_1) \to (A_2, \mu_2)$ the (P, P)-isogeny, it then follows that $H_{\phi^*\mu_2} = P^2H_{\mu_1}$, hence $\rho(A_1, \phi^*\mu_2) = P\rho(A_1, \mu_1)$.

From [GR14b, Lemma 3.4], we have the following relation

$$\rho(A_2, \mu_2) \leq \rho(A_1, \phi^*\mu_2) \leq \deg(\phi)\rho(A_2, \mu_2),$$

and as $\deg\phi = P^2$, combining both we have

$$\rho(A_2, \mu_2) \leq P\rho(A_1, \mu_1) \leq P^2\rho(A_2, \mu_2). \tag{8.22}$$

We relate $\rho(A_i, \mu_i)$ to $\operatorname{Im}\boldsymbol{\tau}_i$. Let us provisionally drop the subindex i of the notation. As $\boldsymbol{\tau}$ is Siegel reduced, in particular $Y = \operatorname{Im}\boldsymbol{\tau}$ is Minkowski reduced. That implies that the diagonal elements correspond to the successive minima of the positive definite quadratic

form induced by Y_i on $\mathbb{Z}^2$. Writing λ_1 for the first minimum of Y^{-1}, and noticing that it corresponds to the restriction to $\mathbb{Z}^2 \subset \mathbb{Z}^2 + \mathbb{Z}^2\boldsymbol{\tau}$, it follows that

$$\rho^2 \leq \lambda_1.$$

Now we follow the proof of [Aut13, Lemme 3.2], which states $\min(\rho^2(A, L), \sqrt{3}/2) = \min(\lambda_1, \sqrt{3}/2)$, so if λ_1 is large enough, there is equality above. We give the details of the proof here to specialize further in dimension two. For $a + b\boldsymbol{\tau} \in \Lambda_{\boldsymbol{\tau}}$, with $a, b \in \mathbb{Z}^2$ and writing $\boldsymbol{\tau} = X + Yi$:

$$\begin{aligned}(a + (X - iY)b)^t Y^{-1}(a + (X + iY)b) &= a^t Y^{-1} a + b^t X Y^{-1} a + a^t Y^{-1} X b + b^t (X Y^{-1} X + Y) b \\ &= b^t Y b + \underbrace{(a + Xb)^t Y^{-1}(a + Xb)}_{\geq 0}\end{aligned}$$

Then if $b \neq 0$, $||a + \boldsymbol{\tau} b|| \geq b^t Y b \geq \lambda_1(Y)$, and if $b = 0$, $||a + \boldsymbol{\tau} b|| \geq a^t Y^{-1} a \geq \lambda_1(Y^{-1})$, hence $\rho^2 \geq \min\{\lambda_1(Y), \lambda_1(Y^{-1})\}$.

Now we use explicitly that we are in dimension two: if $Y = \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix}$, with associated quadratic form $f_Y(u, v) = y_1 u^2 + 2y_2 uv + y_3 v^2$, then $Y^{-1} = \frac{1}{\det Y}\begin{pmatrix} y_3 & -y_2 \\ -y_2 & y_1 \end{pmatrix}$, with quadratic form $f_{Y^{-1}}(u, v) = (y_3 u^2 - 2y_2 uv + y_1 v^2)/\det(Y)$. It follows that $\min_{\mathbb{Z}^2} f_Y = \min_{\mathbb{Z}^2}(\det(Y) f_{Y^{-1}})$, so

$$\lambda_1(Y^{-1}) = \frac{\lambda_1(Y)}{\det Y},$$

in dimension two. As Y is Minkowski reduced, $\lambda_1(Y) = y_1$. Hence we have:

$$\frac{y_1}{\det Y} \geq \rho^2 \geq y_1 \min\left\{1, \frac{1}{\det Y}\right\} = \frac{y_1}{\det Y} \tag{8.23}$$

as we assumed that $\det Y \geq 1$.

We can give lower and upper bounds on $y_1/\det Y$ via the Hermite constant in dimension two as follows. First, as Y is Siegel reduced, $y_1 \geq \sqrt{3}/2$ so we first have $y_1/\det Y \geq \sqrt{3}/2 \det Y$. For the upper bounds, $\frac{1}{\sqrt{\det Y}} Y$ has determinant one. Recall that the Hermite constant in dimension n is defined as the longest shortest vector (the maximum of the first minimum) for unimodular lattice in dimension n. As the Hermite constant in dimension two is $\frac{2}{\sqrt{3}}$, it follows:

$$\frac{2}{\sqrt{3}} \frac{1}{\sqrt{\det Y}} \geq \frac{1}{\sqrt{\det Y}} \frac{y_1}{\sqrt{\det Y}} = \frac{y_1}{\det Y} \geq \frac{\sqrt{3}}{2} \frac{1}{\det Y} \tag{8.24}$$

Coming back to our matrices $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2$, by (8.23) and (8.22) above,

$$\frac{1}{P^2} \leq \frac{y^{(1)}(\det \operatorname{Im} \boldsymbol{\tau}_1)^{-1}}{y^{(2)}(\det \operatorname{Im} \boldsymbol{\tau}_2)^{-1}} \leq P^2.$$

which is our first statement. For the second, use (8.24) to bound $\frac{y^{(1)}(\det \operatorname{Im} \boldsymbol{\tau}_1)^{-1}}{y^{(2)}(\det \operatorname{Im} \boldsymbol{\tau}_2)^{-1}}$ and get

$$\begin{aligned}\frac{3}{4} \frac{\sqrt{\det \operatorname{Im} \boldsymbol{\tau}_2}}{\det \operatorname{Im} \boldsymbol{\tau}_1} &\leq P^2, \\ \frac{4}{3} \frac{\det \operatorname{Im} \boldsymbol{\tau}_2}{\sqrt{\det \operatorname{Im} \boldsymbol{\tau}_1}} &\geq \frac{1}{P^2}\end{aligned}$$

from where we deduce our second statement. □

Fin de la temporada, adiós queridos,
guardaos bien que viene el frío,
ya nos veremos por allá.

Puntos suspensivos, Vetusta Morla

Bibliography

[AB04] Montserrat Alsina and Pilar Bayer. *Quaternion orders, quadratic forms, and Shimura curves*. Vol. 22. CRM Monograph Series. American Mathematical Society, Providence, RI, 2004, pp. xvi+196. ISBN: 0-8218-3359-6. DOI: 10.1090/crmm/022. URL: https://doi.org/10.1090/crmm/022.

[And04] Yves André. *Une introduction aux motifs (motifs purs, motifs mixtes, périodes)*. Vol. 17. Panoramas et Synthèses [Panoramas and Syntheses]. Société Mathématique de France, Paris, 2004, pp. xii+261. ISBN: 2-85629-164-3.

[Aut03] Pascal Autissier. "Hauteur des correspondances de Hecke". In: *Bull. Soc. Math. France* 131.3 (2003), pp. 421–433. ISSN: 0037-9484,2102-622X. DOI: 10.24033/bsmf.2449. URL: https://doi.org/10.24033/bsmf.2449.

[Aut13] Pascal Autissier. "Un lemme matriciel effectif". In: *Math. Z.* 273.1-2 (2013), pp. 355–361. ISSN: 0025-5874,1432-1823. DOI: 10.1007/s00209-012-1008-x. URL: https://doi.org/10.1007/s00209-012-1008-x.

[BCCK24] Raymond van Bommel, Shiva Chidambaram, Edgar Costa, and Jean Kieffer. "Computing isogeny classes of typical principally polarized abelian surfaces over the rationals". In: *LuCaNT: LMFDB, computation, and number theory*. Vol. 796. Contemp. Math. Amer. Math. Soc., [Providence], RI, [2024] ©2024, pp. 187–214. ISBN: 978-1-4704-7260-3. DOI: 10.1090/conm/796/16002. URL: https://doi.org/10.1090/conm/796/16002.

[BDGP96] Katia Barré-Sirieix, Guy Diaz, François Gramain, and Georges Philibert. "Une preuve de la conjecture de Mahler-Manin". In: *Invent. Math.* 124.1-3 (1996), pp. 1–9. ISSN: 0020-9910,1432-1297. DOI: 10.1007/s002220050044. URL: https://doi.org/10.1007/s002220050044.

[Ber97] Daniel Bertrand. "Theta functions and transcendence". In: vol. 1. 4. International Symposium on Number Theory (Madras, 1996). 1997, pp. 339–350. DOI: 10.1023/A:1009749608672. URL: https://doi.org/10.1023/A:1009749608672.

[BG06] Enrico Bombieri and Walter Gubler. *Heights in Diophantine geometry*. Vol. 4. New Mathematical Monographs. Cambridge University Press, Cambridge, 2006, pp. xvi+652. ISBN: 978-0-521-84615-8; 0-521-84615-3. DOI: 10.1017/CBO9780511542879. URL: https://doi.org/10.1017/CBO9780511542879.

[BG08] Srinath Baba and Håkan Granath. "Genus 2 curves with quaternionic multiplication". In: *Canad. J. Math.* 60.4 (2008), pp. 734–757. ISSN: 0008-414X,1496-4279. DOI: 10.4153/CJM-2008-033-7. URL: https://doi.org/10.4153/CJM-2008-033-7.

[BGP25] Florian Breuer, Desirée Gijón Gómez, and Fabien Pazuki. "Explicit bounds on the coefficients of modular polynomials and the size of $X_0(N)$". In: *Proc. Lond. Math. Soc. (3)* 130.1 (2025), Paper No. e70020, 25. ISSN: 0024-6115,1460-244X. DOI: 10.1112/plms.70020. URL: https://doi.org/10.1112/plms.70020.

[BL04] Christina Birkenhake and Herbert Lange. *Complex abelian varieties*. Second. Vol. 302. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 2004, pp. xii+635. ISBN: 3-540-20488-1. DOI: 10.1007/978-3-662-06307-1. URL: https://doi.org/10.1007/978-3-662-06307-1.

[BL09] Reinier Bröker and Kristin Lauter. "Modular polynomials for genus 2". In: *LMS J. Comput. Math.* 12 (2009), pp. 326–339. ISSN: 1461-1570. DOI: 10.1112/S1461157000001546. URL: https://doi.org/10.1112/S1461157000001546.

[BL14] Reinier Bröker and Kristin Lauter. "Evaluating Igusa functions". In: *Math. Comp.* 83.290 (2014), pp. 2977–2999. ISSN: 0025-5718,1088-6842. DOI: 10.1090/S0025-5718-2014-02816-0. URL: https://doi.org/10.1090/S0025-5718-2014-02816-0.

[BL99] Christina Birkenhake and Herbert Lange. *Complex tori*. Vol. 177. Progress in Mathematics. Birkhäuser Boston, Inc., Boston, MA, 1999, pp. xvi+251. ISBN: 0-8176-4103-3. DOI: 10.1007/978-1-4612-1566-0. URL: https://doi.org/10.1007/978-1-4612-1566-0.

[BLS13] Reinier Bröker, Kristin Lauter, and Marco Streng. "Abelian surfaces admitting an (l, l)-endomorphism". In: *J. Algebra* 394 (2013), pp. 374–396. ISSN: 0021-8693,1090-266X. DOI: 10.1016/j.jalgebra.2013.07.011. URL: https://doi.org/10.1016/j.jalgebra.2013.07.011.

[BMM90] Jean-Benoît Bost, Jean-François Mestre, and Laurent Moret-Bailly. "Sur le calcul explicite des "classes de Chern" des surfaces arithmétiques de genre 2". In: *Astérisque* 183. Séminaire sur les Pinceaux de Courbes Elliptiques (Paris, 1988). 1990, pp. 69–105.

[BMOR18] Michael A. Bennett, Greg Martin, Kevin O'Bryant, and Andrew Rechnitzer. "Explicit bounds for primes in arithmetic progressions". In: *Illinois J. Math.* 62.1-4 (2018), pp. 427–532. ISSN: 0019-2082,1945-6581. DOI: 10.1215/ijm/1552442669. URL: https://doi.org/10.1215/ijm/1552442669.

[BP24] Florian Breuer and Fabien Pazuki. "Explicit bounds on the coefficients of modular polynomials for the elliptic j-invariant". In: *Proc. Amer. Math. Soc. Ser. B* 11 (2024), pp. 277–286. ISSN: 2330-1511. DOI: 10.1090/bproc/179. URL: https://doi.org/10.1090/bproc/179.

[Bru08] Jan Hendrik Bruinier. "Hilbert modular forms and their applications". In: *The 1-2-3 of modular forms*. Universitext. Springer, Berlin, 2008, pp. 105–179. ISBN: 978-3-540-74117-6. DOI: 10.1007/978-3-540-74119-0_2. URL: https://doi.org/10.1007/978-3-540-74119-0_2.

[BS10] Reinier Bröker and Andrew V. Sutherland. "An explicit height bound for the classical modular polynomial". In: *Ramanujan J.* 22.3 (2010), pp. 293–313. ISSN: 1382-4090,1572-9303. DOI: 10.1007/s11139-010-9231-8. URL: https://doi.org/10.1007/s11139-010-9231-8.

[Bug23] Yann Bugeaud. "B'". In: *Publ. Math. Debrecen* 103.3-4 (2023), pp. 499–533. ISSN: 0033-3883,2064-2849. DOI: 10.5486/pmd.2023.9661. URL: https://doi.org/10.5486/pmd.2023.9661.

[BW03] Christina Birkenhake and Hannes Wilhelm. "Humbert surfaces and the Kummer plane". In: *Trans. Amer. Math. Soc.* 355.5 (2003), pp. 1819–1841. ISSN: 0002-9947,1088-6850. DOI: 10.1090/S0002-9947-03-03238-0. URL: https://doi.org/10.1090/S0002-9947-03-03238-0.

[BZ01] Daniel Bertrand and Wadim Zudilin. "Derivatives of Siegel modular forms, and exponential functions". In: *Izv. Ross. Akad. Nauk Ser. Mat.* 65.4 (2001), pp. 21–34. ISSN: 1607-0046,2587-5906. DOI: 10.1070/IM2001v065n04ABEH000345. URL: https://doi.org/10.1070/IM2001v065n04ABEH000345.

[BZ20] François Brunault and Wadim Zudilin. *Many variations of Mahler measures—a lasting symphony*. Vol. 28. Australian Mathematical Society Lecture Series. Cambridge University Press, Cambridge, 2020, pp. xv+167. ISBN: 978-1-108-79445-9.

[CF96] J. W. S. Cassels and E. V. Flynn. *Prolegomena to a middlebrow arithmetic of curves of genus* 2. Vol. 230. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 1996, pp. xiv+219. ISBN: 0-521-48370-0. DOI: 10.1017/CBO9780511526084. URL: https://doi.org/10.1017/CBO9780511526084.

[CFN20] Guy Casale, James Freitag, and Joel Nagloo. "Ax-Lindemann-Weierstrass with derivatives and the genus 0 Fuchsian groups". In: *Ann. of Math. (2)* 192.3 (2020), pp. 721–765. ISSN: 0003-486X,1939-8980. DOI: 10.4007/annals.2020.192.3.2. URL: https://doi.org/10.4007/annals.2020.192.3.2.

[Cla03] Pete L. Clark. *Rational points on Atkin-Lehner quotients of Shimura curves*. Thesis (Ph.D.)–Harvard University. ProQuest LLC, Ann Arbor, MI, 2003, p. 184. ISBN: 978-0496-39246-9. URL: http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:3091537.

[Coh84] Paula Cohen. "On the coefficients of the transformation polynomials for the elliptic modular function". In: *Math. Proc. Cambridge Philos. Soc.* 95.3 (1984), pp. 389–402. ISSN: 0305-0041,1469-8064. DOI: 10.1017/S0305004100061697. URL: https://doi.org/10.1017/S0305004100061697.

[Coh96] Paula Beazley Cohen. "Humbert surfaces and transcendence properties of automorphic functions". In: vol. 26. 3. Symposium on Diophantine Problems (Boulder, CO, 1994). 1996, pp. 987–1001. DOI: 10.1216/rmjm/1181072032. URL: https://doi.org/10.1216/rmjm/1181072032.

[Cox22] David A. Cox. *Primes of the form $x^2 + ny^2$—Fermat, class field theory, and complex multiplication*. Third. With contributions by Roger Lipsett. AMS Chelsea Publishing, Providence, RI, [2022] ©2022, pp. xv+533. ISBN: [9781470470289]; [9781470471835].

[CQ05] Gabriel Cardona and Jordi Quer. "Field of moduli and field of definition for curves of genus 2". In: *Computational aspects of algebraic curves*. Vol. 13. Lecture Notes Ser. Comput. World Sci. Publ., Hackensack, NJ, 2005, pp. 71–83. DOI: 10.1142/9789812701640_0006. URL: https://doi.org/10.1142/9789812701640_0006.

[CS86] Gary Cornell and Joseph H. Silverman, eds. *Arithmetic geometry*. Papers from the conference held at the University of Connecticut, Storrs, Connecticut, July 30–August 10, 1984. Springer-Verlag, New York, 1986, pp. xvi+353. ISBN: 0-387-96311-1. DOI: 10.1007/978-1-4613-8655-1. URL: https://doi.org/10.1007/978-1-4613-8655-1.

[CU04] Laurent Clozel and Emmanuel Ullmo. "Équidistribution des points de Hecke". In: *Contributions to automorphic forms, geometry, and number theory*. Johns Hopkins Univ. Press, Baltimore, MD, 2004, pp. 193–254. ISBN: 0-8018-7860-8.

[CW90] Paula Cohen and Jürgen Wolfart. "Modular embeddings for some nonarithmetic Fuchsian groups". In: *Acta Arith.* 56.2 (1990), pp. 93–110. ISSN: 0065-1036. DOI: 10.4064/aa-56-2-93-110. URL: https://doi.org/10.4064/aa-56-2-93-110.

[DM24] Paolo Dolce and Pietro Mercuri. "Intersection matrices for the minimal regular model of $X_0(N)$ and applications to the Arakelov canonical sheaf". In: *J. Lond. Math. Soc. (2)* 110.2 (2024), Paper No. e12964, 30. ISSN: 0024-6107,1469-7750. DOI: 10.1112/jlms.12964. URL: https://doi.org/10.1112/jlms.12964.

[DO21] Christopher Daw and Martin Orr. "Unlikely intersections with $E \times$ CM curves in A_2". In: *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* 22.4 (2021), pp. 1705–1745. ISSN: 0391-173X,2036-2145.

[Dup06] Régis Dupont. "Moyenne arithmético-géométrique, suites de Borchardt et applications." Available at https://www.lix.polytechnique.fr/Labo/Regis.Dupont/these_soutenance.pdf. PhD Thesis. École Polytechnique, 2006.

[EK14] Noam Elkies and Abhinav Kumar. "K3 surfaces and equations for Hilbert modular surfaces". In: *Algebra Number Theory* 8.10 (2014), pp. 2297–2411. ISSN: 1937-0652,1944-7833. DOI: 10.2140/ant.2014.8.2297. URL: https://doi.org/10.2140/ant.2014.8.2297.

[FKRS12] Francesc Fité, Kiran S. Kedlaya, Víctor Rotger, and Andrew V. Sutherland. "Sato-Tate distributions and Galois endomorphism modules in genus 2". In: *Compos. Math.* 148.5 (2012), pp. 1390–1442. ISSN: 0010-437X,1570-5846. DOI: 10.1112/S0010437X12000279. URL: https://doi.org/10.1112/S0010437X12000279.

[FN98] Naum Il'ich Feld'man and Yuri V. Nesterenko. "Transcendental numbers". In: *Number theory, IV*. Vol. 44. Encyclopaedia Math. Sci. Springer, Berlin, 1998, pp. 1–345. ISBN: 3-540-61467-2.

[Fon23] Tiago J. Fonseca. "Higher Ramanujan equations and periods of abelian varieties". In: *Mem. Amer. Math. Soc.* 281.1391 (2023), pp. xxv+125. ISSN: 0065-9266,1947-6221. DOI: 10.1090/memo/1391. URL: https://doi.org/10.1090/memo/1391.

[Fre90] Eberhard Freitag. *Hilbert modular forms*. Springer-Verlag, Berlin, 1990, pp. viii+250. ISBN: 3-540-50586-5. DOI: 10.1007/978-3-662-02638-0. URL: https://doi.org/10.1007/978-3-662-02638-0.

[FT93] A. Fröhlich and M. J. Taylor. *Algebraic number theory*. Vol. 27. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1993, pp. xiv+355. ISBN: 0-521-43834-9.

[Gal12] Steven D. Galbraith. *Mathematics of public key cryptography*. Cambridge University Press, Cambridge, 2012, pp. xiv+615. ISBN: 978-1-107-01392-6. DOI: 10.1017/CBO9781139012843. URL: https://doi.org/10.1017/CBO9781139012843.

[Gee08] Gerard van der Geer. "Siegel modular forms and their applications". In: *The 1-2-3 of modular forms*. Universitext. Springer, Berlin, 2008, pp. 181–245. ISBN: 978-3-540-74117-6. DOI: 10.1007/978-3-540-74119-0_3. URL: https://doi.org/10.1007/978-3-540-74119-0_3.

[Gee82] G. van der Geer. "On the geometry of a Siegel modular threefold". In: *Math. Ann.* 260.3 (1982), pp. 317–350. ISSN: 0025-5831,1432-1807. DOI: 10.1007/BF01461467. URL: https://doi.org/10.1007/BF01461467.

[Gee88] Gerard van der Geer. *Hilbert modular surfaces*. Vol. 16. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin, 1988, pp. x+291. ISBN: 3-540-17601-2. DOI: 10.1007/978-3-642-61553-5. URL: https://doi.org/10.1007/978-3-642-61553-5.

[GHKRW06] 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*. Vol. 4284. Lecture Notes in Comput. Sci. Springer, Berlin, 2006, pp. 114–129. ISBN: 978-3-540-49475-1; 3-540-49475-8. DOI: 10.1007/11935230_8. URL: https://doi.org/10.1007/11935230_8.

[Gij25] Desirée Gijón Gómez. *On the CM exception to a generalization of the Stéphanois theorem*. 2025. arXiv: 2505.08570 [math.NT]. URL: https://arxiv.org/abs/2505.08570.

[Gor02] Eyal Z. Goren. *Lectures on Hilbert modular varieties and modular forms*. Vol. 14. CRM Monograph Series. With the assistance of Marc-Hubert Nicole. American Mathematical Society, Providence, RI, 2002, pp. x+270. ISBN: 0-8218-1995-X. DOI: 10.1090/crmm/014. URL: https://doi.org/10.1090/crmm/014.

[GR14a] Éric Gaudron and Gaël Rémond. "Polarisations et isogénies". In: *Duke Math. J.* 163.11 (2014), pp. 2057–2108. ISSN: 0012-7094,1547-7398. DOI: 10.1215/00127094-2782528. URL: https://doi.org/10.1215/00127094-2782528.

[GR14b] Éric Gaudron and Gaël Rémond. "Théorème des périodes et degrés minimaux d'isogénies". In: *Comment. Math. Helv.* 89.2 (2014), pp. 343–403. ISSN: 0010-2571,1420-8946. DOI: 10.4171/CMH/322. URL: https://doi.org/10.4171/CMH/322.

[Gru08] David Gruenewald. "Explicit Algorithms for Humbert Surfaces". Available at https://www.maths.usyd.edu.au/u/davidg/thesis.pdf. PhD Thesis. University of Sydney, 2008.

[GRV24] Elisa Lorenzo García, Christophe Ritzenthaler, and Fernando Rodríguez Villegas. *An arithmetic intersection for squares of elliptic curves with complex multiplication*. 2024. arXiv: 2412.08738 [math.NT]. URL: https://arxiv.org/abs/2412.08738.

[Gun63] Karl-Bernhard Gundlach. "Die Bestimmung der Funktionen zur Hilbertschen Modulgruppe des Zahlkörpers $Q(\sqrt{5})$". In: *Math. Ann.* 152 (1963), pp. 226–256. ISSN: 0025-5831,1432-1807. DOI: 10.1007/BF01470882. URL: https://doi.org/10.1007/BF01470882.

[GY19] Jia-Wei Guo and Yifan Yang. *Class number relations arising from intersections of Shimura curves and Humbert surfaces*. 2019. arXiv: 1903.07225 [math.NT]. URL: https://arxiv.org/abs/1903.07225.

[GZ86] Benedict H. Gross and Don B. Zagier. "Heegner points and derivatives of L-series". In: *Invent. Math.* 84.2 (1986), pp. 225–320. ISSN: 0020-9910,1432-1297. DOI: 10.1007/BF01388809. URL: https://doi.org/10.1007/BF01388809.

[Har77] Robin Hartshorne. *Algebraic geometry*. Vol. No. 52. Graduate Texts in Mathematics. Springer-Verlag, New York-Heidelberg, 1977, pp. xvi+496. ISBN: 0-387-90244-9.

[Has95] Ki-ichiro Hashimoto. "Explicit form of quaternion modular embeddings". In: *Osaka J. Math.* 32.3 (1995), pp. 533–546. ISSN: 0030-6126. URL: http://projecteuclid.org/euclid.ojm/1200786264.

[Hir71] F. Hirzebruch. "The Hilbert modular group, resolution of the singularities at the cusps and related problems". In: *Séminaire Bourbaki, 23ème année (1970/1971)*. Vol. Vol. 244. Lecture Notes in Math. Springer, Berlin-New York, 1971, Exp. No. 396, pp. 275–288.

[HM95] Ki-ichiro Hashimoto and Naoki Murabayashi. "Shimura curves as intersections of Humbert surfaces and defining equations of QM-curves of genus two". In: *Tohoku Math. J. (2)* 47.2 (1995), pp. 271–296. ISSN: 0040-8735,2186-585X. DOI: 10.2748/tmj/1178225596. URL: https://doi.org/10.2748/tmj/1178225596.

[HP17] Philipp Habegger and Fabien Pazuki. "Bad reduction of genus 2 curves with CM jacobian varieties". In: *Compos. Math.* 153.12 (2017), pp. 2534–2576. ISSN: 0010-437X,1570-5846. DOI: 10.1112/S0010437X17007424. URL: https://doi.org/10.1112/S0010437X17007424.

[HS02] Klaus Hulek and G. K. Sankaran. "The geometry of Siegel modular varieties". In: *Higher dimensional birational geometry (Kyoto, 1997)*. Vol. 35. Adv. Stud. Pure Math. Math. Soc. Japan, Tokyo, 2002, pp. 89–156. ISBN: 4-931469-19-1. DOI: 10.2969/aspm/03510089. URL: https://doi.org/10.2969/aspm/03510089.

[Hum01] Georges Humbert. "Sur les fonctions abéliennes singulières. I, II, III". French. In: *Journ. de Math. (5)* 5, 6, 7 (1899, 1900, 1901), pp. 233–350, 279–386, 97–123.

[Hur98] A. Hurwitz. "Ueber die Entwickelungscoefficienten der lemniscatischen Functionen". In: *Math. Ann.* 51.2 (1898), pp. 196–226. ISSN: 0025-5831,1432-1807. DOI: 10.1007/BF01453637. URL: https://doi.org/10.1007/BF01453637.

[HW08] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Sixth. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles. Oxford University Press, Oxford, 2008, pp. xxii+621. ISBN: 978-0-19-921986-5.

[HZ77] F. Hirzebruch and D. Zagier. "Classification of Hilbert modular surfaces". In: *Complex analysis and algebraic geometry*. Iwanami Shoten Publishers, Tokyo, 1977, pp. 43–77.

[Ibu82] Tomoyoshi Ibukiyama. "On maximal orders of division quaternion algebras over the rational number field with certain optimal embeddings". In: *Nagoya Math. J.* 88 (1982), pp. 181–195. ISSN: 0027-7630,2152-6842. DOI: 10.1017/S002776300002016X. URL: https://doi.org/10.1017/S002776300002016X.

[Igu60] Jun-ichi Igusa. "Arithmetic variety of moduli for genus two". In: *Ann. of Math. (2)* 72 (1960), pp. 612–649. ISSN: 0003-486X. DOI: 10.2307/1970233. URL: https://doi.org/10.2307/1970233.

[Igu62] Jun-ichi Igusa. "On Siegel modular forms of genus two". In: *Amer. J. Math.* 84 (1962), pp. 175–200. ISSN: 0002-9327. DOI: 10.2307/2372812. URL: https://doi.org/10.2307/2372812.

[Igu64] Jun-ichi Igusa. "On Siegel modular forms genus two. II". In: *Amer. J. Math.* 86 (1964), pp. 392–412. ISSN: 0002-9327,1080-6377. DOI: 10.2307/2373172. URL: https://doi.org/10.2307/2373172.

[Igu67] Jun-ichi Igusa. "Modular forms and projective invariants". In: *Amer. J. Math.* 89 (1967), pp. 817–855. ISSN: 0002-9327,1080-6377. DOI: 10.2307/2373243. URL: https://doi.org/10.2307/2373243.

[Igu72] Jun-ichi Igusa. *Theta functions*. Vol. Band 194. Die Grundlehren der mathematischen Wissenschaften. Springer-Verlag, New York-Heidelberg, 1972, pp. x+232.

[JK09] Jay Jorgenson and Jürg Kramer. "Bounds on Faltings's delta function through covers". In: *Ann. of Math. (2)* 170.1 (2009), pp. 1–43. ISSN: 0003-486X,1939-8980. DOI: 10.4007/annals.2009.170.1. URL: https://doi.org/10.4007/annals.2009.170.1.

[Kan16] Ernst Kani. "The moduli spaces of Jacobians isomorphic to a product of two elliptic curves". In: *Collect. Math.* 67.1 (2016), pp. 21–54. ISSN: 0010-0757,2038-4815. DOI: 10.1007/s13348-015-0148-9. URL: https://doi.org/10.1007/s13348-015-0148-9.

[Kan19] Ernst Kani. *Generalized Humbert schemes and intersections of Humbert surfaces*. 2019. URL: https://mast.queensu.ca/~kani/papers/interHum11.pdf.

[Kan94] Ernst Kani. "Elliptic curves on abelian surfaces". In: *Manuscripta Math.* 84.2 (1994), pp. 199–223. ISSN: 0025-2611,1432-1785. DOI: 10.1007/BF02567454. URL: https://doi.org/10.1007/BF02567454.

[Kie21] Jean Kieffer. "Higher-dimensional modular equations, applications to isogeny computations and point counting". Available at https://theses.hal.science/tel-03346032. PhD Thesis. Université de Bordeaux, 2021.

[Kie22] Jean Kieffer. "Degree and height estimates for modular equations on PEL Shimura varieties". In: *J. Lond. Math. Soc. (2)* 105.2 (2022), pp. 1314–1361. ISSN: 0024-6107,1469-7750. DOI: 10.1112/jlms.12540. URL: https://doi.org/10.1112/jlms.12540.

[Kir22] Harun Kir. *The Refined Humbert Invariant for Imprimitive Ternary Forms*. 2022. arXiv: 2211.04396 [math.NT]. URL: https://arxiv.org/abs/2211.04396.

[Kli90] Helmut Klingen. *Introductory lectures on Siegel modular forms*. Vol. 20. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1990, pp. x+162. ISBN: 0-521-35052-2. DOI: 10.1017/CBO9780511619878. URL: https://doi.org/10.1017/CBO9780511619878.

[KPR25] Jean Kieffer, Aurel Page, and Damien Robert. "Computing isogenies from modular equations in genus two". In: *J. Algebra* 666 (2025), pp. 331–386. ISSN: 0021-8693,1090-266X. DOI: 10.1016/j.jalgebra.2024.11.029. URL: https://doi.org/10.1016/j.jalgebra.2024.11.029.

[Lag68] Joseph Louis Lagrange. "Sur la solution des problèmes indéterminés du second degré". In: *Oeuvres. Tome 2*. Publiées par les soins de J.-A. Serret, Nachdruck der Ausgabe Paris 1868. Georg Olms Verlag, Hildesheim-New York, 1868, pp. 377–535.

[Lan06] Herbert Lange. "Principal polarizations on products of elliptic curves". In: *The geometry of Riemann surfaces and abelian varieties*. Vol. 397. Contemp. Math. Amer. Math. Soc., Providence, RI, 2006, pp. 153–162. ISBN: 0-8218-3855-5. DOI: 10.1090/conm/397/07470. URL: https://doi.org/10.1090/conm/397/07470.

[Lan87] Serge Lang. *Elliptic functions*. Second. Vol. 112. Graduate Texts in Mathematics. With an appendix by J. Tate. Springer-Verlag, New York, 1987, pp. xii+326. ISBN: 0-387-96508-4. DOI: 10.1007/978-1-4612-4752-4. URL: https://doi.org/10.1007/978-1-4612-4752-4.

[LMFDB] The LMFDB Collaboration. *The L-functions and modular forms database*. https://www.lmfdb.org. [Online; accessed 23 September 2025]. 2025.

[LY11] Kristin Lauter and Tonghai Yang. "Computing genus 2 curves from invariants on the Hilbert moduli space". In: *J. Number Theory* 131.5 (2011), pp. 936–958. ISSN: 0022-314X,1096-1658. DOI: 10.1016/j.jnt.2010.05.012. URL: https://doi.org/10.1016/j.jnt.2010.05.012.

[LY20] Yi-Hsuan Lin and Yifan Yang. "Quaternionic loci in Siegel's modular threefold". In: *Math. Z.* 295.1-2 (2020), pp. 775–819. ISSN: 0025-5874,1432-1823. DOI: 10.1007/s00209-019-02372-z. URL: https://doi.org/10.1007/s00209-019-02372-z.

[MAGMA] Wieb Bosma, John Cannon, and Catherine Playoust. "The Magma algebra system. I. The user language". In: *J. Symbolic Comput.* 24.3-4 (1997). Computational algebra and number theory (London, 1993), pp. 235–265. ISSN: 0747-7171. DOI: 10.1006/jsco.1996.0125. URL: http://dx.doi.org/10.1006/jsco.1996.0125.

[Mar18] Daniel A. Marcus. *Number fields*. Second. Universitext. With a foreword by Barry Mazur. Springer, Cham, 2018, pp. xviii+203. ISBN: 978-3-319-90232-6; 978-3-319-90233-3. DOI: 10.1007/978-3-319-90233-3. URL: https://doi.org/10.1007/978-3-319-90233-3.

[Mar20] Chloe Martindale. "Hilbert modular polynomials". In: *J. Number Theory* 213 (2020), pp. 464–498. ISSN: 0022-314X,1096-1658. DOI: 10.1016/j.jnt.2019.11.019. URL: https://doi.org/10.1016/j.jnt.2019.11.019.

[Mes91] Jean-François Mestre. "Construction de courbes de genre 2 à partir de leurs modules". In: *Effective methods in algebraic geometry (Castiglioncello, 1990)*. Vol. 94. Progr. Math. Birkhäuser Boston, Boston, MA, 1991, pp. 313–334. ISBN: 0-8176-3546-7.

[Mil15] Enea Milio. "A quasi-linear time algorithm for computing modular polynomials in dimension 2". In: *LMS J. Comput. Math.* 18.1 (2015), pp. 603–632. ISSN: 1461-1570. DOI: 10.1112/S1461157015000170. URL: https://doi.org/10.1112/S1461157015000170.

[Mil86a] J. S. Milne. "Abelian varieties". In: *Arithmetic geometry (Storrs, Conn., 1984)*. Springer, New York, 1986, pp. 103–150. ISBN: 0-387-96311-1.

[Mil86b] J. S. Milne. "Jacobian varieties". In: *Arithmetic geometry (Storrs, Conn., 1984)*. Springer, New York, 1986, pp. 167–212. ISBN: 0-387-96311-1.

[MO13] Ben Moonen and Frans Oort. "The Torelli locus and special subvarieties". In: *Handbook of moduli. Vol. II*. Vol. 25. Adv. Lect. Math. (ALM). Int. Press, Somerville, MA, 2013, pp. 549–594. ISBN: 978-1-57146-258-9.

[Mor72] Yasuo Morita. "On transcendency of special values of arithmetic automorphic functions". In: *J. Math. Soc. Japan* 24 (1972), pp. 268–274. ISSN: 0025-5645,1881-1167. DOI: 10.2969/jmsj/02420268. URL: https://doi.org/10.2969/jmsj/02420268.

[MR20] Enea Milio and Damien Robert. "Modular polynomials on Hilbert surfaces". In: *J. Number Theory* 216 (2020), pp. 403–459. ISSN: 0022-314X,1096-1658. DOI: 10.1016/j.jnt.2020.04.014. URL: https://doi.org/10.1016/j.jnt.2020.04.014.

[MU98] P. Michel and E. Ullmo. "Points de petite hauteur sur les courbes modulaires $X_0(N)$". In: *Invent. Math.* 131.3 (1998), pp. 645–674. ISSN: 0020-9910,1432-1297. DOI: 10.1007/s002220050216. URL: https://doi.org/10.1007/s002220050216.

[Mum08] David Mumford. *Abelian varieties*. Vol. 5. Tata Institute of Fundamental Research Studies in Mathematics. With appendices by C. P. Ramanujam and Yuri Manin, Corrected reprint of the second (1974) edition. Tata Institute of Fundamental Research, Bombay; by Hindustan Book Agency, New Delhi, 2008, pp. xii+263. ISBN: 978-81-85931-86-9; 81-85931-86-0.

[MV07] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative number theory. I. Classical theory*. Vol. 97. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2007, pp. xviii+552. ISBN: 978-0-521-84903-6; 0-521-84903-9.

[Nag83] S. Nagaoka. "On the ring of Hilbert modular forms over $\mathbb{Z}$". In: *J. Math. Soc. Japan* 35.4 (1983), pp. 589–608. ISSN: 0025-5645,1881-1167. DOI: 10.2969/jmsj/03540589. URL: https://doi.org/10.2969/jmsj/03540589.

[Nes96] Yu. V. Nesterenko. "Modular functions and transcendence questions". In: *Mat. Sb.* 187.9 (1996), pp. 65–96. ISSN: 0368-8666,2305-2783. DOI: 10.1070/SM1996v187n09ABEH000158. URL: https://doi.org/10.1070/SM1996v187n09ABEH000158.

[Neu99] Jürgen Neukirch. *Algebraic number theory*. Vol. 322. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder. Springer-Verlag, Berlin, 1999, pp. xviii+571. ISBN: 3-540-65399-6. DOI: 10.1007/978-3-662-03983-0. URL: https://doi.org/10.1007/978-3-662-03983-0.

[Orr17] Martin Orr. "On compatibility between isogenies and polarizations of abelian varieties". In: *Int. J. Number Theory* 13.3 (2017), pp. 673–704. ISSN: 1793-0421,1793-7310. DOI: 10.1142/S1793042117500348. URL: https://doi.org/10.1142/S1793042117500348.

[OZ95] Frans Oort and Yuri Zarhin. "Endomorphism algebras of complex tori". In: *Math. Ann.* 303.1 (1995), pp. 11–29. ISSN: 0025-5831,1432-1807. DOI: 10.1007/BF01460976. URL: https://doi.org/10.1007/BF01460976.

[Paz10] Fabien Pazuki. "Remarques sur une conjecture de Lang". In: *J. Théor. Nombres Bordeaux* 22.1 (2010), pp. 161–179. ISSN: 1246-7405,2118-8572. DOI: 10.5802/jtnb.709. URL: https://doi.org/10.5802/jtnb.709.

[Paz19a] Fabien Pazuki. "Modular invariants and isogenies". In: *Int. J. Number Theory* 15.3 (2019), pp. 569–584. ISSN: 1793-0421,1793-7310. DOI: 10.1142/S1793042119500295. URL: https://doi.org/10.1142/S1793042119500295.

[Paz19b] Fabien Pazuki. "Décompositions en hauteurs locales". In: *Arithmetic geometry: computation and applications*. Vol. 722. Contemp. Math. Amer. Math. Soc., [Providence], RI, [2019] ©2019, pp. 121–140. ISBN: 978-1-4704-4212-5. DOI: 10.1090/conm/722/14529. URL: https://doi.org/10.1090/conm/722/14529.

[Pil13] Jonathan Pila. "Modular Ax-Lindemann-Weierstrass with derivatives". In: *Notre Dame J. Form. Log.* 54.3-4 (2013), pp. 553–565. ISSN: 0029-4527,1939-0726. DOI: 10.1215/00294527-2143853. URL: https://doi.org/10.1215/00294527-2143853.

[PT16] Jonathan Pila and Jacob Tsimerman. "Ax-Schanuel for the j-function". In: *Duke Math. J.* 165.13 (2016), pp. 2587–2605. ISSN: 0012-7094,1547-7398. DOI: 10.1215/00127094-3620005. URL: https://doi.org/10.1215/00127094-3620005.

[Rém17] Gaël Rémond. "Variétés abéliennes et ordres maximaux". In: *Rev. Mat. Iberoam.* 33.4 (2017), pp. 1173–1195. ISSN: 0213-2230,2235-0616. DOI: 10.4171/RMI/967. URL: https://doi.org/10.4171/RMI/967.

[Rob55] Herbert Robbins. "A remark on Stirling's formula". In: *Amer. Math. Monthly* 62 (1955), pp. 26–29. ISSN: 0002-9890,1930-0972. DOI: 10.2307/2308012. URL: https://doi.org/10.2307/2308012.

[Rot03a] V. Rotger. "Abelian varieties with quaternionic multiplication and their moduli". Available at https://web.mat.upc.edu/victor.rotger/docs/Tesi.pdf. PhD Thesis. Universitat de Barcelona, 2003.

[Rot03b] V. Rotger. "Quaternions, polarization and class numbers". In: *J. Reine Angew. Math.* 561 (2003), pp. 177–197. ISSN: 0075-4102,1435-5345. DOI: 10.1515/crll.2003.065. URL: https://doi.org/10.1515/crll.2003.065.

[Rot04a] Victor Rotger. "Modular Shimura varieties and forgetful maps". In: *Trans. Amer. Math. Soc.* 356.4 (2004), pp. 1535–1550. ISSN: 0002-9947,1088-6850. DOI: 10.1090/S0002-9947-03-03408-1. URL: https://doi.org/10.1090/S0002-9947-03-03408-1.

[Rot04b] Victor Rotger. "Shimura curves embedded in Igusa's threefold". In: *Modular curves and abelian varieties*. Vol. 224. Progr. Math. Birkhäuser, Basel, 2004, pp. 263–276. ISBN: 3-7643-6586-2.

[Run99] Bernhard Runge. "Endomorphism rings of abelian surfaces and projective models of their moduli spaces". In: *Tohoku Math. J. (2)* 51.3 (1999), pp. 283–303. ISSN: 0040-8735,2186-585X. DOI: 10.2748/tmj/1178224764. URL: https://doi.org/10.2748/tmj/1178224764.

[Sage] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.1)*. https://www.sagemath.org. 2025.

[Sch37] Theodor Schneider. "Arithmetische Untersuchungen elliptischer Integrale". In: *Math. Ann.* 113.1 (1937), pp. 1–13. ISSN: 0025-5831,1432-1807. DOI: 10.1007/BF01571618. URL: https://doi.org/10.1007/BF01571618.

[Sch57] Theodor Schneider. *Einführung in die transzendenten Zahlen*. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957, pp. v+150.

[Ser72] Jean-Pierre Serre. "Propriétés galoisiennes des points d'ordre fini des courbes elliptiques". In: *Invent. Math.* 15.4 (1972), pp. 259–331. ISSN: 0020-9910,1432-1297. DOI: 10.1007/BF01405086. URL: https://doi.org/10.1007/BF01405086.

[Ser73] J.-P. Serre. *A course in arithmetic*. Vol. No. 7. Graduate Texts in Mathematics. Translated from the French. Springer-Verlag, New York-Heidelberg, 1973, pp. viii+115.

[Shi63] Goro Shimura. "On analytic families of polarized abelian varieties and automorphic functions". In: *Ann. of Math. (2)* 78 (1963), pp. 149–192. ISSN: 0003-486X. DOI: 10.2307/1970507. URL: https://doi.org/10.2307/1970507.

[Shi94] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions*. Vol. 11. Publications of the Mathematical Society of Japan. Reprint of the 1971 original, Kanô Memorial Lectures, 1. Princeton University Press, Princeton, NJ, 1994, pp. xiv+271. ISBN: 0-691-08092-5.

[Sil90] Joseph H. Silverman. "Hecke points on modular curves". In: *Duke Math. J.* 60.2 (1990), pp. 401–423. ISSN: 0012-7094,1547-7398. DOI: 10.1215/S0012-7094-90-06016-8. URL: https://doi.org/10.1215/S0012-7094-90-06016-8.

[SSW08] R. E. Sawilla, A. K. Silvester, and H. C. Williams. "A new look at an old equation". In: *Algorithmic number theory*. Vol. 5011. Lecture Notes in Comput. Sci. Springer, Berlin, 2008, pp. 37–59. ISBN: 978-3-540-79455-4; 3-540-79455-7. DOI: 10.1007/978-3-540-79456-1_2. URL: https://doi.org/10.1007/978-3-540-79456-1_2.

[Str10] Marco Streng. "Complex multiplication of abelian surfaces". Available at https://pub.math.leidenuniv.nl/~strengtc/thesis.pdf. PhD Thesis. Universiteit Leiden, 2010.

[SW95] Hironori Shiga and Jürgen Wolfart. "Criteria for complex multiplication and transcendence properties of automorphic functions". In: *J. Reine Angew. Math.* 463 (1995), pp. 1–25. ISSN: 0075-4102,1435-5345. DOI: 10.1515/crll.1995.463.1. URL: https://doi.org/10.1515/crll.1995.463.1.

[Voi21] John Voight. *Quaternion algebras*. Vol. 288. Graduate Texts in Mathematics. Springer, Cham, [2021] ©2021, pp. xxiii+885. ISBN: 978-3-030-56692-0; 978-3-030-56694-4. DOI: 10.1007/978-3-030-56694-4. URL: https://doi.org/10.1007/978-3-030-56694-4.

[Wal00] Michel Waldschmidt. *Diophantine approximation on linear algebraic groups*. Vol. 326. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Transcendence properties of the exponential function in several variables. Springer-Verlag, Berlin, 2000, pp. xxiv+633. ISBN: 3-540-66785-7. DOI: 10.1007/978-3-662-11569-5. URL: https://doi.org/10.1007/978-3-662-11569-5.

[Wal96] Michel Waldschmidt. "Transcendance et indépendance algébrique de valeurs de fonctions modulaires". In: *Number theory (Ottawa, ON, 1996)*. Vol. 19. CRM Proc. Lecture Notes. Amer. Math. Soc., Providence, RI, 1996, pp. 353–375. ISBN: 0-8218-0964-4. DOI: 10.1090/crmp/019/32. URL: https://doi.org/10.1090/crmp/019/32.

[Wei57] André Weil. "Zum Beweis des Torellischen Satzes". In: *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. IIa.* 1957 (1957), pp. 33–53.