# Subcovers of Curves and Moduli Spaces

**Ernst Kani**

To Herbert Lange on his $75^{th}$ birthday

**Abstract**

This article studies certain questions pertaining to the moduli space $M_g(g', n)$ which was introduced and studied by H. Lange in the 1970's. This space classifies the isomorphism classes of curves $X$ of genus $g$ which admit a subcover $f : X \to Y$ of degree $n$ to a curve $Y$ of genus $g'$.

Here we focus on the case that $g = 2$ and $g' = 1$, in which case the irreducible components of $M_2(1, n)$ are certain *Humbert surfaces*. It is shown here that $M_2(1, n)$ is always connected, and that it has $d(n) - 1$ irreducible components.

One of the questions discussed here in some detail concerns the intersection $M_2(1, n) \cap M_2(1, m)$ of two such moduli spaces. It turns out that the irreducible components of such an intersection are images of modular curves, and these can be described explicitly by using the *generalized Humbert schemes* $H(q)$ and the theory of binary quadratic forms.

## 1 Introduction

It is a pleasure to dedicate this paper[1] to Herbert Lange for his $75^{th}$ birthday, not only because we have known each other for a long time, but also (and more importantly) because the results which are presented here are closely connected with some of his early work.

Starting with his thesis, much of Herbert Lange's early work (cf. Lange[L1] – [L4]) concerns the following topic:

**Problem.** Describe the set of curves $X/\mathbb{C}$ which admit a non-rational subcover.

Thus, he was interested in studying the set $M_g(g', n)$ of (isomorphism classes of) curves $X$ of genus $g$ which have a subcover $f : X \to Y$ of degree $n$ to some curve $Y$ of genus $g' \geq 1$ inside the *moduli space* $M_g$ of curves genus $g$, i.e.,

$$M_g(\mathbb{C}) \ = \ \{\text{isomorphism classes } \langle X \rangle \text{ of curves } X/\mathbb{C} \text{ of genus } g\}.$$

This was the subject both of his *dissertation* (the case $g = 2$) and of his *habilitation* (for arbitrary $g \geq 2$), and of at least 4 publications.

The following result is (a part of) Satz I of Lange[L4].

[1] This is an expanded version of my lecture at the *Langefest*

**Theorem 1.1 (Lange, 1977)** (a) *If $g > g' \geq 1$ and $n \geq 2$, then the set $M_g(g', n)$ is a closed subset of $M_g$.*

(b) *The subscheme $M_g(g', n)$ is equidimensional of dimension*

$$\dim M_g(g', n) \ = \ g - (n-2)(g'-1),$$

*provided that either*

$$g' \geq 2 \ \text{ and } \ \frac{g+1}{g'+1} \leq n \leq \frac{g+1}{g'-1} \quad \text{or that} \quad g' = 1 \ \text{ and } \ n > \frac{g+1}{2}.$$

*Moreover, $M_g(g', n) = \emptyset$ in all other cases, except possibly in the case that $n = \frac{g+1}{2}$ and $g' = 1$.*

These beautiful results of Lange naturally lead to further questions about the *geometric structure* of the subschemes $M_g(g', n)$. For example:

**Question 1** *How many irreducible components does $M_g(g', n)$ have? When is $M_g(g', n)$ irreducible?*

**Question 2** *What is the "geometric type" of each irreducible component? When are they all rational or of* general type*?*

**Question 3** *Is $M_g(g', n)$ connected?* (This is a question of Accola/Previato[AP] in the case that $g = 2$.)

**Question 4** *What can be said about the* intersection *of $M_g(g', n)$ with one or more $M_g(g'', n')$'s?*

These and other related questions will be investigated in the case that $g = 2$.

**Acknowledgment.** I would like to thank the organizers of the *Langefest* and of the conference *Geometry at the Frontier III* for inviting me and for providing travel funds and for their kind hospitality at the conference. In addition, I would like to thank the referee for his/her useful comments and suggestions which improved this paper. Moreover, I gratefully acknowledge receipt of funding from the Natural Sciences and Research Council of Canada (NSERC).

# 2 Humbert surfaces

When $g = 2$, the irreducible components of $M_2(1, n)$ turn out to be open subsets of certain *Humbert surfaces;* these naturally live in the *moduli space* $A_2$ of *principally polarized* abelian surfaces, i.e.,

$A_2(\mathbb{C}) = \{$isomorphism classes $\langle A, \lambda \rangle$ of principally polarized abelian surfaces $(A, \lambda)\}$,

where the pair $(A, \lambda)$ consists of an abelian surface $A/\mathbb{C}$ together with a *principal polarization* $\lambda = \phi_\theta : A \xrightarrow{\sim} \hat{A}$; cf. [LB] or [M2].

**Remark 2.1** Via the *Torelli map* $\langle X\rangle \mapsto \langle J_X, \lambda_X\rangle$ we can (and will) view the moduli space $M_2$ as a subset of $A_2$, i.e. $M_2 \subset A_2$. Here, $J_X$ is the Jacobian surface of $X$, and

$$\lambda_X = \phi_{\theta_X} : J_X \xrightarrow{\sim} \hat{J}_X$$

is the polarization induced by the *theta-divisor* $\theta_X$ on $J_X$.

The Humbert surfaces $H_\Delta$ are defined via the *Humbert invariant* $\Delta$, which is a number which Humbert[Hu] attached in 1899 to a so-called "singular relation" (Humbert's terminology). Such a relation is one which is satisfied by the entries of a (normalized) period matrix of a principally polarized abelian surface $(A, \lambda)$ and only exists for special abelian surfaces. (See §6 below for an algebraic description of this invariant.) Thus, the Humbert surface with Humbert invariant $\Delta$ is defined by

$$H_\Delta = \{(A,\lambda) \in A_2 : (A,\lambda) \text{ has a singular relation with Humbert invariant } \Delta\}$$

In van der Geer[vdG], Ch. IX, one finds a modern treatment of Humbert surfaces. Humbert's results are summarized in Kap. I, §1, of Hecke's dissertion [He].

The following result follows (more or less) from what Humbert[Hu] proved about his invariant.

**Theorem 2.2 (Humbert)** *For each positive integer $n \equiv 0, 1 \pmod 4$, there exists an irreducible (analytic) surface $H_n \subset A_2$ (now called a* Humbert surface*) such that:*

(a) *If $\langle A, \lambda\rangle \in A_2(\mathbb{C})$, then* $\mathrm{End}(A) \neq \mathbb{Z} \Leftrightarrow \langle A, \lambda\rangle \in H_n(\mathbb{C})$, *for some $n$;*

(b) $M_2 = A_2 \setminus H_1$*;*

(c) *If $\langle X\rangle \in M_2(\mathbb{C})$, then $\exists$ a surjection $f : X \to E$, where $E/\mathbb{C}$ is some elliptic curve $\Leftrightarrow \langle J_X, \lambda_X\rangle \in H_{N^2}(\mathbb{C})$, for some $N \geq 2$.*

*Proof.* For convenience of the reader,[2] here is a sketch of how these results follow from the results of Humbert[Hu]. (See also Remark 6.1 below.)

First of all, Humbert[Hu] explains in §3 (p. 237) that if one normalized period matrix of $(A, \lambda)$ has a singular relation with invariant $\Delta$, then every normalized period matrix of $(A, \lambda)$ also has a singular relation with invariant $\Delta$. Thus, this property is an invariant of the isomorphism class of $(A, \lambda)$.

In §11 (p. 245) he shows that his invariant satisfies $\Delta \equiv 0, 1 \pmod 4$ and that $\Delta > 0$ (Théorème 14, p. 246).

In Théorème 12 (p. 245) he shows that one can choose the period matrix in such a way that the singular relation assumes a simple form. This means (in modern language) that for each $\Delta \equiv 0, 1 \pmod 4$ there is a linear relation $f_\Delta = 0$ in the

[2] This follows a suggestion of the referee, who asked me to explain where these results are proved in Humbert's paper

Siegel half-space $\mathfrak{H}_2$ such that its image in $A_2 = \mathrm{Sp}_2(\mathbb{Z})\backslash\mathfrak{H}_2$ is $H_\Delta$. Thus, $H_\Delta$ is an irreducible analytic surface in $A_2$; cf. also [vdG], Theorem (IX.2.4) on p. 212.

(a) In §138 Humbert shows (on p. 285) that if $(A, \lambda) \in H_\Delta$ for some $\Delta$, the $A$ is "singular" in his (and Hurwitz's) terminology, i.e., $\mathrm{End}(A) \neq \mathbb{Z}$.

I wasn't able to find that the converse implication was stated (or proved) in [Hu], except in special cases; cf. Théorème I in §104 and Théorème II in §105. However, Hecke[He] states at the end of §I.1 (without proof or reference) that one can deduce that the converse holds.

Note that part (a) follows directly from what was proven in [K1], Corollary 5.5.

(b) In Remarque II in §17 on p. 248, Humbert points out that $\Delta = 1$ if and only if $(A, \lambda) \simeq (E_1 \times E_2, \lambda_1 \otimes \lambda_2)$, where $(E_i, \lambda_i)$ is a (principally polarized) elliptic curve, for $i = 1, 2$, and $\lambda_1 \otimes \lambda_2$ is the product polarization. This, together with Satz 2 of Weil[We], proves the assertion.

Note that in view of [K1], §5, this also follows Proposition 6 in [K3].

(c) This is Théorème 15 in §15 on p. 247. □

**Remark 2.3** (a) Each Humbert surface $H_n$ is a Zariski-closed subset of $A_2$. This does not follow directly from Humbert[Hu].

(b) Part (c) of Theorem 2.2 had already been stated and proved by Biermann in 1883; cf. Krazer[Kr], Satz V on p. 485. Perhaps Humbert did not know this.

(c) The result of Theorem 2.2(c) was refined in Theorem 1.9 of [K1] in the following way:

$$\langle J_X, \lambda_X \rangle \in H_{N^2} \Leftrightarrow \exists \text{ a minimal } f : X \to E \text{ with } \deg(f) = N. \tag{1}$$

Here and below, a cover $f : X \to E$ is called *minimal* if it does not factor over an isogeny of $E$ of degree $> 1$.

**Corollary 2.4** *For any $n \geq 2$ we have*

$$M_2(1, n) = \bigcup_{1<N|n} H^*_{N^2}, \tag{2}$$

*where $H^*_{N^2} = H_{N^2} \cap M_2$. Thus, $M_2(1, n)$ is equidimensional of dimension 2, and has $d(n) - 1$ irreducible components, where $d(n)$ is the number of divisors of $n$. In particular, $M_2(1, n)$ is irreducible if and only if $n$ is a prime number.*

*Proof.* Each subcover $f : X \to E$ factors as $f = h \circ f_{min}$, where $f_{min} : X \to E'$ is minimal, so the formula (2) follows from (1). The other assertions follow from (2) and Theorem 2.2, together with the fact that $H^*_{N^2} = H_{N^2} \cap M_2 \neq \emptyset$ when $N > 1$ (see Theorem 4.1 below). □

Note that this answers Question 1 of the introduction. We next turn to the study of Question 2.

## 3 The geometry of Humbert surfaces

To construct the moduli space $M_g(g', n)$, Lange[L4] adapts and generalizes Mumford's construction[M1] of the moduli space $M_g$.

Thus, he first constructs a suitable subvariety $H = H_{g,g',n}$ of some Hilbert scheme, and then shows that $M_g(g', n)$ is the quotient scheme of $H$ under the action of the group $\mathrm{PGL}(m)$, where $m \geq 3$ is suitable.

Unfortunately, this construction sheds little light on the geometry of $M_g(g', n)$.

Since it suffices to construct the irreducible components of these schemes, and since for $M_2(1, n)$ these components are known by Corollary 2.4 to be Humbert surfaces, it suffices to construct these surfaces. For them, there is another construction method available, which is based on what G. Frey and I call the *the basic construction*; cf. [FK1], [K2], and [FK2].

**The basic construction.** Fix an integer $N \geq 2$, and let $(E, E', \psi)$ be a triple consisting of two elliptic curves $E$ and $E'$ and an isomorphism

$$\psi : E[N] \xrightarrow{\sim} E'[N]$$

which is an *anti-isometry* with respect to the $e_N$-pairings (called Weil-pairings in [Si], p. 96) on the groups of $N$-torsion points of $E$ and of $E'$, respectively. This condition means that we have

$$e_N(\psi(P), \psi(Q)) \ = \ e_N(P, Q)^{-1}, \quad \forall P, Q \in E[N].$$

Given such a triple $(E, E', \psi)$, construct the *abelian surface*

$$J_\psi := (E \times E')/\mathrm{Graph}(-\psi).$$

This surface has a "canonical" *principal polarization* $\lambda_\psi$, so $\langle J_\psi, \lambda_\psi \rangle \in A_2(\mathbb{C})$.

This principal polarization $\lambda_\psi$ is "canonical" in the following sense. If we let $\pi_\psi : E \times E' \to J_\psi$ denote the quotient morphism, then $\lambda_\psi$ is uniquely determined by the condition that

$$\hat{\pi}_\psi \circ \lambda_\psi \circ \pi_\psi = N(\lambda \otimes \lambda'),$$

where $\hat{\pi}_\psi : \hat{J}_\psi \to (E \times E')^\wedge$ is the dual isogeny and $\lambda \otimes \lambda'$ denotes the product polarization on $E \times E'$.

Moreover, if the theta-divisor $X_\psi \in \mathrm{Div}^+(J_\psi)$ of $\lambda_\psi$ is irreducible, then we have a (minimal) cover $f_\psi : X_\psi \to E$ of degree $N$, and every (minimal) cover $f : X \to E$ of degree $N$ arises this way.

Since this construction also works for families of elliptic curves over a base (cf. [K2]), it has a *modular interpretation*, i.e., it induces a map between the associated functors. To explain this in more detail, we first note the following result from [FK2]; cf. [FK2], Proposition 4.5.

**Theorem 3.1** *Let $\mathcal{Z}_N$ denote the functor which classifies isomorphism classes of triples $(E, E', \psi)$, where $\psi : E[N] \to E'[N]$ is an anti-isometry. If $N \geq 3$, then $\mathcal{Z}_N$ is coarsely represented by the* modular diagonal quotient surface

$$Z_N \;:=\; \Delta_N^* \backslash (X(N) \times X(N)),$$

*where $X(N) = \Gamma(N)\backslash\mathfrak{H}$ is the usual modular curve of level $N$ on which the group $G_N := \Gamma(1)/(\pm\Gamma(N)) \simeq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/(\pm 1)$ acts, and*

$$\Delta_N^* \;=\; \{(g, \beta g \beta^{-1}) : g \in G_N\} \;\leq\; G_N \times G_N, \quad \textit{where } \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

**Remark 3.2** (a) By construction, the modular diagonal quotient surface $Z_N$ is a quotient of the product surface $Y(N) := X(N) \times X(N)$ of the modular curve $X(N)$. We thus have a finite surjective morphism

$$\Phi_N : Y(N) = X(N) \times X(N) \;\to\; \Delta_N^* \backslash Y(N) = Z_N.$$

(b) Since the curve $X(N)$ has a natural compactification (by adding cusps), the surfaces $Y(N)$ and $Z_N$ also admit natural compactifications $\bar{Y}(N)$ and $\bar{Z}_N$, respectively. The geometry of the (desingularization of the) modular diagonal quotient surfaces $\bar{Z}_N$ was determined by Hermann[H1] and also by Kani and Schanz[KS].

The "basic construction" leads to the following result.

**Theorem 3.3** *The rule $(E, E', \psi) \mapsto (J_\psi, \lambda_\psi)$ defines a finite morphism*

$$\beta_N : Z_N \;\to\; A_2$$

*whose image is the Humbert surface $H_{N^2}$. Thus, $H_{N^2}$ is an irreducible surface which is a Zariski-closed subset of $A_2$. Moreover, the normalization of $H_{N^2}$ is the* symmetric modular diagonal quotient surface

$$Z_N^{sym} := \langle w_N \rangle \backslash Z_N,$$

*where $w_N \in \mathrm{Aut}(Z_N)$ is the involution which is induced via $\Phi_N$ from the involution $\tau_N \in \mathrm{Aut}(Y(N))$ which interchanges the factors of $Y(N) = X(N) \times X(N)$. In particular,* $\deg(\beta_N : Z_N \to H_{N^2}) = 2$.

*Proof.* As is shown in [K8], the first assertions can be deduced from Theorem 15 of [K7], which is a refinement of [K1], Corollary 1.8 and of [FK2], Proposition 4.11. The assertion about the normalization of $H_{N^2}$ was stated (without proof) in [FK2], loc. cit., and is proved in Theorem 7.3 of [K8].

The geometric structure of the desingularization of the natural compactification of $Z_N^{sym}$ was determined by Hermann[H2] in the prime case.

**Theorem 3.4 (Hermann)** *If $N \geq 3$ is prime, then the surface $Z_N^{sym}$ is:*

- *rational, if $N \leq 13$,*
- *an elliptic K3-surface, if $N = 17$,*
- *an honest elliptic surface, if $N = 19$,*
- *of general type, if $N \geq 23$.*

**Remark 3.5** Thus, by Theorems 3.3 and 3.4 one has a precise description of the geometric type of the components $H_{N^2}$ of $M_2(1,n)$, at least when $N$ is prime. In particular, this gives the birational description of $M_2(1,n)$, when $n$ is prime. It is to be expected that a similar description exists for all $N$'s, but this has not been worked out (as far as I know).

On the other hand, the geometric description of $Z_N$, which is a double cover of $Z_N^{sym}$, has been given for all $N$ by Hermann[H1] and by Kani and Schanz[KS], as was mentioned above.

# 4 Connectedness properties

We now turn to discuss Question 3, which was (implicitly) raised by Accola and Previato[AP]. Recall that this question asks whether $M_2(1,n)$ is *connected* for all $n \geq 2$. That this is indeed the case follows from the following much stronger assertion (which also partially answers Question 4).

**Theorem 4.1** *If $n > m \geq 2$, then $H_{n^2} \cap H_{m^2} \cap M_2 \neq \emptyset$. Moreover, $H_{n^2} \cap H_{m^2}$ is a finite union of (images of)* modular curves.

**Remark 4.2** The first part of Theorem 4.1 is a special case of Corollary 8.2 below, which will be proven in §8. The second part follows from the more precise assertion (8) in Theorem 10.5.

The first part of this theorem was also proven by Franciosi/Pardini/Rollenske[FPR] by another method.

**Corollary 4.3** *The moduli space $M_2(1,n)$ is connected for any $n \geq 2$.*

*Proof.* By Corollary 2.4 (and Theorem 3.3) we know that the Humbert surfaces $H^*_{N^2} = H_{N^2} \cap M_2$ with $N|n$ and $N \geq 2$ are the irreducible components of $M_2(1,n)$. Thus, by Theorem 4.1 we see that any two irreducible components of $M_2(1,n)$ have a non-empty intersection, so $M_2(1,n)$ is connected. □

## 5 Intersections of Humbert surfaces

We now come to the discussion of Question 4, which concerns the intersection of different moduli spaces. In our situation ($g = 2$) this question reduces to the following.

**Question 4*** *What are the irreducible components of the intersection* $M_2(1, n) \cap M_2(1, m)$*, when* $\gcd(n, m) = 1$*?*

In view of the above structure theorem (Corollary 2.4) for the $M_2(1, n)$'s, it suffices to answer the following question.

**Question 5** *How can we describe the irreducible components of the intersection of two Humbert surfaces?*

**Remark 5.1** The intersection $H_{n^2} \cap H_{m^2} \cap M_2$ classifies curves $X$ of genus 2 with two minimal morphisms $f_1 : X \to E_1$ and $f_2 : X \to E_2$ of degrees $n$ and $m$.

To understand the components of this intersection, we will generalize Humbert's construction of the Humbert surfaces $H_n$ in $A_2$ in order to construct certain one-dimensional subschemes of $A_2$.

These generalized Humbert schemes will be defined by considering a *refinement* of the Humbert invariant, which Humbert introduced in his work.

However, while Humbert's invariant is a number, the *refined Humbert invariant* is a *quadratic form*, as will be explained presently.

**The basic idea.** As will be explained below, each integral, positive definite quadratic form $q$ defines a subset

$$H(q) \subset A_2,$$

called a *generalized Humbert scheme.* Its definition relies on the notion of the refined Humbert invariant.

As a preview, we state some of the basic properties of the generalized Humbert schemes. These will be verified in §7.

**Proposition 5.2** (a) *The set* $H(q)$ *depends only on the* $\mathrm{GL}_r(\mathbb{Z})$*-equivalence class of the quadratic form* $q = q(x_1, \ldots, x_r)$*. Moreover,*

$$q_1 \to q_2 \quad \Rightarrow \quad H(q_1) \subset H(q_2). \tag{3}$$

*where the symbol* $q_1 \to q_2$ *means that* $q_1$ *primitively represents the form* $q_2$*; cf. §7.*

(b) *The usual Humbert surface is* $H_n = H(nx^2)$*.*

(c) *We have that* $H(q) \neq A_2$*, but* $H(q)$ *may be empty. Indeed,* $H(q) = \emptyset$*, if* $q$ *is not positive-definite or if* $r > 3$*.*

(d) *If $n \neq m$, then*

$$H_n \cap H_m = \bigcup_{q \to n,m} H(q), \tag{4}$$

*where the union is over all integral, positive definite* binary *quadratic forms $q$ which* represent *both $n$ and $m$* primitively*; cf. §8.*

**Remark 5.3** If $m \neq n$, then up to $\mathrm{GL}_2(\mathbb{Z})$-equivalence there are only finitely many positive binary forms $q$ satisfying the condition $q \to n, m$ because this condition implies that $|\mathrm{disc}(q)| \leq 4mn$. Thus, by Proposition 5.2(a), the union on the right hand side of (4) consists only of finitely many $H(q)$'s.

The *intersection formula* (4) constitutes a *first step* towards understanding the components of the intersection of two Humbert surfaces.

However, further work is necessary to obtain a better picture, and this leads to the following questions.

**Question 6** *For which integral* binary *quadratic forms $q$ is $H(q) \neq \emptyset$?*

**Question 7** *What is the geometric structure of $H(q)$? Is $H(q)$ irreducible?*

**Question 8** *Can (the components of) $H(q)$ be described by images of modular curves? If so, which modular curves map to $H(q)$?*

These questions will be addressed below.

# 6 The refined Humbert invariant

The *refined Humbert invariant* $q_{(A,\lambda)}$ of a principally polarized abelian surface $(A, \lambda) \in A_2(\mathbb{C})$ is defined as follows.

**Definition.** Let $A$ be an abelian surface with a principal polarization $\lambda : A \xrightarrow{\sim} \hat{A}$, and let $\theta = \theta_\lambda \in \mathrm{Div}(A)$ be an associated theta-divisor, i.e., $\lambda = \phi_\theta$ in the notation of [M2]. Let $\mathrm{NS}(A)$ denote the Néron-Severi group of $A$, i.e., $\mathrm{NS}(A) := \mathrm{Div}(A)/\equiv$, where $\equiv$ denotes numerical equivalence of divisors, and put

$$\mathrm{NS}(A, \lambda) := \mathrm{NS}(A)/\mathbb{Z}\theta_\lambda.$$

Consider the quadratic form on $\mathrm{NS}(A)$ defined by

$$q_{(A,\lambda)}(D) = (D.\theta_\lambda)^2 - 2(D.D), \quad \forall D \in \mathrm{NS}(A),$$

where ( . ) denotes the intersection pairing on the Néron-Severi group $\mathrm{NS}(A)$. It then follows from the definition and the Hodge index theorem that $q_{(A,\lambda)}$ defines a *positive-definite quadratic form* on the module $\mathrm{NS}(A,\lambda)$; cf. [K1]. This induced quadratic form on $\mathrm{NS}(A,\lambda)$ is called *refined Humbert invariant* of $(A,\lambda)$ and is also denoted by $q_{(A,\lambda)}$.

**Remark 6.1** If $\bar{D} \in \mathrm{NS}(A,\lambda)$ is *primitive*, i.e., if $\mathrm{NS}(A,\lambda)/\mathbb{Z}\bar{D}$ is torsionfree (and $\bar{D} \neq 0$), then $\overline{D}$ gives rise to a "singular relation" of $A$ (in the sense of Humbert), and conversely every such relation arises in this way, as was shown in [K1], §5. Moreover, $\Delta := q_{(A,\lambda)}(\bar{D})$ is equal to the *Humbert invariant* $\Delta$ attached to the singular relation; cf. [K1], Corollary 5.5.

# 7 Generalized Humbert schemes

Using the *refined Humbert invariant* $q_{(A,\lambda)}$, which was defined in the previous section, we can generalize the notion of a Humbert surface as follows.

**Notation.** If $q : \mathbb{Z}^r \to \mathbb{Z}$ is an integral, positive-definite quadratic form in $r$ variables, then we put

$$H(q) := \{\langle A,\lambda\rangle \in A_2(\mathbb{C}) : q_{(A,\lambda)} \to q\}.$$

Here the symbol $q_{(A,\lambda)} \to q$ indicates that the form $q_{(A,\lambda)}$ *primitively represents* the form $q$. This means that there exists an injective homomorphism

$$h : \mathbb{Z}^r \hookrightarrow \mathrm{NS}(A,\lambda)$$

such that $\mathrm{NS}(A,\lambda)/h(\mathbb{Z}^r)$ is torsionfree and such that the $h$-restriction of $q_{(A,\lambda)}$ to $\mathbb{Z}^r$ is $q$, i.e., $q_{(A,\lambda)} \circ h = q$.

If $q_{(A,\lambda)} \to q$ and $h$ is surjective, then we say that $q_{(A,\lambda)}$ is *equivalent* to $q$ and write $q_{(A,\lambda)} \sim q$.

*Proof of* Proposition 5.2. (a) This is clear from the definition. Indeed, if $q_1$ is $\mathrm{GL}_r(\mathbb{Z})$-equivalent to $q_2$, then $\forall\langle A,\lambda\rangle \in A_2(\mathbb{C})$ we have that $q_{(A,\lambda)} \to q_1 \Leftrightarrow q_{(A,\lambda)} \to q_2$, and so $H(q_1) = H(q_2)$.

Similarly, if $\langle A,\lambda\rangle \in H(q_1)$, then $q_{(A,\lambda)} \to q_1$, and if $q_1 \to q_2$, then $q_{(A,\lambda)} \to q_2$, so $\langle A,\lambda\rangle \in H(q_2)$. Thus $H(q_1) \subset H(q_2)$. This proves (3).

(b) This follows from Remark 6.1.

(c) If $q$ is any (non-zero) quadratic form, then $q \to nx^2$, for some $n \neq 0$, and then by parts (a) and (b) we have that $H(q) \subset H(nx^2) = H_n$. By Humbert (Theorem 2.2), $H_n$ is either empty or a surface, so $H_n \neq A_2$ because $\dim A_2 = 3$. This proves the first statement.

Since $q_{(A,\lambda)}$ is positive-definite, then the same is true for any $q$ with $q_{(A,\lambda)} \to q$, and so it follows that $H(q) = \emptyset$, if $q$ is not positive-definite.

Moreover, since $\mathrm{NS}(A) \simeq \mathbb{Z}^\rho$, where $\rho \le h^{1,1}(A) = 4$ (cf. [vdG], p. 207), we see that $\mathrm{NS}(A,\lambda) \simeq \mathbb{Z}^r$, with $r \le 3$. Thus, the same is true for any $q$ with $q_{(A,\lambda)} \to q$, and so it follows that $H(q) = \emptyset$, if $q$ has rank $> 3$.

(d) The fact that the right hand side of (4) is contained in the left hand side follows immediately from (3).

Conversely, suppose that $\langle A, \lambda \rangle \in H_m \cap H_n$. Then there exist primitive vectors $v, w \in M := \mathrm{NS}(A,\lambda)$ such that $q_{(A,\lambda)}(v) = m$ and $q_{(A,\lambda)}(w) = n$. If $v$ and $w$ were linearly dependent, then $v = \pm w$ and hence $q_{(A,\lambda)}(v) = q_{(A,\lambda)}(w)$, contrary to the hypothesis. Thus, $v$ and $w$ are linearly independent and hence $M_0 := \mathbb{Z}v + \mathbb{Z}w$ has rank 2. Let $M_1$ be the saturation of $M_0$ in $M$. Then the restriction $q$ of $q_{(A,\lambda)}$ to $M_1$ is a positive definite, binary quadratic form which is primitively represented by $q_{(A,\lambda)}$, and so $\langle A, \lambda \rangle \in H(q)$. Moreover, $m = q(v)$ is primitively represented by $q$ (because $v$ is primitive in $M$, hence also in $M_1$. Similarly, $n = q(w)$ is primitively represented by $q$. Thus $q$ is one of the forms of the right side of (4), so $\langle A, \lambda \rangle \in \cup H(q)$. □

**Theorem 7.1** *If $q$ is an integral quadratic form with $q \to N^2$, for some integer $N \ge 1$, then $H(q)$ is a closed subset of $A_2$. Moreover, if $H(q) \ne \emptyset$ then $H(q)$ is equidimensional of dimension* $\dim H(q) = 3 - r$, *where $r$ is the number of variables of $q$.*

**Remark 7.2** This theorem follows easily from the results below; cf. §10, where this is proved. It is to be expected that the same result holds for all integral quadratic forms, but the methods below do not prove this.

**Definition.** A (non-empty) subset of $A_2$ of the form $H(q)$, where $q$ is a quadratic form, will be called a *generalized Humbert scheme.*

# 8 Special quadratic forms

We now want to classify those binary quadratic forms $q$ for which $H(q) \ne \emptyset$, and thus answer Question 6. Here we will only consider those forms which primitively represent a square because only those lead to elliptic subcovers of genus 2 curves; cf. Theorem 2.2(c) and Proposition 5.2(b).

**Notation.** Write $q = [a, b, c]$ for the binary quadratic form

$$q(x, y) = ax^2 + bxy + cy^2$$

on $\mathbb{Z}^2$. Moreover, let $Q$ denote the set of integral binary quadratic forms $q$ which satisfy the following two conditions:

(i) $q$ is positive-definite;

(ii) $q(x,y) \equiv 0,1 \pmod 4$, $\forall x,y \in \mathbb{Z}$.

In addition, for $n \in \mathbb{N}$, let

$$Q(n) = \{q \in Q : q \to n\}$$

denote the set of forms $q \in Q$ which *primitively represent* $n$, i.e.,

$$q(x,y) = n, \quad \text{for some } x,y \in \mathbb{Z} \text{ with } \gcd(x,y) = 1.$$

We can now classify the binary $q$'s for which $H(q) \neq \emptyset$ (and such that $H(q)$ lies on some Humbert surface $H_{N^2}$) in the following way.

**Theorem 8.1** *Let $q$ be an integral binary quadratic form which primitively represents a square, and let $N \geq 1$. Then:*

$$H(q) \neq \emptyset \ \textit{ and } \ H(q) \subset H_{N^2} \quad \Leftrightarrow \quad q \in Q(N^2). \tag{5}$$

*If this is the case, then there exists $\langle A, \lambda \rangle \in H(q)$ such that $q_{(A,\lambda)}$ is equivalent to $q$.*

*Proof.* If $q \in Q(N^2)$, then the Existence Theorem 32 of [K5] shows that that there exists $\langle A, \lambda \rangle \in H(q)$ with $q_{(A,\lambda)} \sim q$, so the last assertion holds. Since $q \to N^2x^2$, we have by (3) and Proposition 5.2(b) that $H(q) \subset H_{N^2}$.

Conversely, suppose that $\langle A, \lambda \rangle \in H(q)$. Then $q_{(A,\lambda)} \to q$, so $q \in Q$ because properties (i) and (ii) hold for $q_{(A,\lambda)}$ and hence are inherited by $q$. By hypothesis, $q \to M^2$, for some $M \geq 1$. Thus $q \in Q(M^2)$ and so by the Existence Theorem 32 of [K5] there exists $\langle A', \lambda' \rangle \in H(q)$ such that $q_{(A',\lambda')} \sim q$. Since $\langle A', \lambda' \rangle \in H(q)$ and since $H(q) \subset H_{N^2}$ by our hypothesis, it follows that $q_{(A',\lambda')} \to N^2$ (cf. Proposition 5.2(b)), and hence also $q \to N^2$ because $q \sim q_{(A',\lambda')}$. Thus $q \in Q(N^2)$, as desired. □

Note that Theorem 8.1 answers Question 6 for the $H(q)$'s that appear in the intersection $H_{N^2} \cap H_{m^2}$. Furthermore, this theorem allows us to give a simple proof of the first part of Theorem 4.1. Indeed, one can prove just as easily the following more general result.

**Corollary 8.2** *If $m \equiv \varepsilon \pmod 4$, where $\varepsilon \in \{0,1\}$, and if $m, N > 1$, then*

$$H_m \cap H_{N^2} \cap M_2 \neq \emptyset.$$

*Proof.* Note that $q := [N^2, 2\varepsilon N, m] \in Q(N^2) \cap Q(m)$, so $H(q) \neq \emptyset$ by Theorem 8.1. Since $q \to N^2$ and $q \to m$, we see that $H(q) \subset H_m \cap H_{N^2}$ by (4).

Moreover, since $q(x,y) = (Nx + \varepsilon y)^2 + (m - \varepsilon^2)y^2 > 1$, we see that $q \not\to 1$. Thus $H(q) \not\subset H_1$ by (5), so $H(q) \not\subset H_1 = A_2 \setminus M_2$ by Theorem 2.2(b), and hence $H(q) \cap M_2 \neq \emptyset$. Thus also $H_m \cap H_{N^2} \cap M_2 \neq \emptyset$. □

## 9 The components of $H(q)$

We now consider Question 7 which in part asks about the irreducibility of $H(q)$. The following two results give a partial answer to this question.

**Theorem 9.1** *If $q = [a, b, c] \in Q(N^2)$ and if $8 \nmid c(q) := \gcd(a, b, c)$, then $H(q)$ has at most $2^{\omega(c(q))}$ irreducible components, where $\omega(c) = |\{p \mid c\}|$ denotes the number of distinct prime divisors of $c = c(q)$.*

*In particular, if $q$ is* primitive*, i.e., if $c(q) = 1$, then $H(q)$ is an irreducible curve.*

**Theorem 9.2** *Let $q \in Q(N^2)$. If $|\mathrm{disc}(q)| > 4N^4$ and if $N$ is odd, then $H(q)$ has precisely $2^{\omega(c(q))}$ irreducible components, except when $q \sim [N^2, 0, 4d]$, for some integer $d \geq 1$.*

The proofs of these two theorems will be discussed below in §11.

**Remark 9.3** (a) If $8|c(q)$, then $H(q)$ has at most $2^{\omega(c(q))+1}$ irreducible components, and an analogue of Theorem 9.2 holds (but there are more exceptions). Moreover, the number of components can also be determined in the exceptional cases.

(b) If $q \in Q(N^2)$ has discriminant $\mathrm{disc}(q) = -16d$ with $\gcd(d, N) = 1$, then $H(q)$ is irreducible, as was proved in Theorem 4 of [K3]. In fact, that theorem also gives the *birational structure* of $H(q)$ in this case. The result is that $H(q)$ is birational to the modular curve $X_0^+(d) = X_0(d)/\langle w_d \rangle$, where $w_d$ is the *Fricke involution* on $X_0(d) = \Gamma_0(d)\backslash\mathfrak{H}$, except when $q$ is an *ambiguous form* (see below). In latter case, $H(q)$ is birational to a (non-trivial) quotient of $X_0^+(d)$, which is given explicitly in Theorem 57 of [K3].

When $N$ is an odd prime, then the number of components of the $H(q)$'s can be completely determined. However, this number depends on whether $N \equiv 1 \pmod 4$ or not, as the following result shows. To state it, recall (from Gauss) that a binary integral form $q$ is said to be *ambiguous* if $q \sim [a, ak, c]$, for some integers $a, k, c \in \mathbb{Z}$.

**Theorem 9.4** *Let $N > 2$ be a prime and let $q \in Q(N^2)$. Then $H(q)$ has at most 2 irreducible components. Moreover, if $N \equiv 1 \,(mod\, 4)$, then we have that*

$$H(q) \text{ is irreducible} \quad \Leftrightarrow \quad q \text{ is primitive, i.e., } c(q) = 1,$$

*whereas if $N \equiv 3 \,(mod\, 4)$, then we have that*

$$H(q) \text{ is irreducible} \Leftrightarrow q \text{ is primitive or ambiguous.}$$

## 9.1 Numerical examples

By using the *reduction theory* of binary quadratic forms and the above results (and others from [K7]), one can work out the intersections of Humbert surfaces with small invariants, and hence also the intersections of Lange's moduli spaces $M_2(1,n)$.

**Proposition 9.5** *The intersection of some the Humbert surfaces are as follows:*

$$\begin{aligned}
H_1 \cap H_4 &= H[1,0,4],\\
H_1 \cap H_5 &= H[1,0,4],\\
H_4 \cap H_9 &= H[4,0,5] \cup H[4,0,9] \cup H[4,4,9],\\
H_4 \cap H_{16} &= H[4,0,12] \cup H[4,0,16] \cup H[4,4,5] \cup H[4,4,8] \cup H[4,4,16],\\
H_4 \cap H_{25} &= H[1,0,4] \cup H[4,0,9] \cup H[4,0,21] \cup H[4,0,25] \cup H[4,4,17] \cup H[4,4,25].
\end{aligned}$$

*Moreover, all the $H(q)$'s listed above are irreducible. Thus, if we put as before $H^*(q) = H(q) \cap M_2$, then the intersection of some of Lange's moduli spaces are as follows:*

$$\begin{aligned}
M_2(1,2) \cap M_2(1,3) &= H_4 \cap H_9 \cap M_2\\
&= H^*[4,0,5] \cup H^*[4,0,9] \cup H^*[4,4,9],\\
M_2(1,2) \cap M_2(1,5) &= H_4 \cap H_{25} \cap M_2\\
&= H^*[4,0,9] \cup H^*[4,0,21] \cup H^*[4,0,25] \cup H^*[4,4,17] \cup H^*[4,4,25].
\end{aligned}$$

The above proposition suggests that the number of irreducible components of such intersections becomes quite large as we increase the parameters. This is also borne out by the following table which gives the number of irreducible components of $H_{N^2} \cap H_{M^2}$ for small values of $N$ and $M$:

| $N \backslash M$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| 2 | $*$ | 3 | 5 | 6 |
| 3 | 3 | $*$ | 6 | 9 |
| 4 | 5 | 6 | $*$ | 12 |
| 5 | 6 | 9 | 12 | $*$ |

**Remark 9.6** In the above table of intersections, each $H(q)$ is irreducible. But this will not be true in general, as Theorems 9.2 and 9.4 and the following table show. The latter gives the number of irreducible components of $H_{N^2} \cap H_m$, when $m \equiv 0, 1 \bmod 4$:

| $N^2 \backslash m$ | 1 | 4 | 5 | 8 | 9 | 12 | 13 | 16 | 17 | 20 | 21 | 24 | 25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | $*$ | 1 | 1 | 2 | 1 | 2 | 2 | 2 | 3 | 3 | 2 | 3 | 3 |
| 4 | 1 | $*$ | 3 | 4 | 3 | 4 | 5 | 5 | 5 | 6 | 5 | 6 | 6 |
| 9 | 1 | 3 | 3 | 5 | $*$ | 6 | 5 | 6 | 8 | 7 | $8^*$ | $10^*$ | 9 |
| 16 | 2 | 5 | 5 | 6 | 6 | 9 | 9 | $*$ | 9 | 12 | 10 | 11 | 12 |
| 25 | 3 | 6 | $7^*$ | 8 | 9 | 9 | 10 | 12 | 15 | $16^*$ | 11 | 13 | $*$ |

Here the starred entries $8^*$, $10^*$, $7^*$ and $16^*$ are those for which the intersection $H_{N^2} \cap H_m$ contains reducible $H(q)$'s. This table is discussed in more detail in [K7].

## 9.2 Application to curves with automorphisms

Let $G$ be a finite group and $0 \leq g' < g$. Then it is known that the set

$$M_g(G, g') := \{\langle X \rangle \in M_g(\mathbb{C}) : G \leq \mathrm{Aut}(X) \text{ and } g_{X/G} = g'\}$$

is a closed subset of the moduli space $M_g$; cf. Baily[Ba] and Kuribayashi[Ku].

In the case that $g = 2$, these subschemes turn out to be generalized Humbert schemes $H^*(q) = H(q) \cap M_2$, except possibly when $\dim M_2(G, g') = 0$.

**Theorem 9.7** *The G-loci $M_2(G, g')$ in $M_2$ of dimension $\geq 1$ are as follows:*

$$\begin{aligned}
M_2(C_2, 0) &= M_2 \\
M_2(C_2, 1) &= H_4^* \\
M_2(V_4, 0) &= H_4^* \\
M_2(D_4, 0) &= H^*[4, 0, 4] \\
M_2(D_6, 0) &= H^*[4, 4, 4]
\end{aligned}$$

*where $C_2 = \mathbb{Z}/2\mathbb{Z}, V_4 = C_2 \times C_2$, and $D_n$ is the dihedral group of order $2n$.*

*Proof.* The first equality is obvious since every curve $\langle X \rangle \in M_2(\mathbb{C})$ is hyperelliptic, and the second is clear from (1). The uniqueness of the hyperelliptic involution shows that $M_2(C_2, 1) = M_2(C_4, 0)$, so the third equality holds. The last 2 equalities follow directly from Theorem 4(a) of [K5].

Moreover, it follows from the discussion on p. 141 of [AP] that $M_2(G, 0)$ is either empty or consists of single point, if $G$ is not one of the above cases. Furthermore, it easy to see that $M_2(G, 1) = \emptyset$ if $G \neq C_2$. Thus, the above list is complete. □

**Remark 9.8** The curves belonging to $H_4^*$, $H^*[4, 0, 4]$ and to $H^*[4, 4, 4]$ have the following explicit equations:

(a) $y^2 = x(x-1)(x-\alpha)(x-\beta)(x-\alpha\beta)$ (Jacobi, 1832)

(b) $y^2 = x(1-x^2)(1-\kappa^2 x^2)$ (Legendre, 1832)

(c) $y^2 = x^6 + ax^3 + 1$ (Bolza, 1888)

Note that Accola and Previato [AP] give on p. 142 the associated period matrices for these curves (and others).

## 10 The structure of $H(q)$

The aim of this section is to describe all the irreducible components of the generalized Humbert schemes $H(q)$ in the case that $q \in Q(N^2)$ and thus answer Question 8. As an application, we can then prove the second part of Theorem 4.1 and also Theorem 7.1.

For this, we'll use the (modified) "basic construction map"

$$\tilde{\beta}_N = \beta_N \circ \Phi_N : Y(N) = X(N) \times X(N) \overset{\Phi_N}{\to} Z_N \overset{\beta_N}{\to} H_{N^2},$$

where, as above, $\Phi_N : X(N) \times X(N) \to \Delta_N^* \backslash (Y(N) = Z_N$ is the quotient map and $\beta_N$ is the basic construction map of §3.

Thus, since $H(q) \subset H_{N^2}$, we can expect that $H(q)$ can be described by the images of suitable curves on the product surface $X(N) \times X(N)$. These "suitable curves" are the *modular correspondences* on $Y(N) = X(N) \times X(N)$.

### 10.1 The modular correspondences $T_\alpha^N$

We begin by recalling the classical definition of the modular correspondences on the product surface $X(N) \times X(N)$ which were introduction by Klein[K] in 1880.

**Notation.** For $d \geq 1$, let $\mathcal{M}_d$ denote the set of *primitive* (integral) matrices of determinant $d$. Thus,

$$\mathcal{M}_d \;=\; \Gamma(1)\alpha_d\Gamma(1), \quad \text{where } \Gamma(1) = \mathrm{SL}_2(\mathbb{Z}), \text{ and } \alpha_d = \left(\begin{smallmatrix} 1 & 0 \\ 0 & d \end{smallmatrix}\right).$$

**Proposition 10.1 (Klein)** *For each $\alpha \in \mathcal{M}_d$ there is an irreducible curve*

$$T_\alpha^N \;\subset\; X(N) \times X(N)$$

*which depends only on the double coset $\pm\Gamma(N)\alpha\Gamma(N)$.*

*Proof.* Klein/Fricke[KF], p. 597, or Shimura[Sh], p. 170. □

**Remark 10.2** Analytically $T_\alpha^N$ is constructed as follows. Recall first that

$$X(N) = \Gamma(N)\backslash\mathfrak{H}, \quad \text{where } \Gamma(N) = \mathrm{Ker}(\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z}))$$

is the principal congruence subgroup of level $N$. Now if $\alpha \in \mathcal{M}_d$, and if $\Gamma_\alpha \subset \mathfrak{H} \times \mathfrak{H}$ denotes the graph of $\alpha$, where we view $\alpha$ as a fractional linear transformation on the upper half-plane $\mathfrak{H}$, then the modular correspondence $T_\alpha^N$ is the image of $\Gamma_\alpha$ with respect to the quotient map

$$\mathcal{H} \times\, \mathcal{H} \;\to (\Gamma(N)\backslash\mathfrak{H}) \times (\Gamma(N)\backslash\mathfrak{H}) = X(N) \times X(N).$$

## 10.2 The quadratic form of a modular correspondence

In order to connect the modular correspondences $T_\alpha^N$ with the $H(q)$'s, it is useful to associate to each $\alpha \in \mathcal{M}_d$ and $N \geq 1$ a binary quadratic form $q_\alpha^N$ in the following way.

**Notation.** For $\alpha \in \mathcal{M}_d$ and $N \geq 1$ put

$$q_\alpha^N := [N^2, 2mt, m^2(t^2+4d)/N^2].$$

Here, $t = \mathrm{trace}(\beta\alpha)$, where $\beta = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $m|N$ is given by

$$N/m = \gcd(x-w, y, z, N), \quad \text{if } \beta\alpha = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right). \tag{6}$$

**Proposition 10.3 (Key Lemma)** (a) *If $\alpha \in \mathcal{M}_d$, and $N \geq 1$, then $q_\alpha^N \in Q(N^2)$. Moreover, if $m$ is defined by* (6), *then*

$$\mathrm{disc}(q_\alpha^N) = -16m^2 d \quad \textit{and} \quad \gcd(N/m, d) = 1, \tag{7}$$

(b) *If $q \in Q(N^2)$, then there are unique positive integers $m|N$ and $d$ such that* (7) *holds, and there is a (primitive) matrix $\alpha \in \mathcal{M}_d$ such that $q \sim q_\alpha^N$.*

*Proof.* (a) This is Corollary 12 of [K5]. (Note that the statement of that corollary has a typo: "$N$-primitive" should be a replaced by "primitive".)

(b) The existence and uniqueness of $m$ and $d$ satisfying (7) is proven in Proposition 7 of [K5], and the existence of $\alpha \in \mathcal{M}_d$ with $q \sim q_\alpha^N$ follows from Proposition 15 (or from Theorem 16) of [K5]. $\square$

**Definition.** The forms $q \in Q(N^2)$ which satisfy (7) will be called *forms of type* $(N, m, d)$.

## 10.3 The structure theorem

We now come to the main structure theorem for the $H(q)$'s in the case that $q \in Q(N^2)$. This states that $H(q)$ is a union of certain *modular curves* which are obtained as the images of modular correspondences with respect to the "basic construction map" $\tilde{\beta}_N$. To justify the term "modular curve", we first observe:

**Proposition 10.4** *For each $\alpha \in \mathcal{M}_d$ and $N \geq 1$, the image*

$$\overline{T}_\alpha^N := \tilde{\beta}_N(T_\alpha^N) \subset H_{N^2} \subset A_2$$

*of the modular correspondence $T_\alpha^N$ is a closed irreducible curve on $H_{N^2}$ and on $A_2$.*

*Proof.* By Proposition 10.1 we know that $T_\alpha^N$ is a closed irreducible curve on $Y(N) = X(N) \times X(N)$, hence $\tilde{\beta}_N(T_\alpha^N)$ is also a closed and irreducible curve on $H_{N^2}$ and on $A_2$ because $\beta_N$ and hence $\tilde{\beta}_N$ is a finite morphism; cf. Theorem 3.3. $\square$

**Theorem 10.5 (Structure Theorem)** *If $q$ is a binary form of type $(N, m, d)$, then*

$$H(q) = \bigcup_\alpha \overline{T}_\alpha^N, \tag{8}$$

*where the union is over all $\alpha \in \mathcal{M}_d$ such that $q_\alpha^N \sim q$. This is a finite union because for $\alpha_1, \alpha_2 \in \mathcal{M}_d$ we have that*

$$\gamma\beta\alpha_1\gamma^{-1} \equiv \pm\beta\alpha_2 \pmod N, \text{ for some } \gamma \in \Gamma(1) \quad \Rightarrow \quad \overline{T}_{\alpha_1}^N = \overline{T}_{\alpha_2}^N. \tag{9}$$

The proof of the Structure Theorem 10.5 will be discussed in more detail in the next section. Here we observe that we can use it to complete the proofs of Theorems 4.1 and 7.1. For this, we first note the following consequence of Theorem 10.5.

**Corollary 10.6** *Let $q_i \in Q(N_i^2)$, for $i = 1, 2$. Then*

$$|H(q_1) \cap H(q_2)| = \infty \quad \Leftrightarrow \quad H(q_1) = H(q_2) \quad \Leftrightarrow \quad q_1 \sim q_2. \tag{10}$$

*Proof.* If $q_1 \sim q_2$, then $H(q_1) = H(q_2)$, and if $H(q_1) = H(q_2)$, then $H(q_1) \cap H(q_2) = H(q_1)$ is infinite, since $H(q_1)$ is a (non-empty) union of irreducible curves by Theorems 8.1 and 10.5. (Note that $q_i$ has type $(N_i, m_i, d_i)$ for some $m_i | N_i$ by Proposition 10.3(b).) Thus, it suffices to show that $|H(q_1) \cap H(q_2)| = \infty \Rightarrow q_1 \sim q_2$.

Now since $H(q_i)$ is a union of irreducible curves by Theorem 10.5, it follows that $|H(q_1) \cap H(q_2)| < \infty$ except when $H(q_1)$ and $H(q_2)$ have a common component. By Theorem 10.5, such a component has the form $\overline{T}_\alpha^N$ with $q_\alpha^N \sim q_i$, for $i = 1, 2$. But then $q_1 \sim q_2$, which proves (10). $\square$

*Proof of* Theorem 4.1. The first assertion follows from Corollary 8.2, and the second follows from (8). $\square$

*Proof of* Theorem 7.1. Since $H(q) = \emptyset$, if $r > 3$ (cf. Proposition 5.2(c)), we may assume that $r \leq 3$.

If $r = 1$, then $q = N^2x^2$ because $q \to N^2$. Thus $H(q) = H_{N^2}$ by Proposition 5.2(b), and so $H(q)$ is a closed subset of $A_2$ by Theorem 3.3 because $\beta_N$ is a finite and hence a proper morphism. Moreover, $H_{N^2}$ is irreducible (hence equidimensional) of dimension $2 = 3 - r$ by Theorem 3.3 (or by Theorem 2.2).

If $r = 2$, i.e., if $q$ is a binary quadratic form, and if $H(q) \neq \emptyset$, then $q \in Q(N^2)$, and then by Theorem 10.5 we know that $H(q)$ is a finite union of modular curves of the form $\overline{T}_\alpha^N = \tilde{\beta}_N(T_\alpha^N)$. By Proposition 10.4, these are all closed and irreducible curves on $H_{N^2}$, so $H(q)$ is a closed subset of $H_{N^2}$ which is equidimensional of dimension $1 = 3 - r$.

Finally, if $r = 3$, then by Lemma 10.7 below we have that $H(q)$ is a finite set, so $H(q)$ is equidimensional of dimension $0 = 3 - r$, if $H(q) \neq \emptyset$. $\square$

In the above proof we had used the following fact which is interesting in itself since it does not require the hypothesis that $q$ primitively represents a square.

**Lemma 10.7** *If $q$ is a ternary form, then there exists an integer $N \geq 1$ and binary quadratic forms $q_i \in Q(N^2)$ with $q_1 \not\sim q_2$ such that $H(q) \subset H(q_1) \cap H(q_2)$. In particular, $H(q)$ is a finite set.*

*Proof.* The last assertion follows from (10), so it is enough to prove the first. Moreover, if $H(q) = \emptyset$, then this assertion holds trivially by taking any two forms $q_i \in Q(N^2)$ with $q_1 \not\sim q_2$; for example, we can take $q_i = [N^2, 0, 4d_i]$, with $d_1 \neq d_2$.

Thus, assume that there exists $(A, \lambda) \in H(q)$. Then $q_{(A,\lambda)} \sim q$ (cf. the proof of Proposition 5.2(c)), so $\rho(A) = r + 1 = 4$. By the structure theorems for $\mathrm{End}(A)$, this implies that $A \sim E \times E$, where $E/\mathbb{C}$ is a CM elliptic curve; cf. [vdG], p. 207.

Thus, $A$ has an elliptic subgroup $E_1 \sim E$; put $N = (E_1.\theta_\lambda)$ as in §6. Then by Proposition 10(a) of [K4] there exists an elliptic curve $E_2/\mathbb{C}$ and an anti-isometry $\psi : E_1[N] \to E_2[N]$ such that $(A, \lambda) \simeq (J_\psi, \lambda_\psi)$ in the notation of Theorem 3.3. This means that $(E_1, E_2, \psi)$ is an $N$-presentation of $(A, \lambda)$.

To construct $q_1$ and $q_2$, let $d_i$ be two values which are primitively represented by the degree form $q_{E_1,E_2}$ of $\mathrm{Hom}(E_1, E_2)$, which is a positive binary quadratic form because $E_1 \sim E_2 \sim E$ are CM elliptic curves. By elementary quadratic form theory (cf. Cox[Co], Lemma 2.25), we can choose $d_1$ and $d_2$ such there exists a prime $p|d_2$ with $p \nmid 2Nd_1$. Then by Theorem 25 of [K5], there exists $m_i|N$ such that $q_{(A,\psi)} \to q_i$, where $q_i$ is a form of type $(N, m_i, d_i)$, so in particular, $q_i \in Q(N^2)$. Since $\mathrm{disc}(q_i) = -16m_i^2 d_i$, we see that $\mathrm{disc}(q_1) \neq \mathrm{disc}(q_2)$ because $p|\mathrm{disc}(q_2)$ but $p \nmid \mathrm{disc}(q_1)$. Thus $q_1 \not\sim q_2$. Since $q \sim q_{(A,\lambda)} \to q_i$, we have by Proposition 5.2(a) that $H(q) \subset H(q_1) \cap H(q_2)$, as desired. $\square$

# 11 Method of proof

The first main technique for studying the points lying on the Humbert surface $H_{N^2}$ is closely related to the basic construction (cf. §3). This gives rise to the following terminology which was introduced in [K4]:

**Definition.** An *$N$-presentation* of a principally polarized abelian surface $(A, \lambda)$ is a 4-tuple $(E_1, E_2, \psi, \pi)$ where $E_i/\mathbb{C}$ are elliptic curves, $\psi : E_1[N] \to E_2[N]$ is an anti-isometry, and

$$\pi : E_1 \times E_2 \to A$$

is an isogeny such that $\mathrm{Ker}(\pi) = \mathrm{Graph}(-\psi)$ and such that

$$\pi^*\theta_\lambda \equiv N(\theta_1 + \theta_2),$$

where $\theta_\lambda$ is the theta-divisor of $(A, \lambda)$ and $\theta_i = pr_i^*(0_{E_i})$, where $pr_i : E_1 \times E_2 \to E_i$ is the $i^{th}$ projection for $i = 1, 2$.

**Remark 11.1** It follows from the basic construction (cf. §3 or [K4], Proposition 10) that

$$(A, \lambda) \text{ has an } N\text{-presentation} \quad \Leftrightarrow \quad \langle A, \lambda\rangle \in H_{N^2}.$$

A second ingredient for studying points on $H_{N^2}$ is a useful *modular description* of the modular correspondences $T_\alpha^N$ and hence of their images $\overline{T}_\alpha^N$. This allows us to find an $N$-presentation of the elements $\langle A, \lambda\rangle \in \overline{T}_\alpha^N(\mathbb{C})$.

Such a modular description was determined in [K6]. Specializing this to our situation, we obtain:

**Proposition 11.2** *Let $\alpha \in \mathcal{M}_d$, and let $z := \langle E, E', \psi\rangle \in Z_N(\mathbb{C})$, where $N \geq 2$. Then $z \in \Phi_N(T_\alpha^N)$ if and only if there exists a cyclic isogeny $f : E \to E'$ of degree $d$ and a (symplectic) level $N$-structure $\phi : (\mathbb{Z}/N\mathbb{Z}) \times (Z/N\mathbb{Z}) \xrightarrow{\sim} E[N]$ such that*

$$\psi^{-1} \circ f_{|E[N]} \;=\; \phi \circ [\beta\alpha]_N \circ \phi^{-1}, \tag{11}$$

*where $[\beta\alpha]_N \in \mathrm{End}((\mathbb{Z}/N\mathbb{Z})^2)$ denotes the endomorphism of $(\mathbb{Z}/N\mathbb{Z})^2$ defined by the matrix $\beta\alpha$ mod $N$. Here, as before, $\beta = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$.*

*Proof.* This follows easily from the modular description given in Theorem 10 of [K6]. □

*Proof of* Theorem 10.5. (Sketch; full details in [K7]). Let $z = \langle E, E', \psi\rangle \in \Phi_N(T_\alpha^N)$, where $\alpha \in \mathcal{M}_d$, so $\beta_N(z) = \langle J_\psi, \lambda_\psi\rangle \in \overline{T}_\alpha^N$. Thus, , if $\pi : E \times E' \to J_\psi$ is the quotient map, then $(E, E', \psi, \pi)$ is an $N$-presentation of $(J_\psi, \lambda_\psi)$. Moreover, by Proposition 11.2 we know there exists a cyclic isogeny $f : E \to E'$ of degree $d$ such that (11) holds. It thus follows from Proposition 28 and Proposition 11 of [K5] that $q_{(J_\psi,\lambda_\psi)} \to q_\alpha^N$, and so $\overline{T}_\alpha^N \subset H(q_\alpha^N) = H(q)$, for any $q \sim q_\alpha^N$; cf. Proposition 5.2(a). This shows that the right hand side of (8) is contained in the left hand side.

Conversely, let $\langle A, \lambda\rangle \in H(q)$, so $q_{(A,\lambda)} \to q$. By using Theorem 31 of [K5] (together with [K5], Lemma 30) we can find an $N$-presentation $(E, E', \psi, \pi)$ of $(A, \lambda)$ and a cyclic isogeny $f : E \to E'$ of degree $d$ such that (11) holds for some $\alpha \in \mathcal{M}_d$ and some level $N$-structure $\phi : (\mathbb{Z}/N\mathbb{Z})^2 \to E[N]$. Thus $z = \langle E, E', \psi\rangle \in \Phi_N(T_\alpha^N)$ by Proposition 11.2 and so $\langle A, \lambda\rangle = \beta_N(z) \in \overline{T}_\alpha^N$. This proves (8).

To prove the assertion (9), note that if $\alpha_1, \alpha_2 \in \mathcal{M}_d$, then by [K6], Proposition 15 and Lemma 11, we have that

$$T_{\alpha_1}^N = T_{\alpha_2}^N \quad \Leftrightarrow \quad \alpha_2 \equiv \pm\alpha_2 \pmod N.$$

Using this and the modular description of $\Phi_N$, we see that

$$(12)\quad \Phi_N(T_{\alpha_1}^N) = \Phi_N(T_{\alpha_2}^N) \quad \Leftrightarrow \quad \gamma\beta\alpha_1\gamma^{-1} \equiv \pm\beta\alpha_2 \pmod N, \text{ for some } \gamma \in \Gamma(1),$$

and from this (9) follows because $\overline{T}_{\alpha_i}^N = \beta_N(\Phi_N(T_{\alpha_i}^N))$, for $i = 1, 2$. □

In order to deduce Theorems 9.1 and 9.2 from the Structure Theorem 10.5, further work is necessary. In view of (8) and (9), this leads to the study the following three problems:

1. Determine the $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$-conjugacy classes of the matrices $\alpha \pmod N$.
2. Study the $\pm$-action on the conjugacy classes.
3. Examine the converse of implication (9).

The solutions of these problems will now be discussed in turn.

**1. Conjugacy classes of matrices.** By an easy extension of the work of Nobs[No], who treated the case of a prime power $N = p^r$, it is possible to determine explicit representatives for the $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$-conjugacy classes of $2 \times 2$ matrices mod $N$.

**Notation.** For an integral $2 \times 2$ matrix $\alpha = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \in M_2(\mathbb{Z})$ and integer $N \geq 2$ put

$$g_{\alpha,N} = \gcd(x - w, y, z, N).$$

Note that $g_{\alpha,N}$ depends only on $\alpha \pmod N$ and that $g_{\alpha,N}$ is invariant under conjugation by $\mathrm{SL}_2(\mathbb{Z})$; cf. [K5], Corollary 13. Furthermore, for $t, d \in \mathbb{Z}$ put

$$S_N(t, d) = \{a \in (Z/N\mathbb{Z})^\times : \exists(\xi, \eta) \in \mathbb{Z}^2 \text{ with } \xi^2 + t\xi\eta + d\eta^2 \equiv a \pmod N\}.$$

As in [No], it is easy to see that $S_N(t, d)$ is a subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$.

The following result gives an overview of the $\mathrm{SL}(\mathbb{Z}/N\mathbb{Z})$-conjugacy classes of matrices $\alpha \pmod N$.

**Proposition 11.3** (a) *Let $\alpha \in M_2(\mathbb{Z})$ and $N \geq 2$, and put $g = g_{\alpha,N}$ and $m = \frac{N}{g}$. Then there exist integers $x, t, d, z, z' \in \mathbb{Z}$ with $0 \leq x < g$ and $zz' \equiv 1 \pmod m$ such that*

$$(13)\qquad \gamma\alpha\gamma^{-1} \equiv xI + g\begin{pmatrix} 0 & -dz' \\ z & t \end{pmatrix} \pmod N, \quad \textit{for some } \gamma \in \mathrm{SL}_2(\mathbb{Z}).$$

(b) *Let $g|N$ and put $m = \frac{N}{g}$. If $\alpha_1, \alpha_2 \in M_2(\mathbb{Z})$ have the form*

$$\alpha_i = x_i + g \begin{pmatrix} 0 & -d_i z_i' \\ z_i & t_i \end{pmatrix},$$

*where $0 \le x_i < g$ and $z_i z_i' \equiv 1 \pmod{m}$, for $i = 1, 2$, then the following conditions are equivalent:*

(i) *There exists $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma\alpha_1\gamma^{-1} \equiv \alpha_2 \pmod{N}$.*

(ii) *We have that $x_1 = x_2$, $t_1 \equiv t_2 \pmod{m}$, $d_1 \equiv d_2 \pmod{m}$, and $a \in S_m(t_1, d_1)$, where $a \equiv z_1 z_2' \pmod{m}$.*

This leads to the following result which refines the Structure Theorem 10.5.

**Theorem 11.4** *Let $q$ be a form of type $(N, m, d)$ and put $g = \frac{N}{m}$. Then there exist integers $x$, $t$ and $\delta$ with $0 \le x < g$ such that if $z_1, \ldots, z_s$ is a system of representatives of $(\mathbb{Z}/m\mathbb{Z})^\times / S_m(t, \delta)$, and if $\alpha_1, \ldots, \alpha_s \in \mathcal{M}_d$ are primitive matrices such that*

$$\beta\alpha_i \equiv xI + g \begin{pmatrix} 0 & -\delta z_i' \\ z_i & t \end{pmatrix} \pmod{N},$$

*where $z_i z_i' \equiv 1 \pmod{m}$, for $1 \le i \le s$, then*

$$H(q) = \bigcup_{i=1}^{s} \overline{T}_{\alpha_i}^N. \tag{14}$$

*In particular, the number of irreducible components of $H(q)$ satisfies the estimate*

$$|\mathrm{Irr}(H(q))| \le i_m(t, \delta) := |(\mathbb{Z}/m\mathbb{Z})^\times : S_m(t, \delta)|. \tag{15}$$

**Remark 11.5** The above is a shortened version of Theorem 40 of [K7], which also gives the precise recipe of how $x, t$, and $\delta$ are determined from $q$.

Thus, to find an upper bound for the number of irreducible components of $H(q)$, it suffices by (15) to bound the index $i_m(t, \delta)$. For prime powers $m$, Nobs[No] computed $S_m(t, d)$ and hence also $i_m(t, d)$. Using this, we obtain:

**Proposition 11.6** *Let $t, \delta \in \mathbb{Z}$ be integers and put $\Delta_m(t, \delta) = \gcd(t^2 - 4\delta, m)$. Then*

$$i_m(t, \delta) \le 2^{\omega(\Delta_m(t,\delta))}, \tag{16}$$

*except when $8|\Delta_m(t, \delta)$, in which case $i_m(t, \delta) \le 2^{\omega(\Delta_m(t,\delta))+1}$. Furthermore, equality holds in (16) if $2 \nmid \Delta_m(t, \delta)$.*

**Remark 11.7** In [K7] it is shown that in the situation of Theorem 11.4 we have that

$$\Delta_m(t,\delta) = \gcd(c(q), m), \tag{17}$$

where, as before, $c(q)$ is the content of the form $q$. Using this, we see that Theorem 9.1 follows immediately from (15) and (16).

**2. The $\pm$-action on conjugacy classes.** In view of the explicit representatives of the $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$-conjugacy classes of matrices given in Proposition 11.3, it is an easy exercise to determine when the $\pm$-action identifies some of the $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$-conjugacy classes or not. For example, the following result follows easily from Proposition 11.3.

**Lemma 11.8** *Let $\alpha_1, \ldots, \alpha_s \in \mathcal{M}_d$ be as in* Theorem 11.4*, and suppose that $N$ is odd. If $x \not\equiv 0$ or if $t \not\equiv 0 \pmod{m}$, then we have that*

$$-\beta\alpha_i \not\equiv \gamma\beta\alpha_j\gamma^{-1} \pmod{N}, \quad \textit{for } 1 \le i, j \le s \textit{ and all } \gamma \in \mathrm{SL}_2(\mathbb{Z}). \tag{18}$$

**Remark 11.9** (a) If $x$ and $t$ do not satisfy the conditions of Lemma 11.8, then $q \sim [N^2, 0, 4d]$, so $q$ has type $(N, N, d)$. This explains why this form plays a special role in Theorem 9.2.

(b) If $N$ is even, then there is an analogue of Lemma 11.8, but then there are more exceptional cases. In this case the exceptional $q$'s are those of the form $[N^2, 0, 4d]$, $[N^2, N^2, (\frac{N}{2})^2 + 4d]$ and $[N^2, \varepsilon N^2, (\frac{\varepsilon N}{2})^2 + d]$, where $\varepsilon \in \{0, 1\}$ and $d \equiv 1 + \varepsilon N \pmod{4}$ in the last case; cf. [K7].

**3. The converse of the implication (9).** This task is much more difficult than the previous two, and complete results are not yet available, except when $N$ is an odd prime.

To understand the difficulties, recall from the proof of Theorem 10.5 that the converse of (9) does hold for the modular curves on $Z_N$; cf. (12). However, since the basic construction map $\beta_N : Z_N \to H_{N^2}$ is generically $2:1$ (cf. Theorem 3.3), we can expect that distinct modular curves on $Z_N$ are mapped to the same modular curve $\overline{T}_\alpha^N$ on $H_{N^2}$. However, we have the following result.

**Proposition 11.10** *Every modular curve $\Phi_N(T_\alpha^N)$ on $Z_N$ is stable under the involution $w_N$ of* Theorem 3.3*. Thus, if*

$$\pi_N : Z_N \;\to\; Z_N^{sym} \;=\; \langle w_N \rangle \backslash Z_N$$

*denotes the quotient morphism, then we have for $\alpha_1, \alpha_2 \in \mathcal{M}_d$ that*

(19)

$$\pi_N\beta_N(T_{\alpha_1}^N) = \pi_N\beta_N(T_{\alpha_2}^N) \;\Leftrightarrow\; \gamma\beta\alpha_1\gamma^{-1} \equiv \pm\beta\alpha_2 \pmod{N}, \textit{ for some } \gamma \in \mathrm{SL}_2(\mathbb{Z}).$$

*Proof.* (Sketch; details in [K7].) It is not difficult to see (cf. [K6]) that $\tau_N(T_N^\alpha) = T_{\alpha^*}^N$, where $\tau$ denotes the involution of $X(N) \times X(N)$ which interchanges the factors (cf. Theorem 3.3) and $\alpha^* = \det(\alpha)\alpha^{-1}$ denotes the adjoint of $\alpha$. Thus

$$w_N(\Phi_N(T_\alpha^N)) \;=\; \Phi_N(T_{\alpha^*}^N).$$

Using Proposition 11.3 we see that $\beta\alpha^*$ is conjugate mod $N$ to $\beta\alpha$, and so we obtain that $w_N(\Phi_N(T_\alpha^N)) = \Phi_N(T_\alpha^N)$ by (12). Thus $\Phi_N(T_\alpha^N)$ is $w_N$-stable, as claimed.

From this it follows that the map $\pi_N$ maps the modular curves on $Z_N$ bijectively to those of $Z_N^{sym}$, and so (19) follows from (12). $\square$

**Remark 11.11** Since $\beta_N : Z_N \to H_{N^2}$ is $w_N$-invariant, it factors over $\pi_N$ as

$$\beta_N \;=\; \nu_N \circ \pi_N,$$

where $\nu_N : Z_N^{sym} \to H_{N^2}$ is finite and birational; cf. Theorem 3.3. Thus, $\nu_N$ is generically an isomorphism, and so only finitely many of the modular curves on $Z_N^{sym}$ are not mapped bijectively to $H_{N^2}$; these then lie in the inverse image of the *conductor locus* of $H_{N^2}$. The following result and Remark 11.13 below make this precise.

**Theorem 11.12** *Let $q$ be a form of type $(N, m, d)$ which satisfies the condition*

$$|\{(x,y) \in \mathbb{Z}^2 : q(x,y) = N^2, \gcd(x,y) = 1\}| \;=\; 2. \tag{20}$$

*Then the converse of (9) holds for the matrices $\alpha_i \in \mathcal{M}_d$ with $q_{\alpha_i}^N \sim q$.*

**Remark 11.13** If $d > N^4/(4m^2)$, then the reduction theory of binary quadratic forms shows that (20) holds. Thus, there are only finitely many $H(q)$'s on $H_{N^2}$ which fail to satisfy the condition (20) because for them the absolute value of the discriminant of $q$ is bounded by $4N^4$, and so there are only finitely equivalence classes of such $q$'s.

*Proof of* Theorem 11.12. (Sketch; details in [K7]). Suppose that $\overline{T}_{\alpha_1}^N = \overline{T}_{\alpha_2}^N$, and let $\langle A, \lambda\rangle \in \overline{T}_{\alpha_i}^N \subset H(q)$ be a non-CM point, so $q_{(A,\lambda)} \sim q$. Then the hypothesis $\overline{T}_{\alpha_1}^N = \overline{T}_{\alpha_2}^N$ implies that there exist $z_i \in \Phi_N(T_{\alpha_i}^N)$ such that $\beta(z_1) = \beta_N(z_2) = \langle A, \lambda\rangle$. Using the hypothesis (20), one then shows that in fact $z_1 = z_2$. Since $T_{\alpha_i}^N$ has infinitely many non-CM points, it follows that $\Phi_N(T_{\alpha_1}^N) = \Phi_N(T_{\alpha_2}^N)$, and so the converse of (9) follows by using (12). $\square$

*Proof of* Theorem 9.2. If $q \in Q(N^2)$, then $q$ has type $(N, m, d)$ for some unique $m$, $d$ by Proposition 10.3(b), so by (7) and the hypothesis we have $16m^2 d = |\mathrm{disc}(q)| > 4N^4$, or $d > N^4/(4m^2)$. Thus, by Remark 11.13 we have that $q$ satisfies (20). Thus, by

Theorem 11.12 we know that the converse of (9) holds for the $\alpha_i \in \mathcal{M}_d$'s such that $q^N_{\alpha_i} \sim q$.

Moreover, since by hypothesis $q \not\sim [N^2, 0, 4d]$, we have by Remark 11.9(a) and Lemma 11.8 and the converse of (9) that the $\alpha_i$'s of Theorem 11.4 give distinct modular curves $\overline{T}^N_{\alpha_i}$ on $H_{N^2}$, and so by Theorem 11.4 we see that $|\mathrm{Irr}(H(q))| = i_m(t, \delta)$, and so the desired formula follows from Proposition 11.6 and Remark 11.7. $\square$

In the case that $q$ does not satisfy condition (20), the analysis of the converse of (9) is much more delicate. In the case of prime level $N > 2$, we have the following technical result which (almost) suffices to prove Theorem 9.4.

**Proposition 11.14** *Let $q \in Q(N, N, d)$, where $N > 2$ is prime. If condition* (20) *does not hold for $q$, then $N \nmid d$ and $N \mid c(q)$, so $\beta_N^{-1}(H(q)) \subset Z_N$ consists of two irreducible components. On the other hand, we have that*

$$(21) \qquad H(q) \ \text{is irreducible} \quad \Leftrightarrow \quad N \equiv 3 \,(\mathrm{mod}\, 4).$$

*Proof.* (Sketch; details in [K8].) This is rather technical. The first two assertions are easily verified. Thus $\omega(\gcd(c(q), N)) = 1$, so we have by (12) and a refinement of Theorem 11.4 and Proposition 11.6 that $\beta_N^{-1}(H(q))$ consists of two distinct components $\Phi_N(T^N_{\alpha_i})$, $i = 1, 2$. Thus, the matrices $\alpha_i \in \mathcal{M}_d$ are such that $q \sim q^N_{\alpha_1} \sim q^N_{\alpha_2}$, but $\beta\alpha_1$ is not mod $N$-conjugate to $\pm\beta\alpha_2$.

To prove (21), suppose that $z_i \in \Phi_N(T^N_{\alpha_i})$ are non-CM points such that $\langle A, \lambda\rangle = \beta(z_1) = \beta(z_2)$, so $(A, \lambda)$ has two distinct $N$-presentations. A long and delicate computation using these two $N$-presentations shows that $\beta\alpha_1$ is mod $N$-conjugate to $\pm\beta\alpha_2$ if and only $p \equiv 1 \,(\mathrm{mod}\, 4)$. Thus, if $p \equiv 1 \,(\mathrm{mod}\, 4)$, then we have a contradiction, so in particular $\beta_N(T^N_{\alpha_1})$ and $\beta_N(T^N_{\alpha_2})$ are distinct (and intersect only in CM-points). On the other hand, if $p \equiv 3 \,(\mathrm{mod}\, 4)$, then reversing the above argument shows that each $\langle A, \lambda\rangle \in H(q)$ (which is not a CM point) is the image of a point $z_i \in \Phi_N(T^N_{\alpha_i})$ on each of the two components, so $\overline{T}^N_{\alpha_1} = \overline{T}^N_{\alpha_2} = H(q)$. $\square$

To connect the the above hypotheses with those of Theorem 9.4, we observe the following fact which is proved in [K8].

**Lemma 11.15** *If $q$ be an imprimitive form of type $(N, m, d)$, where $N > 2$ is prime, then $m = N$. If in addition $q \not\sim [N^2, 0, 4d]$, then condition* (20) *does not hold for $q$ if and only if $q$ is ambiguous.*

**Example 11.16** The generalized Humbert scheme $H[9, 6, 9] \subset H_9$.

The form $q = [9, 6, 9]$ has type $(3, 3, 2)$ and condition (20) fails for $q$ because $\pm(1, 0)$ and $\pm(0, 1)$ are primitive solutions of $q(x, y) = 9$. (Note also that $q \sim [24, 24, 9]$, so $q$

is an ambiguous form.) Thus, Proposition 11.14 shows that the converse of (9) does not hold for $q$, and that $H(q)$ is an irreducible curve; cf. also Theorem 9.4.

Note that (the proof of) Proposition 11.14 implies that $H(q)$ lies in the *conductor locus* of $H_9$,i.e., in the (closed) subset of non-normal points of $H_9$, so in particular $Z_3^{sym} \not\subset H_9$, and so $H_9$ is *not normal.* In fact, it turns out that $H_{N^2}$ is not normal for any $N \geq 3$; cf. [K8].

# References


[AP] R. Accola, E. Previato, Covers of Tori: Genus 2. *Letters for Math. Phys.* **76** (2006), 135–161.

[Ba] W. Baily, On the automorphism group of a generic curve of genus 2. *J. Math. Kyoto Univ.* **1** 1961/1962, 101-108; correction p. 325.

[Co] D. Cox, *Primes of the Form $x^2 + ny^2$*. Wiley & Sons, New York, 1989.

[FPR] M. Franciosi, R. Pardini, S. Rollenske, $(p, d)$-elliptic curves of genus 2. Preprint, 16pp. ArXiv:1611.06756v1.

[FK1] G. Frey, E. Kani, Curves of genus 2 covering elliptic curves and an arithmetical application. In: *Arithmetic Algebraic Geometry* (G. van der Geer, F. Oort, J. Steenbrink, eds.), Progress In Math. vol. 89, Birkhäuser, Boston, 1991, pp. 153-176.

[FK2] G. Frey, E. Kani, Curves of genus 2 and associated Hurwitz spaces. *Contemp. Math.* **487** (2009), 33–81.

[He] E. Hecke, Höhere Modulfunktionen und ihre Anwendung auf die Zahlentheorie. *Math. Ann.* **71** (1912), 1–37 = Gesammelte Werke, pp. 21–57.

[H1] C. F. Hermann, Modulflächen quadratischer Diskriminante. *Manuscr. Math.* **72** (1991), 95–110.

[H2] C. F. Hermann, Symmetrische Modulflächen der Diskriminante $p^2$. *Manuscr. Math.* **76** (1992), 169–179.

[Hu] G. Humbert, Sur les fonctions abéliennes singulières. I–III. *J. de Math.* (ser. 5) **5** (1899), 233–350; **6** (1900), 279–386; **7** (1901), 97–123.

[K1] E. Kani, Elliptic curves on abelian surfaces. *Manusc. math.* **84** (1994), 199–223.

[K2] E. Kani, The Hurwitz space of genus 2 covers of an elliptic curve. *Collect. Math.* **54** (2003), 1–51.

[K3] E. Kani, The moduli spaces of Jacobians isomorphic to a product of two elliptic curves. *Collect. Math.* **67** (2016), 21–54.

[K4] E. Kani, Elliptic subcovers of a curve of genus 2 I. The isogeny defect. *Ann. Sci. Math. Québec* **43** (2019), 281–303.

[K5] E. Kani, Elliptic subcovers of a curve of genus 2 II. The refined Humbert invariant. *J. Number Theory* **193** (2018), 302–335.

[K6] E. Kani, Modular correspondences on $X(N)$. Preprint, 22 pages.

[K7] E. Kani, Generalized Humbert schemes and intersections of Humbert surfaces. Preprint, 35pp.

[K8] E. Kani, Modular curves on Humbert Surfaces. Preprint (in preparation).

[KS] E. Kani and W. Schanz, Modular diagonal quotient surfaces. *Math. Z.* **227** (1998), 337–366.

[K] F. Klein, Zur Theorie der elliptischen Modulfunctionen. *Math. Ann.* **17** (1880), 62–70.

[KF] F. Klein, R. Fricke, *Vorlesungen über die Theorie der elliptischen Modulfunktionen II.* Teubner, Leipzig, 1892.

[Kr] A. Krazer, *Lehrbuch der Thetafunktionen.* Teubner, Leipzig, 1903; Chelsea Reprint, New York, 1970.

[Ku] A. Kuribayashi, On analytic families of compact Riemann surfaces with non-trivial automorphisms. *Nagoya Math. J.* **28** (1966), 119-165.

[L1] H. Lange, Über die Modulschemata der Kurven vom Geschlecht 2 mit 1, 2 oder 3 Weierstraßpunkten. *J. reine angew. Math.* **277** (1975), 27–36.

[L2] H. Lange, Moduli spaces of algebraic curves with rational maps. *Math. Proc. Cambridge Philos. Soc.* **78** (1975), 283–292.

[L3] H. Lange, Über die Modulvarietät der Kurven vom Geschlecht 2. *J. reine angew. Math.* **281** (1976), 80–96.

[L4] H. Lange, Kurven mit rationaler Abbildung. *J. reine angew. Math.* **295** (1977), 80–115.

[LB] H. Lange, Ch. Birkenhake, *Complex Abelian Varieties.* Springer-Verlag, Berlin, 1992.

[M1] D. Mumford, *Geometric Invariant Theory.* Springer-Verlag, Berlin, 1966.

[M2] D. Mumford, *Abelian Varieties.* Oxford U Press, Oxford, 1970.

[No] A. Nobs, Die irreduziblen Darstellungen der Gruppen $\mathrm{SL}_2(\mathbb{Z}_p)$, insbesondere $\mathrm{SL}_2(\mathbb{Z}_2)$. 1. Teil. *Comment. Math. Helvetici* **39** (1977), 465–489.

[Sh] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions.* Princeton University Press, Princeton, NJ, 1971.

[Si] J. Silverman, *The Arithmetic of Elliptic Curves.* Springer-Verlag, New York, 1986.

[vdG] G. van der Geer, *Hilbert Modular Surfaces.* Springer-Verlag, Berlin, 1988.

[We] A. Weil, Zum Beweis des Torellischen Satzes. *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Klasse* 1954, = Œuvres II, pp. 307–327.