Modular Diagonal Quotient Surfaces

E. Kani and W. Schanz

Abstract. The main objective of this paper is to analyze the geometry of the *modular diagonal quotient surface* $Z_{N,\varepsilon} = \Delta_\varepsilon \backslash (X(N) \times X(N))$ which classifies pairs of elliptic curves E_1 and E_2 together with an isomorphism of "determinant ε" between their associated modular representations mod N. In particular, we calculate some of the numerical invariants of its minimal desingularization $\tilde{Z}_{N,\varepsilon}$ such as its Betti and Chern numbers and determine its place in the Enriques-Kodaira classification table.

Introduction

Let $X(N) = \Gamma(N)\backslash\mathfrak{H}^*$ denote the modular curve of level N which admits the group $G_N = \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ as a subgroup of its automorphism group. Then $G_N \times G_N$ acts on the product surface $Y_N = X(N) \times X(N)$, and hence so does the "graph subgroup" $\Delta_\varepsilon = \{(g, \alpha_\varepsilon(g)) : g \in G_N\}$ associated to the automorphism $\alpha_\varepsilon \in \mathrm{Aut}(G)$ which is defined by conjugation by the element $Q_\varepsilon = \left(\begin{smallmatrix} \varepsilon & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Gl}_2(\mathbb{Z}/N\mathbb{Z})$. We propose to call the resulting quotient surface $Z_{N,\varepsilon} = \Delta_\alpha \backslash Y$ a *(twisted) modular diagonal quotient surface*; it is a special case of the general diagonal quotient surfaces studied in [10].

The modular diagonal quotient surfaces occur naturally as the compactifications of the (coarse) moduli spaces associated to certain moduli problems, as will be explained below. Furthermore, there is a close analogy between these surfaces and the Hilbert modular surfaces studied by Hirzebruch and others (cf. van der Geer[3]).

Since Δ_ε has finitely many fixed points on Y_N, the quotient surfaces $Z_{N,\varepsilon}$ have finitely many isolated quotient singularities which may be described explicitly (cf. Theorem 2.1 and Corollary 2.4 below); in particular, we have:

Theorem 1 *If $N \geq 5$, the modular diagonal quotient surface $Z_{N,\varepsilon}$ has precisely $r_0 + r_1 + r_\infty$ singularities, where $r_0 = h(-4N^2)$ and $r_1 = h(-3N^2)$ are certain class numbers of binary quadratic forms, and $r_\infty = \sum_{1<d|N} \phi(d)\phi(\frac{N}{d})$ denotes the number of cusps of $X_1(N)$. All are cyclic quotient singularities.*

The above theorem is an easy consequence of the general results of [10], once certain *local invariants* have been calculated; this is done in section 1. In the same way we can determine some of the numerical invariants of its minimal desingularization $\tilde{Z}_\varepsilon$ (cf. Theorem 2.6):

Theorem 2 *The desingularization $\tilde{Z}_{N,\varepsilon}$ of the modular diagonal quotient surface $Z_{N,\varepsilon}$ is a regular surface whose geometric genus $p_{g,\varepsilon}$ and Chern numbers $K_\varepsilon^2 = c_1^2(\tilde{Z}_{N,\varepsilon})$ and $\chi_{top,\varepsilon} = c_2(\tilde{Z}_{N,\varepsilon})$ are given by*

$$\begin{aligned} p_{g,\varepsilon} &= \frac{m(N-12)}{144N} + \frac{\phi(N)}{8} + \frac{r_0}{8} + \frac{2r_1}{9} - \frac{s_{1,1,\varepsilon}}{9} + \frac{r_\infty}{3} + \frac{\mathbb{L}_{\infty,\varepsilon}}{12} - \frac{R_{\infty,\varepsilon}}{12} - 1, \\ K_\varepsilon^2 &= \frac{m(N-12)}{18N} + \phi(N) - \frac{s_{1,1,\varepsilon}}{3} + 3r_\infty - R_{\infty,\varepsilon}, \\ \chi_{top,\varepsilon} &= \frac{m(N-12)}{36N} + \frac{\phi(N)}{2} + \frac{3r_0}{2} + \frac{8r_1}{3} - s_{1,1,\varepsilon} + r_\infty + \mathbb{L}_{\infty,\varepsilon}, \end{aligned}$$

provided that $N \geq 5$. Here, $m = |G_N|$, and r_0, r_1 and r_∞ are as in Theorem 1. In addition, $s_{1,1,\varepsilon} = \frac{1}{2}r_1$ if $3 \not| N$, and $s_{1,1,\varepsilon} = \frac{1}{2}(1+(\frac{-3}{\varepsilon}))r_1$, if $3|N$, and $\mathbb{L}_{\infty,\varepsilon}$ and $R_{\infty,\varepsilon}$ are certain sums involving the lengths of the finite continued fraction expansions of $\frac{N}{q}$, where $0 < q < N$.

It is, however, much more difficult to determine the Kodaira dimension κ of $\tilde{Z}_{N,\varepsilon}$ since there does not seem to be an easy method for computing it (cf. the discussion of [10], section 4.) Nevertheless, it turns out rather surprisingly that κ and even the type of the surface $\tilde{Z}_{N,\varepsilon}$ is completely characterized by its geometric genus (cf. Theorem 2.11):

Theorem 3 *The Kodaira dimension of the modular diagonal quotient surface is given by*

$$\kappa(\tilde{Z}_{N,\varepsilon}) = \min(2, p_{g,\varepsilon} - 1).$$

Despite the simplicity of the result, its proof is not so direct, for we succeded in proving it only as a consequence of the following much more precise result. To state it, we adopt the convention that we view ε as an element of the quotient group $(\mathbb{Z}/N\mathbb{Z})^\times/((\mathbb{Z}/N\mathbb{Z})^\times)^2$: this is permissible because if $\varepsilon' \equiv \varepsilon d^2 \pmod N$, then $\alpha_{\varepsilon'}$ and α_ε differ by an inner automorphism of G, and so the resulting surfaces are isomorphic.

Theorem 4 a) *$\tilde{Z}_{N,\varepsilon}$ is a rational surface if and only if $p_g(\tilde{Z}_{N,\varepsilon}) = 0$, and this is the case precisely for $N \leq 5$ or for $(N,\varepsilon) = (6,1)$, $(7,1)$ or $(8,1)$.*

b) *$\tilde{Z}_{N,\varepsilon}$ is a (blown-up) elliptic K3-surface if and only if $p_g(\tilde{Z}_{N,\varepsilon}) = 1$, i.e. if and only if $(N,\varepsilon) = (6,5)$, $(7,3)$, $(8,3)$, $(8,5)$, $(9,1)$ or $(12,1)$.*

c) *$\tilde{Z}_{N,\varepsilon}$ is a (blown-up) elliptic surface with $\kappa = 1$ if and only if $p_g(\tilde{Z}_{N,\varepsilon}) = 2$. This is the case for $(N,\varepsilon) = (8,7)$, $(9,2)$, $(10,1)$, $(10,3)$ or $(11,1)$.*

d) *$\tilde{Z}_{N,\varepsilon}$ is a surface of general type if and only if $p_g(\tilde{Z}_{N,\varepsilon}) \geq 3$, or equivalently, if $N \geq 13$ or if $(N,\varepsilon) = (11,2)$, $(12,5)$, $(12,7)$ or $(12,11)$.*

It is interesting to observe that almost all of Theorem 4 follows relatively easily from the general criteria of [10], with the exception of the two cases $(N,\varepsilon) = (10,3)$ and $(11,1)$ which require a more detailed analysis (cf. Proposition 2.14).

As was already mentioned above, the modular diagonal quotient surfaces $Z_{N,\varepsilon}$ have a natural modular interpretation. Indeed, by using the well-known modular interpretation of $X'(N) = \Gamma\backslash\mathfrak{H} = X(N)\backslash\{cusps\}$, it is easy to see that the points of $Z'_{N,\varepsilon} := \Delta_\varepsilon\backslash(X'(N) \times X'(N)) \subset Z_{N,\varepsilon}$ can naturally be identified with triplets (E_1, E_2, ψ), where E_1 and E_2 are elliptic curves and $\psi : E_1[N] \xrightarrow{\sim} E_2[N]$ is an isomorphism of the N-torsion subgroups $E_i[N]$ of determinant ε; the latter means that the Weil pairings are related by the formula

$$e_{E_1,N} \circ (\psi \times \psi) = e^{\varepsilon}_{E_2,N}.$$

Furthermore, by using Galois descent it is not difficult to see that the surface $Z_{N,\varepsilon}$ has a "canonical model" over $\mathbb{Q}$ (and even over $\mathrm{Spec}(\mathbb{Z}[\frac{1}{N}])$ which serves as a coarse moduli scheme for the moduli functor $\mathcal{Z}_{N,\varepsilon}$ defined by

$$\begin{aligned}\mathcal{Z}_{N,\varepsilon}(S) \;=\; \{&(E_1, E_2, \psi) : E_1,\, E_2 \text{ are elliptic curves over } S \text{ and } \psi : E_1[N] \xrightarrow{\sim} E_1[N] \\ &\text{is an } S\text{-isomorphism with } \det(\psi) = \varepsilon\}/(\text{isomorphisms}).\end{aligned}$$

Note that if $S = \mathrm{Spec}(K)$, K a number field, then this set may be identified with the set of isomorphism classes of pairs of elliptic curves together with an isomorphism ψ of the associated modular Galois representations $\bar{\rho}_{E_i,N} :\ \mathrm{Gal}(\bar{K}/K) \to\ \mathrm{Aut}(E_i[N])$.

As a result of this modular interpretation, the above theorem has several interesting applications which are worth mentioning.

First of all, the above theorem has an interesting connection with a *question of Mazur* which was studied by Kraus and Oesterlé [12]. In this question Mazur [15] poses the problem of finding examples of pairs (E, E') of non-isogenous elliptic curves over $\mathbb{Q}$ such that for some integer $N \geq 7$ their associated Galois representations on their N-torsion points are *symplectically* isomorphic, and in [12] such an example is exhibited for $N = 7$. In view of the above modular interpretation, Mazur's question can be rephrased as asking for the existence of $\mathbb{Q}$-rational points on modular diagonal quotient surface $Z_{N,1}$ which do not lie on any "obvious" curves on $Z_{N,1}$ (such as the curves at infinity or those coming from the relation of isogeny, i.e. on "graphs of Hecke correspondences"). Viewed in this light, the above theorem therefore shows that the Kraus-Oesterlé example should come as no surprise: indeed, since $Z_{7,1}$ is rational (over $\bar{\mathbb{Q}}$) by Theorem 4, there should exist infinitely many such examples, at least over a sufficiently large number field K. More generally, one might expect the following conjecture to be true:

Conjecture 5 *For any number field K and any $N \leq 12$, there are infinitely many pairs (E_1, E_2) of non-isogenous elliptic curves defined over K whose associated Galois representations mod N are isomorphic.*

Theorem 4 gives partial evidence for this: since for $N \leq 12$, at least one of the surfaces $\tilde{Z}_{N,\varepsilon}$ is either rational or an elliptic fibring, the above conjecture is true if we replace "any number field" by "any sufficiently large number field" and "non-isogeneous" by "not isogenous by an isogeny of degree $\leq d$" (for any fixed d).

It is an intriguing and important question to determine to what extent the converse of Conjecture 5 is valid, i.e.

Question 6 *For which integers N is it true that there are only finitely many pairs (E, E') (up to $\bar{\mathbb{Q}}$-isomorphism) of non-isogenous elliptic curves defined over a number field K whose associated Galois representations mod N over K are isomorphic?*

Indeed, H. Darmon and others have conjectured that this question has a positive answer for all sufficiently large N, and this would have far-reaching diophantine consequences such as the Asymtotic Fermat Conjecture (cf. Frey[2]).

Again, Theorem 4 gives some partial evidence for this, particularly if we admit the conjecture of S. Lang [14] that almost all of the K-rational points of an algebraic surface Z of general type are contained in the "exceptional locus" which by definition is the union of the (conjecturally finitely many) curves of genus ≤ 1 lying on Z. What remains to be done, however, is to identify these curves as Hecke correspondences (at least when N is large), but this seems to be a rather formidable problem!

Yet another application concerns the moduli space $\mathcal{M}_2^{ell}(N)$ of curves of genus 2 with an elliptic differential of degree N which was defined and studied in [8] (see also [7]). By definition, $\mathcal{M}_2^{ell}(N)$ classifies isomorphism classes of pairs $\langle C, f\rangle$ consisting of a curve C of genus 2 and a (minimal) covering $f : C \to E$ of degree N to an elliptic curve E, and it can be shown (cf. [9]) that this space is a double cover of the Humbert surface with invariant $\Delta = N^2$ (as defined in van der Geer [3], p. 210ff; cf. also [9]). More precisely, one can prove (cf. [8]) that $\mathcal{M}_2^{ell}(N)$ is an open subvariety of the twisted diagonal quotient surface $Z_{N,-1}$, and so we obtain from Theorem 4:

Theorem 7 *The (desingularization of the) compactification $\overline{\mathcal{M}}_2^{ell}(N) = Z_{N,-1}$ of the moduli space $\mathcal{M}_2^{ell}(N)$ is rational if $N \leq 5$, is a (blown-up) elliptic K3-surface if $N = 6$ or 7, is an elliptic surface (with $\kappa = 1$) if $8 \leq N \leq 10$, and is of general type if $N \geq 11$.*

After we had proved a preliminary version of Theorems 4 and 7, R. Weissauer drew our attention to the article of C. F. Hermann [5] who had also determined the classification type of the surfaces $Z_{N,\varepsilon}$. Unfortunately, the interesting connection between the geometric genus p_g of $\tilde{Z}_{N,\varepsilon}$ and the Enriques-Kodaira classification type is not mentioned there; in addition, the (more subtle) case $(N, \varepsilon) = (10, 3)$ is missing in Hermann's list of elliptic surfaces.[1] However, it should be emphasized that our method is substantially different from that of Hermann in that most of Theorem 6 is deduced from the general classification Theorems 1 – 5, and only in two cases do we need to analyze the surfaces in more detail. This should be compared to the 14 cases or so which would have to be studied in detail in Hermann's article (but were not since virtually all details of his proof were suppressed). In particular, we do not need Hermann's curves F_n (which admittedly are of interest in connection with Mazur's problem) for the proof of the above theorems, with the possible exception of the case $(N, \varepsilon) = (11, 1)$. Furthermore, there are a number of typographical errors and other inaccuracies in his paper; cf. Remarks 1.7, 2.2 and 2.12 below.

Acknowledgements: The authors are very thankful for helpful and stimulating discussions with G. Frey, A. Geramita, E. Viehweg, R. Weissauer, and J. Zarhin. The second author wants to thank Queen's University for its hospitality during a visit supported by NSERC and ARC research grants held by the first author.

This research was supported in part by the Natural Science and Engineering Research Council of Canada (NSERC) and also by the Advisory Research Council (ARC) of Queen's University.

[1] After this manuscript was completed, Hermann showed the first author a copy of his (unpublished) *Habilitationsschrift* in which both these *lacunae* were correctly addressed.

1 Invariants associated to the modular curve $X(N)$

1.1 The genus of X_k

For each $N \geq 1$, the modular curve $X(N)$ is the compactification of the quotient $\Gamma(N)\backslash\mathfrak{H}$ of the upper half plane $\mathfrak{H}$ by the action of the principal congruence subgroup $\Gamma(N) = \{A \in \mathrm{SL}_2(\mathbb{Z}) : A \equiv 1 \bmod N\}$, where $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(1)$ acts on $z \in \mathfrak{H}$ by the linear transformation $z \mapsto \frac{az+b}{cz+d}$. For the basic properties of $X(N)$ we refer to Schoeneberg [18] or Miyake[16]. The action of $\Gamma(1)$ on $\mathfrak{H}$ induces an action of $G = G_N = \Gamma(1)/\pm\Gamma(N) \simeq \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ on $X(N)$, and the quotient $\bar{X} = G_N\backslash X(N) = X(1) \simeq \mathbb{P}^1$ is rational.

The quotient map $\pi : X(N) \to \bar{X}$ is ramified precisely at the points $\bar{P}_0, \bar{P}_1$ and $\bar{P}_\infty \in \bar{X}$ which are the images of the points $i, \rho = e^{\frac{2\pi i}{3}}$ and $\infty \in \mathfrak{H}^* = \mathfrak{H} \cup \mathbb{Q} \cup \{\infty\}$, respectively. The corresponding ramification indices $e_k = e_{\bar{P}_k}$ are $e_0 = 2$, $e_1 = 3$ and $e_\infty = N$. More precisely, if $P_k = P_{k,N}$ $(k = 0, 1, \infty)$ denote (respectively) the images of i, ρ and $\infty \in \mathfrak{H}^*$ in $X(N)$, then the stabilizers $G_k = G_{P_k} = \langle\sigma_{k,N}\rangle$ of $P_{k,N}$ are cyclic groups of order e_k with generators

$$\sigma_0 = \sigma_{0,N} := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_1 = \sigma_{1,N} := \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}, \quad \sigma_\infty = \sigma_{\infty,N} := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Since the level N will be fixed from now on , we shall suppress the dependence on N in our notation, as long as this does not lead to any confusion.

For $\varepsilon \in (\mathbb{Z}/N\mathbb{Z})^\times$ we let α_ε denote the automorphism of G_N induced by conjugation with the matrix $Q_\varepsilon = \left(\begin{smallmatrix} \varepsilon & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Gl}_2(\mathbb{Z}/N\mathbb{Z})$; i.e.

$$\alpha_\varepsilon(g) \;=\; Q_\varepsilon g Q_\varepsilon^{-1} \;=\; \begin{pmatrix} a & \varepsilon b \\ \varepsilon^{-1}c & d \end{pmatrix}, \quad \text{if } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

As was explained in the introduction, we want to study the geometry of the diagonal quotient surface $Z = (X(N) \times X(N))/G_N$, as well as that of its twists $Z_\varepsilon = Z_{\alpha_\varepsilon}$. From [10] we know that the fundamental invariants (Betti, Hodge and Chern numbers) of the desingularization $\tilde{Z}_\varepsilon$ of Z_ε are determined by three basic invariants $\mathbb{G}_\varepsilon = \mathbb{G}_{\alpha_\varepsilon}, \mathbb{S}_\varepsilon = \mathbb{S}_{\alpha_\varepsilon}$ and $\mathbb{L}_\varepsilon = \mathbb{L}_{\alpha_\varepsilon}$ (together with the genera $g = g(X)$ and $\bar{g} = g(\bar{X}) = 0$), which therefore have to be computed here. Since all three are defined by invariants involving the quotient curves $X_{k,\varepsilon} = \langle\alpha_\varepsilon^{-1}(\sigma_{k,N})\rangle\backslash X(N)$, we begin by calculating their genera $g_{k,\varepsilon} = g(X_{k,\varepsilon})$. For this, let $r_{k,\varepsilon}$ denote the number of ramification points on $X_{k,\varepsilon}$ of the covering $\pi_{k,\varepsilon} : X \to X_{k,\varepsilon}$ and $t_{k,\varepsilon}$ the number of fixed points of the group $G_{k,\varepsilon} = \langle\alpha_\varepsilon^{-1}(\sigma_k)\rangle$ on X. Then by the Riemann-Hurwitz formula (together with Schoeneberg [10], p. 76) we have:

Proposition 1.1 *If $N \geq 3$, the genus $g = g(X(N))$ of $X(N)$ and the genera $g_{k,\varepsilon}$ of its quotients $X_{k,\varepsilon} = G_{k,\varepsilon}\backslash X(N)$ are given by*

$$\begin{aligned} g &= 1 + m\frac{N-6}{12N}, \qquad \text{where } m = |G_N| = \frac{N^3}{2}\prod_{p|N}\left(1 - \frac{1}{p^2}\right), \\ g_{0,\varepsilon} &= \tfrac{1}{2}\left(g + 1 - \tfrac{1}{2}r_{0,\varepsilon}\right), \qquad g_{1,\varepsilon} = \tfrac{1}{3}\left(g + 2 - r_{1,\varepsilon}\right), \\ g_{\infty,\varepsilon} &= \frac{g-1}{N} + 1 - \frac{1}{2}r_{\infty,\varepsilon} + \frac{1}{2N}t_{\infty,\varepsilon}. \end{aligned}$$

We next want to compute the numbers $r_{k,\varepsilon}$ and $t_{k,\varepsilon}$, as well as the related invariant $r_{n,\varepsilon}(\bar{P}_k)$, which denotes the number of points $x \in X_{k,\varepsilon}$ which have ramification degree $e_x(\pi_{k,\varepsilon}) = n$. It turns out that these are closely related to the following expressions which may be interpreted as class numbers $h(d)$ of binary quadratic forms of discriminant d (cf. Hua [6], pp. 321-2):

Notation 1.2 For a positive integer $N > 1$ put

$$\begin{aligned} h(-4N^2) &= \tfrac{1}{2}N \prod_{p|N} \left(1 - \frac{1}{p}\left(\frac{-4}{p}\right)\right), \\ h(-3N^2) &= \tfrac{1}{3}N \prod_{p|N} \left(1 - \frac{1}{p}\left(\frac{-3}{p}\right)\right), \end{aligned}$$

where $\left(\frac{d}{p}\right)$ denotes the Legendre/Kronecker symbol; in particular, $\left(\frac{-4}{2}\right) = 0$ and $\left(\frac{-3}{2}\right) = -1$.

The relation between these expressions and the invariants $r_{k,\varepsilon}$, $t_{k,\varepsilon}$ and $r_{n,\varepsilon}(\bar{P}_k)$ is given by the following result which will be proved in the next section as a corollary of the computation of the finer "local invariants" $s_{\nu,\varepsilon}(\bar{P}_l, \bar{P}_k)$.

Proposition 1.3 *For $N \geq 5$, the isomorphism class of the curve $X_k = X_{k,\varepsilon}$ does not depend on $\varepsilon \in (\mathbb{Z}/N\mathbb{Z})^\times$ and hence neither does its genus $g_k = g_{k,\varepsilon}$. Furthermore, the same is true for the invariants $r_k = r_{k,\varepsilon}$, $t_k = t_{k,\varepsilon}$ and $r_n(\bar{P}_k) = r_{n,\varepsilon}(\bar{P}_k)$; these are given by*

a) $r_0 = t_0 = h(-4N^2)$,

b) $r_1 = t_1 = h(-3N^2)$,

c) $r_n(\bar{P}_\infty) = \frac{1}{2}\phi(n)\phi(\frac{N}{n})$, *for* $1 < n \mid N$,

d) $r_\infty = \frac{1}{2}\sum_{n|N} \phi(n)\phi(\frac{N}{n}) - \frac{1}{2}\phi(N)$ *and* $t_\infty = \frac{m}{N} - \frac{1}{2}N\phi(N)$.

Remark 1.4 Some of the above formulae can be found in the literature albeit in a different guise. For example, since the stabilizer subgroup of P_∞ in G_N is $G_\infty = \pm\Gamma_1(N)/\pm\Gamma(N)$, where $\Gamma_1(N) = \{A \in \Gamma(1) : A \equiv \left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right) \bmod N\}$, the quotient curve X_∞ is the modular curve usually denoted by $X_1(N) = \pm\Gamma_1(N)\backslash\mathfrak{H}^*$. Thus, the above two propositions yield the well-known relation for the genus of $X_1(N)$,

$$g_\infty = g(X_1(N)) = 1 + \frac{m}{12N} - \frac{1}{2}\nu_\infty,$$

where $\nu_\infty = \frac{1}{2}\sum_{\nu|N} \phi(\nu)\phi(\frac{N}{\nu})$ for $N > 4$; cf. Miyake[16], Theorem 4.2.9.

Similarly, the genus of X_0 and X_1 was determined by Hecke [4] for the case of prime level N; however, the presentation of his result is somewhat more indirect and does not easily generalize to composite N's.

1.2 The local invariants

As was explained in [10], much of the structure of the twisted diagonal quotient surface Z_ε associated to $X = X(N)$ is known once we have determined the local invariants $s_{\nu,\varepsilon}(\bar{P}_l, \bar{P}_k) := s_{\nu,\alpha_\varepsilon}(\bar{P}_l, \bar{P}_k)$ which count the number of points $x \in X_{k,\varepsilon}$ lying above $\bar{P}_l$ whose "ramification characters" $\lambda_{x,\varepsilon}$ are of a certain type (cf. [10], Notation 1.4 and Remark 1.5). Here we want to compute these numbers and hence derive in particular the values for r_k, t_k, etc., which were asserted in Proposition 1.3. In order to state the result, we first introduce the following notation.

Notation 1.5 For any pair of natural numbers $N, k \in \mathbb{N}$ with $(k, N) = 1$, let

$$\rho(k, N) = |\{x \in \mathbb{Z}/N\mathbb{Z} : x^2 = k\}|$$

denote the number of square roots of k mod N. Thus, $\rho(k, N) = 0$ if $k \notin ((\mathbb{Z}/N\mathbb{Z})^\times)^2$ and otherwise $\rho(k, N) = \rho(N)$, where

$$\rho(N) \;:=\; |(\mathbb{Z}/N\mathbb{Z})_2^\times| \;=\; [(\mathbb{Z}/N\mathbb{Z})^\times : ((\mathbb{Z}/N\mathbb{Z})^\times)^2] \;=\; 2^{s+t}, \tag{1}$$

where s denotes the number of odd primes dividing N, and $t = 0$ if $4 \nmid N$, $t = 2$ if $8|N$ and $t = 1$ otherwise (cf. Landau [13], Satz 88).

Theorem 1.6 *The twisted local invariants $s_{\nu,\varepsilon} = s_{\nu,\alpha_\varepsilon}$ are related to the standard (untwisted) invariants $s_\nu = s_{\nu,id}$ by the formula*

$$s_{\nu,\varepsilon}(\bar{P}_l, \bar{P}_k) = s_{(\kappa\nu)}(\bar{P}_l, \bar{P}_k), \qquad \text{for } \; 1 \le \nu \le e_k, \tag{2}$$

where $\kappa = \kappa_{k,\varepsilon}$ is defined by

$$\kappa_{k,\varepsilon} = \begin{cases} 1 & \text{if } k = 0 \text{ or } k = 1 \text{ and } 3 \nmid N \\ \varepsilon & \text{if } k = \infty \text{ or } k = 1 \text{ and } 3|N, \end{cases} \tag{3}$$

and $(\kappa\nu)$ denotes the least positive residue of $\kappa\nu \pmod{e_k}$. In particular,

$$r_n(\bar{P}_l, \bar{P}_k) = r_{n,\alpha_\varepsilon}(\bar{P}_l, \bar{P}_k), \;\; \text{where } \; n|e_k, \tag{4}$$

is independent of ε. Furthermore, with the notation introduced in Notations 1.2 *and* 1.5 *we have*

$$\begin{array}{lllll}
a) & s_{\nu,\varepsilon}(\bar{P}_l, \bar{P}_k) & = \delta_{\nu e_k} r_{1,\varepsilon}(\bar{P}_l, \bar{P}_k) & = \delta_{\nu e_k} \frac{m}{e_l e_k}, & \text{if } \; l \ne k; \\
b) & s_{1,\varepsilon}(\bar{P}_0, \bar{P}_0) & = r_{2,\varepsilon}(\bar{P}_0, \bar{P}_0) & = r_0 \;\; = \;\; h(-4N^2), & \\
 & s_{2,\varepsilon}(\bar{P}_0, \bar{P}_0) & = r_{1,\varepsilon}(\bar{P}_0, \bar{P}_0) & = \frac{m}{4} - \frac{1}{2} r_0; & \\
c) & s_{\nu,\varepsilon}(\bar{P}_1, \bar{P}_1) & = \frac{1}{2} h(-3N^2) & = \frac{1}{2} r_1, & \text{if } \nu = 1, 2 \text{ and } 3 \nmid N, \\
 & s_{\nu,\varepsilon}(\bar{P}_1, \bar{P}_1) & = h(-3N^2) & = r_1, & \text{if } \nu = 1, 2, \; \nu \equiv \varepsilon \pmod 3 \text{ and } 3|N, \\
 & s_{\nu,\varepsilon}(\bar{P}_1, \bar{P}_1) & = 0, & & \text{if } \nu = 1, 2, \; \nu \not\equiv \varepsilon \pmod 3 \text{ and } 3|N, \\
 & s_{3,\varepsilon}(\bar{P}_1, \bar{P}_1) & = r_1(\bar{P}_1, \bar{P}_1) & = \frac{m}{9} - \frac{1}{3} r_1 \;=\; \frac{m}{9} - \frac{1}{3} h(-3N^2); & \\
d) & s_{\nu,\varepsilon}(\bar{P}_\infty, \bar{P}_\infty) & = \frac{1}{2} \rho(\frac{\varepsilon\nu}{d}, \frac{N}{d}) \phi(d), & \text{where } d = (\nu, N), & \\
 & r_{d,\varepsilon}(\bar{P}_\infty, \bar{P}_\infty) & = \frac{1}{2} \phi(d) \phi(\frac{N}{d}), & \text{if } d|N. &
\end{array}$$

Remark 1.7 As the explicit formulae in Theorem 1.6 show, the local invariants $s_{\nu,\varepsilon}(\bar{P}_l, \bar{P}_k)$ depend only on the square class $\varepsilon \cdot ((\mathbb{Z}/N\mathbb{Z})^\times)^2$ of ε modulo N. This may also be seen directly as follows. First note that if $\varepsilon = e^2$ is a square modulo N, then conjugation by Q_ε is the same as conjugation by $\left(\begin{smallmatrix} e & 0 \\ 0 & e^{-1} \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z}/n\mathbb{Z})$, and hence the map $\varepsilon \mapsto \alpha_\varepsilon$ induces a homomorphism

$$\bar{\alpha} : (\mathbb{Z}/N\mathbb{Z})^\times \,/\, \left((\mathbb{Z}/N\mathbb{Z})^\times\right)^2 \to \mathrm{Aut}(G_N)/\mathrm{Inn}(G_N) = \mathrm{Out}(G_N).$$

Thus, since the local invariants $s_{\nu,\alpha}$ depend only on the image of α in $\mathrm{Out}(G)$ (cf. [10], Remark 1.5d), the assertion follows.

As regards to the map $\bar{\alpha}$, note that $\bar{\alpha}$ is injective for $N > 4$ (as is easy to see), but that this is false for $n = 4$ since G_4 is isomorphic to the symmetric group S_4 which has no outer automorphisms whereas $\rho(4) = |(\mathbb{Z}/4\mathbb{Z})^\times/((\mathbb{Z}/4\mathbb{Z})^\times)^2| = 2$.

In Hermann [5], p. 96, it is asserted that $\bar{\alpha}$ is in fact an isomorphism (even for $N = 4$!). For a proof, Hermann refers to Praetorius [17] §2, who actually only considers the case $N = p^r$ with a prime number $p > 3$. By showing that G_{p^r} is a characteristic subgroup of G_N, if $p^r \parallel N$ and $p > 3$, one can deduce from Praetorius that this is true if $(N, 6) = 1$. In the general case, however, it is still unclear which automorphisms actually exist.

The proof of Theorem 1.6 is based on the group-theoretical method presented in [10], Proposition 1.7. There it was shown that the local invariants $s_{\nu,\alpha}(\bar{x}, \bar{y})$ are determined by the orders of certain "normalizing subsets" $N^*_{\nu,i}(\sigma, \tau)$ of G attached to generators σ and τ of the ramification groups G_x and $G_{y,\alpha} = \alpha^{-1}(G_y)$, where $x \in \pi^{-1}(\bar{x})$ and $y \in \pi^{-1}(\bar{y})$. Recall from [10], Notation 1.6, that these normalizing sets are defined as follows. If $r = (e_x, e_y)$ and $k = (\nu, r)$, $\bar{\sigma} = \sigma^{e_x/r}, \bar{\tau} = \tau^{e_y/r}$, then for any $i \in \mathbb{Z}$:

$$\begin{aligned}
N_\nu(\sigma,\tau) \;=\; N_k(\sigma,\tau) \;&:=\; \{g \in G : g\bar{\sigma}^\nu g^{-1} \in \langle \bar{\tau}^\nu \rangle\} \;=\; \{g \in G : g\langle \bar{\sigma}^\nu \rangle g^{-1} = \langle \bar{\tau}^\nu \rangle\},\\
N^i_\nu(\sigma,\tau) \;=\; N^i_k(\sigma,\tau) \;&:=\; \{g \in G : g\bar{\sigma}^k g^{-1} = \bar{\tau}^{ki}\},\\
N^*_\nu(\sigma,\tau) \;=\; N^*_k(\sigma,\tau) \;&:=\; N_k(\sigma,\tau) \setminus \bigcup_{d|k, d\neq k} N_d(\sigma,\tau),\\
N^*_{\nu,i}(\sigma,\tau) = N^*_{k,i}(\sigma,\tau) \;&:=\; N^*_\nu(\sigma,\tau) \cap N^i_\nu(\sigma,\tau).
\end{aligned}$$

In the case that $\sigma = \sigma_k$ and $\tau = \alpha_\varepsilon^{-1}(\sigma_k)$, these normalizing sets are closely related to the following sets, as we shall see in Lemma 1.9.

Notation 1.8 For $k = 0, 1$ and $\varepsilon \in (\mathbb{Z}/N\mathbb{Z})^\times$ let

$$\begin{aligned}
\tilde{M}_{k,\varepsilon}(N) \;&=\; \left\{ A \in \mathrm{Gl}_2(\mathbb{Z}/N\mathbb{Z}) : A = \begin{pmatrix} x & -y \\ y & x + \delta_{1k} y \end{pmatrix} \;\text{ and }\; \det(A) = \varepsilon \right\},\\
M_{k,\varepsilon}(N) \;&=\; \{(x, y) \in \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} : q_k(x, y) = \varepsilon\},
\end{aligned}$$

where $q_k(x, y) = x^2 + \delta_{k1} xy + y^2$. These two sets are related by the map $p : \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \mapsto (x, z)$ which induces a bijection

$$p_{k,\varepsilon} : \tilde{M}_{k,\varepsilon}(N) \xrightarrow{\sim} M_{k,\varepsilon}(N) \;.$$

Furthermore, for $d|N$ and $(n, \frac{N}{d}) = 1$ we put

$$\begin{aligned}
\tilde{M}_{\infty,d,n}(N) &= \{\left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z}) : x \equiv n \pmod{\tfrac{N}{d}}, (z, N) = \tfrac{N}{d}\}, \\
M_{\infty,d,n}(N) &= \{(x, y) \in \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} : x \equiv n \pmod{\tfrac{N}{d}} \text{ and } (y, N) = \tfrac{N}{d}\}, \\
\tilde{M}^{(2)}_{\infty,d,n}(N) &= \bigcup_{\nu^2=n} \tilde{M}_{\infty,d,\nu}(N) \quad \text{and} \quad M^{(2)}_{\infty,d,n}(N) = \bigcup_{\nu^2=n} M_{\infty,d,\nu}(N).
\end{aligned}$$

Here the above map $p : \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \mapsto (x, z)$ induces a bijection

$$p_{\infty,d,n} : \tilde{M}_{\infty,d,n}(N)/\langle\sigma_{\infty,N}\rangle \xrightarrow{\sim} M_{\infty,d,n}(N).$$

Lemma 1.9 *If $d|N \geq 5$ and $(n, \frac{N}{d}) = (\varepsilon, N) = 1$, then*

$$\begin{aligned}
a)\quad & N^*_{1,1}(\sigma_{0,N}, \alpha_\varepsilon^{-1}(\sigma_{0,N})) &&= Q_\varepsilon^{-1}\left(\tilde{M}_{0,\varepsilon}(N) \cup Q_{-1}\tilde{M}_{0,-\varepsilon}(N)\right)/\{\pm 1\}; \\
b)\quad & N^*_{1,1}(\sigma_{1,N}, \alpha_\varepsilon^{-1}(\sigma_{1,N})) &&= Q_\varepsilon^{-1}\tilde{M}_{1,\varepsilon}(N)/\{\pm 1\}, \\
& N^*_{1,2}(\sigma_{1,N}, \alpha_\varepsilon^{-1}(\sigma_{1,N})) &&= Q_\varepsilon^{-1}\left(\tilde{M}_{1,-\varepsilon}(N)\left(\begin{smallmatrix} 1 & 1 \\ 0 & -1 \end{smallmatrix}\right)\right)/\{\pm 1\}; \\
c)\quad & N^*_{d,n}(\sigma_{\infty,N}, \alpha_\varepsilon^{-1}(\sigma_{\infty,N})) &&= Q_\varepsilon^{-1}\tilde{M}^{(2)}_{\infty,d,\varepsilon n}(N)/\{\pm 1\}.
\end{aligned}$$

Proof. Let $f : \tilde{G} := \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z}) \to \tilde{G}/\{\pm 1\} = G_N$ denote the projection map and put

$$\tilde{N}^i_\varepsilon(\sigma_k) := f^{-1}(N^i_1(\sigma_k, \alpha_\varepsilon^{-1}(\sigma_k))), \qquad \text{for} \quad k \in \{0, 1, \infty\};$$

thus we have $N^i_1(\sigma_k, \alpha_\varepsilon^{-1}(\sigma_k)) = \tilde{N}^i_\varepsilon(\sigma_k)/\{\pm 1\}$.

a) Since σ_0 has prime order (in G_N), we have $N^*_{1,1}(\sigma_0, \alpha_\varepsilon^{-1}(\sigma_0)) = N^1_1(\sigma_0, \alpha_\varepsilon^{-1}(\sigma_0)) = \tilde{N}^1_\varepsilon(\sigma_0)/\{\pm 1\}$, and so it is enough to calculate the "twisted centralizer" $\tilde{N}^1_\varepsilon(\sigma_0)$. For this, let $A = \left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})$. Then

$$\begin{aligned}
A \in \tilde{N}^1_\varepsilon(\sigma_0) &\Leftrightarrow \begin{pmatrix} x & y \\ z & w \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \pm\begin{pmatrix} 0 & -\varepsilon^{-1} \\ \varepsilon & 0 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} \text{ and } xw - yz = 1 \\
&\Leftrightarrow\ y = -\varepsilon^{-1}z, w = \varepsilon x, \varepsilon x^2 + \varepsilon^{-1}z^2 = 1 \text{ or } y = \varepsilon^{-1}z, w = -\varepsilon x, -\varepsilon x^2 - \varepsilon^{-1}z^2 = 1 \\
&\Leftrightarrow\ A \in Q_\varepsilon^{-1}\tilde{M}_{0,\varepsilon}(N) \cup Q_{-\varepsilon}^{-1}\tilde{M}_{0,-\varepsilon}(N).
\end{aligned}$$

b) To determine the twisted centralizer of σ_1 we observe that

$$\begin{aligned}
A \in \tilde{N}^1_\varepsilon(\sigma_1) &\Leftrightarrow \begin{pmatrix} x & y \\ z & w \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = \pm\begin{pmatrix} 0 & -\varepsilon^{-1} \\ \varepsilon & 1 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} \text{ and } xw - yz = 1 \\
&\Leftrightarrow \begin{pmatrix} y & -x + y \\ w & -z + w \end{pmatrix} = \pm\begin{pmatrix} -\varepsilon^{-1}z & -\varepsilon^{-1}w \\ \varepsilon x + z & \varepsilon y + w \end{pmatrix} \text{ and } xw - yz = 1 \\
&\Leftrightarrow\ y = -\varepsilon^{-1}z, w = \varepsilon x + z, xw - yz = 1 \ \Leftrightarrow\ A \in Q_\varepsilon^{-1}\tilde{M}_{1,\varepsilon}(N).
\end{aligned}$$

Here, the second last equivalence is valid because the case of the minus sign implies $2w = 2y = 0$, so $2 = 2(xw - yz) = 0$, which is false for $N > 2$, and the last because $Q_\varepsilon^{-1}\tilde{M}_{1,\varepsilon} = \{\left(\begin{smallmatrix} \varepsilon^{-1}x & -\varepsilon^{-1}z \\ z & x+z \end{smallmatrix}\right) : x^2 + xz + z^2 = \varepsilon\} = \{\left(\begin{smallmatrix} x & -\varepsilon^{-1}z \\ z & \varepsilon x+z \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})\}$.

In the same way one can determine $\tilde{N}^2_\varepsilon(\sigma_{1,N}) = \{A \in \tilde{G} : A\sigma_1 = \pm\left(\begin{smallmatrix} -1 & -\varepsilon^{-1} \\ \varepsilon & 0 \end{smallmatrix}\right)A\}$.

c) For a divisor μ of N we have the following equivalences which will be justified below:

$$\begin{aligned} & A \in N_{\tilde{G}}(\langle \pm\sigma_\infty^\mu \rangle) \\ & \Leftrightarrow \text{ there is a k: } \begin{pmatrix} x & y \\ z & w \end{pmatrix}\begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix} = \pm\begin{pmatrix} 1 & k\mu \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x & y \\ z & w \end{pmatrix} \text{ and } xw - yz = 1 \\ & \Leftrightarrow \text{ there is a } k: \begin{pmatrix} x & x\mu + y \\ z & z\mu + w \end{pmatrix} = \begin{pmatrix} x + kz\mu & y + kw\mu \\ z & w \end{pmatrix} \text{ and } xw - yz = 1 \\ & \Leftrightarrow \ z \equiv 0 \bmod \frac{N}{\mu} \text{ and } xw - yz = 1. \end{aligned}$$

To justify the second equivalence, note that $A\sigma_\infty^\mu = -\sigma_\infty^\mu A$ leads to $z = -z$ and $2w = z\mu = -z\mu = -2w$, hence $4 = 4(xw - yz) = 0$ which is a contradiction since $N > 4$.

The last equivalence may be derived as follows. First of all, if $z\mu + w = w$ (in $\mathbb{Z}/N\mathbb{Z}$), then $z \equiv 0(\bmod \frac{N}{\mu})$, which proves one implication. Conversely, assume that $z \equiv 0 \pmod{\frac{N}{\mu}}$ and $xw - yz = 1$. Thus $1 \equiv xw - yz \equiv xw \pmod{\frac{N}{\mu}}$, i.e. $(w, \frac{N}{\mu}) = 1$ and consequently there are k', l such that $1 = k'w + l\frac{N}{\mu}$. By multiplying with $x\mu$, it follows that $x\mu = xk'\mu + lxN \equiv wk\mu \pmod N$ (with $k = xk'$), which is the condition of the third line. Therefore

$$N_d^*(\sigma_{\infty,N}) = \left\{A \in G_N : z \equiv 0 \bmod \frac{N}{d}, z \not\equiv 0 \bmod \frac{N}{\mu}, \mu | d, \mu < d\}\right\} = \left\{A : (z, N) = \tfrac{N}{d}\right\}.$$

Moreover, if $A \in N_d^*(\sigma_\infty, \sigma_\infty)$, then the above computation shows that in fact $A \in N_d^k(\sigma_\infty, \sigma_\infty)$ for $wkd \equiv xd \pmod N$ or, equivalently, for $k \equiv x^2 \pmod{\frac{N}{d}}$, and so the assertion follows for $\varepsilon = 1$.

From this the general case follows easily. Indeed, if $\varepsilon^*\varepsilon \equiv 1 \pmod N$, then $\alpha_\varepsilon^{-1}(\sigma_\infty) = \sigma_\infty^{\varepsilon^*}$ and $\tilde{M}^{(2)}_{\infty,d,\varepsilon^* n}(N) = Q_\varepsilon^{-1}\tilde{M}^{(2)}_{\infty,d,\varepsilon n}(N)$, and so we see that

$$N^*_{d,n}(\sigma_\infty, \alpha_\varepsilon^{-1}(\sigma_\infty)) = N^*_{d,n}(\sigma_\infty, \sigma_\infty^{\varepsilon^*})) = N^*_{d,\varepsilon^* n}(\sigma_\infty, \sigma_\infty) = \tilde{M}^{(2)}_{\infty,d,\varepsilon^* n}(N) = Q_\varepsilon^{-1}\tilde{M}^{(2)}_{\infty,d,\varepsilon n}(N).$$

We now count the number of elements in each of the above sets.

Lemma 1.10 *For any N, d, ε and $n \in \mathbb{N}$ with $d \mid N$ and $(\varepsilon, N) = (n, \frac{N}{d}) = 1$ we have*

$$\begin{array}{lllll} a) & |\tilde{M}_{0,\varepsilon}(N)| & = |M_{0,\varepsilon}(N)| & = & 2h(-4N^2), \text{ if } 4 \not| N, \\ & |\tilde{M}_{0,\varepsilon}(N)| & = |M_{0,\varepsilon}(N)| & = & 2h(-4N^2)\left(1 + \left(\frac{-1}{\varepsilon}\right)\right), \text{ if } 4|N; \\ b) & |\tilde{M}_{1,\varepsilon}(N)| & = |M_{1,\varepsilon}(N)| & = & 3h(-3N^2), \text{ if } 3 \not| N, \\ & |\tilde{M}_{1,\varepsilon}(N)| & = |M_{1,\varepsilon}(N)| & = & 3h(-3N^2)\left(1 + \left(\frac{-3}{\varepsilon}\right)\right), \text{ if } 3|N; \\ c) & |\tilde{M}_{\infty,d,n}(N)| & = |M_{\infty,d,n}(N)| \cdot N & = & d\phi(d)N, \\ & |\tilde{M}^{(2)}_{\infty,d,n}(N)| & = |M^{(2)}_{\infty,d,n}(N)| \cdot N & = & \rho\left(n, \frac{N}{d}\right) d\phi(d)N. \end{array}$$

Proof. a) Keller [11] (The proof is similar to b) below.)

b) As was already remarked in Notation 1.8, we have a bijection $\tilde{M}_{1,\varepsilon} \xrightarrow{\sim} M_{1,\varepsilon}$, so the the first equality in each case is clear. Since $\tilde{M} := \cup_\varepsilon \tilde{M}_{1,\varepsilon}$ is a subgroup of $\mathrm{Gl}_2(\mathbb{Z}/N\mathbb{Z})$, and since $\tilde{M}_{1,1} = \mathrm{Ker}(\det_{|\tilde{M}})$, we see that $\tilde{M}_{1,\varepsilon}$ is either empty or an $\tilde{M}_{1,1}$-coset, and that the latter occurs precisely when $\varepsilon \in D := \det(\tilde{M}) \subset (\mathbb{Z}/N\mathbb{Z})^\times$. Thus, the assertion follows once we have determined D and have shown that $|\tilde{M}_{1,1}(N)| = 3h(-3N^2)$; for this we shall use the quadratic form interpretation of $\tilde{M}_{1,\varepsilon}(N) \xrightarrow{\sim} M_{1,\varepsilon}$.

Since by Chinese remainder theorem the number of solutions of the given equation is multiplicative (as a function of N), as is $3h(-3N^2)$, it suffices to consider powers of primes $N = p^r$. If $p \neq 2$, we can transform the quadratic form $a^2 + ac + b^2$ to $x^2 + 3y^2$ by putting $a = x + y$ and $c = -2y$. Thus, if $p > 3$ then we see that $D = (\mathbb{Z}/N\mathbb{Z})^\times$ (cf. [6], p. 310), and so by using the same arguments as in Keller[11] we get the result in this case.

If $p = 3$, then clearly $-1 \notin D$, so $D = ((\mathbb{Z}/N\mathbb{Z})^\times)^2$ since D contains the squares which have index 2 in $(\mathbb{Z}/N\mathbb{Z})^\times$. Furthermore, since $\left(\frac{1-3y^2}{3^r}\right) = \left(\frac{1}{3}\right) = 1$, the equation $x^2 = 1-3y^2$ has exactly two solutions for each $y \in \mathbb{Z}/3^r\mathbb{Z}$, which means that $|M_{1,1}| = 2 \cdot 3^r$. Thus, the formula is true in this case.

Finally, if $p = 2$ then $D = (\mathbb{Z}/N\mathbb{Z})^\times$. Indeed, D certainly contains all squares, i.e. all $x \equiv 1 \bmod 8$. Choosing $a \equiv c \equiv 1 \pmod 2$ we have $a^2 + ac + c^2 \equiv 2 + ac \pmod 8$ and thus all other residue classes modulo 8 appear in D, i.e. $D = (\mathbb{Z}/N\mathbb{Z})^\times$. Since $a^2 + ac + c^2 \equiv 0 \pmod 2 \Leftrightarrow a \equiv b \equiv 0 \pmod 2$ we obtain $|\tilde{M}_{1,1}(N)| = |\tilde{M}|/|D| = (2^{2r} - 2^{2(r-1)})/2^{(r-1)} = 2^{r-1} \cdot 3 = 3h(-3 \cdot 2^{2r})$.

c) It is clear that the second formula follows from the first. To prove the latter, we may assume by the Chinese remainder theorem that $N = p^r$ and $d = p^s$ are prime powers. We now distinguish two cases:

Case 1: $r = s$. Here we have only the condition $(z, N) = 1$. The number of choices for z is therefore $\phi(N)$, and for each choice of $x, w \pmod N$, there is a unique $y \pmod N$ such that $xw - yz \equiv 1 \pmod N$. Thus, we have in total $\phi(N)N^2$ matrices in $\tilde{M}_{\infty,d,n}(N)$ and $\phi(N)N$ vectors in $M_{\infty,d,n}(N)$.

Case 2: $r > s$. The number of $x \pmod N$ with $x \equiv n \pmod{\frac{N}{d}}$ is clearly d. Note that for each such x we have $(x, p) = 1$ because $(n, p^{r-s}) = 1$ by hypothesis. Thus, for each y, z there is a unique $w \pmod N$ such that $xw - yz \equiv 1 \pmod N$. Of these, the number of z satisfying $(z, N) = \frac{N}{d}$ is $\phi(d)$, so in total we have $d\phi(d)N$ matrices and $d\phi(d)$ vectors.

For the proof of Theorem 1.6 we also need the following fact.

Lemma 1.11 *If $N \geq 5$, $\varepsilon \in (\mathbb{Z}/N\mathbb{Z})^\times$ and $\bar{x} \neq \bar{y} \in \bar{X}$, then $G_x \cap G_{y,\varepsilon} = \{1\}$, for all $x \in \pi^{-1}(\bar{x})$ and $y \in \pi^{-1}(\bar{y})$, and hence $s_{\nu,\varepsilon}(\bar{x}, \bar{y}) = 0$, for $1 \leq \nu < e_{\bar{y}}$.*

Proof. Since $\pm\Gamma_1(N)$ has no elliptic elements for $N > 4$ (cf. [16], Th. 4.2.9), we have $G_x \cap G_y = \{1\}$ for $y = P_\infty$ and for all $x \in \pi^{-1}(\bar{P}_k)$, $k = 0, 1$. Now since $\alpha_\varepsilon(\sigma_\infty) = \sigma_\infty^\varepsilon$, we see that $G_{y,\varepsilon} = \alpha_\varepsilon^{-1}(G_y) = G_y$, and hence also $G_x \cap G_{y,\varepsilon} = \{1\}$. Furthermore, if $y' = gy \in \pi^{-1}(\bar{y})$, then $G_x \cap G_{y',\varepsilon}$ is a conjugate of $G_{x'} \cap G_{y,\varepsilon}$ for a suitable $x' = g'x$, and hence is also trivial. Thus, the assertion holds for $\bar{x} = \bar{P}_k, k = 0, 1$ and $\bar{y} = \bar{P}_\infty$, and also for $\bar{x} = \bar{P}_\infty$ and $\bar{y} = \bar{P}_k, k = 0, 1$, since $G_y \cap G_{x,\varepsilon} = \alpha_\varepsilon^{-1}(G_x \cap G_{y,\varepsilon^{-1}})$. This proves the first statement since the assertion is trivial for all other pairs of distinct points, and the second follows immediately from the first from the definition of $s_{\nu,\varepsilon}(\bar{x}, \bar{y})$; cf. [10], Notation 1.4.

Proof of Theorem 1.6. It is enough to verify the formulae in a) – d), for (2) clearly follows from these, and (4) follows from (2) since by Remark 1.5 of [10] we have (with $e = e_{\bar{y}}$)

$$r_{n,\alpha}(\bar{x},\bar{y}) \;=\; \sum_{\substack{\nu=1\\(\nu,e)=\frac{e}{n}}}^{e} s_{\nu,\alpha}(\bar{x},\bar{y}), \tag{5}$$

and since multiplication by κ permutes the ν with $(\nu, e_k) = \frac{e_k}{n}$.

a) For $\nu \neq e_k$, this follows directly from Lemma 1.11, and hence, in view of (5), the result also holds for $\nu = e_k$ because we always have (cf. [10], Remark 1.5):

$$\sum_{\nu=1}^{e_k} s_{\nu,\varepsilon}(\bar{P}_l,\bar{P}_k)(\nu,e_k) \;=\; \sum_{d|e_k} r_{d,\varepsilon}(\bar{P}_l,\bar{P}_k)\frac{e_k}{d} \;=\; \frac{m}{e_l}. \tag{6}$$

b) Fix ε and put $\sigma_k' := \alpha_\varepsilon^{-1}(\sigma_k)$. Then we have trivially $\lambda_{P_k,\alpha_\varepsilon}(\sigma_k') = \lambda_{P_k}(\sigma_k)$, where $\lambda_{P_k,\alpha_\varepsilon} = \lambda_{P_k} \circ \alpha_\varepsilon$ is the twisted local ramification character (cf. [10], Notation 1.2, 1.3), and hence it follows from [10], Proposition 1.7c) that

$$s_{\nu,\varepsilon}(\bar{P}_k,\bar{P}_k) \;=\; \frac{1}{de_k}|N^*_{\nu,\bar{\nu}^*}(\sigma_k,\sigma_k')| \;=\; \frac{1}{de_k}|N^*_{d,\bar{\nu}^*}(\sigma_k,\sigma_k')|, \tag{7}$$

where $d = (\nu, e_k)$ and $\bar{\nu}^*\frac{\nu}{d} \equiv 1 \pmod{\frac{e_k}{d}}$. Thus, since the combination of Lemmas 1.9a) and 1.10a) yields $|N^*_{1,1}(\sigma_0,\sigma_0')| = 2h(-4N^2)$, we obtain from (7) that $s_{1,\varepsilon}(\bar{P}_0,\bar{P}_0) = \frac{1}{2}|N^*_{1,1}(\sigma_0,\sigma_0')| = h(-4N^2)$, which does not depend on ε. Furthermore, since $e_0 = 2$ we have by (5) and a) that $s_{1,\varepsilon}(\bar{P}_0,\bar{P}_0) = r_{2,\varepsilon}(\bar{P}_0,\bar{P}_0) = \sum_{\bar{x}\in\bar{X}} r_{2,\varepsilon}(\bar{x},\bar{P}_0) =: r_{0,\varepsilon} = r_0$. The second equation of b) follows from the first by using (5) and (6).

c) From Lemmas 1.9b) and 1.10b) we have $|N^*_{1,1}(\sigma_1,\sigma_1')| = \frac{3}{2}h(-3N^2)$, if $3 \nmid N$ (and $|N^*_{1,1}(\sigma_1,\sigma_1')| = 3h(-3N^2)$, if $3|N$ and $\varepsilon \equiv 1 \pmod 3$, and $|N^*_{1,1}(\sigma_1,\sigma_1')| = 0$, if $3|N$ and $\varepsilon \not\equiv 1 \pmod 3$), and so by the same argument as in a) we obtain the indicated values for $s_{1,\varepsilon}(\bar{P}_1,\bar{P}_1) = r_{1,\varepsilon}(\bar{P}_1,\bar{P}_1)$. The formula for $s_{2,\varepsilon}(\bar{P}_1,\bar{P}_1)$ is proven similarly, and the formula for $s_{3,\varepsilon}(\bar{P}_1,\bar{P}_1)$ follows from (5) and (6).

d) Let $\bar{\nu}^*$ be such that $\bar{\nu}^*\frac{\nu}{d} \equiv 1 \pmod{\frac{N}{d}}$, where $d = (\nu, N)$. Then by Lemmas 1.9c) and 1.10c) we have $|N^*_{d,\bar{\nu}^*}(\sigma_\infty,\sigma_\infty')| = \frac{1}{2}\rho(\varepsilon\bar{\nu}^*,\frac{N}{d})d\phi(d)N = \frac{1}{2}\rho(\frac{\varepsilon\nu}{d},\frac{N}{d})d\phi(d)N$, the latter since $(\frac{\nu}{d})/\bar{\nu}^*$ is a square mod $\frac{N}{d}$. Thus, by (7) we obtain $s_{\nu,\varepsilon}(\bar{P}_\infty,\bar{P}_\infty) = \frac{1}{2}\rho(\frac{\varepsilon\nu}{d},\frac{N}{d})\phi(d)$, as asserted. The last formula of d) follows from the previous one by using (5) and the fact that

$$\sum_{\substack{\nu=1\\(\nu,n)=1}}^{n} \rho(\varepsilon\nu,n) = \phi(n).$$

Proof of Proposition 1.3. The first assertion follows directly from Lemma 1.12 below, and the independence assertions follow from (4). It thus remains to prove the asserted formulae.

a), b) If $e_{\bar{x}} = p$ is prime, then $r_{\bar{x}} = t_{\bar{x}} = r_p(\bar{x}) = \sum s_\nu(\bar{x})$, where the sum is over all ν with $1 \leq \nu < e_{\bar{x}}$ and $(\nu, e_{\bar{x}}) = 1$, and so both a) and b) follow directly from Theorem 1.6.

c) Since $r_{n,\varepsilon}(\bar{P}_\infty) = \sum_{\bar{x}\in\bar{X}} r_{n,\varepsilon}(\bar{x},\bar{P}_\infty)$ by definition, we have by Theorem 1.6a), d) that $r_{n,\varepsilon}(\bar{P}_\infty) = r_{n,\varepsilon}(\bar{P}_\infty,\bar{P}_\infty) = \frac{1}{2}\phi(n)\phi(\frac{N}{d})$.

d) The formula for r_∞ follows directly from part c), and that for t_∞ follows from c) together with the following (well-known) identity which, in view of Theorem 1.6d), follows from formula (6) above:

$$\tfrac{1}{2}\sum_{\nu|N}\nu\phi(\nu)\phi\left(\frac{N}{\nu}\right) \;=\; \frac{m}{N} \;=\; \tfrac{1}{2}N^2\prod_{p|N}\left(1-\frac{1}{p^2}\right). \tag{8}$$

In the above proof we had used the following result.

Lemma 1.12 *For each $k = 0, 1, \infty$ the stabilizer $G_{k,\varepsilon} = \alpha_\varepsilon^{-1}(G_k)$ is conjugate to G_k; more precisely, we have $\alpha_\varepsilon(\sigma_k) = \tau_k^{-1}\sigma_k^{\kappa_\varepsilon}\tau_k$, where $\kappa_\varepsilon = \kappa_{k,\varepsilon}$ is as in (3) and $\tau_k = \tau_{k,\varepsilon} \in \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$; in fact, we can take $\tau_k = 1$ if $k = \infty$ and $\tau_k = g_{k,\varepsilon}^{\pm}Q_{\pm\varepsilon}^{-1}$ with $g_{k,\varepsilon}^{\pm} \in \tilde{M}_{k,\pm\varepsilon}$, if $k = 0, 1$.*

Proof. Since $\alpha_\varepsilon(\sigma_\infty) = \sigma_\infty^\varepsilon$, the assertions are clear for $k = \infty$.

For $k = 0, 1$,we first observe that either $\tilde{M}_{k,\varepsilon}(N)$ or $\tilde{M}_{k,-\varepsilon}(N)$ is non-empty (cf. Lemma 1.10), so there exists an element $g_{k,\varepsilon}^{\pm} \in \tilde{M}_{k,\pm\varepsilon}$. Clearly $\tau_k := g_{k,\varepsilon}^{\pm}Q_{\pm\varepsilon}^{-1} \in \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})$ and $\tau_k Q_{\pm\varepsilon} = g_{k,\varepsilon}^{\pm} \in \bigcup_\mu \tilde{M}_{k,\mu}(N) = C_{\mathrm{Gl}_2(\mathbb{Z}/N\mathbb{Z})}(\sigma_k)$. Thus $Q_{\pm\varepsilon}^{-1}\tau_k^{-1}\sigma_k\tau_k Q_{\pm\varepsilon} = \sigma_k$, or $\tau_k^{-1}\sigma_k\tau_k = Q_{\pm\varepsilon}\sigma_k Q_{\pm\varepsilon}^{-1} = \alpha_\varepsilon(\sigma_k^{\pm 1})$. This proves the assertions, for the (-1)-case is only plays a role if $k = 1$ and in this case is only necessary if $3|N$ and $\varepsilon \equiv 2 \pmod 3$ (cf. Lemma 1.10).

Corollary 1.13 *The character $h^1_{\alpha_\varepsilon}$ afforded by the G-module $H^1(X(N), \mathbb{C})$ (endowed with the twisted G-action) is independent of ε, i.e. $h^1_{\alpha_\varepsilon} = h^1$.*

Proof. The above Lemma 1.12 implies that $(1_{\alpha^{-1}(G_k)})^G = (1_{G_k})^G$, for $k = 0, 1, \infty$, and so the assertion follows directly from [10], Proposition 1.8, Equation (16).

Although $h^1_\varepsilon = \omega_\varepsilon + \overline{\omega_\varepsilon}$, where $\omega_\varepsilon = \omega \circ \alpha_\varepsilon$ denotes the character afforded by the G-module $H^0(X, \omega_X)$ of holomorphic differentials (with the twisted G-action), it is not true in general that the character ω_ε is independent of ε. However, we have:

Proposition 1.14 *For any character χ of G and any $\varepsilon \in (\mathbb{Z}/N\mathbb{Z})^\times$ we have $\chi_{-\varepsilon} = \overline{\chi_\varepsilon}$, where $\chi_\varepsilon := \chi \circ \alpha_\varepsilon$, and hence also $\omega_{-\varepsilon} = \overline{\omega_\varepsilon}$. In particular, if -1 is a square mod N, then $\omega_\varepsilon = \frac{1}{2}h^1$ is independent of ε.*

Proof. In view of the above corollary and the remarks after it (together with Remark 1.7), the last assertions clearly follow from the first. Furthermore, since $\alpha_{-\varepsilon} = \alpha_\varepsilon \circ \alpha_{-1}$, it is enough to verify the first assertion for $\varepsilon = 1$, and for this it suffices to prove that $\alpha_{-1}(A) = \left(\begin{smallmatrix} a & -b \\ -c & d \end{smallmatrix}\right)$ is conjugate to A^{-1} for all $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})$, where $N = p^r$ is a prime power. For the p-adic valuation we have w.l.o.g. $v_p(a-d) \geq \min\{v_p(b), v_p(c)\} =: s$ (otherwise consider $A' = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)A^{-1} = \left(\begin{smallmatrix} a-b & b \\ a-d+b+c & b+d \end{smallmatrix}\right)$). If $s = v_p(c)$ then $\frac{c}{p^s} \in (\mathbb{Z}/p^r\mathbb{Z})^*$ and hence $S := \left(\begin{smallmatrix} 1 & \frac{a-d}{p^s}(\frac{c}{p^s})^{-1} \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z}/p^r\mathbb{Z})$. Observing that $c\frac{a-d}{p^s}(\frac{c}{p^s})^{-1} = a - d$ (in $\mathbb{Z}/p^r\mathbb{Z}$) we find $S\alpha_{-1}(A)S^{-1} = A^{-1}$. In the case that $s = v_p(b)$ we use $S = \left(\begin{smallmatrix} 1 & 0 \\ \frac{d-a}{p^s}(\frac{b}{p^s})^{-1} & 1 \end{smallmatrix}\right)$.

2 The diagonal quotient surfaces $Z_{N,\varepsilon}$ and $\tilde{Z}_{N,\varepsilon}$

2.1 The singularities of $Z_{N,\varepsilon}$

Now that we have calculated the local invariants $s_{\nu,\alpha_\varepsilon}$ of the modular curve $X = X(N)$, we can apply the general results of [10] to determine the geometry of the *modular diagonal quotient surface* $Z_\varepsilon = Z_{N,\varepsilon} = \Delta_\varepsilon \backslash (X(N) \times X(N))$, where $\Delta_\varepsilon = \{(g, \alpha_\varepsilon(g)) : g \in G_N\}$.

We begin by summarizing the nature of the singularities of $Z_{N,\varepsilon}$. Recall from [10] that every diagonal quotient surface has only cyclic quotient singularities and that these may be classified according to their type (n, q); this means that the singularity in question has the form $A_{n,q} = C_n \backslash \mathbb{C}^2$, where the cyclic group C_n acts on $\mathbb{C}^2$ via $\sigma(z_1, z_2) = (\chi_1(\sigma)z_1, \chi_2(\sigma)z_2)$ where the "weights" $\chi_i : C_n \to \mathbb{C}^\times$ have order n and satisfy $\chi_1 = \chi_2^q$.

To state the result, we let φ and ψ denote the canonical quotient maps induced by the inclusion of subgroups $\{1\} \leq \Delta_\varepsilon \leq G_N \times G_N$:

$$Y = X(N) \times X(N) \xrightarrow{\varphi} Z_\varepsilon \xrightarrow{\psi} \bar{Y} = X(1) \times X(1) \cong \mathbb{P}^1 \times \mathbb{P}^1.$$

Theorem 2.1 *a) For $N \geq 5$ the set S_ε of singularities of Z_ε decomposes into the disjoint union of the sets $S_{k,\varepsilon} = S_\varepsilon \cap \psi^{-1}(\bar{P}_k, \bar{P}_k)$ of singularities lying over $(\bar{P}_k, \bar{P}_k) \in \bar{Y}$, where $k = 0, 1, \infty$.*

b) The total number of singularities over $(\bar{P}_k, \bar{P}_k)$ is given by $|S_{k,\varepsilon}| = r_k$, where r_k is as in Proposition 1.3; in particular, this number does not depend on ε. Moreover:

1. *All r_0 singularities in $S_{0,\varepsilon}$ are of type $(2, 1)$.*
2. *We have $r_1 = s_{1,1,\varepsilon} + s_{1,2,\varepsilon}$, where $s_{1,q,\varepsilon}$ denotes the number of singularities in $S_{1,\varepsilon}$ of type $(3, q)$, $q = 1, 2$. Moreover, if $3 \not| N$, then both types occur equally often (so $s_{1,q,\varepsilon} = \frac{r_1}{2}$), whereas if $3|N$, then all r_1 singularities are of type $(3, q)$, where $q \equiv \varepsilon \pmod 3$.*
3. *For each $d|N$, $d \neq N$, and each $q \in ((\mathbb{Z}/\frac{N}{d}\mathbb{Z})^\times)^2$, there are $\frac{1}{2}\rho(\frac{N}{d})\phi(d)$ singularities of type $(\frac{N}{d}, \varepsilon q)$ in $S_{\infty,\varepsilon}$, and other types do not occur. In particular, if N is prime, Z_ε has exactly one singularity of type (N, q) for every $1 \leq q < N$ with $\left(\frac{q}{N}\right) = \left(\frac{\varepsilon}{N}\right)$.*

Remark 2.2 In [5], Hilfssatz 1, Hermann claims that $|S_{0,\varepsilon}| = \frac{N}{2}\prod_{p|N}\left(1 - \left(\frac{-1}{p}\right)\right)$ and $|S_{1,\varepsilon}| = \frac{N}{3}\prod_{p|N}\left(1 - \left(\frac{-3}{p}\right)\right)$, rather than the above values $|S_{0,\varepsilon}| = \frac{N}{2}\prod_{p|N}\left(1 - \frac{1}{p}\left(\frac{-1}{p}\right)\right)$ and $|S_{1,\varepsilon}| = r_1 = \frac{N}{3}\prod_{p|N}\left(1 - \frac{1}{p}\left(\frac{-3}{p}\right)\right)$. Presumably this is a typographical error in [5].

Proof of Theorem 2.1. a) In view of [10], Theorem 2.3b), this is equivalent to the assertion that $s_{\nu,\alpha_\varepsilon}(\bar{x}, \bar{y}) = 0$ if $\bar{x} \neq \bar{y}$ and $1 \leq \nu < e_{\bar{y}}$, which is true by Theorem 1.6a).

b) By [10], Theorem 2.3c), we have $|S_{k,\varepsilon}| = r_{k,\alpha_\varepsilon}$, which by (4) does not depend on ε (and hence is given by Proposition 1.3). The last assertions follow immediately from the fact that if $d = (\nu, e_k)$, then the number of singularities of type $(\frac{e_k}{d}, \frac{\nu}{d})$ in $S_{k,\varepsilon}$ is $s_{\nu,\alpha_\varepsilon}(\bar{P}_k, \bar{P}_k)$ (cf. [10], Theorem 2.3) which by Theorem 1.6 has the indicated values.

The above theorem describes mainly the *number* of singularities of a given type. However, by analyzing the above proof, we see that we have also determined the singularities themselves as well in the process. To make this more precise, let us first introduce the following notation.

Notation 2.3 For $\varepsilon \in (\mathbb{Z}/N\mathbb{Z})^\times$, let

$$\begin{aligned} M^*_{k,\varepsilon}(N) &= \{(x,y) \in \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} : q_k(x,y) \equiv \pm\varepsilon \pmod N\}, \quad \text{if } k = 0, 1, \\ M^*_{\infty,\varepsilon}(N) &= \{(x,y) \in \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} : \gcd(x,y,N) = 1, \gcd(y,N) \neq 1\}, \end{aligned}$$

where $q_0(x,y) = x^2 + y^2$ and $q_1(x,y) = x^2 + xy + y^2$; cf. Notation 1.8. Furthermore, if $d|e_k$, $d \neq e_k$ and $1 \leq n < e_k$, $(n, \frac{e_k}{d}) = 1$, then let $M^*_{k,\varepsilon,d,n} := M^*_{k,\varepsilon,d,n}(N) \subset M^*_{k,\varepsilon} := M^*_{k,\varepsilon}(N)$ be defined as follows:

$$\begin{array}{rclcrcl} M^*_{0,\varepsilon,1,1} &=& M^*_{0,\varepsilon} &=& M_{0,\varepsilon} \cup M_{0,-\varepsilon}, & & \\ M^*_{1,\varepsilon,1,1} &=& M_{1,\varepsilon}, & & M^*_{1,\varepsilon,1,2} &=& M_{1,-\varepsilon}, \\ M^*_{\infty,\varepsilon,d,n} &=& M^{(2)}_{\infty,d,\varepsilon n}. & & & & \end{array}$$

It thus follows from the definitions that

$$M^*_{k,\varepsilon} = \dot{\bigcup_{d,n}} M^*_{k,\varepsilon,d,n},$$

and from Lemma 1.9 that

$$p_\varepsilon(N^*_{d,n}(\sigma_k, \alpha_\varepsilon^{-1}(\sigma_k))) = M^*_{k,d,n,\varepsilon}/\{\pm 1\},$$

where $p_\varepsilon(\left(\begin{smallmatrix} x & y \\ z & w \end{smallmatrix}\right)) = (\varepsilon x, z)$. In particular, for each $(x,y) \in M^*_{k,\varepsilon,d,n}$ there is an element $\sigma_{k,\varepsilon,x,y} \in N^*_{d,n}(\sigma_k, \alpha_\varepsilon^{-1}(\sigma_k)) \subset \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ such that $p_\varepsilon(\sigma_{k,\varepsilon,x,y}) = (x,y)$. (For example, $\sigma_{0,\varepsilon,x,y} = Q_\varepsilon^{-1}\left(\begin{smallmatrix} x & -y \\ y & x \end{smallmatrix}\right) = \left(\begin{smallmatrix} \varepsilon^{-1}x & -\varepsilon^{-1}y \\ y & x \end{smallmatrix}\right)$, if $(x,y) \in M_{0,\varepsilon}$, and $\sigma_{0,\varepsilon,x,y} = Q_{-\varepsilon}^{-1}\left(\begin{smallmatrix} -x & y \\ y & -x \end{smallmatrix}\right) = \left(\begin{smallmatrix} \varepsilon^{-1}x & \varepsilon^{-1}y \\ y & -x \end{smallmatrix}\right)$, if $(x,y) \in M_{0,-\varepsilon}$, etc.)

Note that if we view the elements of $M^*_{k,\varepsilon}$ as column vectors, then $\pm\sigma_k$ acts on $M^*_{k,\varepsilon,d,n}$ and hence also on $M^*_{k,\varepsilon}$; we shall denote the quotient space of this action by $\overline{M}^*_{k,\varepsilon,d,n} = \langle \pm\sigma_k \rangle \backslash M^*_{k,\varepsilon,d,n}$, and by $\overline{M}^*_{k,\varepsilon} = \langle \pm\sigma_k \rangle \backslash M^*_{k,\varepsilon}$.

Corollary 2.4 *If $N \geq 5$, then the map $(a,c) \mapsto z_{k,\varepsilon,a,c} := \varphi(\sigma_{k,\varepsilon,a,c} P_k, P_k)$ induces a natural bijection*

$$\zeta_{k,\varepsilon} : \overline{M}^*_{k,\varepsilon} \quad \tilde{\to} \quad S_{k,\varepsilon}$$

between the set $\overline{M}^*_{k,\varepsilon} := \langle \pm\sigma_k \rangle \backslash M^*_{k,\varepsilon}$ of $\pm\sigma_k$-orbits of $M^*_{k,\varepsilon}$ and the set $S_{k,\varepsilon}$ of singularities of Z_ε which lie above $(\bar{P}_k, \bar{P}_k)$. Furthermore, if $(a,c) \in M^*_{k,\varepsilon,d,n}$, then $z_{k,\varepsilon,a,c} = \zeta_{k,\varepsilon}(a,c)$ is a singularity of type $(\frac{e_k}{d}, q)$, where $qn \equiv 1 \pmod{\frac{e_k}{d}}$. In particular, if $a, c \in \mathbb{Z}$ satisfy the conditions $d := (c, N) > 1$ and $(a,d) = 1$, then $z_{\infty,\varepsilon,\varepsilon a,c} = (\tilde{\pi}(\frac{a}{c}), P_\infty)$ is a singularity of type (d,q), where $q\varepsilon a^2 \equiv 1\,(d)$ and $\tilde{\pi} : \mathfrak{H}^* \to X(N) = \Gamma(N)\backslash\mathfrak{H}^*$ denotes the quotient map.

Proof. We first note that it follows from Lemma 1.9 that the above map $p_\varepsilon : \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto (\varepsilon a, c)$ induces a bijection

$$\bar{p}_\varepsilon : G_{k,\varepsilon} \backslash N^*_{d,n}(\sigma_k, \sigma'_k)/G_k \quad \tilde{\to} \quad \overline{M}^*_{k,\varepsilon,d,n},$$

where $\sigma'_k = \alpha_\varepsilon^{-1}(\sigma_k)$. On the other hand, since $\lambda_k(\sigma_k) = \lambda_{k,\varepsilon}(\sigma'_k)$ (cf. the proof of Theorem 1.6), it follows from [10], Corollary 2.5c), that the map $g \mapsto \varphi(gP_k, P_k)$ induces a bijection

$$G_{k,\varepsilon} \backslash N^*_{d,n}(\sigma_k, \sigma'_k)/G_k \quad \tilde{\to} \quad S_{dq,\alpha_\varepsilon}(\bar{P}_k, \bar{P}_k),$$

where $S_{dq,\alpha_\varepsilon}(\bar{P}_k, \bar{P}_k)$ denotes the set of singularities lying above $(\bar{P}_k, \bar{P}_k)$ which are of type $(\frac{e_k}{d}, q)$, where $nq \equiv 1 \pmod{\frac{e_k}{d}}$. Combining these two bijections, we see that the assignment $\zeta_{k,\varepsilon} : (a,c) \mapsto \varphi(\bar{p}_\varepsilon^{-1}(a,c)P_k, P_k) = \varphi(\sigma_{k,\varepsilon,a,c} P_k, P_k)$ yields the desired bijection.

2.2 The configuration of exceptional curves on $\tilde{Z}_{N,\varepsilon}$

We now turn our attention to the minimal desingularization $\sigma : \tilde{Z}_\varepsilon \to Z_\varepsilon$ of Z_ε, which is obtained from Z_ε by replacing each singularity $s \in S_\varepsilon$ by an "exceptional curve" E_s. Here $E_s = C_{s,1} + \ldots + C_{s,r_s}$ is a chain of $\mathbb{P}^1$'s whose length r_s and self-intersection numbers $C_{s,j}^2 = -c_{s,j}$ are determined by the continued fraction expansion $\frac{n_s}{q_s} = [[c_{s,1}, \ldots, c_{s,r_s}]]$ of $\frac{n_s}{q_s}$, where (n_s, q_s) denotes the type of the singularity $s \in S_\varepsilon$; cf. [1], III.2 or [10], section 3.2.

We can view the surface $\tilde{Z}_\varepsilon$ as an $X(N)$-fibration over $\bar{X} \simeq \mathbb{P}^1$ in two ways. Indeed, if $\tilde{\psi}_i : \tilde{Z}_\varepsilon \to \bar{X}$ denotes the composition of $\tilde{\psi} = \psi \circ \sigma : \tilde{Z}_\varepsilon \to \bar{Y}$ with the i-th projection map $pr_i : \bar{Y} = \bar{X} \times \bar{X} \to \bar{X}$, where $i = 1, 2$, then all except three of the fibres $\tilde{\psi}_i^{-1}(\bar{x})$ of $\tilde{\psi}_i$ are isomorphic to $X(N)$ (cf. [10], Proposition 2.1) whereas those over $\bar{P}_k$, $k = 0, 1, \infty$, are reducible: their components consist of the above exceptional curves $C_{s,j}$ together with the curves $\tilde{C}_{k,i}$ which are the proper transforms under σ of the curves $C_{k,1} = \varphi(\bar{P}_k \times X)$ and $C_{k,2} = \varphi(X \times \bar{P}_k)$ on Z_ε. More precisely, we have:

Proposition 2.5 *The divisor $E = \sum_{s \in S_\varepsilon} E_s$ of exceptional curves on $\tilde{Z}_\varepsilon$ is the sum of the three disjoint curves $E_k = \sum_{s \in S_{k,\varepsilon}} E_s = \sum_{s \in S_{k,\varepsilon}} \sum_{j=1}^{r_s} C_{s,j}$, where $k = 0, 1, \infty$. If we put*

$$D_k \;=\; \tilde{C}_{k,1} + \tilde{C}_{k,2} + E_k, \quad D_k^* \;=\; \tilde{\psi}_1^*(\bar{P}_k) + \tilde{\psi}_2^*(\bar{P}_k), \quad \text{and} \quad D_k' \;=\; D_k^* - D_k,$$

then we have:

a) E_0 is the sum of r_0 disjoint (-2)-curves joining $\tilde{C}_{0,1}$ to $\tilde{C}_{0,2}$, and we have

$$\tilde{\psi}_i^*(\bar{P}_0) = 2\tilde{C}_{0,i} + E_0, \quad \text{for } i = 1, 2, \quad \text{and hence} \quad D_0^* = 2D_0, \quad D_0' = D_0. \tag{9}$$

b) $E_1 = E_{11} + E_{12}$ decomposes into two disjoint curves E_{11} and E_{12}, where E_{11} consists of $s_{1,1,\varepsilon}$ disjoint (-3)-curves joining $\tilde{C}_{1,1}$ to $\tilde{C}_{1,2}$, whereas E_{12} consists of $s_{1,2,\varepsilon}$ disjoint (-2)-chains of length 2 joining $\tilde{C}_{1,1}$ to $\tilde{C}_{1,2}$. Moreover, if E_{12i} denotes the sum of the components of E_{12} which meet $\tilde{C}_{1,i}$, then

$$\tilde{\psi}_i^*(\bar{P}_1) = 3\tilde{C}_{1,i} + E_1 + E_{12i}, \text{ for } i = 1, 2, \text{ and thus } D_1^* = 3D_1 - E_{11},\ D_1' = 2D_1 - E_{11}. \tag{10}$$

c) $E_\infty = \sum_{s \in S_\infty} E_s$ decomposes into r_∞ disjoint $\mathbb{P}^1$-chains $E_s = C_{s,1} + \ldots + C_{s,r_s}$ indexed by $s \in S_{\infty,\varepsilon} = \zeta_{\infty,\varepsilon}(M_{\infty,\varepsilon}^*)$. Moreover, if $s = z_{\infty,\varepsilon,a,c}$, where $n_s := (c, N) \neq 1$, and $(a, n_s) = 1$, then s is a singularity of type (n_s, q_s), where $a^2 q_s \equiv \varepsilon \pmod{n_s}$ and hence $C_{s,j}^2 = -c_{s,j}$, where $\frac{n_s}{q_s} = [[c_{s,1}, \ldots, c_{s,r_s}]]$ is the continued fraction expansion of $\frac{n_s}{q_s}$. Furthermore,

$$\tilde{\psi}_i^*(\bar{P}_\infty) = N\tilde{C}_{\infty,i} + \sum_{s \in S_{\infty,\varepsilon}} \sum_{j=1}^{r_s} a_{s,i,j} C_{s,j}, \tag{11}$$

where $i = 1, 2$ and the coefficients $a_{s,i,j}$ are determined from the continued fraction expansion of $\frac{n_s}{q_s}$ by the recursion relations $a_{s,i,j+1} = c_{s,j} a_{s,i,j} - a_{s,i,j-1}, 1 \leq j \leq r_s$, together with the boundary conditions $a_{s,1,0} = a_{s,2,r_s+1} = N$, $a_{s,1,r_s+1} = a_{s,2,0} = 0$.

d) Each $\tilde{C}_{k,i}$ is a smooth curve of genus g_k whose self-intersection number is given by:

$$\tilde{C}_{0,i}^2 = -r_0/2,\ \tilde{C}_{1,i}^2 = -(2r_1 - s_{1,1,\varepsilon})/3, \text{ and } \tilde{C}_{\infty,i}^2 = -\sum_{\nu=1}^{N-1} \phi((\nu, N)) \left\langle \frac{\nu^2 \varepsilon}{N(\nu, N)} \right\rangle,$$

where $\langle x\rangle = x - [x]$ *denotes the fractional part. Furthermore, we have* $(\tilde{C}_{l,1}.\tilde{C}_{k,2}) = r_1(\bar{P}_l, \bar{P}_k) > 0$*, where* $r_1(\bar{P}_l, \bar{P}_k)$ *is as in Theorem 1.6.*

e) The canonical divisor K_ε *of* $\tilde{Z}_\varepsilon$ *is given by*

$$K_\varepsilon \ \sim\ -3D_0 + D'_1 + D'_\infty \ \sim\ D_0 - D_1 - D_\infty. \tag{12}$$

Proof. The first assertion follows from Theorem 2.1a). Moreover, since $e_0 = 2$, $e_1 = 3$, and $e_\infty = N$, a), b) and c) follow easily from Propositions 3.5, 3.13 and 3.12 of [10], together with Theorem 2.1b). (For c), use also Corollary 2.4.) Moreover, d) follows from Corollary 3.15 of [10] by using Theorem 1.6. Finally, for e) we note that the different divisor $D(\tilde{\psi})$ (as defined in [10], Theorem 3.10) is by definition $D(\tilde{\psi}) = D'_0 + D'_1 + D'_\infty$ and that $\tilde{\psi}^*(K_{\bar{Y}}) \sim -2D^*_k$, for any $k = 0, 1, \infty$, and so the first equivalence in (12) follows from [10], Theorem 3.10, and (9). The second follows from the first by noting that $D'_1 + D'_\infty \sim 2D^*_0 - D_1 - D_\infty \sim 4D_0 - D_1 - D_\infty$ by (9) again.

2.3 The invariants of the surface $\tilde{Z}_\varepsilon$

We now turn to study the numerical invariants of $\tilde{Z}_\varepsilon$. Recall that by Theorem 3.7 of [10], the Betti, Hodge and Chern numbers of $\tilde{Z}_\varepsilon$ can all be expressed in terms of the three fundamental invariants $\mathbb{G}_\varepsilon$, $\mathbb{S}_\varepsilon$ and $\mathbb{L}_\varepsilon$. Here these invariants have the following explicit values.

Theorem 2.6 *The invariants* $\mathbb{G}_\varepsilon$*,* $\mathbb{S}_\varepsilon$ *and* $\mathbb{L}_\varepsilon$ *are given by*

$$\begin{aligned}
\mathbb{G}_\varepsilon &= \frac{m(N-12)}{144N} - 1 + \frac{1}{8}\phi(N) + \frac{1}{8}r_0 + \frac{1}{6}r_1 + \frac{1}{4}r_\infty,\\
\mathbb{S}_\varepsilon &= \frac{1}{18}\left(2s_{1,1,\varepsilon} - r_1\right) + \mathbb{S}_{\infty,\varepsilon}\\
\mathbb{L}_\varepsilon &= r_0 + 2r_1 - s_{1,1,\varepsilon} + \mathbb{L}_{\infty,\varepsilon},
\end{aligned}$$

where r_0, r_1, r_∞ *and* $s_{1,1,\varepsilon}$ *are as in Theorem 2.1 and*

$$\mathbb{S}_{\infty,\varepsilon} \ =\ \sum_{\substack{d|N\\ d\neq N}} \frac{\phi(d)}{2} \sum_{\substack{n=1\\ (n,\frac{N}{d})=1}}^{\frac{N}{d}} \mathbb{S}\left(\varepsilon n^2, \frac{N}{d}\right) \quad \text{and} \quad \mathbb{L}_{\infty,\varepsilon} \ =\ \sum_{\substack{d|N\\ d\neq N}} \frac{\phi(d)}{2} \sum_{\substack{n=1\\ (n,\frac{N}{d})=1}}^{\frac{N}{d}} \mathbb{L}\left(\frac{N}{\varepsilon n^2 d}\right),$$

where $\mathbb{S}(q,n) = \sum_{k=1}^{n-1} \left(\left(\frac{k}{n}\right)\right)\left(\left(\frac{kq}{n}\right)\right)$ *denotes the usual Dedekind sum and* $\mathbb{L}(\frac{n}{q})$ *the length of the continued fraction expansion of* $\frac{n}{(q)}$*, where* (q) *is the least positive residue of* q *mod* n*. Thus, if we put*

$$R_{\infty,\varepsilon} \ :=\ 12\mathbb{S}_{\infty,\varepsilon} + \mathbb{L}_{\infty,\varepsilon} + r_\infty \ =\ \sum_{\substack{d|N\\ d\neq N}} \frac{\phi(d)}{2} \sum_{\substack{n=1\\ (n,\frac{N}{d})=1}}^{\frac{N}{d}} \left(\left\langle \frac{\varepsilon n^2 + \varepsilon^*(n^*)^2}{N/d}\right\rangle + \mathbb{L}\left(\frac{N}{N-n^2\varepsilon}\right)\right),$$

where $n^* n \equiv 1 \pmod{\frac{N}{d}}$ *and* $0 < n^* < \frac{N}{d}$*, then the geometric genus* $p_{g,\varepsilon}$*, and the Chern numbers* $c_1^2 = K_\varepsilon^2$ *and* $c_2 = \chi_{top,\varepsilon}$ *of* $\tilde{Z}_\varepsilon$ *are given by*

$$
\begin{aligned}
p_{g,\varepsilon} &= \mathbb{G}_\varepsilon - \mathbb{S}_\varepsilon = \frac{m(N-12)}{144N} + \frac{\phi(N)}{8} + \frac{r_0}{8} + \frac{2r_1}{9} - \frac{s_{1,1,\varepsilon}}{9} + \frac{r_\infty}{3} + \frac{\mathbb{L}_{\infty,\varepsilon}}{12} - \frac{R_{\infty,\varepsilon}}{12} - 1, \\
K_\varepsilon^2 &= 8 + 8\mathbb{G}_\varepsilon - \mathbb{L}_\varepsilon - 12\mathbb{S}_\varepsilon = \frac{m(N-12)}{18N} + \phi(N) - \frac{s_{1,1,\varepsilon}}{3} + 3r_\infty - R_{\infty,\varepsilon}, \\
\chi_{top,\varepsilon} &= 4 + 4\mathbb{G}_\varepsilon + \mathbb{L}_\varepsilon = \frac{m(N-12)}{36N} + \frac{\phi(N)}{2} + \frac{3r_0}{2} + \frac{8r_1}{3} - s_{1,1,\varepsilon} + r_\infty + \mathbb{L}_{\infty,\varepsilon}.
\end{aligned}
$$

Proof. The formula for $\mathbb{G}_\varepsilon := \frac{1}{2}(g - g_{0,\varepsilon} - g_{1,\varepsilon} - g_{\infty,\varepsilon})$ (cf. [10], Remark 1.5), follows immediately from Propositions 1.1 and 1.3. Furthermore, by definition (cf. [10], Theorem 3.7) we have $\mathbb{S}_\varepsilon = \mathbb{S}_{0,\varepsilon} + \mathbb{S}_{1,\varepsilon} + \mathbb{S}_\varepsilon$ where $\mathbb{S}_{k,\varepsilon} = \sum_{\nu=1}^{e_k-1} s_{\nu,\varepsilon}(\bar{P}_k)\mathbb{S}(\frac{\nu}{(\nu,e_k)}, \frac{e_k}{(\nu,e_k)})$ denotes the contribution from $\bar{P}_k$. Since $\mathbb{S}(1,2) = 0$ and $\mathbb{S}(1,3) = -\mathbb{S}(2,3) = \frac{1}{18}$, it follows that $\mathbb{S}_{0,\varepsilon} = 0$ and $\mathbb{S}_{1,\varepsilon} = \frac{1}{18}s_{1,1,\varepsilon} - \frac{1}{18}s_{1,2,\varepsilon} = \frac{1}{18}(2s_{1,1,\varepsilon} - r_1)$ (cf. Theorem 2.1b)). Moreover, by Theorem 2.1b) we have

$$
\mathbb{S}_{\infty,\varepsilon} = \sum_{\substack{d|N \\ d\neq N}} \frac{\phi(d)}{2} \sum_{\substack{n=1 \\ (n,\frac{N}{d})=1}}^{\frac{N}{d}} \rho\left(\varepsilon n, \frac{N}{d}\right) \mathbb{S}\left(n, \frac{N}{d}\right).
$$

This equals the asserted expression because the inner sum runs only over the n satisfying $\varepsilon n \equiv x^2 \pmod{\frac{N}{d}}$ for some x, and the number of x satisfying $\varepsilon x^2 \equiv n \pmod{\frac{N}{d}}$ is precisely $\rho(\varepsilon n, \frac{N}{d}) = \rho(\frac{N}{d})$.

The proof for $\mathbb{L}_\varepsilon$ is similar, using the fact that $\mathbb{L}(\frac{2}{1}) = \mathbb{L}(\frac{3}{1}) = 1$ and $\mathbb{L}(\frac{3}{2}) = 2$.

The indicated expression for $R_{\infty,\varepsilon}$ follows directly from the above expressions for $\mathbb{S}_{\infty,\varepsilon}$ and $\mathbb{L}_{\infty,\varepsilon}$, together with the identity

$$
12\mathbb{S}(q,n) + \mathbb{L}\left(\frac{n}{q}\right) + 1 = \mathbb{L}\left(\frac{n}{n-q}\right) + \frac{q}{n} + \frac{q^*}{n} \leq n, \tag{13}
$$

which was established in [10], Proposition 1.15. Finally, the first expression for $p_{g,\varepsilon}$ (and similarly, those for K_ε^2 and $\chi_{top,\varepsilon}$) is that given in Theorem 3.7 of [10], and the second follows from the first by substituting the above values of $\mathbb{G}_\varepsilon$, $\mathbb{S}_\varepsilon$ and $\mathbb{L}_\varepsilon$.

Remark 2.7 Some of above invariants have a natural character–theoretic interpretation: we have by Theorem 3.7 of [10] that

$$
\mathbb{G}_\varepsilon = \frac{1}{4}(h^1, h^1_\varepsilon),\ \mathbb{S}_\varepsilon = \frac{1}{4}(h^1, h^1_\varepsilon) - (\omega, \overline{\omega_\varepsilon}),\ \text{ and }\ p_{g,\varepsilon} = (\omega, \overline{\omega_\varepsilon}),
$$

where h^1_ε and ω_ε are as in Corollary 1.13 and Proposition 1.14. It thus follows from [10], Corollary 3.3, together with Proposition 1.14, that $p_g(\tilde{Z}_\varepsilon) + p_g(\tilde{Z}_{-\varepsilon}) = 2\mathbb{G}_\varepsilon$ is independent of ε, and hence, if -1 is a square modulo N, then the same is true for $p_g(\tilde{Z}_\varepsilon) = \mathbb{G}_\varepsilon$ itself.

In order to be able to classify the surfaces according to the Enriques–Kodaira classification scheme, we require some good lower bounds for K_ε^2 and p_g.

Proposition 2.8 *The following inequalities hold for all $N \geq 5$ and all ε:*

$$
K_\varepsilon^2 > \frac{m}{18N^2}(N-1)(N-30) > \frac{1}{60}N(N-1)(N-30). \tag{14}
$$

Moreover, we have $K_\varepsilon^2 > 0$ whenever $N \geq 17$ or $N \geq 13$ and $\varepsilon \not\equiv 1 \mod ((\mathbb{Z}/N\mathbb{Z})^\times)^2$.

Proof. We shall bound each of the terms appearing in the formula for K_ε^2. For $s_{1,1,\varepsilon}$ we have the estimate:

$$s_{1,1,\varepsilon} \le r_1 = \frac{N}{3}\prod_{p|N}\left(1-\frac{1}{p}\left(\frac{-3}{p}\right)\right) \le \frac{N}{3}\prod_{p|N}\left(1+\frac{1}{p}\right) = \frac{2m}{3N\phi(N)} \le \frac{m}{6N},$$

since $\phi(N) \ge 4$ if $N \ge 5$. Moreover, by Proposition 1.3d) and identity (8) we have

$$r_\infty \;=\; \frac{1}{2}\sum_{d|N,d\neq N}\phi(d)\phi(\frac{N}{d}) \;\ge\; \frac{1}{2N}\sum_{d|N,d\neq N} d\phi(d)\phi(\frac{N}{d}) \;=\; \frac{m}{N^2}-\frac{\phi(N)}{2}. \tag{15}$$

Furthermore, using the inequality (13) and Theorem 1.6d) we obtain

$$R_{\infty,\varepsilon} \;\le\; \sum_{1<d|N} r_{d,\varepsilon}(\bar{P}_\infty,\bar{P}_\infty)d \;=\; \frac{1}{2}\sum_{d|N,d\neq N}\frac{N}{d}\phi(d)\phi(\frac{N}{d}) \;=\; \frac{m}{N}-\frac{\phi(N)}{2}, \tag{16}$$

the latter by (8) again. Substituting these estimates in the formula for K_ε^2 in Theorem 2.6 yields

$$\begin{aligned} K_\varepsilon^2 &> \frac{m(N-12)}{18N}+\phi(N)-\frac{m}{18N}+3\left(\frac{m}{N^2}-\frac{\phi(N)}{2}\right)-\frac{m}{N}+\frac{\phi(N)}{2} \\ &= \frac{m}{18N^2}(N^2-31N+54) \\ &> \frac{m}{18N^2}(N^2-31N+30) = \frac{m}{18N^2}(N-1)(N-30), \end{aligned}$$

which proves the first of the desired inequalities (14). The second follows from this because

$$\frac{m}{N^3} \;=\; \frac{1}{2}\prod_{p|N}\left(1-\frac{1}{p^2}\right) \;\ge\; \frac{1}{2}\prod_{p}\left(1-\frac{1}{p^2}\right) \;=\; \frac{1}{2\zeta(2)} \;=\; \frac{3}{\pi^2} \;>\; \frac{3}{10}. \tag{17}$$

From (14) we see that $K_\varepsilon^2 > 0$ if $N \ge 30$. By computing the values K_ε^2 for $N \le 29$ (cf. Tables 1 and 2 below), one verifies that this is also true if $N \ge 16$ or if $N \ge 13$ and $\varepsilon \not\equiv 1 \pmod{((\mathbb{Z}/N\mathbb{Z})^\times)^2}$.

Proposition 2.9 *For $N \ge 9$ or for $N \ge 6$ and $\varepsilon \not\equiv 1 \bmod ((\mathbb{Z}/N\mathbb{Z})^\times)^2$, the geometric genus of $\tilde{Z}_\varepsilon$ is positive: $p_{g,\varepsilon} > 0$. Moreover,*

$$p_{g,\varepsilon} > \frac{1}{480}N(N-1)(N-23), \tag{18}$$

and hence we have $p_{g,\varepsilon} \ge 3$ if $N \ge 13$.

Proof. We first derive the inequality (18). Since r_0, $\mathbb{L}_{\infty,\varepsilon}$ and $r_1 \ge s_{1,1,\varepsilon}$ are positive, substituting the estimates (15) and (16) in the formula for $p_{g,\varepsilon}$ of Theorem 2.6 gives

$$\begin{aligned} p_g(\tilde{Z}_\varepsilon)+1 &> \frac{m(N-12)}{144N}+\frac{1}{8}\phi(N)+\frac{1}{3}\left(\frac{m}{N^2}-\frac{\phi(N)}{2}\right)-\frac{1}{12}\left(\frac{m}{N}-\frac{\phi(N)}{2}\right) \\ &= \frac{m}{144N^2}(N^2-24N+48). \end{aligned}$$

Using the lower bound (17) we obtain

$$p_g \;>\; \frac{3N}{1440}(N^2-24N+48)-1 \;=\; \frac{1}{480}N(N-1)(N-23)+\frac{5}{96}N-1,$$

from which (18) follows readily.

If $N \geq 25$, then (18) shows that $p_g > \frac{1}{480}25\cdot 24\cdot 2 = \frac{5}{2} > 2$, which proves the last assertion for $N \geq 25$. For small N, the corresponding statement follows by explicit computations of the values (cf. Tables 1 and 2). Similarly, the first assertions may be verified in this way.

Remark 2.10 We can compute the geometric genus of $\tilde{Z}_\varepsilon$ not only by the Noether formula as above but also with the help of the character ω. This yields a formula involving the class numbers $h(-d)$ of the imaginary quadratic fields $\mathbb{Q}(\sqrt{-d})$, where $d|N$. (As usual, we adopt the convention that $h(-3) = \frac{1}{3}$ and $h(-4) = \frac{1}{2}$; cf. Notation 1.2.) To be precise, we have for $N \geq 4$ the following formula:

$$p_g(\tilde{Z}_\varepsilon) = \mathbb{G}_\varepsilon + \frac{1}{18}(s_{1,1,\varepsilon}-s_{1,2,\varepsilon}) - \frac{1}{2}\sum_{\nu|N}\phi(\nu)\phi(\frac{N}{\nu})\sum_{\mu|\nu}\frac{1}{\phi(\mu)}\sum_{\substack{d|\mu \\ -d \text{ fund}}}\chi_d(\varepsilon)\prod_{\substack{p|\mu \\ p\nmid d}}(1-\chi_d(p))^2 h(-d)^2,$$

where the above sum runs only over those $d|N$ such that $-d$ is a fundamental discriminant, and χ_d denotes the usual quadratic character on $(\mathbb{Z}/N\mathbb{Z})^\times$ defined by $\chi_d(x) = \left(\frac{-d}{x}\right)$. Since we do not require this formula in the sequel, we shall not present its proof here.

2.4 Classification

We shall now classify the surfaces $\tilde{Z}_\varepsilon$ according to their types in the Enriques–Kodaira classification table. For this, we may and shall view $\varepsilon \in (\mathbb{Z}/N\mathbb{Z})^\times/((\mathbb{Z}/N\mathbb{Z})^\times)^2$ since the isomorphism type of the diagonal quotient surfaces $\mathbb{Z}_\varepsilon$ depends only on the image of α_ε in the outer isomorphism group $\mathrm{Out}(G_N)$ (cf. Remark 1.7 and [10], Remark 3.9). It turns out rather surprisingly that the type of $\tilde{Z}_\varepsilon$ is completely determined by its geometric genus p_g:

Theorem 2.11 *a) $\tilde{Z}_\varepsilon$ is a rational surface if and only if $p_g(\tilde{Z}_\varepsilon) = 0$, and this is the case precisely for $N \leq 5$ or for $(N,\varepsilon) = (6,1)$, $(7,1)$ or $(8,1)$.*

b) $\tilde{Z}_\varepsilon$ is a (blown-up) elliptic K3-surface if and only if $p_g(\tilde{Z}_\varepsilon) = 1$, i.e. if and only if $(N,\varepsilon) = (6,5)$, $(7,3)$, $(8,3)$, $(8,5)$, $(9,1)$ or $(12,1)$.

c) $\tilde{Z}_\varepsilon$ is a (blown-up) elliptic surface with $\kappa = 1$ if and only if $p_g(\tilde{Z}_\varepsilon) = 2$. This is the case for $(N,\varepsilon) = (8,7)$, $(9,2)$, $(10,1)$, $(10,3)$ or $(11,1)$.

d) $\tilde{Z}_\varepsilon$ is a surface of general type if and only if $p_g(\tilde{Z}_\varepsilon) \geq 3$, or equivalently, if $N \geq 13$ or if $(N,\varepsilon) = (11,2)$, $(12,5)$, $(12,7)$ or $(12,11)$.

Remark 2.12 While this theorem is closely related to Satz 3 of Hermann[5], it is not identical with the latter. First of all, the interesting role of p_g in the classification of $\tilde{Z}_\varepsilon$ is not mentioned by Hermann. Secondly, the somewhat tenacious case $(N,\varepsilon) = (10,3)$ seems to be missing in Satz 3c) of [5]. Thirdly, Hermann claims on p. 96 that the residue classes of $(\mathbb{Z}/N\mathbb{Z})^\times/((\mathbb{Z}/N\mathbb{Z})^\times)^2$ are represented by the elements $\varepsilon \in (\mathbb{Z}/N\mathbb{Z})^\times$ with $\varepsilon^2 = 1$, which is false e.g. for primes $p \equiv 1 \bmod 4$ (in which case the only solutions of $\varepsilon^2 = 1$ are ± 1 which are both squares). Thus the list of cases considered by Hermann seems to be incomplete.

The proof of this theorem is based on the classification criteria of [10], section 4. Here we apply these criteria to the curves $C = \tilde{C}_{\infty,i}$, whose properties seem to control the type of $\tilde{Z}_\varepsilon$, as the following proposition shows. It is interesting to note that virtually all the cases of Theorem 2.11 are covered by this result.

Proposition 2.13 *a) If $g_\infty = 0$ and $\tilde{C}^2_{\infty,1} = -1$, then $\tilde{Z}_\varepsilon$ is rational.*

b) If $g_\infty = 0$ and $\tilde{C}^2_{\infty,1} = -2$, then $\tilde{Z}_\varepsilon$ is not of general type (i.e. $\kappa(\tilde{Z}_\varepsilon) \leq 1$). If, in addition, $p_g(\tilde{Z}_\varepsilon) = 1$ and either $\phi(N) > 4$ or $\phi(N) = 4$ and $-\varepsilon \in ((\mathbb{Z}/d\mathbb{Z})^\times)^2$ for some $d|N$, $d \neq 1$, then $\tilde{Z}_\varepsilon$ is a (blown-up) K3-surface; in the latter case, the minimal model of $\tilde{Z}_\varepsilon$ is also an elliptic surface.

c) If $g_\infty = 0$ and $\tilde{C}^2_{\infty,1} = -3$ and

$$s(N,\varepsilon) \;:=\; \frac{1}{2} \sum_{\substack{1<d|N \\ -\varepsilon\in((\mathbb{Z}/d\mathbb{Z})^\times)^2}} \rho(d)\phi\left(\frac{N}{d}\right) \;>\; 5 - \frac{1}{2}\phi(N),$$

then $\tilde{Z}_\varepsilon$ is not of general type.

d) If $g_\infty \geq 1$ and $d_\varepsilon := 2p_g(\tilde{Z}_\varepsilon) - 2g_\infty + \tilde{C}^2_{\infty,1} \geq 1$, then $\tilde{Z}_\varepsilon$ is of general type.

Proof. Most of this proposition follows from Theorem 6 of [10], applied to $x = P_\infty$. By Proposition 2.5d) we have, with the notation there:

$$c_{x,\varepsilon} \;=\; -\tilde{C}^2_{\infty,1} \;=\; -\tilde{C}^2_{\infty,2} \qquad \text{and} \qquad r_{1,x,\varepsilon} \;=\; (\tilde{C}_{\infty,1}.\tilde{C}_{\infty,2}) \;=\; \frac{1}{2}\phi(N); \tag{19}$$

in particular, the curves $\tilde{C}_{\infty,1}$ and $\tilde{C}_{\infty,2}$ always meet. Note that both are smooth of genus g_∞ (cf. Proposition 2.5d). In addition we note that by Theorem 1.6d) the invariant $s_{x,\varepsilon}$ of Theorem 6 is given by $s_{x,\varepsilon} := \sum_{d|e_x, d\neq e_x} s_{e_x-d,\varepsilon}(\bar{x},\bar{x}) = s(N,\varepsilon)$.

a) (cf. [10], Theorem 6a)) The hypothesis means that $\tilde{C}_{\infty,1}$ and $\tilde{C}_{\infty,2}$ are two (-1)-curves which meet, so $\tilde{Z}_\varepsilon$ is rational (cf. van der Geer[3], VII.2.2).

b) If $g \leq 1$, i.e. $N \leq 6$, then $\kappa(\tilde{Z}_\varepsilon) \leq 0$ by the classification Theorem 4.1c) of [10]. Thus, assume $N \geq 7$, so $\phi(N) \geq 4$. But then $\tilde{C}_{\infty,1}$ and $\tilde{C}_{\infty,2}$ meet in at least 2 points (counting multiplicities), so $\{\tilde{C}_{\infty,1}, \tilde{C}_{\infty,2}\}$ is a (-2)-configuration which cannot exist on a surface of general type (cf. vdG[3], VII.2.7). Thus $\kappa(\tilde{Z}_\varepsilon) \leq 1$.

If, in addition, $p_g = 1$ etc., then we can apply Theorem 6b) of [10]. Here the hypotheses imply that $k_{x,\varepsilon} := 4g_\infty - 2\tilde{C}^2_{\infty,1} - 4 = 0$ and $i_{x,\varepsilon} = 2r_{1,x,\varepsilon} + 2\tilde{C}_{\infty,\varepsilon} = \phi(N) - 4$. Thus, the hypotheses on N and ε guarantee that $s_{x,\varepsilon} > (k^2_{x,\varepsilon} - i_{x,\varepsilon})/2$, and so it follows from Theorem 6b) of [10] that $\tilde{Z}_\varepsilon$ is a K3-surface. Finally, to verify the last assertion we note that in the latter situation the hypotheses of [3], Proposition VII.2.9 apply (since $\{\tilde{C}_{\infty,1}, \tilde{C}_{\infty,2}\}$ forms an elliptic configuration), and so the last assertion follows.

c) In this case $k_{x,\varepsilon} = 2$ and $i_{x,\varepsilon} = \phi(N) - 6$, so again $s_{x,\varepsilon} > (k^2_{x,\varepsilon} - i_{x,\varepsilon})/2$, and the assertion follows from Theorem 6b) as before.

d) This follows directly from Theorem 6c) of [10].

Proof of Theorem 2.11. We first note that p_g has the asserted values. Indeed, if $N \geq 17$, then $p_g \geq 3$ by Proposition 2.9, whereas if $N \leq 16$, then Table 1 below shows that p_g is as stated in the theorem. Thus, since the cases a) – d) cover all possibilities and are mutually

exclusive, it is enough to show that $\tilde{Z}_\varepsilon$ has the indicated type for each of the values (N, ε) listed.

a) If $N \leq 5$, then $g = 0$, so $\tilde{Z}_\varepsilon$ is rational since Z_ε is dominated by $Y \simeq \mathbb{P}^1 \times \mathbb{P}^1$; cf. also [10], Theorem 4.1a). For the other three cases the hypotheses of Proposition 2.13a) are satisfied (cf. Table 1), so $\tilde{Z}_\varepsilon$ is rational in all these cases.

b) If $(N, \varepsilon) = (6, 1)$, then $\tilde{Z}_\varepsilon$ is an elliptic K3-surface by [10], Theorem 4.2c) because $g = 1$, $\bar{g} = 0$ and $p_g = 1$. Next, if $(N, \varepsilon) = (8, 3)$, $(8, 5)$ or $(12, 1)$, then $g_\infty = 0$ and $\tilde{C}^2_{\infty,i} = -2$, and $\phi(N) = 4$. Since $2|N$, the last criterion of Proposition 2.13b) applies and so $\tilde{Z}_\varepsilon$ is an elliptic K3-surface. Finally, if $(N, \varepsilon) = (7, 3)$ or $(9, 1)$, then we have $g_\infty = 0$, $\tilde{C}^2_{\infty,i} = -2$ and $\phi(N) = 6 > 4$, so $\tilde{Z}_\varepsilon$ is a (blown-up) K3-surface by Proposition 2.13b). It remains to show that the minimal models $\bar{Z}_\varepsilon$ of these two surfaces are elliptic.

Suppose first that $(N, \varepsilon) = (7, 3)$. Since $K^2_\varepsilon = -1$ and $\bar{Z}_\varepsilon$ is a K3-surface, we see that $K_\varepsilon = E$, where E is the unique blow-down curve. Consider $C := \tilde{C}_{1,1}$, which is a smooth elliptic curve on $\tilde{Z}_\varepsilon$. Thus, by the adjunction formula $(C.E) = (C.K_\varepsilon) = -C^2 = 1$, so the image $\bar{C}$ of C on $\bar{Z}_\varepsilon$ is a smooth elliptic curve with $\bar{C}^2 = 0$. Thus, $\bar{Z}_\varepsilon$ has an elliptic configuration and is therefore an elliptic K3-surface (cf. [3], Proposition VII.2.9).

Finally, suppose that $(N, \varepsilon) = (9, 1)$. Here $K^2_\varepsilon = -6$, so we have 6 blow-down curves $E_1, \ldots, E_6$. Three of these are: $E_1 = \tilde{\Gamma}$, the image of the diagonal (cf. [10], Remark 4.9), which meets a unique (-2)-curve E_2 lying in the fibre (of $\tilde{\psi}_1$) over $\bar{P}_0$ and a unique (-3)-curve E_3 in the fibre over $\bar{P}_1$. Furthermore, E_1 meets the fibre over $\bar{P}_\infty$ only in the (-9)-component. Now the latter fibre contains four (-3)-curves: three meet $\tilde{C}_{\infty,1}$, the fourth does not but is connected to $\tilde{C}_{\infty,1}$ by a (-2)-chain. Since the K3-surface $\bar{Z}_\varepsilon$ does not contain any curves with odd self-intersection number (by the adjunction formula), we see that each of these four (-3)-curves has to be met transversely by a blow-down curve. But we have only three such curves (since E_1, E_2 and E_3 cannot meet them), so one blow-down curve (say E_4) has to meet two of them. Thus, since $\tilde{C}_{\infty,1}$ is a (-2)-curve, we obtain, after blowing down E_4, an elliptic configuration consisting of a cycle of (-2)-curves, and hence $\bar{Z}_\varepsilon$ is again elliptic.

c) If $(N, \varepsilon) = (10, 3)$ or $(11, 1)$, then $\tilde{Z}_\varepsilon$ is an elliptic surface by Proposition 2.14 below. For the other three cases we have $s(N, \varepsilon) + \frac{1}{2}\phi(N) = 6$, so by Proposition 2.13c) we have $\kappa(\tilde{Z}_\varepsilon) \leq 1$. Since $p_g \geq 2$, we must have $\kappa = 1$, and so $\tilde{Z}_\varepsilon$ is a (blown-up) elliptic surface, as asserted.

d) By Proposition 2.8 we have $K^2_\varepsilon > 0$ if $N \geq 17$ or if $N \geq 13$ and $\varepsilon \not\equiv 1 \bmod ((\mathbb{Z}/N\mathbb{Z})^\times)^2$, so $\tilde{Z}_\varepsilon$ is of general type in those cases (cf. [10], Corollary 4.6). We are thus left with the cases

$$(11, 2),\ (14, 1),\ (15, 1),\ (16, 1).$$

But for each of these cases we have $g_\infty \geq 1$ and $d_\varepsilon := 2p_g(\tilde{Z}_\varepsilon) - 2g_\infty + C^2_{\infty,1} \geq 1$ (cf. Table 1), and so Proposition 2.13d) shows that these surfaces are all of general type as well.

To conclude the proof of Theorem 2.11, we still have to study the following two cases which seem to be much more subtle than the others in that they cannot be treated by general principles but instead require a detailed study of certain special curve configurations lying on them.

Proposition 2.14 *If $(N,\varepsilon) = (10,3)$ or $(11,1)$, then $\tilde{Z}_\varepsilon$ is a (blown-up) elliptic surface with $\kappa = 1$.*

Proof. We first observe that it is enough to show that $\kappa(\tilde{Z}_\varepsilon) \leq 1$ because $p_g(\tilde{Z}_\varepsilon) = 2$ and hence $\kappa(\tilde{Z}_\varepsilon) \geq 1$ in both cases.

a) Let us first consider the case that $(N,\varepsilon) = (10,3)$. Then (with the notation of 2.5) we have $D_\infty = \tilde{C}_{\infty,1} + \tilde{C}_{\infty,2} + E_{\infty,1} + \ldots + E_{\infty,6}$, where $E_{\infty,1}, \ldots, E_{\infty,6}$ are 6 disjoint chains joining $\tilde{C}_{\infty,1}$ to $\tilde{C}_{\infty,2}$. More explicitly, $E_{\infty,1}$ is a $(-4,-2,-2)$-chain, $E_{\infty,2}$ and $E_{\infty,3}$ are (-2)-curves, $E_{\infty,4}$ is a $(-2,-3)$-chain, $E_{\infty,5}$ a $(-3,-2)$-chain, and finally $E_{\infty,6}$ is a $(-2,-2,-4)$-chain; these correspond under the correspondence of Proposition 2.5c) to the points $z_{\infty,\varepsilon,1,0}, z_{\infty,\varepsilon,1,2}, z_{\infty,\varepsilon,1,4}, z_{\infty,\varepsilon,1,5}, z_{\infty,\varepsilon,2,5}$, and $z_{\infty,\varepsilon,3,0}$, respectively. (Here, $\varepsilon = 3$.) Thus, using the terminology of [10], Definition 4.16, we see that $D = D(\bar{P}_\infty, \bar{P}_\infty) = \tilde{C}_{\infty,1} + \tilde{C}_{\infty,2} + E_{\infty,2} + E_{\infty,3}$ is a (-2)-join of breadth 2 of the two (-3)-curves $\tilde{C}_{\infty,1}$ and $\tilde{C}_{\infty,2}$ with intersection number $(\tilde{C}_{\infty,1}.\tilde{C}_{\infty,2}) = \frac{1}{2}\phi(10) = 2$.

Claim 1: If $\tilde{Z}_\varepsilon$ is of general type and C is a (-1)-curve on $\tilde{Z}_\varepsilon$, then C meets D_∞ in each of its two (-4)-components transversely and in no other components, i.e. $(C.D_\infty) = 2$.

First note that since none of the fibre components of $\tilde{\psi}_i$ are (-1)-curves, C is not a component of a fibre and so $(C.\tilde{\psi}_i(\bar{P}_k)) > 0$; in particular, $(C.D_\infty) > 0$.

Since $\tilde{Z}_\varepsilon$ is of general type by hypothesis, we are in the situation of [10], Corollary 4.18 (with $D = D(\bar{P}_\infty, \bar{P}_\infty)$), from which we conclude that C can only meet the (-4)-components of D_∞. Moreover, C has to meet any such component C' transversely, for otherwise the image $\bar{C}'$ of C' on the blow-down surface $\bar{Z}$ would be a singular curve with $(K_{\bar{Z}}.\bar{C}') \leq (K_{\tilde{Z}_\varepsilon}.C') - 2 = 0$, which is impossible (cf. [1], VII.2.3).

We thus have $(C.D_\infty) = 1$ or 2, depending on whether C meets one or both of the (-4)-components. Suppose the former were possible, i.e. that C meets only $E_{\infty,1}$ (or $E_{\infty,6}$). Since the multiplicity of the (-4)-component of $E_{\infty,1}$ in $\tilde{\psi}_i^*(\bar{P}_\infty)$ is 3 respectively 1 for $i = 1$ respectively $i = 2$ (cf. 2.5), we have $(C.\tilde{\psi}_1^*(\bar{P}_\infty)) = 3$ and $(C.\tilde{\psi}_2^*(\bar{P}_\infty)) = 1$. Then also $(C.\tilde{\psi}_1^*(\bar{P}_1)) = 3$ and $(C.\tilde{\psi}_2^*(\bar{P}_1)) = 1$, which is impossible: on the one hand C has to meet a unique component $C_{s,j}$ of D_1 transversely, but then $(C.\tilde{\psi}_1^*(\bar{P}_1)) \neq 3$ (cf. 2.5). This, therefore, shows that the case $(C.D_\infty) = 1$ is impossible, and so we must must have the situation of the claim.

Claim 2: Let $f : \tilde{Z}_\varepsilon \to \bar{Z}$ denote the blow-down map with respect to a (-1)-curve C as in claim 1. Then $\bar{Z}$ is minimal.

If not, then there is a (-1)-curve $\bar{C}$ on $\bar{Z}$. Since $\bar{D} = f_*(D_\infty)$ consists entirely of (-2)- and $(-3,-2)$-chains (etc.) which join the (-3)-curves $f(\tilde{C}_{\infty,1})$ and $f(\tilde{C}_{\infty,2})$, we conclude from [10], Corollary 4.18 that $\bar{C}$ does not meet $\bar{D}$. Thus, if $\tilde{C}$ denotes the proper transform of $\bar{C}$ with respect to f, then $(\tilde{C}.(D_\infty + 2C)) = (\tilde{C}.f^*\bar{D}) = (\bar{C}.\bar{D}) = 0$, which means that $\tilde{C}$ meets neither C nor D_∞. But then $\tilde{C}$ is a (-1)-curve on $\tilde{Z}_\varepsilon$ which doesn't meet D_∞, which is impossible by claim 1. Thus, $\bar{C}$ cannot exist, which verifies claim 2.

From claims 1 and 2 we can readily see that $\kappa(\tilde{Z}_\varepsilon) \leq 1$. Indeed, since $K_\varepsilon^2 = -1$ and $p_g = 2 > 0$, $\tilde{Z}_\varepsilon$ is not minimal, so there exists a (-1)-curve C on $\tilde{Z}_\varepsilon$. Thus, if $\tilde{Z}_\varepsilon$ were of general type, then by claims 1 and 2 the blow-down $\bar{Z}$ is minimal. But $K_{\bar{Z}}^2 = K_{\tilde{Z}}^2 + 1 = 0$, so $\bar{Z}$ cannot be of general type, contradiction. Thus $\kappa(\tilde{Z}_\varepsilon) \leq 1$ as desired.

b) We now turn to the much more difficult case that $(N, \varepsilon) = (11, 1)$. Here we have (with the notation of 2.5) that E_0, resp. E_{11}, resp. E_{12} consists of 6 disjoint (-2)-curves, resp. of 2 disjoint (-3)-curves, resp. of 2 disjoint (-2)-chains of length 2. Moreover, $E_\infty = E_{\infty,1} + \ldots + E_{\infty,5}$, where $E_{\infty,i}$ is the $\mathbb{P}^1$-chain associated to the singularity $z_{\infty,1,i,0} \in S_{\infty,1}$; cf. Prop. 2.5. Thus, $E_{\infty,1}$ is a (-11)-chain, $E_{\infty,2}$ is a $(-4,-3)$-chain, $E_{\infty,3}$ is a $(-3,-2,-2,-2,-2)$-chain, $E_{\infty,4}$ is a $(-2,-2,-2,-2,-3)$-chain, and $E_{\infty,5}$ is a $(-3,-4)$-chain.

As in Remark 4.9 of [10], let $\tilde{\Gamma}$ denote the image of the diagonal of Y, and let Γ_k denote the unique component of E_k which meets $\tilde{\Gamma}$. From that remark we easily deduce:

Claim 3: $\tilde{\Gamma}$ is a (-1)-curve which meets the (-2)-curve Γ_0 and the (-3)-curve Γ_1 transversely; in particular, there is a blow-down map $f_1 : \tilde{Z} \to \bar{Z}_1$ which contracts each of Γ_0, Γ_1 and $\Gamma_2 := \tilde{\Gamma}$ to a point. Moreover, we have $\tilde{\tau}(\Gamma_k) = \Gamma_k$ for $k = 0, 1, 2$, where $\tilde{\tau}$ denotes the involution of $\tilde{Z}$ induced by the one on $Y = X \times X$ which interchanges the two factors.

Claim 4: There exist (-1)-curves Γ_3 and Γ_4 on $\tilde{Z}$ with the property that $\tilde{\tau}(\Gamma_k) = \Gamma_k$ and that Γ_k meets the $(-k)$-components of both $E_{\infty,2}$ and $E_{\infty,5}$ transversely for $k = 3, 4$.

To construct these Γ_k we shall use certain twists of the *Hecke correspondences* $T_n = T'(1, n)$ on Y (cf. Shimura[19], §3.3, §7.3). Explicitly, for $n, k \in \mathbb{N}$ with $(nk, N) = 1$, let $T_{n,k} := \varphi((\tau_k^{-1} \times id_X)T_n) \subset Z_\varepsilon$, where $\tau_k = \left(\begin{smallmatrix} k^{-1} & 0 \\ 0 & k \end{smallmatrix}\right) \in \mathrm{Sl}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$. Then $T_{n,k}$ is an irreducible curve on Z_ε which is birationally isomorphic to the modular curve $X_0(n)$ if and only if $k^2 n\varepsilon \equiv 1 \pmod{N}$. If this is the case, then it follows from the definitions that its proper transform $\tilde{T}_{n,k}$ on $\tilde{Z}_\varepsilon$ is one of the curves $F_n^{(i)}$ studied by Hermann[5], pp. 104-7. Moreover, it is easily verified that in that case we have

$$T_{n,k} \cap S_{\infty,\varepsilon} = \{z_{\infty,\varepsilon,\varepsilon kn/d,0} : d|n\}. \tag{20}$$

In addition, if $\varepsilon = 1$, then we also have $\tau(T_{n,k}) = T_{n,k}$ and hence $\tilde{\tau}(\tilde{T}_{n,k}) = \tilde{T}_{n,k}$.

Applying this to the case $(N, \varepsilon) = (11, 1)$, we shall now show that $\Gamma_3 = \tilde{T}_{4,5}$ and $\Gamma_4 = \tilde{T}_{3,2}$ satisfy the properties of the claim. Indeed, Γ_3 and Γ_4 are (-1)-curves by Hermann[5], Hilfssatz 16. Moreover, (20) shows that $T_{4,5}$ passes through $z_{\infty,1,2,0}$ and $z_{\infty,1,5,0}$, so Γ_3 meets $E_{\infty,2}$ and $E_{\infty,5}$ (and no other components of E_∞), and similarly, Γ_4 meets $E_{\infty,1} = \Gamma_\infty$, $E_{\infty,2}$ and $E_{\infty,5}$. Next we note that $(\Gamma_3.\tilde{\psi}_i^*(\bar{P}_\infty)) = 6$ and $(\Gamma_4.\tilde{\psi}_i^*(\bar{P}_\infty)) = 4$ because in general $(\tilde{T}_{n,k}.\tilde{\psi}_i^*(\bar{P})) = \psi(n)$ (by the projection formula). Thus, since the multiplicities of the components of $E_{\infty,2}$ in $\tilde{\psi}_1^*(\bar{P}_\infty)$ and in $\tilde{\psi}_2^*(\bar{P}_\infty))$ are $(3, 1)$ and $(1, 4)$, respectively, and those of $E_{\infty,5}$ are $(4, 1)$ and $(1, 3)$, respectively, one easily sees that Γ_k must meet the $(-k)$-components of $E_{\infty,2}$ and $E_{\infty,5}$ transversely, which verifies the claim.

Let us now put $\Gamma_5 = \tilde{T}_{5,3}$ and $\Gamma_6 = \tilde{T}_{9,4}$. By exactly the same method as that of the proof of claim 4 one shows:

Claim 5: There exist (-2)-curves Γ_5 and Γ_6 on $\tilde{Z}$ such that Γ_5 (resp. Γ_6) meets the (-3)-components (resp. the extremal (-2)-components) of both $E_{\infty,3}$ and $E_{\infty,4}$ transversely.

Claim 6: Let $f : \tilde{Z} \to \bar{Z}$ denote the blow-down map, where $\bar{Z}$ is the surface obtained by blowing down $\bar{Z}_1$ (cf. claim 3) via the images of Γ_3 and Γ_4 on $\bar{Z}_1$. Then $K_{\bar{Z}}^2 = 0$ and the involution $\tilde{\tau}$ on $\tilde{Z}$ induces an involution $\bar{\tau}$ on $\bar{Z}$. Moreover, $\bar{E}_0 := f_*(E_0)$ and $\bar{E}_1 := f_*(E_1)$ consist entirely of (-2)-curves, whereas $\bar{E}_\infty := f_*(E_\infty)$ consists of (-2) and (-3)-curves, together with the (-7)-curve $\bar{\Gamma}_\infty = f_*(\Gamma_\infty)$.

The first assertion is clear since $K^2_{\tilde{Z}} = -5$ (cf. Table 1) and five curves are blown down, as is the second since $\tilde{\tau}(\Gamma_k) = \Gamma_k$ for $0 \leq k \leq 4$. To prove the third, note first that Γ_0, Γ_1 and Γ_2 meet E_0 (resp. E_1) only in Γ_0 (resp. in Γ_1), which is blown down to a point by f. Moreover, by Hermann[5], Hilfssatz 11, Γ_3 and Γ_4 do not meet E_0 (because $\nu_2(3) = \nu_2(4) = 0$),[2] and so all the components of E_0 other than Γ_0 remain (-2)-curves. Similarly, Γ_3 does not meet E_1 ($\nu_3(4) = 0$) whereas Γ_4 meets a (-3)-component of E_1 (because $\nu_3(3) = 1$), which has to be different from Γ_1, for otherwise two (-1) curves would intersect on a suitable blow-down. Thus, this (-3)-component becomes a (-2)-component after blowing down Γ_4, and so the third assertion follows since all other components are (-2)-chains which do not meet Γ_k, $0 \leq k \leq 4$ (cf. [5], Hilfssatz 11). Finally, the last assertion is immediate by claim 4 and the fact that Γ_3 meets Γ_∞ but Γ_4 does not.

Claim 7: For the minimal model Z_{min} of $\tilde{Z}$ (and of $\bar{Z}$) we have $K^2_{Z_{min}} \leq 1$.

We may assume that Z_{min} is of general type for otherwise $K^2_{Z_{min}} = 0$. The two curves $\bar{C}_i = f(\tilde{C}_{\infty,i})$, $i = 1, 2$, on $\bar{Z}$ are joined by the (-2)-chain $\bar{E}_1$ arising from the images of the (-3)-components of $E_{\infty,2}$ and $E_{\infty,5}$, and also by the (-2)-chain $\bar{E}_2$ obtained from the images of the extremal (-2)-curves of $E_{\infty,3}$ and $E_{\infty,4}$ together with Γ_6 (cf. claim 5). Since no Γ_k, $0 \leq k \leq 4$, meets $\tilde{C}_{\infty,1} + \tilde{C}_{\infty,2}$, we see that $\bar{D} := \bar{C}_1 + \bar{C}_2 + \bar{E}_1 + \bar{E}_2$ is a (-2)-join of breadth 2 with invariants $k(\bar{C}_1, \bar{C}_2) := (K_{Z_{min}}.(\bar{C}_1 + \bar{C}_2)) = 4$ and $i(\bar{C}_1, \bar{C}_2) := (\bar{C}_1 + \bar{C}_2)^2 = 6$, and so we have $K^2_{Z_{min}} \leq 4^2/(2 \cdot 2 + 6) < 2$ by Proposition 4.17 of [10].

Claim 8: $\bar{Z}$ is minimal and hence $\kappa(\tilde{Z}) \leq 1$.

Suppose not; then there is a (-1)-curve $\bar{C}$ on $\bar{Z}$, and so by claims 7 and 6, the blown-down $\bar{Z}'$ of $\bar{Z}$ via $\bar{C}$ is a minimal surface of general type. Thus $\bar{C}$ is unique (so $\bar{\tau}(\bar{C}) = \bar{C}$) and $\bar{C}$ cannot meet any (-2)-curve on $\bar{Z}$. From claim 6 we therefore see that $\bar{C}$ cannot meet $\bar{E}_0$ and $\bar{E}_1$, and hence the total transform C of $\bar{C}$ does not meet $E_0 + E_1$. Thus, $2(C.\tilde{C}_{0,1}) = (C.\tilde{\psi}_1^*(\bar{P}_0)) = (C.\tilde{\psi}_1^*(\bar{P}_1)) = 3(C.\tilde{C}_{1,1})$, and so it follows that $(C.\tilde{\psi}_1^*(\bar{P}_k)) \equiv 0 \pmod 6$, and hence $(C.\tilde{\psi}_1^*(\bar{P}_k)) \geq 6$ because $\bar{C}$ cannot be a component of $\bar{E}_k$ by claim 6.

Next we note that $(C.\tilde{C}_{\infty,i}) = 0$, for $i = 1, 2$. If not, then $(C.\tilde{C}_{\infty,i}) \geq 1$ for $i = 1, 2$ (because $\tilde{\tau}$ fixes C and interchanges $\tilde{C}_{\infty,1}$ and $\tilde{C}_{\infty,2}$), and so the image $\bar{D}' = \bar{C}'_1 + \bar{C}'_2 + \bar{E}'_1 + \bar{E}'_2$ on $\bar{Z}'$ of the (-2)-join $\bar{D}$ constructed in the proof of claim 7 would have invariants $k(\bar{C}'_1, \bar{C}'_2) \leq 2$ and $i(\bar{C}'_1, \bar{C}'_2) > 6$, which violates the inequality of Proposition 4.17 of [10].

In addition, $\bar{C}$ cannot meet any (-3)-components of $\bar{E}_\infty = \bar{E}_{\infty,1} + \ldots + \bar{E}_{\infty,5}$, where $\bar{E}_{\infty,j} := f_*(E_{\infty,j})$. Indeed, if $\bar{C}$ meets that of $\bar{E}_{\infty,2}$ or of $\bar{E}_{\infty,5}$, then it also meets both because $\bar{\tau}$ interchanges these components, and then the images of $\bar{E}_{\infty,2}$ and $\bar{E}_{\infty,5}$ on $\bar{Z}'$ would form an elliptic (-2)-configuration, which is impossible (cf. [3], VII.2.7). Similarly, if $\bar{C}$ meets one of the (-3)-components of $\bar{E}_{\infty,3}$ and $\bar{E}_{\infty,4}$, then it meets both, and the images of $\bar{E}_{\infty,3}, \bar{E}_{\infty,4}$ and Γ_6 on $\bar{Z}'$ would form an elliptic (-2)-configuration, contradiction.

From claim 6 we thus see that $\bar{C}$ can meet $\tilde{\psi}_1^*(\bar{P}_\infty)$ only in $\Gamma_\infty = E_{\infty,1}$, which has multiplicity 1 in $\tilde{\psi}_1^*(\bar{P}_\infty)$. Thus $(\bar{C}.\bar{E}_{\infty,1}) \geq (C.E_{\infty,1}) = (C.\tilde{\psi}_1^*(\bar{P}_\infty)) \geq 6$, and so the image $\bar{E}'$ of $\bar{E}_{\infty,1}$ on $\bar{Z}'$ is a (singular) curve with $(\bar{E}'.K_{\bar{Z}'}) = (\bar{E}_{\infty,1}.K_{\bar{Z}}) - (\bar{C}.\bar{E}_{\infty,1}) \leq 5 - 6 = -1$, which is impossible since $\bar{Z}'$ is minimal. Thus, no such curve $\bar{C}$ exists and so $\bar{Z}$ is minimal.

[2] Note that the formulae for $\nu_2(n)$ and $\nu_3(n)$ on p. 104 of [5] are incorrect when $4|n$ and $9|n$, and have to be modified as in Miyake [16], p. 108

N	ε	m	g	g_0	g_1	g_∞	$\tilde{C}_0^2$	$\tilde{C}_1^2$	$\tilde{C}_\infty^2$	d	$2\mathbb{G}$	$\mathbb{L}$	$2\mathbb{S}$	p_g	c_2	K^2	κ
5	1	60	0	0	0	0	−1	−1	−1	−1	0	10	0	0	14	−2	−1
5	2	60	0	0	0	0	−1	−1	−1	−1	0	9	0	0	13	−1	−1
6	1	72	1	0	0	0	−2	−1	−1	−1	1	10	1	0	16	−4	−1
6	5	72	1	0	0	0	−2	−2	−2	0	1	18	−1	1	24	0	0
7	1	168	3	1	1	0	−2	−1	−1	−1	1	12	1	0	18	−6	−1
7	3	168	3	1	1	0	−2	−1	−2	0	1	19	−1	1	25	−1	0
8	1	192	5	2	1	0	−2	−2	−1	−1	2	14	2	0	22	−10	−1
8	3	192	5	2	1	0	−2	−2	−2	0	2	18	0	1	26	−2	0
8	5	192	5	2	1	0	−2	−2	−2	0	2	18	0	1	26	−2	0
8	7	192	5	2	1	0	−2	−2	−3	1	2	28	−2	2	36	0	1
9	1	324	10	4	3	0	−3	−1	−2	0	3	20	1	1	30	−6	0
9	2	324	10	4	3	0	−3	−2	−3	1	3	28	−1	2	38	−2	1
10	1	360	13	6	3	0	−2	−3	−3	1	4	30	0	2	42	−6	1
10	3	360	13	6	3	0	−2	−3	−3	1	4	25	0	2	37	−1	1
11	1	660	26	12	8	1	−3	−2	−2	0	5	27	1	2	41	−5	1
11	2	660	26	12	8	1	−3	−2	−3	1	5	34	−1	3	48	0	2
12	1	576	25	11	7	0	−4	−2	−2	0	7	22	5	1	40	−16	0
12	5	576	25	11	7	0	−4	−4	−4	4	7	38	−1	4	56	4	2
12	7	576	25	11	7	0	−4	−2	−4	2	7	30	1	3	48	0	2
12	11	576	25	11	7	0	−4	−4	−6	6	7	58	−5	6	76	8	2
13	1	1092	50	24	16	2	−3	−2	−3	1	8	39	0	4	59	1	2
13	2	1092	50	24	16	2	−3	−2	−3	1	8	34	0	4	54	6	2
14	1	1008	49	23	15	1	−4	−3	−4	2	10	36	2	4	60	0	2
14	3	1008	49	23	15	1	−4	−3	−5	5	10	49	−2	6	73	11	2
15	1	1440	73	35	23	1	−4	−2	−4	2	14	34	6	4	66	−6	2
15	2	1440	73	35	23	1	−4	−4	−6	4	14	44	2	6	76	8	2
15	7	1440	73	35	23	1	−4	−2	−6	8	14	54	−2	8	86	22	2
15	11	1440	73	35	23	1	−4	−4	−8	10	14	76	−6	10	108	24	2
16	1	1536	81	39	25	2	−4	−4	−3	1	15	34	7	4	68	−8	2
16	3	1536	81	39	25	2	−4	−4	−5	5	15	44	1	7	78	18	2
16	5	1536	81	39	25	2	−4	−4	−5	7	15	50	−1	8	84	24	2
16	7	1536	81	39	25	2	−4	−4	−7	11	15	80	−7	11	114	30	2
17	1	2448	133	65	43	5	−4	−3	−4	6	20	62	0	10	106	26	2
17	3	2448	133	65	43	5	−4	−3	−4	6	20	47	0	10	91	41	2
18	1	1944	109	52	34	2	−6	−3	−5	7	21	46	5	8	92	16	2
18	5	1944	109	52	34	2	−6	−6	−8	14	21	82	−5	13	128	40	2
19	1	3420	196	96	64	7	−5	−3	−4	10	29	60	1	14	122	58	2
19	2	3420	196	96	64	7	−5	−3	−5	11	29	67	−1	15	129	63	2

Table 1: The invariants of $\tilde{Z}_{N,\varepsilon}$ for $5 \leq N \leq 19$

N	ε	m	g	g_0	g_1	g_∞	$\tilde{C}_0^2$	$\tilde{C}_1^2$	$\tilde{C}_\infty^2$	d	$2\mathbb{G}$	$\mathbb{L}$	$2\mathbb{S}$	p_g	c_2	K^2	κ
20	1	2880	169	83	53	3	−4	−6	−6	12	30	64	6	12	128	28	2
20	3	2880	169	83	53	3	−4	−6	−8	14	30	62	2	14	126	54	2
20	11	2880	169	83	53	3	−4	−6	−10	20	30	104	−6	18	168	60	2
20	13	2880	169	83	53	3	−4	−6	−8	18	30	74	−2	16	138	66	2
21	1	4032	241	117	79	5	−8	−2	−6	14	40	56	10	15	140	52	2
21	2	4032	241	117	79	5	−8	−4	−8	16	40	64	6	17	148	68	2
21	5	4032	241	117	79	5	−8	−4	−12	28	40	124	−10	25	208	104	2
21	10	4032	241	117	79	5	−8	−2	−10	26	40	100	−6	23	184	104	2
22	1	3960	241	118	77	6	−6	−6	−6	16	40	63	6	17	147	69	2
22	7	3960	241	118	77	6	−6	−6	−9	25	40	102	−6	23	186	102	2
23	1	6072	375	185	123	12	−6	−4	−4	18	55	55	9	23	169	119	2
23	5	6072	375	185	123	12	−6	−4	−7	33	55	112	−9	32	226	170	2
24	1	4608	289	141	93	5	−8	−4	−4	12	50	48	24	13	152	16	2
24	5	4608	289	141	93	5	−8	−8	−10	30	50	88	0	25	192	120	2
24	7	4608	289	141	93	5	−8	−4	−10	30	50	88	0	25	192	120	2
24	11	4608	289	141	93	5	−8	−8	−12	36	50	120	−8	29	224	136	2
24	13	4608	289	141	93	5	−8	−4	−8	24	50	64	8	21	168	96	2
24	17	4608	289	141	93	5	−8	−8	−10	30	50	92	0	25	196	116	2
24	19	4608	289	141	93	5	−8	−4	−10	30	50	84	0	25	188	124	2
24	23	4608	289	141	93	5	−8	−8	−16	48	50	208	−24	37	312	144	2
25	1	7500	476	236	156	12	−5	−5	−9	39	72	106	0	36	254	190	2
25	2	7500	476	236	156	12	−5	−5	−9	39	72	99	0	36	247	197	2
26	1	6552	421	208	137	10	−6	−6	−9	37	66	111	0	33	247	161	2
26	5	6552	421	208	137	10	−6	−6	−9	37	66	84	0	33	220	188	2
27	1	8748	568	280	187	13	−9	−3	−9	49	88	102	4	42	282	234	2
27	2	8748	568	280	187	13	−9	−6	−12	54	88	132	−4	46	312	252	2
28	1	8064	529	261	173	10	−8	−6	−9	47	85	102	9	38	276	192	2
28	3	8064	529	261	173	10	−8	−6	−15	59	85	162	−9	47	336	240	2
28	5	8064	529	261	173	10	−8	−6	−11	55	85	110	−1	43	284	244	2
28	11	8064	529	261	173	10	−8	−6	−13	51	85	106	1	42	280	236	2
29	1	12180	806	400	266	22	−7	−5	−7	67	118	118	0	59	358	362	2
29	2	12180	806	400	266	22	−7	−5	−7	67	118	109	0	59	349	371	2
30	1	8640	577	285	187	9	−8	−6	−10	50	96	94	18	39	290	190	2
30	7	8640	577	285	187	9	−8	−6	−12	64	96	110	2	47	306	270	2
30	11	8640	577	285	187	9	−8	−12	−18	78	96	216	−18	57	412	284	2
30	17	8640	577	285	187	9	−8	−12	−16	64	96	132	−2	49	328	272	2

Table 2: The invariants of $\tilde{Z}_{N,\varepsilon}$ for $20 \leq N \leq 30$

References

[1] *W. Barth, C. Peters, A. Van de Ven:* Compact Complex Surfaces. Springer-Verlag, Berlin, 1984.

[2] *G. Frey:* On ternary equations of Fermat type and relations with elliptic curves. To appear.

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

[4] *E. Hecke:* Über ein Fundamentalproblem aus der Theorie der elliptischen Modulfunktionen. *Abh. Math. Sem. Hamburg* **6** (1928), 235-257 = Math. Werke, pp. 525-547.

[5] *C.F. Hermann:* Modulflächen quadratischer Diskriminante. *Manuscr. math.* **72** (1991), 95-110.

[6] *Hua Loo Keng:* Introduction to Number Theory. Springer-Verlag, Berlin, 1982.

[7] *E. Kani:* Curves of genus 2 with elliptic differentials and the height conjecture for elliptic curves. Proc. Conf. Number Theory and Arithmetic Geometry (G. Frey, ed.) Univ. Essen, Preprint No. 18 (1991), 30-39.

[8] *E. Kani:* Curves with elliptic differentials. Preprint.

[9] *E. Kani:* Elliptic curves on abelian surfaces. *Manuscr. math.* **84** (1994), 199-223.

[10] *E. Kani, W. Schanz:* Diagonal quotient surfaces. *Manuscr. math.* (to appear)

[11] *O.-H. Keller:* Darstellungen von Restklassen (mod n) als Summe von zwei Quadraten. *Acta Sci. Math. (Szeged)* **25** (1964), 191-192.

[12] *A. Kraus, O. Oesterlé:* Sur une question de B. Mazur. *Math. Ann.* **293** (1992), 259-275.

[13] *E. Landau:* Vorlesungen über Zahlentheorie. Chelsea Publ. Co., New York, 1950.

[14] *S. Lang:* Number Theory III. Ency. Math. Sci. vol. 60, Springer-Verlag, Berlin, 1991.

[15] *B. Mazur:* Rational isogenies of prime degree. *Invent. math.* **44** (1978), 129-162.

[16] *T. Miyake:* Modular Forms, Springer-Verlag, Berlin, 1989.

[17] *H.W. Praetorius:* Die Charaktere der Modulgruppe der Stufe q^2. *Abh. Math. Sem. Univ. Hamburg* **9** (1933), 365-394.

[18] *B. Schoeneberg:* Elliptic Modular Functions. Springer-Verlag, Berlin, 1974.

[19] *G. Shimura:* Introduction to the Arithmetic Theory of Automorphic Functions. Iwanami Shoten & Princeton University Press, 1971.

E. Kani
Department of Mathematics
and Statistics
Queen's University
Kingston, Ontario, Canada
K7L 3N6

W. Schanz
Schuppstr. 2
65191 Wiesbaden
Germany