KLEIN FORMS AND THE GENERALIZED SUPERELLIPTIC EQUATION

MICHAEL A. BENNETT AND SANDER R. DAHMEN

ABSTRACT. If $F(x, y) \in \mathbb{Z}[x, y]$ is an irreducible binary form of degree $k \geq 3$ then a theorem of Darmon and Granville implies that the generalized superelliptic equation

$$F(x, y) = z^l$$

has, given an integer $l \geq \max\{2, 7 - k\}$, at most finitely many solutions in coprime integers x, y and z. In this paper, for large classes of forms of degree $k = 3, 4, 6$ and 12 (including, heuristically, "most" cubic forms), we extend this to prove a like result, where the parameter l is now taken to be variable. In the case of irreducible cubic forms, this provides the first examples where such a conclusion has been proven. The method of proof combines classical invariant theory, modular Galois representations, and properties of elliptic curves with isomorphic mod n Galois representations.

CONTENTS

Date: June 2010.

1991 *Mathematics Subject Classification.* Primary 11D41, 11D59, Secondary 11G05, 11G30, 11J68.

Key words and phrases. Superelliptic equations, Galois representations, Thue-Mahler equations, elliptic curves with isomorphic n-torsion.

1. Introduction

A classic result of Siegel [48], from 1929, is that the set of K-integral points on a smooth algebraic curve of positive genus, defined over a number field K, is finite. As an application of this to Diophantine equations, Leveque [29] showed that if $f(x) \in \mathbb{Z}[x]$ is a polynomial of degree $k \geq 2$ with, say, no repeated roots, and $l \geq \max\{2, 5-k\}$ is an integer, then the *superelliptic* equation

$$f(x) = y^l \tag{1}$$

has at most finitely many solutions in integers x and y. Already in a 1925 letter from Siegel to Mordell (partly published in 1926 under the pseudonym X [47]), Siegel had proved precisely this result in case $l = 2$ (and had remarked that his argument readily extends to all exponents $l \geq 2$). Via lower bounds for linear forms in logarithms, Schinzel and Tijdeman [42] deduced that, in fact, equation (1) has at most finitely many solutions in integers x, y and *variable* $l \geq \max\{2, 5-k\}$ (where we count the solutions with $y^l = \pm 1, 0$ only once). This latter result has the additional advantage over Leveque's theorem in that it is effective (the finite set of values for x is effectively computable).

If we replace the polynomial $f(x)$ with a binary form $F(x, y)$ over $\mathbb{Z}$ (i.e. a homogeneous polynomial in $\mathbb{Z}[x, y]$) of degree $k \geq 3$, with no repeated roots, then work of Darmon and Granville (Theorem 1 of [8]) ensures that the corresponding *generalized superelliptic equation*

$$F(x, y) = z^l, \quad \gcd(x, y) = 1 \tag{2}$$

has, for fixed $l \geq \max\{2, 7-k\}$, at most finitely many solutions in integers x and y. That the restriction of coprimality is a necessary one is readily observed; in case $(k, l) = (3, 3)$ or $(4, 2)$, equation (2) corresponds, generically, to a curve of genus 1. In the proof of the result of Darmon and Granville, Faltings' theorem [18] plays an analogous role to that of Siegel's theorem for equation (1).

In what follows, our goal is to investigate the degree to which it is possible to generalize the theorem of Darmon and Granville to prove a result akin to that of Schinzel and Tijdeman, valid for binary forms rather than polynomials. That is, we wish to address the following

Question. *For a given binary form $F(x,y) \in \mathbb{Z}[x,y]$ of degree k, is the number of integer 4-tuples (x,y,z,l) satisfying equation (2) and $l \geq \max\{2, 7-k\}$ finite?*

For many quadratic forms F (and for certain forms with repeated factors in $\mathbb{Z}[x,y]$; see Theorem 1 of [8]), the answer to this question is clearly "no". In general, for forms of higher degree, the problem rapidly becomes more complicated – to illustrate its difficulty, it is worthwhile noting that the affirmative answer to this question, in case

$$F(x,y) = xy(x+y) \text{ or } x(x^2-y^2),$$

and $z \neq 0, \pm 1$, is equivalent to the asymptotic version of Wiles' theorem [58], née Fermat's Last Theorem, and to a special case of a result of Darmon and Merel [9], respectively.

For irreducible forms (over $\mathbb{Q}[x,y]$), the only results in the literature answering our question affirmatively are for certain diagonal quartic forms, such as $F(x,y) = x^4 + y^4$ (together with "covering" forms of higher degree), due to Ellenberg [15] and to Dieulefait and Jiménez Urroz [12]) (see also [2]). In particular, there are hitherto no primitive, irreducible cubic forms for which the above question has been addressed. A special case of the main result of this paper (Theorem 8.4 of Section 8) is the following:

Theorem 1.1. *Suppose that $F(x,y) \in \mathbb{Z}[x,y]$ is an irreducible binary cubic form and that S_F is the set of primes dividing $2\,\Delta_F$, together with a prime at ∞, where Δ_F is the discriminant of F. If the Thue-Mahler equation*

$$F(x,y) \in \mathbb{Z}^*_{S_F} \tag{3}$$

has no solutions in integers x and y, then there exists an effectively computable constant l_0 (depending on F), such that there are no solutions to equation (2) in integers x, y and z, and prime $l > l_0$. Consequently, equation (2) has at most finitely many solutions in integers x, y, z and integer $l \geq 4$.

Here, we denote by $\mathbb{Z}^*_{S_F}$, the set of S_F-units, that is, the nonzero integers u with the property that if a prime p satisfies $p \mid u$, then $p \in S_F$.

At first blush, the conditions imposed upon F by the insolubility of (3) appear to be extremely restrictive. In particular, since $1 \in \mathbb{Z}^*_{S_F}$, the above theorem fails to apply to any $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of monic forms. Indeed, of the 190 classes of primitive, irreducible cubic forms with $|\Delta_F| \leq 1000$, equation (3) has integer solutions in every case! Despite this, in Section 12, we will sketch a heuristic which indicates that this is the law of small numbers at play and, in fact, "almost all" cubic forms F have the property that (3) has no solutions in integers. In Section 9, we will demonstrate this for an infinite family of cubic forms. A representative form of smallest absolute discriminant to which we may apply Theorem 1.1 is

$$F(x,y) = 3x^2 + 2x^2y + 5xy^2 + 3y^3,$$

with discriminant $\Delta_F = -2063$; there are 24 such classes of forms with (3) insoluble and $|\Delta_F| \leq 10^4$. It is worth remarking, at this juncture, that there are effective, indeed efficient, algorithms for determining all solutions to equation (3) (see e.g. Tzanakis and de Weger [53]).

We derive analogues to Theorem 1.1 for forms of degrees $4, 6$ and 12 which are somewhat less general, in that they apply only to Klein forms of these degrees. As we shall see, the Klein forms are a density zero subset of the set of all binary

forms of degrees $4, 6$ and 12. For these forms, the finiteness of solutions to (2) is a consequence of the *abc*-conjecture (over $\mathbb{Q}$) of Masser and Oesterlé. For more general irreducible forms, this does not seem to be the case and one must apparently combine a version of the *abc*-conjecture, valid in number fields (as in Elkies [16]), with some careful descent arguments.

The results of this paper may be viewed as an attempt to develop methods for solving Diophantine equations, arising from the modularity of Galois representations, which exhibit greater flexibility than those currently in the literature. In particular, we wish to illustrate that such techniques are not simply limited to ternary equations akin to the generalized Fermat equation. Our method of proof of Theorem 1.1 and related results proceeds via construction of Frey or, if you like, Frey-Hellegouarch curves over $\mathbb{Q}$, corresponding to putative solutions to equation (2) for *Klein forms* $F(x, y)$ of *index* n. Here, $2 \leq n \leq 5$ is an integer; the case $n = 2$ is that of nondegenerate cubic forms. Typically, in such matters, one is led to study arithmetic properties of weight 2 cuspidal newforms of fixed level N. Previously, the combination of the presence of one-dimensional forms at level N (equivalently, elliptic curves $E/\mathbb{Q}$ of conductor N), with the absence of rational isogenies for the given Frey-Hellegouarch curve, has provided a substantial barrier to progress. What is novel in the present paper is that we are able to overcome these difficulties by exploiting the fact that the Frey-Hellegouarch curves corresponding to a fixed Klein form have constant n-torsion (equivalently, isomorphic mod-n Galois representations). In a certain sense, this property is stronger than the existence of a rational n-isogeny, and plays an analogous role in our proofs to that of the latter in, for instance, [9].

The outline of this paper is as follows. In Section 2, we discuss the necessary background information from classical invariant theory and define the notion of a Klein forms. Section 3 details the connection between Klein forms and elliptic curves. In Section 4, we describe how to attach families of Frey-Hellegouarch curves to Klein forms. Section 5 is a brief summary of (well-known) results on Galois representations arising from elliptic curves.

In Section 6, we introduce the fundamental new observation, key to the proof of Theorem 1.1, that the families of Frey-Hellegouarch curves defined in Section 4 possess, for a fixed Klein form F of index n, constant n-torsion. Together with technical hypotheses, this enables us to eliminate the possibility of the corresponding mod l Galois representations arising from modular forms of dimension 1 at certain levels N, for suitably large prime l. Section 7 contains the information we require from an effective version of the Chebotarev density theorem to quantify this last statement. In Section 8, we appeal to the so-called *modular method*, based upon modularity of the canonical mod l Galois representations associated to our Frey-Hellegouarch curves, together with level-lowering, to state our main results.

The remainder of the paper is devoted to discussing the applicability of the results of Section 8. In particular, in Sections 9, 10 and 11, we appeal to our main theorem to answer our question on the finiteness of solutions to (2) affirmatively for families of binary cubic, quartic, sextic and duodecic (degree 12) forms. As a sample of the results from these sections, we have

Theorem 1.2. *If a is an integer such that $a^2 + 9a + 81$ is squarefree and neither a nor $-a - 9$ is in the set $\{-4, 8, 22, 31\}$, then the Diophantine equation*

$$3x^3 - ax^2y - (a + 9)xy^2 - 3y^3 = z^n$$

has at most finitely many solutions in nonzero integers x, y, z and n, with $\gcd(x, y) = 1$ *and* $n \geq 4$.

This provides an infinite family of forms F satisfying the hypotheses of Theorem 1.1. The proof of Theorem 1.2 is by no means straightforward. Indeed, it requires the solution of an infinite family of Thue-Mahler equations; as far as we are aware, this is the first "nontrivial" such family to be solved where the corresponding set of nonarchimedean primes is unbounded (in both cardinality and height). Additionally, the techniques of Section 9 are quite distinct from those of the remainder of the paper and may be of independent interest.

In Section 12, we discuss heuristics which indicate that Theorem 1.1 is applicable to a density one subset of the set of all cubic forms. The remaining sections are of a less speculative nature, being concerned with the specialization of our arguments to diagonal forms (Section 13), and with explicit computations and examples (Section 14), chosen to illustrate the utility of our methods.

2. Klein forms

2.1. **Some invariant theory.** We will begin by describing a number of results from classical invariant theory; these are well covered in Klein [25] or Hilbert [23]. More recent references for what we require along these lines are the book of Olver [37] and, from a more arithmetic perspective, the thesis of Edwards [13], corresponding article [14], and the survey of Beukers [3].

Our starting point is a topic central to invariant theory in the 19th century, namely the theory of invariants and covariants for binary forms. Let $F \in \overline{\mathbb{Q}}[x, y]$ be such a form, say

$$F(x, y) = \sum_{i=0}^{k} \alpha_i \, x^{k-i} y^i,$$

and $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\overline{\mathbb{Q}})$. Then by taking

$$(F \circ M)(x, y) = F(ax + by, cx + dy),$$

we obtain a right action of $\mathrm{GL}_2(\overline{\mathbb{Q}})$ on the set of nondegenerate binary forms over $\overline{\mathbb{Q}}$ of given degree k. Suppose that C is a form in x and y with coefficients that are homogeneous polynomials over $\overline{\mathbb{Q}}$ in the coefficients of F, $(\alpha_0, \alpha_1, \ldots, \alpha_k)$; we write $C = C(F)$ to emphasize its dependence on the form F. Call $C(F)$ a *covariant* of F if there exists an integer $p \geq 0$ such that

$$C(F \circ M) = \det(M)^p C(F) \circ M$$

for all $M \in \mathrm{GL}_2(\overline{\mathbb{Q}})$. We define p to be the *weight* of the covariant C. If a covariant I depends only on the coefficients $(\alpha_0, \alpha_1, \ldots, \alpha_k)$ and not upon x and y, so that

$$I(F \circ M) = \det(M)^p I(F),$$

we call I an *invariant of weight* p for the form F.

A theorem of Gordan [20] from 1868 asserts, given a fixed degree k, the existence of a finite set $C_1, C_2, \ldots, C_t$ of covariants such that every covariant of a binary form F of degree k is a polynomial over $\overline{\mathbb{Q}}$ in the C_j's (such a set is nowadays termed

a *Hilbert basis*). For binary forms of degrees $3, 4, 6$ and 12, there are in fact the following numbers of independent invariants and covariants:

degree	3	4	6	12
# invariants	1	2	5	109
# covariants	4	5	26	949

In the simplest case (at least for our purposes), that of cubic forms, the invariants are, up to scaling, just powers of the discriminant of the binary form:

$$\Delta_F = 18\,\alpha_0\alpha_1\alpha_2\alpha_3 + \alpha_1^2\alpha_2^2 - 27\,\alpha_0^2\alpha_3^2 - 4\,\alpha_0\alpha_2^3 - 4\,\alpha_1^3\alpha_3.$$

The independent covariants may be taken as Δ_F (which has weight 6), the form F itself (of weight 0), the *Hessian* of F, i.e. the quadratic form

$$H(x,y) = \frac{1}{4}\begin{vmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{vmatrix} = (3\alpha_0\alpha_2 - \alpha_1^2)x^2 + (9\alpha_0\alpha_3 - \alpha_1\alpha_2)xy + (3\alpha_1\alpha_3 - \alpha_2^2)y^2$$

(of weight 2) and the *Jacobian determinant* of F and H, in this case the cubic form

$$G(x,y) = \begin{vmatrix} F_x & F_y \\ H_x & H_y \end{vmatrix}$$

(of weight 3). Here, as is customary, F_x, F_y, etc, refer to corresponding partial derivatives. Connecting all four covariants, one has the syzygy

$$4\,H(x,y)^3 + G(x,y)^2 = -27\,\Delta_F\,F(x,y)^2. \tag{4}$$

It is this last identity (and its analogues) that we will exploit for Diophantine purposes.

For forms F of degree $k > 3$, we must narrow our focus considerably as, in general, we are unable to deduce a ternary relationship akin to (4). Let us define binary forms $F_n(x,y)$ for $n = 2, 3, 4$ and 5 as follows:

$$\begin{array}{rcl} F_2(x,y) & = & xy(x+y), \\ F_3(x,y) & = & y(x^3+y^3), \\ F_4(x,y) & = & xy(x^4+y^4), \\ F_5(x,y) & = & xy(x^{10} - 11x^5y^5 - y^{10}). \end{array} \tag{5}$$

These forms arise in Klein's [25] classification of the finite subgroups of $\mathrm{Aut}_{\overline{\mathbb{Q}}}(\mathbb{P}_1)$, where F_2, F_3, F_4 and F_5 are homogenizations of polynomials whose roots correspond to the vertices of an equilateral triangle, tetrahedron, octahedron and icosahedron, respectively, after projection to the complex plane. We call F a *Klein form* if $F = F_n \circ M$ for some $n \in \{2,3,4,5\}$ and $M \in \mathrm{GL}_2(\overline{\mathbb{Q}})$; the integer n is uniquely determined by the Klein form and is termed its *index*.

As it transpires, via a classic theorem of Gordan [21], if F is a binary form of degree $k \geq 3$ with nonzero discriminant, then F is a Klein form precisely when the *fourth transvectant* $\tau_4(F)$ vanishes. Transvectants are a class of covariants which are typically defined via either direct appeal to the representation theory of SL_2 and the Clebsch-Gordan formulae, or, through the use of differential operators (see e.g. Olver [37]). We will not provide an explicit definition for $\tau_4(F)$ in general (the interested reader is directed to Appendix A of [13]), but instead note that $\tau_4(F) \equiv 0$ is equivalent to the simultaneous vanishing of certain quadratic forms in

the coefficients of F. If F is cubic, then $\tau_4(F)$ is, in fact, identically zero. Otherwise, writing

$$F(x,y) = \sum_{i=0}^{k} \alpha_i x^{k-i} y^i, \tag{6}$$

we have $\tau_4(F) \equiv 0$ for a form of degree 4, precisely when

$$12\,\alpha_0\alpha_4 - 3\,\alpha_1\alpha_3 + \alpha_2^2 = 0. \tag{7}$$

For forms of degree 6, if $\alpha_0 \neq 0$, then $\tau_4(F) \equiv 0$ is equivalent to

$$\begin{aligned} 10\,\alpha_0\alpha_4 - 5\,\alpha_1\alpha_3 + 2\,\alpha_2^2 &= 0, \\ 25\,\alpha_0\alpha_5 - 5\,\alpha_1\alpha_4 + \alpha_2\alpha_3 &= 0, \\ 50\,\alpha_0\alpha_6 - 2\,\alpha_2\alpha_4 + \alpha_3^2 &= 0. \end{aligned} \tag{8}$$

Finally, for forms of degree 12, again if $\alpha_0 \neq 0$, then $\tau_4(F)$ being identically zero is equivalent to the simultaneous vanishing of the following 9 quadratic forms:

$$\begin{aligned} 44\,\alpha_0\alpha_4 - 33\,\alpha_1\alpha_3 + 15\,\alpha_2^2, \\ 55\,\alpha_0\alpha_5 - 22\,\alpha_1\alpha_4 + 6\,\alpha_2\alpha_3, \\ 330\,\alpha_0\alpha_6 - 55\,\alpha_1\alpha_5 - 20\,\alpha_2\alpha_4 + 18\,\alpha_3^2, \\ 55\,\alpha_0\alpha_7 - 5\,\alpha_2\alpha_5 + 2\,\alpha_3\alpha_4, \\ 440\,\alpha_0\alpha_8 + 55\,\alpha_1\alpha_7 - 30\,\alpha_2\alpha_6 - 5\,\alpha_3\alpha_5 + 8\,\alpha_4^2, \\ 198\,\alpha_0\alpha_9 + 44\,\alpha_1\alpha_8 - 5\,\alpha_2\alpha_7 - 6\,\alpha_3\alpha_6 + 3\,\alpha_4\alpha_5, \\ 660\,\alpha_0\alpha_{10} + 198\,\alpha_1\alpha_9 + 16\,\alpha_2\alpha_8 - 19\,\alpha_3\alpha_7 - 2\,\alpha_4\alpha_6 + 5\,\alpha_5^2, \\ 3630\,\alpha_0\alpha_{11} + 1320\,\alpha_1\alpha_{10} + 270\,\alpha_2\alpha_9 - 52\,\alpha_3\alpha_8 - 45\,\alpha_4\alpha_7 + 25\,\alpha_5\alpha_6, \end{aligned} \tag{9}$$

and

$$21780\alpha_0\alpha_{12} + 9075\alpha_1\alpha_{11} + 2670\alpha_2\alpha_{10} + 171\alpha_3\alpha_9 - 284\alpha_4\alpha_8 - 25\alpha_5\alpha_7 + 75\alpha_6^2.$$

For forms of degree 6 and 12, there are additional conditions in case $\alpha_0 = 0$ – we have little need for them here. We direct the reader to Appendix A of [13] for details.

It follows from (7), (8) and (9) that a Klein form with $\alpha_0 \neq 0$ is entirely determined by the values of its first four coefficients $\alpha_0, \alpha_1, \alpha_2$ and α_3. We will refer to this fact frequently in the sequel and, here and henceforth, adopt the shorthand

$$(\alpha_0, \alpha_1, \alpha_2, \alpha_3)_k = \sum_{i=0}^{k} \alpha_i x^{k-i} y^i,$$

for a Klein form of degree k. One may note, if $\alpha_0 \neq 0$, that a quadruple of rationals $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)$ corresponds, via scalar multiplication, to precisely one Klein form $F(x,y) \in \mathbb{Z}[x,y]$ of each degree $k \in \{3,4,6,12\}$, with positive leading coefficient. Supposing $\alpha_i \in \mathbb{Z}$ for each $i = 0,1,2,3$, however, does not guarantee that the associated Klein form $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)_k$ is actually in $\mathbb{Z}[x,y]$.

2.2. **The syzygy.** Our main purpose in studying Klein forms is that certain of their covariants satisfy a ternary relationship, closely analogous to (4). This will prove important in the sequel, for our construction of Frey-Hellegouarch curves corresponding to solutions to (2). Let F be a Klein form of degree $k \in \{3,4,6,12\}$ and index $n = 6 - 12/k$. As in the case $k = 3$, we define the Hessian of F as

$$H(x,y) = \frac{1}{(k-1)^2} \begin{vmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{vmatrix} \tag{10}$$

and the Jacobian determinant of F and H by

$$G(x,y) = \frac{1}{k-2}\begin{vmatrix} F_x & F_y \\ H_x & H_y \end{vmatrix}. \tag{11}$$

Then, analogous to (4), one has the classical syzygy

$$4H(x,y)^3 + G(x,y)^2 = d_n F(x,y)^n, \tag{12}$$

for certain $d_n \in \overline{\mathbb{Q}}^*$. If $F(x,y) \in \mathbb{Z}[x,y]$, then the same is true of $G(x,y)$ and $H(x,y)$, and d_n is, in fact, a nonzero integer. Specifically, we can write

$$d_n = -2^{i_n} 3^{j_n} \delta_n \tag{13}$$

where δ_n is an integer satisfying

$$\Delta_F = c_n \delta_n^{k(k-1)/6}, \tag{14}$$

and i_n, j_n and c_n are as follows:

n	i_n	j_n	c_n
2	0	3	1
3	8	0	-3^3
4	4	3	-2^8
5	8	3	5^{25}

Equation (14) determines δ_n and hence d_n uniquely, if $n = 2$ or $n = 4$, and up to sign if $n = 3$ or $n = 5$. For the sake of completeness, we will provide an expression for δ_n in terms of the coefficients of F (see e.g. Appendix A of Edwards [13]). Write our Klein form $F(x,y) \in \mathbb{Z}[x,y]$ as

$$\begin{aligned} F(x,y) &= ax^3 + bx^2y + cxy^2 + dy^3, \\ F(x,y) &= ax^4 + bx^3y + 3cx^2y^2 + dxy^3 + ey^4, \\ F(x,y) &= ax^6 + bx^5y + 5cx^4y^2 + 10dx^3y^3 + 5ex^2y^4 + fxy^5 + gy^6, \end{aligned}$$

and

$$\begin{aligned} F(x,y) = \; & ax^{12} + bx^{11}y + 11cx^{10}y^2 + 55dx^9y^3 + 165ex^8y^4 + 66fx^7y^5 + 11gx^6y^6 \\ & + 66hx^5y^7 + 165ix^4y^8 + 55jx^3y^9 + 11kx^2y^{10} + lxy^{11} + my^{12}, \end{aligned}$$

for index $2, 3, 4$ and 5 respectively, where $a, b, c, \ldots, m$ are integers (the local conditions for index $3, 4$ and 5, such as the fact that α_2 is divisible by 11, in case $n = 5$, are immediate consequences of (7), (8) and (9)). Then we have

$$\begin{aligned} \delta_2 &= b^2c^2 - 4ac^3 - 4b^3d + 18abcd - 27a^2d^2 = \Delta_F, \\ \delta_3 &= 8ace + bcd - ad^2 - eb^2 - 2c^3, \\ \delta_4 &= -15d^2 + 10ce - bf + 6ag. \end{aligned}$$

There is no such polynomial expression for δ_5, but since we have $(a,b) \neq (0,0)$ for a Klein form, δ_5 can always be obtained from the identities

$$\begin{aligned} 5a\delta_5 &= 2ag - 7bf + 140ce - 105d^2, \\ 5b\delta_5 &= -420de + 126cf - 5bg + 84ah. \end{aligned}$$

For future use, we will have need of the following result.

Proposition 2.1. *The resultant of a binary form F of degree k with its Hessian H satisfies*

$$\mathrm{Res}(H(x,y), F(x,y)) = (-1)^k \, \Delta_F^2. \tag{15}$$

Proof. To prove this, we begin by considering the identity

$$(k-1)\,(F_x\,F_y - x\,y\,H) = k\,F_{xy}\,F,$$

readily established by equating coefficients. It follows that

$$\mathrm{Res}(F, F_x\,F_y) = \mathrm{Res}(F, x\,y\,H)$$

and hence

$$\mathrm{Res}(F, F_x)\,\mathrm{Res}(F, F_y) = \mathrm{Res}(F, x)\,\mathrm{Res}(F, y)\,\mathrm{Res}(F, H).$$

Without loss of generality, we may assume that $F(1,0)F(0,1) \neq 0$. Since

$$\mathrm{Res}(F, F_x) = (-1)^{k(k-1)/2}F(1,0)\Delta_F \quad \text{and} \quad \mathrm{Res}(F, F_y) = (-1)^{k(k-1)/2}F(0,1)\Delta_F,$$

while

$$\mathrm{Res}(F, x)\,\mathrm{Res}(F, y) = (-1)^k F(1,0)\,F(0,1),$$

we obtain (15). □

3. From elliptic curves to Klein forms (and back)

There is a strong connection between Klein forms and elliptic curves, whose arithmetic consequences we wish to exploit. Let us consider an elliptic curve in Weierstrass form

$$E \;:\; y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \tag{16}$$

where the a_i lie in $\overline{\mathbb{Q}}$, say. Define, as usual,

$$\begin{aligned} b_2 &= a_1^2 + 4a_2, \\ b_4 &= 2a_4 + a_1a_3, \\ b_6 &= a_3^2 + 4a_6, \\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2 \end{aligned}$$

and, for integer $n \geq 2$, let

$$\psi_n^E(x) = \prod_{\substack{\{P,-P\}\subset E[n],\\ \mathrm{ord}(\pm P)=n}} (x - x_P)$$

denote the monic (division) polynomial with roots the x-coordinates of the points of E with exact order n. One readily checks that

$$\begin{aligned} 4\psi_2^E(x) &= 4x^3 + b_2x^2 + 2b_4x + b_6, \\ 3\psi_3^E(x) &= 3x^4 + b_2x^3 + 3b_4x^2 + 3b_6x + b_8, \\ 2\psi_4^E(x) &= 2x^6 + b_2x^5 + 5b_4x^4 + 10b_6x^3 + 10b_8x^2 + (b_2b_8 - b_4b_6)x + b_4b_8 - b_6^2, \end{aligned}$$

and

$$5\,\psi_5^E(x) = 32\,\psi_2^2\psi_4 - 27\,\psi_3^3.$$

For technical reasons, we will have use of a slightly modified version of $\psi_5^E(x)$:

$$\tilde{\psi}_5^E(x) = x^{12} + b_2x^{11} + 11b_4x^{10} + 55b_6x^9 + 165b_8x^8 - 66(b_4b_6 - b_2b_8)x^7$$
$$+11(-b_2b_4b_6 - 15b_6^2 + b_2^2b_8 + 10b_4b_8)x^6 - 66(b_2b_6^2 - b_2b_4b_8 + b_6b_8)x^5$$
$$-165(b_4b_6^2 - b_2b_6b_8 + 5b_8^2)x^4 - 55(5b_6^3 - 6b_4b_6b_8 + b_2b_8^2)x^3$$
$$-11(b_2b_6^3 - 2b_2b_4b_6b_8 + 21b_6^2b_8 + b_2^2b_8^2 - 25b_4b_8^2)x^2$$
$$+(-b_2^2b_6^3 + 20b_4b_6^3 + 2b_2^2b_4b_6b_8 - 46b_2b_6^2b_8 - b_2^3b_8^2 + 30b_2b_4b_8^2 + 95b_6b_8^2)x$$
$$-b_2b_4b_6^3 + 25b_6^4 + 2b_2^2b_6^2b_8 - 56b_4b_6^2b_8 - b_2^2b_4b_8^2 + 27b_2b_6b_8^2 - 125b_8^3.$$

The polynomials ψ_5^E and $\tilde{\psi}_5^E$ are not unrelated; one can in fact show that they have the same splitting field. We associate to the polynomials $\psi_n^E(x)$ ($n = 2, 3$ and 4) and to $\tilde{\psi}_5^E(x)$, binary forms, via homogenization:

$$K_n^E(x,y) = \begin{cases} \psi_n^E(x/y)\, y^k & \text{if } n = 2, 3 \text{ and } 4, \text{ for } k = 3, 4 \text{ and } 6, \text{ respectively;} \\ \tilde{\psi}_5^E(x/y)\, y^{12} & \text{if } n = 5. \end{cases}$$

For a singular curve E given by (16), we define forms K_n^E by the the same polynomial identities. If we denote by

$$\Delta_E = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6$$

the discriminant of E and by Δ_n the discriminant of the form $(6-n)K_n^E$ (for $n = 2, 3, 4, 5$), then a direct computation yields

$$\Delta_n = \begin{cases} 2^4\Delta_E & \text{if } n = 2, \\ -3^3\Delta_E^2 & \text{if } n = 3, \\ -2^8\Delta_E^5 & \text{if } n = 4, \\ 5^{25}\Delta_E^{22} & \text{if } n = 5. \end{cases} \tag{17}$$

Our motivation for introducing the forms K_n^E is provided by the following result.

Proposition 3.1. *Let E be an elliptic curve over a number field L, and let $n \in \{2, 3, 4, 5\}$. Then $K_n^E \in L[x, y]$ is a Klein form. Conversely, if $F(x, y) \in L[x, y]$ is a Klein form of index n with $F(1, 0) \neq 0$, then there exists an elliptic curve E/L such that*

$$F(x,y) = F(1,0)\, K_n^E(x,y). \tag{18}$$

Proof. To show that K_n^E is a Klein form, one checks that the conditions for the vanishing of the fourth transvectant $\tau_4(K_n^E)$ are satisfied; this is a short calculation, using the relation $4\, b_8 = b_2\, b_6 - b_4^2$. Furthermore, from (17) we immediately obtain that K_n^E is nondegenerate, which concludes the proof that K_n^E is a Klein form.

Conversely, from the explicit equations for the vanishing of the $\tau_4(F)$ (i.e. from (7), (8) and (9)), we observe that the coefficients of $F(x, y) = \sum_{i=0}^k \alpha_i x^{k-i}y^i$ are uniquely determined, via recursion, from the first four, $\alpha_0, \alpha_1, \alpha_2, \alpha_3$, provided $\alpha_0 \neq 0$. We define E as in (16), with $a_1 = a_3 = 0$ and

$$(a_2, a_4, a_6) = \begin{cases} (\frac{\alpha_1}{\alpha_0}, \frac{\alpha_2}{\alpha_0}, \frac{\alpha_3}{\alpha_0}) & \text{if } n = 2, \\ (\frac{3\alpha_1}{4\alpha_0}, \frac{\alpha_2}{2\alpha_0}, \frac{\alpha_3}{4\alpha_0}) & \text{if } n = 3, \\ (\frac{\alpha_1}{2\alpha_0}, \frac{\alpha_2}{5\alpha_0}, \frac{\alpha_3}{20\alpha_0}) & \text{if } n = 4, \\ (\frac{\alpha_1}{4\alpha_0}, \frac{\alpha_2}{22\alpha_0}, \frac{\alpha_3}{220\alpha_0}) & \text{if } n = 5, \end{cases}$$

whereby (18) holds. It remains to check that the discriminant of E is nonzero. This follows directly from (17) and the fact that the discriminant of a Klein form is, by definition, nonzero. □

Under the multiplication by two map on E, the points of order 4 are, naturally enough, mapped to points of order 2. This is expressed, at least at the level of x-coordinates, by the identity

$$K_2^E(x^4 - b_4x^2y^2 - 2b_6xy^3 - b_8y^4, 4yK_2^E(x,y)) = K_4^E(x,y)^2. \tag{19}$$

As a consequence of this, if F is a Klein form of index 4, there necessarily exists a Klein form G of index 2, and quartic forms p and q, all defined over the same field as F, for which $F(x,y)^2 = G(p(x,y), q(x,y))$. The forms G, p and q are readily obtained from (19) (if $F(1,0) = 0$, one needs to apply a suitable linear translation). Note that if the sextic form $F(x,y) \in \mathbb{Z}[x,y]$, the same need not be true for the corresponding cubic form $G(x,y)$.

4. Frey-Hellegouarch curves for Klein forms

We construct Frey-Hellegouarch curves for Klein forms $F(x,y) \in \mathbb{Z}[x,y]$ by exploiting the fact that an elliptic curve in short Weierstrass form

$$Y^2 = X^3 + aX + b$$

has discriminant

$$-2^4(4a^3 + 27b^2).$$

Multiplying identity (12) by 3^3, we find that

$$4\,(3\,H(x,y))^3 + 27\,G(x,y)^2 = 3^3\,d_n\,F(x,y)^n,$$

whereby the Frey-Hellegouarch curve

$$E_{x,y} \;:\; Y^2 = X^3 + 3\,H(x,y)X + G(x,y) \tag{20}$$

has discriminant

$$\begin{aligned} \Delta(x,y) &= -2^4 \cdot 3^3(4H(x,y)^3 + G(x,y)^2) \\ &= -2^4 \cdot 3^3 d_n F(x,y)^n. \end{aligned}$$

From the formulae for d_n (13), we may thus write

$$\Delta(x,y) = 2^{4+i_n} \cdot 3^{3+j_n} \delta_n F(x,y)^n. \tag{21}$$

Other fundamental quantities associated to $E_{x,y}$, which we record for later use, are

$$\begin{aligned} c_4(x,y) &= -2^4 \cdot 3^2 H(x,y), \\ c_6(x,y) &= -2^5 \cdot 3^3 G(x,y) \end{aligned}$$

and

$$j(x,y) = \frac{-2^{8-i_n} \cdot 3^{3-j_n} H(x,y)^3}{\delta_n F(x,y)^n}. \tag{22}$$

When we specialize (x,y) in (20) to integers (x_0, y_0) with $F(x_0, y_0) \neq 0$, we obtain an elliptic curve E_{x_0,y_0} over $\mathbb{Q}$ (given via a short Weierstrass model over $\mathbb{Z}$). For our purposes, it is important to understand arithmetic properties of $E_{x,y}$, for arbitrary (x,y), not just those satisfying the generalized superelliptic equation (2).

From the equation for $\Delta(x,y)$, it is immediate that if F is a Klein form of index n with $F(1,0) \neq 0$, then $E_{1,0}$ is an elliptic curve; the Klein form $K_n^{E_{1,0}}(x,y)$ associated to it is simply the original Klein form F up to scaling and a linear

change of variables. Specifically, we have (as can be verified via any computer algebra package)

$$K_n^{E_{1,0}}(x,y) = \frac{1}{\alpha_0} F\left(x - \alpha_1 y, k\,\alpha_0 y\right), \tag{23}$$

where $k = 12/(6-n)$.

In practice, we will typically consider quadratic twists of (20), chosen to minimize its conductor, when (x,y) are specialized to a solution (x_0,y_0) of (2). The quadratic twist of (20) by $t \in \mathbb{Q}^*$ is denoted by $E_{x,y}^{(t)}$, i.e.

$$E_{x,y}^{(t)} \,:\, Y^2 = X^3 + 3\,H(x,y)t^2X + G(x,y)t^3.$$

We will choose t such that $E_{x_0,y_0}^{(t)}$ has good reduction outside primes dividing $n\,\Delta_F\,F(x_0,y_0)$; to demonstrate the existence of such a twist, we will derive explicit (minimal) models for $E_{x,y}^{(t)}$ in the next two sections and Appendix A.1 (though, for simplicity, in case $n=5$, our models are not necessarily minimal at 2).

4.1. **Twisting and minimization at 3.** For forms of index $n \in \{2,4,5\}$, we can actually remove the factor $3^{3+j_n} = 3^6$ in the discriminant $\Delta(x,y)$ of the Frey-Hellegouarch curve (20) as follows. Let us define $T(x,y)$ to be $\alpha_1 x - \alpha_2 y$, if $n=2$, $\alpha_1 x^4 - \alpha_2 x^3 y + \alpha_4 x y^3 - \alpha_5 y^4$, if $n=4$, and

$$\alpha_1 x^{10} - \alpha_2 x^9 y + \alpha_4 x^7 y^3 - \alpha_5 x^6 y^4 + \alpha_7 x^4 y^6 - \alpha_8 x^3 y^7 + \alpha_{10} x y^9 - \alpha_{11} y^{10},$$

if $n=5$. Then translation by $T(x,y)$, and twisting over $\mathbb{Q}(\sqrt{3})$, transforms (20) into a model for $E_{x,y}^{(3)}$ given by

$$Y^2 = X^3 + TX^2 + \frac{T^2+H}{3}X + \frac{T^3+3TH+G}{27}. \tag{24}$$

It is relatively easy to show that $(T^2+H)/3, (T^3+3TH+G)/27 \in \mathbb{Z}[x,y]$, in each case. Fundamental quantities associated to (24) are

$$\Delta(x,y) = 2^{4+i_n}\,\delta_n\,F(x,y)^n, \tag{25}$$

$$c_4(x,y) = -2^4 H(x,y), \quad c_6(x,y) = -2^5 G(x,y)$$

and

$$j(x,y) = \frac{-2^{8-i_n}\,H(x,y)^3}{\delta_n F(x,y)^n}. \tag{26}$$

Note that if we specialize x and y to (coprime) integers here, the resulting model over $\mathbb{Z}$ need not be minimal at 3.

4.2. **Twisting and minimization at 2.** For Klein forms with index $n \in \{3,5\}$ and odd discriminant Δ_F, we can remove the factor $2^{4+i_n} = 2^{12}$ from $\Delta(x_0,y_0)$, again through choice of suitable twist.

Lemma 4.1. *Suppose that F is a Klein form of degree 4 or 12, with Δ_F odd. Then, for any coprime $x_0, y_0 \in \mathbb{Z}$ with $F(x_0,y_0) \neq 0$ either the model (20) for E_{x_0,y_0} or its twist by -1 is not minimal at 2. In particular, for the corresponding $t = \pm 1$, there exists a Weierstrass model over $\mathbb{Z}$ for $E_{x_0,y_0}^{(t)}$ with*

$$c_4 = -3^2 H(x_0,y_0), \quad \Delta = 3^{3+j_n}\delta_n F(x_0,y_0)^n.$$

Proof. For a form F as in (6), with $k = 4$, we twist (20) by $(-1)^\delta$, where

$$\delta = \begin{cases} 1 & \text{if } \alpha_1 \equiv 1 \pmod 4 \text{ or } \alpha_3 \equiv -1 \pmod 4, \\ 0 & \text{if } \alpha_1 \equiv -1 \pmod 4 \text{ or } \alpha_3 \equiv 1 \pmod 4. \end{cases} \tag{27}$$

Note that assuming Δ_F (and hence δ_3) to be odd is equivalent to

$$\alpha_1\alpha_2\alpha_3 + \alpha_0\alpha_3 + \alpha_1\alpha_4 \equiv 1 \pmod 2, \tag{28}$$

which guarantees that at least one of α_1 or α_3 is also odd, while (7) ensures that α_1 and α_3 are distinct modulo 4. The interested reader can find further information (including minimal models) for Klein forms of degree 4 in Appendix A.1. We omit the (gory) details of the degree 12 case; they are essentially similar (if much longer). □

It is worthwhile to note that the minimization of the valuation of the conductor at the prime 2 can be done independently from that at 3.

4.3. **Frey-Hellegouarch curves: conclusions.** Combining our models from the preceding subsections, we arrive at a general result on primes dividing the minimal discriminants of the $E_{x,y}$. Here and henceforth, let us denote by S_F the set of primes dividing $n\,\Delta_F$, together with the prime at ∞. Furthermore, we define, for a nonzero integer m, $\nu_p(m)$ to be the largest power of p dividing m, and set, for $m, n \in \mathbb{Z} \setminus \{0\}$, $\nu_p(m/n) = \nu_p(m) - \nu_p(n)$.

Proposition 4.2. *Let $F(x, y) \in \mathbb{Z}[x, y]$ be a Klein form of index n, with corresponding family of Frey-Hellegouarch curves $E_{x,y}$ as given by (20) and let x_0, y_0 be coprime integers with $F(x_0, y_0) \neq 0$. Then there exists $t \in \{\pm 1, \pm 3\}$ such that for all primes $p \notin S_F$ we have that $E^{(t)}_{x_0,y_0}$ is semistable at p and*

$$\nu_p(\Delta_{\min}(E^{(t)}_{x_0,y_0})) = n\,\nu_p(F(x_0, y_0)).$$

Proof. This follows from the preceding formulae for c_4 and Δ, together with Proposition 2.1 and the well known fact that if a Weierstrass model over $\mathbb{Z}$ for an elliptic curve E has $p \nmid c_4$ or $p \nmid \Delta$ for a prime p, then E is semistable at p and the model is minimal at p. □

5. Galois representations attached to $E/\mathbb{Q}$

Let $E/\mathbb{Q}$ be an elliptic curve and n a positive integer. By ρ^E_n we denote the mod n Galois representation $G_\mathbb{Q} \to \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ induced by the natural action of the absolute Galois group $G_\mathbb{Q} = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the n-torsion points $E[n]$ (and the choice of a basis for $E[n]$). Recall the following standard result.

Proposition 5.1. *Let $E/\mathbb{Q}$ be an elliptic curve with conductor N, n a positive integer and p a prime. If $p \nmid nN$, then ρ^E_n is unramified at p and*

$$\begin{aligned} \mathrm{Trace}(\rho^E_n(\mathrm{Frob}_p)) &\equiv a_p(E) \pmod n, \\ \mathrm{Det}(\rho^E_n(\mathrm{Frob}_p)) &\equiv p \pmod n. \end{aligned}$$

In case n is prime, combining the theorem above with the Chebotarev density theorem and the Brauer-Nesbitt theorem, we obtain the following well-known result.

Proposition 5.2. *Let E and E' be elliptic curves over $\mathbb{Q}$ and n be prime. Then the semisimplifications of ρ^E_n and $\rho^{E'}_n$ are isomorphic if and only if, for all but finitely many primes p, we have $a_p(E) \equiv a_p(E') \pmod n$.*

Since we desire to apply this result with $n \in \{2,3,4,5\}$, we require something analogous to Proposition 5.2 for the case $n = 4$. More generally, for $n = l^e$ a prime power, we can replace the Brauer-Nesbitt theorem by Carayol's partial generalization [5, Thm. 1] for representations over local rings ($\mathbb{Z}/l^e\mathbb{Z}$ in our case), at the cost of assuming absolute irreducibility for the corresponding residual mod l representation.

Proposition 5.3. *Let E and E' be elliptic curves over $\mathbb{Q}$, and let $n = l^e$ for some prime l and positive integer e. Suppose that ρ_l^E and $\rho_l^{E'}$ are absolutely irreducible. Then ρ_n^E and $\rho_n^{E'}$ are isomorphic if and only if, for all but finitely many primes p, we have $a_p(E) \equiv a_p(E') \pmod{n}$.*

6. Moduli of elliptic curves: constant n-torsion

As it transpires and crucially for our arguments, the Frey-Hellegouarch curves associated to the Klein forms of index n actually describe families of elliptic curves with constant n-torsion. Before we make this more precise, we introduce, following Fisher [19], some terminology.

Definition 6.1. Let K be a number field, E and E' be elliptic curves over K, and n be a positive integer. We call E *n-congruent* to E' is there exists an isomorphism $\psi : E[n] \to E'[n]$ of $\mathrm{Gal}(\overline{K}/K)$-modules that respects the Weil pairing, i.e. for all $P, Q \in E[n]$ we have

$$e_n(\psi(P), \psi(Q)) = e_n(P, Q).$$

We call E *reverse n-congruent* to E' if there exists an isomorphism $\psi : E[n] \to E'[n]$ of $\mathrm{Gal}(\overline{K}/K)$-modules that reverses the Weil pairing, i.e. for all $P, Q \in E[n]$ we have

$$e_n(\psi(P), \psi(Q)) = e_n(P, Q)^{-1}.$$

For our purposes, it is quite straightforward to classify elliptic curves that are n-congruent to our Frey-Hellegouarch curves.

Proposition 6.2. *Suppose that $F \in K[x,y]$ is a Klein form of index $n \in \{2,3,4,5\}$ with $F(1,0) \neq 0$ and corresponding family of Frey-Hellegouarch curves $E_{x,y}$. Let E be an elliptic curve over K. Then E is n-congruent to $E_{1,0}$ if and only if E is isomorphic over K to $E_{x,y}$ for some $x, y \in K$.*

Proof. Up to some scaling and linear changes of variables, this is immediate from Theorem 9.4 of Fisher [19] (see also Rubin and Silverberg [40], [41], and Silverberg [49]). □

In case $n = 2$, this characterization provides us with all we require, since, for elliptic curves E and E' over K, we have $E[2] \simeq E'[2]$ as $\mathrm{Gal}(\overline{K}/K)$ modules precisely when E and E' are 2-congruent. The analogous statement for $n \geq 3$, however, is no longer true in general. Indeed, suppose that $E/\mathbb{Q}$ is an elliptic curve and $n \geq 3$ an integer. Write $\tilde{Y}_E(n)$ for the functor from $\mathbb{Q}$-schemes to sets which is determined by denoting by $\tilde{Y}_E(n)(K)$ the set of isomorphisms classes (E', ψ), where E' is an elliptic curve over K and $\psi : E'[n] \to E[n]$ is an isomorphism of $\mathrm{Gal}(\overline{K}/K)$-modules. There exists, then, an affine curve $Y_E(n)$ over $\mathbb{Q}$ representing the functor $\tilde{Y}_E(n)$, with the property that $Y_E(n)$ decomposes over $\mathbb{Q}$ into $\phi(n)$ smooth absolutely irreducible affine curves $Y_{E^a}(n)$, one for each $a \in (\mathbb{Z}/n\mathbb{Z})^*$. Let e_n denote the Weil pairing. Then the K-rational points on the (moduli) curve

$Y_{E^a}(n)$ correspond to pairs (E', ψ) as above, but with the extra condition that the isomorphism $\psi : E'[n] \to E[n]$ satisfies $e_n(\psi(P), \psi(Q)) = e_n(P, Q)^a$. The multiplication by $b \in (\mathbb{Z}/n\mathbb{Z})^*$ map on $E[n]$ induces an isomorphism between $Y_{E^a}(n)$ and $Y_{E^{ab^2}}(n)$, whereby in order to determine all elliptic curves E' such that $E[n] \simeq E'[n]$ it suffices to restrict attention to the moduli curves $Y_{E^a}(n)$ with $a \in (\mathbb{Z}/n\mathbb{Z})^*$ distinct modulo $((\mathbb{Z}/n\mathbb{Z})^*)^2$. In particular, for $n = 3$ and 4 (that is, for Klein forms of degrees 4 and 6), it suffices to consider both n-congruent and reverse n-congruent curves to a given Frey-Hellegouarch curve. In the case of index $n = 5$, we must treat not only 5-congruent curves, but also curves arising from $Y_{E^a}(5)$ for one nonsquare value of a modulo 5, say $a = 2$.

6.1. **Curves with isomorphic n-torsion.** Suppose that $F \in \mathbb{Q}[x, y]$ is a Klein form of index n with $F(1, 0) \neq 0$. To complete the classification of all elliptic curves $E/\mathbb{Q}$ with n-torsion isomorphic to the Frey-Hellegouarch curve $E_{1,0}$ corresponding to F, we begin by defining a companion (Klein) form $\tilde{F}$ by

$$\tilde{F} = \begin{cases} F & \text{if } n = 2, 4, \\ H & \text{if } n = 3, \\ (A, B, 11C, 55D)_{12} & \text{if } n = 5, \end{cases} \tag{29}$$

where, for $n = 5$, we write $F = (a, b, 11c, 55d)_{12}$ and set

$$\begin{aligned} A &= -27(7b^6 - 504ab^4c + 11072a^2b^2c^2 - 64000a^3c^3 + 1024a^2b^3d \\ &\quad -36864a^3bcd + 110592a^4d^2), \\ B &= 1620(b^2 - 24ac)^2(b^3 - 36abc + 216a^2d), \\ C &= -162(b^2 - 24ac)(3b^6 - 216ab^4c + 4288a^2b^2c^2 - 12800a^3c^3 \\ &\quad +896a^2b^3d - 32256a^3bcd + 96768a^4d^2), \\ D &= 108(b^3 - 36abc + 216a^2d)(b^6 - 72ab^4c + 1472a^2b^2c^2 - 5632a^3c^3 \\ &\quad +256a^2b^3d - 9216a^3bcd + 27648a^4d^2). \end{aligned}$$

Remark 6.3. A more canonical choice, in case $n = 2$, would be to define $\tilde{F} = G$. For our applications, however, this provides no advantages (and introduces some minor complications).

To a companion form $\tilde{F}$, we associate a family of elliptic curves $\tilde{E}_{x,y}$ as follows. If $n = 2$, we simply take $\tilde{E}_{x,y} = E_{x,y}$. For $n = 3$, we define $\tilde{E}_{x,y}$ to be following twist of the standard family of curves associated to $\tilde{F}$:

$$\tilde{E}_{x,y} \ : \ Y^2 = X^3 + 2^4 \cdot 3\, \delta_3^3\, F(x, y)X - 2^3\, \delta_3^4 G(x, y).$$

In this case, the corresponding discriminant and j-invariant are given by

$$\tilde{\Delta}(x, y) = 2^{12} \cdot 3^3\, \delta_3^8\, H(x, y)^3 \quad \text{and} \quad \tilde{j}(x, y) = \frac{-2^{12}\, 3^3\, \delta_3\, F(x, y)^3}{H(x, y)^3}. \tag{30}$$

Note that here we have

$$j(x, y)\tilde{j}(x, y) = 1728^2.$$

For $n = 4$, we define $\tilde{E}_{x,y} = E_{x,y}^{(-\delta_4)}$. Our motivation for introducing $\tilde{E}_{x,y}$ (in case $n = 3$ or 4) is apparent in the following result.

Proposition 6.4. *Let $F \in \mathbb{Q}[x, y]$ be a Klein form of index $n \in \{3, 4\}$ with $F(1, 0) \neq 0$ and corresponding family of curves $E_{x,y}$. If $E/\mathbb{Q}$ is an elliptic curve,*

then E is reverse n-congruent to $E_{1,0}$ if and only if E is isomorphic over $\mathbb{Q}$ to $\tilde{E}_{x,y}$ for some coprime $x, y \in \mathbb{Z}$.

Proof. This can be found, modulo scaling and a linear change of variables, in [19, p. 22] (after observing that $-\Delta(x,y)/\delta_4$ is a square). □

Since -1 is a square modulo 5, the property of being reverse 5-congruent is equivalent to that of being 5-congruent. To complete our classification of curves with isomorphic n-torsion to $E_{1,0}$, it remains, then, to find elliptic curves corresponding to the rational points on $Y_{E^a}(5)$ for some nonsquare a, modulo 5, say $a = 2$. Let j_5 and j_5' denote the j-maps from $X_E(5)$ and $X_{E^2}(5)$, respectively, to $X(1)$. We wish to relate j_5' to j_5. For this purpose, we introduce the map $J : X(1) \to X(1)$ given by

$$\begin{aligned} (31)\, J(j) &= \frac{j(j^2 - 1456j - 3670016)^3}{(-7j + 4096)^5} \\ &= -\frac{(-j + 1728)(j^3 - 1320j^2 + 9043968j + 1073741824)^2}{(-7j + 4096)^5} + 1728. \end{aligned}$$

In [7, p. 86], it is shown that, for a rational point P on $X_E(5)$, we have $J(j_5(P)) = j_5'(P')$ for some rational point P' on $X_{E^2}(5)$.

Now consider the Klein form $F = (a, b, 11c, 55d)_{12}$, where we assume that $a \neq 0$, and as usual write $j(x,y)$ for the j-invariant of the family $E_{x,y}$ associated to F (which we identify with j_5). Denote by $\tilde{E}'_{x,y}$ the corresponding family of curves associated to the companion form $\tilde{F}$, i.e.

$$\tilde{E}'_{x,y} : Y^2 = X^3 + 3\,H(\tilde{F})(x,y)X + G(\tilde{F})(x,y),$$

and let $\tilde{j}(x,y)$ denote its j-invariant. It is straightforward to check that $J(j(1,0))$ is the image of a rational point under $\tilde{j}(x,y)$ (and hence this j-map can be identified with j_5'). It follows that some twist of $\tilde{E}'_{x,y}$ is the family of elliptic curves corresponding to the rational points on the modular curve $X_{E^2_{1,0}}(5)$. We denote this twist by $\tilde{E}_{x,y}$. For any given duodecic Klein form $F \in \mathbb{Z}[x,y]$, it is a simple matter to explicitly write down the correct twist; in practice, we actually only need its j-invariant, which is, of course, given by $\tilde{j}(x,y)$ above. We summarize.

Proposition 6.5. *Let $F \in \mathbb{Q}[x,y]$ be a Klein form of index $n = 5$ with $F(1,0) \neq 0$ and corresponding family of curves $E_{x,y}$ and let $E/\mathbb{Q}$ be an elliptic curve. Then E is isomorphic over $\mathbb{Q}$ to $\tilde{E}_{x,y}$ for some coprime $x, y \in \mathbb{Z}$ if and only if there exists a $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-module isomorphism $\psi : E[5] \to E_{1,0}[5]$ for which, for all $P, Q \in E[5]$, we have $e_n(\psi(P), \psi(Q)) = e_n(P,Q)^2$.*

For $n = 2, 3, 4$ and 5 we can now describe all elliptic curves with isomorphic n-torsion to a given one.

Theorem 6.6. *Let $F \in \mathbb{Q}[x,y]$ be a Klein form of index $n \in \{2,3,4,5\}$ with $F(1,0) \neq 0$ and corresponding families of curves $E_{x,y}$ and $\tilde{E}_{x,y}$. Let $E/\mathbb{Q}$ be an elliptic curve. Then $E[n] \simeq E_{1,0}[n]$ as $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-modules if and only if E is isomorphic over $\mathbb{Q}$ to $E_{x,y}$ or $\tilde{E}_{x,y}$ for some integers x, y (which can be taken to be coprime if $n \neq 2$).*

Next, we define $\tilde{S}_F = S_{\tilde{F}} = \{p : p \mid n\Delta_{\tilde{F}}\} \cup \{\infty\}$. A direct computation gives

$$\Delta_{\tilde{F}}/c_n = \begin{cases} \Delta_F/c_n & \text{if } n = 2, 4 \\ (2^6\Delta_F/c_n)^2 & \text{if } n = 3, \\ (2^{198} \cdot 3^{99} \cdot F(1,0)^{110}\Delta_F/c_n)^2 & \text{if } n = 5. \end{cases}$$

This yields

$$\tilde{S}_F = \begin{cases} S_F & \text{if } n = 2, 4 \\ S_F \cup \{2\} & \text{if } n = 3, \\ S_F \cup \{p : p \mid 6\, F(1,0)\} & \text{if } n = 5. \end{cases} \tag{32}$$

We conclude this subsection with a straightforward but useful result.

Lemma 6.7. *Suppose that $F \in \mathbb{Z}[x, y]$ is a Klein form of index n with corresponding families of curves $E_{x,y}$ and $\tilde{E}_{x,y}$. Write $j(x, y)$ for the j-invariant of $E_{x,y}$ and $\tilde{j}(x, y)$ for the j-invariant of $\tilde{E}_{x,y}$. Let $E/\mathbb{Q}$ be an elliptic curve of conductor $N_E \in \mathbb{Z}^*_{S_F}$ with j-invariant j_E. If $j_E = j(x, y)$ for some coprime integers x, y, then $F(x, y) \in \mathbb{Z}^*_{S_F}$. If $j_E = \tilde{j}(x, y)$ for some coprime integers x, y, then $\tilde{F}(x, y) \in \mathbb{Z}^*_{\tilde{S}_F}$.*

Proof. Let us suppose that $j_E = j(x, y)$ for coprime integers x, y, and assume that $F(x, y) \notin \mathbb{Z}^*_{S_F}$ (the case where $j_E = \tilde{j}(x, y)$ is handled in a similar fashion). There thus exists a prime p for which $p \mid F(x, y)$ and $p \nmid n\Delta_F$. From the explicit equation for $j(x, y)$, together with Proposition 2.1, it follows that $\nu_p(j(x, y)) < 0$ and hence E has bad reduction at p. We conclude that $N_E \notin \mathbb{Z}^*_{S_F}$, a contradiction. □

For future reference, as in the case of $E_{x,y}$, we define, for $t \in \mathbb{Q}^*$, $\tilde{E}^{(t)}_{x,y}$ to be the quadratic twist of $\tilde{E}_{x,y}$ by t.

6.2. **Irreducibility and ramification properties.** The explicit formulae for our families of elliptic curves $E_{x,y}$ associated to a Klein form F of index n imply that the corresponding Klein form $K_n^{E_{1,0}}$ is, up to a linear change of variables, a constant multiple of F (see (23)). This provides useful information about the irreducibility of $\rho_n^{E_{x,y}}$.

Proposition 6.8. *Let $F \in \mathbb{Q}[x, y]$ be a Klein form of index $n \in \{2, 3, 4, 5\}$ with $F(1, 0) \neq 0$. Consider the family of elliptic curves $E_{x,y}$. If F is irreducible, then for any coprime integers x and y, the mod-n Galois representation $\rho_n^{E_{x,y}}$ is irreducible. Moreover, if $n \in \{2, 3, 4\}$, then, for any such x and y, $\rho_n^{E_{x,y}}$ is irreducible precisely when F has no linear factor over $\mathbb{Q}$.*

Proof. Because the mod-n Galois representations are constant on the family $E_{x,y}$, it suffices to check irreducibility for $(x, y) = (1, 0)$. For $n = 2, 3$ and 4, the modular curves $X_0(n)$ and $X_1(n)$ coincide, whereby $\rho_n^{E_{1,0}}$ is irreducible if and only if $\psi_n^{E_{1,0}}$ has no linear factor over $\mathbb{Q}$. By construction, this is equivalent to $K_n^{E_{1,0}}$ having no linear factor over $\mathbb{Q}$. Since this form is a constant multiple of F, up to linear change of variables, the desired result follows for $n = 2, 3$ and 4. In case $n = 5$, the irreducibility of F forces the field $\mathbb{Q}(E[5]_x)$ to be "large", in the sense that $\mathrm{Gal}(\mathbb{Q}(E[5]/\mathbb{Q}))$ is not contained in a Borel subgroup, i.e. $\rho_5^{E_{x,y}}$ is irreducible. □

Remark 6.9. If a Klein form F of index $n = 5$ is reducible, it could still happen that the corresponding mod 5 representation $\rho_5^{E_{x,y}}$ is irreducible. In any particular

case, this is easy to determine by checking whether or not $E_{1,0}$ has a rational 5-isogeny.

Since $\rho_n^{E_{x,y}} \simeq \rho_n^{\tilde{E}_{x,y}}$, we immediately conclude that the above proposition also holds with the families $E_{x,y}$ replaced by $\tilde{E}_{x,y}$. Furthermore, irreducibility is not affected by taking quadratic twists, so $E_{x,y}$ can also be replaced by any quadratic twist of $E_{x,y}$ or $\tilde{E}_{x,y}$.

In light of Proposition 5.3, we would also like to have a criterion for the residual mod 2 representation associated to $\rho_4^{E_{x,y}}$ to be absolutely irreducible.

Proposition 6.10. *Let $F \in \mathbb{Q}[x,y]$ be a Klein form of index 4 with corresponding family of Frey-Hellegouarch curves $E_{x,y}$. If F has no factor of degree ≤ 2 over $\mathbb{Q}$ and δ_4 is not a square in $\mathbb{Q}$, then for any (nondegenerate) x, y, the mod-2 Galois representation $\rho_2^{E_{x,y}}$ is absolutely irreducible.*

Proof. If $E/\mathbb{Q}$ is an elliptic curve with a rational 2-torsion point, then $\psi_4^E(x)$ necessarily has a factor of degree ≤ 2 over $\mathbb{Q}$. Further, if ρ_2^E is irreducible, but not absolutely irreducible, then Δ_E is a square in $\mathbb{Q}$. Arguing as in the proof of Proposition 6.8, the result follows. □

Although the conductor of $E_{x,y}^{(t)}$ associated to a Klein form of index n may contain many primes not dividing $n\,\Delta_F$, the ramification of $\mathbb{Q}(E_{x,y}^{(t)}[n])$ can still be restricted in a satisfactory manner, using Tate curves.

Proposition 6.11. *Let $F \in \mathbb{Z}[x,y]$ be a Klein form of index n, x and y integers with $F(x,y) \neq 0$, and let $t \in \mathbb{Z} \setminus \{0\}$ be such that $E_{x,y}^{(t)}$ is semistable outside S_F. Then the representation $\rho_n^{E_{x,y}^{(t)}}$ is unramified outside primes dividing $n\,\Delta_F$.*

Proof. By Proposition 4.2 the minimal discriminant of $E_{x,y}^{(t)}$ is an n-th power, up to primes dividing $n\,\Delta_F$, and at primes not dividing $n\,\Delta_F$ the elliptic curve has good or multiplicative reduction. A simple application of the theory of Tate curves now gives the proposition. □

7. The Chebotarev density theorem

Our goal in this section is to effectively bound the smallest prime p with $a_p(E_1)$ and $a_p(E_2)$ distinct, given two elliptic curves $E_1/\mathbb{Q}$ and $E_2/\mathbb{Q}$ with $E_1[n] \not\simeq E_2[n]$. To do this, we appeal to an effective version of the Chebotarev density theorem, due to Lagarias, Montgomery and Odlyzko [27].

Theorem 7.1 (Lagarias, Montgomery, Odlyzko)**.** *Let $K/\mathbb{Q}$ be a nontrivial finite Galois extension with Galois group G and denote by d_K the absolute value of the discriminant of K. Let C be a conjugacy class of G. Then there exists a (rational) prime p such that* $\mathrm{Frob}_p = C$ *and*

$$\log p \leq c_1 \log d_K,$$

where c_1 is an absolute effectively computable constant.

Proof. See [27, Theorem 1.1]. □

Note that the prime p referenced in Theorem 7.1 is necessarily unramified in K. For our purposes, however, we require somewhat more, namely a prime p with Frob_p corresponding to a given conjugacy class, but also lying outside a specified finite

set of primes S, containing all the ramified primes. The need for this stems from the fact that if $K = \mathbb{Q}(E[n])$, where $E/\mathbb{Q}$ denotes an elliptic curve with conductor N, then we wish to work with primes $p \nmid Nn$ – but not all primes dividing Nn are necessarily ramified in $\mathbb{Q}(E[n])$.

It is worthwhile observing at this juncture that, under the assumption of the Generalized Riemann Hypothesis, the upper bound in Theorem 7.1 can be improved to

$$p \leq c_2 \, \log^2 d_K,$$

where c_2 is another absolute constant (see Serre [44]; one may take $c_2 = 70$). In fact, in [44], one finds an argument that enables one to obtain (again under the assumption of the Generalized Riemann Hypothesis) a bound of the shape

$$p \leq c_3 \, [K : \mathbb{Q}]^2 \left(\log([K : \mathbb{Q}]) + \sum_{q \in S} \log(q) \right)^2,$$

for the smallest rational prime p with $\mathrm{Frob}_p = C$, satisfying the additional condition that $p \notin S$, for S a finite set of primes containing those primes that ramify in K. Here c_3 is again an absolute constant (one can take $c_3 = 280$). In a similar fashion, we may derive the following variant of Theorem 7.1.

Theorem 7.2. *Let $K/\mathbb{Q}$ be a finite Galois extension with Galois group G of degree $n > 1$ and denote by d_K the absolute value of the discriminant of K. Let S be a finite set of primes, including those primes that ramify in K, and let C be a conjugacy class of G. Then there exists a (rational) prime $p \notin S$ such that $\mathrm{Frob}_p = C$ and*

$$\log p \leq c_1 \left(n \log 2 + n \sum_{\substack{q \in S \\ q \nmid d_K}} \log q + 2 \log d_K \right),$$

where c_1 is the same absolute effectively computable constant as in Theorem 7.1.

Proof. Define $K' = K(\sqrt{\pm D})$, where D is the product of the odd primes in S that fail to ramify in K, and the sign is chosen so that $d_{K'}$ is minimal under the restriction that 2 ramifies in K' if $2 \in S$ and fails to ramify in K. Arguing as in [44, pp. 135–136], we deduce the inequality

$$d_{K'} \leq 2^n \, d_K^2 \prod_{\substack{q \in S \\ q \nmid d_K}} q^n,$$

whereby an application of Theorem 7.1, with K replaced by K', together with standard properties of Frob_p, provides the required bound for p. □

Corollary 7.3. *Let $K/\mathbb{Q}$ be a finite Galois extension with Galois group G of degree $n > 1$ and denote by d_K the absolute value of the discriminant of K. Let S be a finite set of primes, including those primes that ramify in K, and let C be a conjugacy class of G. Then there exists a (rational) prime $p \notin S$ such that $\mathrm{Frob}_p = C$ and*

$$\log p \leq \kappa_n \sum_{q \in S} \log q,$$

where κ_n is an effectively computable constant depending only on $n = [K : \mathbb{Q}]$.

Proof. As is well known, there exists an effectively computable constant C_n, only depending on n, such that for every prime $q \mid d_K$ we have $\nu_q(d_K) \leq C_n$. (Denoting by D_K the different of the extension $K/\mathbb{Q}$, this follows easily from the facts that $d_K = |\mathrm{Norm}_{K/\mathbb{Q}}(D_K)|$ and $\nu_{\mathfrak{q}}(D_K) \leq e - 1 + \nu_{\mathfrak{q}}(e)$ for every prime $\mathfrak{q}$ of K with ramification index e for the extension $K/\mathbb{Q}$). The result is now immediate from Theorem 7.2. □

An application of this corollary leads to our desired result on non-isomorphic mod-n Galois representations associated to elliptic curves.

Proposition 7.4. *Let $E_1/\mathbb{Q}$ and $E_2/\mathbb{Q}$ be elliptic curves with conductors N_1 and N_2, respectively, where $N_1 \mid N_2$. Let $n = l^e$ for some prime l and positive integer e, and write $\rho_i = \rho_n^{E_i}$ for $i = 1$ and 2. Suppose that ρ_2 is unramified outside primes dividing nN_1, that ρ_2 is irreducible and, if $e > 1$, that $\rho_l^{E_2}$ is absolutely irreducible. If $\rho_1 \not\simeq \rho_2$, then there exists a prime p with $p \nmid nN_1$, for which both*

$$\mathrm{Trace}(\rho_1(\mathrm{Frob}_p)) \not\equiv \mathrm{Trace}(\rho_2(\mathrm{Frob}_p)) \pmod{n}$$

and

$$\log p \leq \kappa_n \sum_{q \mid nN_1} \log q, \tag{33}$$

where κ_n is an effectively computable constant only depending on n. In particular, for this prime p, we have $p \mid N_2$ or $a_p(E_1) \neq a_p(E_2)$.

Proof. Consider the (continuous) homomorphism

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}(2, \mathbb{Z}/n\mathbb{Z}) \times \mathrm{GL}(2, \mathbb{Z}/n\mathbb{Z})$$

given by

$$\sigma \mapsto (\rho_1(\sigma), \rho_2(\sigma)),$$

and denote by H its image and by K the fixed field of its kernel. Then $K/\mathbb{Q}$ is Galois, unramified outside $S := \{p : p \mid nN_1\}$, and we can and shall identify its Galois group with H. Propositions 5.2 and 5.3 guarantee the existence of an element $(a, b) \in H$ with

$$\mathrm{Trace}(a) \not\equiv \mathrm{Trace}(b) \pmod{n}. \tag{34}$$

By our version of the effective Chebotarev density theorem, i.e. Corollary 7.3, we conclude that there exists a prime $p \notin S$ such that p satisfies (33) and (a, b) is a Frobenius element above p. Thus, from (34),

$$\mathrm{Trace}(\rho_1(\mathrm{Frob}_p)) \not\equiv \mathrm{Trace}(\rho_2(\mathrm{Frob}_p)) \pmod{n}.$$

The last statement of the proposition follows immediately, since if $p \nmid N_2$ (we already know $p \nmid nN_1$), then $\mathrm{Trace}(\rho_i(\mathrm{Frob}_p)) \equiv a_p(E_i) \pmod{n}$ for $i = 1, 2$. □

8. The modular method

Having constructed Frey-Hellegouarch curves from our Klein forms, with corresponding mod l Galois representations having suitable ramification properties, it is relatively straightforward to translate this information into a statement about congruences between modular forms.

Proposition 8.1. *Let F be a Klein form of index n and suppose that x_0, y_0 and $z_0 \neq 0$ are integers such that*

$$F(x_0, y_0) = u_0\, z_0^l, \quad \gcd(x_0, y_0) = 1, \tag{35}$$

where $u_0 \in \mathbb{Z}_{S_F}^*$ and $l > 163$ is prime. Denote by N_0 the conductor of the corresponding Frey-Hellegouarch curve $E_0 = E_{x_0,y_0}^{(t)}$, with t chosen so that E_0 is semistable outside S_F. Then $\rho_l^{E_0}$ is modular of level

$$N_1 = \prod_{p \mid n\Delta_F} p^{\nu_p(N_0)}. \tag{36}$$

In particular, there exists a newform f of level N_1, weight 2 and trivial character, whose coefficients lie in a number field K_f, and a prime $\mathfrak{L} \subset \mathcal{O}_{K_f}$ lying above l, such that for all primes p with $p \nmid lN_0$,

$$a_p(f) \equiv a_p(E_0) \pmod{\mathfrak{L}}, \tag{37}$$

while for those primes p with $p \nmid lN_1$ and $p \mid N_0$, we have

$$a_p(f) \equiv \pm(p+1) \pmod{\mathfrak{L}}. \tag{38}$$

Proof. Via work of Breuil, Conrad, Diamond and Taylor [4] (building on that of Wiles [58]), we have that $\rho_l^{E_0}$ is modular of level N_0. Since $l > 163$, we may appeal to a result of Mazur [31] to conclude that $\rho_l^{E_0}$ is irreducible. Now by level lowering (Ribet [38], [39]), the representation $\rho_l^{E_0}$ is modular of level N_0/N' where N' is any product of primes p for which E_0 has multiplicative reduction at p and $l \mid \nu_p(\Delta_{\min}(E_0))$. Proposition 4.2, together with (35), now immediately show that we can take N_0/N' to be equal to N_1, as given by (36). One might note that N_1 is not necessarily the Serre level of the representation $\rho_l^{E_0}$; for the applications we have in mind, this is not especially important. Congruences (37) and (38) are well-known and follow, essentially, from comparing traces of Frobenii. □

To apply a result of the flavour of Proposition 8.1, one would like to extract as much arithmetic information as possible from the congruences (37) and (38). In the most optimistic of worlds, they can be used to deduce an outright contradiction, at least for l suitably large. The simpler situation, though paradoxically the one which leads to the worse bound for l, is when the newform f (whose existence is guaranteed by Proposition 8.1) is non-rational, i.e when $K_f \neq \mathbb{Q}$. For such a newform f of level N, we know by a result of Kraus [26, Lemme 1] that $a_q(f) \notin \mathbb{Z}$ for some prime $q \leq (N/6)\prod_{p \mid N}(1 + 1/p)$. In the interests of keeping our exposition reasonably self-contained, and since the result will lead to subsequently cleaner statements for our upper bounds, we will sharpen this slightly.

Lemma 8.2. *Let f be a newform of level N. If for some positive integer n, $a_n(f) \notin \mathbb{Z}$, then $a_q(f) \notin \mathbb{Z}$ for some prime $q \nmid N$ with $q \leq B(N)$, where*

$$B(N) = \frac{N}{6} \prod_{p \mid N} \left(1 + \frac{1}{p}\right) - \prod_{p \mid N} \sum_{a=0}^{\nu_p(N)} \phi\left(p^{[a/2]}\right) + 1. \tag{39}$$

Here, $[a/2]$ denotes the greatest integer $\leq a/2$.

Proof. From the classical theory of modular forms, we know (by appealing to Riemann-Roch or simply by integrating over the boundary of a fundamental region) that for a nonzero modular form g of weight k with respect to a congruence subgroup Γ, we have

$$\sum_P \nu_P(g) = \frac{k}{12}\,[\mathrm{SL}(2,\mathbb{Z}) : \pm\Gamma]. \tag{40}$$

Here, the sum is over a fundamental region (for the action of Γ) of the extended upper half plane (one must be careful in defining the multiplicities ν_P as they can be nonintegral at elliptic points and irregular cusps).

Now suppose that $a_n(f) \notin \mathbb{Z}$ for some positive integer n. Then $g = f - \overline{f}$ is nonzero for some conjugate $\overline{f}$ of f (which is also a newform of level N). Since $\Gamma_0(N)$ has no irregular cusps (or since the weight of g is even) and g is a cuspform, we have $\nu_P(g) \geq 1$ at all cusps P. By applying (40), we obtain

$$\nu_{i\infty}(g) \leq \frac{1}{6}[\mathrm{SL}(2,\mathbb{Z}) : \Gamma_0(N)] - \#\{\text{cusps of } \Gamma_0(\mathrm{N})\} + 1.$$

Further, we have the well known formulae

$$[\mathrm{SL}(2,\mathbb{Z}) : \Gamma_0(N)] = N \prod_{p|N} \left(1 + \frac{1}{p}\right)$$

and

$$\#\{\text{cusps of } \Gamma_0(\mathrm{N})\} = \prod_{p|N} \sum_{a=0}^{\nu_p(N)} \phi\left(p^{[a/2]}\right).$$

It follows that $a_n(g) \neq 0$ (equivalently, $a_n(f) \notin \mathbb{Z}$) for some positive integer n satisfying $n \leq B(N)$. Since f is a newform, we may assume that n is prime and also that $n \nmid N$ (since $a_p(f) \in \{\pm 1, 0\}$ for primes $p \mid N$). □

Combining the preceding two results, we have

Proposition 8.3. *Let F be a Klein form of index n and suppose that x_0, y_0 and $z_0 \neq 0$ are integers satisfying (35) where $u_0 \in \mathbb{Z}_{S_F}^*$ and $l > 163$ is prime. If the newform f whose existence is guaranteed by Proposition 8.1 has the property that $[K_f : \mathbb{Q}] > 1$, then there exists an effectively computable absolute constant c such that*

$$\log l < c \left(\prod_{p \in S_F} p^2\right) \sum_{p \in S_F} \log p. \tag{41}$$

Proof. If q is prime, coprime to lN_1, then it follows from (37) and (38) that the (rational) prime l divides either

$$\mathrm{Norm}_{K/\mathbb{Q}}\left(a_q(f) - a_q(E_0)\right) \quad \text{or} \quad \mathrm{Norm}_{K/\mathbb{Q}}\left(a_q(f) \mp (q+1)\right),$$

depending on whether q does or does not divide N_0, respectively. If these norms are nonzero, the Hasse-Weil bounds thus imply that

$$l \leq (1 + \sqrt{q})^{2\,[K_f:\mathbb{Q}]}. \tag{42}$$

Applying Lemma 8.2, we find that $a_q(f) \notin \mathbb{Z}$ (whereby the norms above are nonzero) for some prime q not dividing N_1, with

$$q \leq \frac{N_1}{6} \prod_{p|N_1} \left(1 + \frac{1}{p}\right) - \prod_{p|N_1} \sum_{a=0}^{\nu_p(N_1)} \phi\left(p^{[a/2]}\right) + 1.$$

Since N_1 necessarily divides $2^8 \cdot 3^5 \prod_{p \in S_F \setminus \{2,3\}} p^2$, it follows, after a little work, that

$$1 + \sqrt{q} \leq 2^4 \cdot 3^2 \prod_{p \in S_F \setminus \{2,3\}} \sqrt{p(p+1)}.$$

If $q = l$, this implies (41), as desired. If $q \neq l$, then combining this inequality with the fact that

$$[K_f : \mathbb{Q}] \leq \dim(S_2^{\text{new}}(N_1)) \leq \frac{1}{12} \left(2^8 \cdot 3^5 \prod_{p \in S_F \setminus \{2,3\}} p^2\right) = 2^6 \cdot 3^4 \prod_{p \in S_F \setminus \{2,3\}} p^2,$$

where the second inequality follows from Martin [30, Theorem 2], we may thus conclude that

$$\log l \leq 2^7 \cdot 3^4 \left(\prod_{p \in S_F \setminus \{2,3\}} p^2\right) \log \left(2^4 \cdot 3^2 \prod_{p \in S_F \setminus \{2,3\}} \sqrt{p(p+1)}\right),$$

and hence (41). □

8.1. **The main theorem.** Collecting what we have proved so far, we obtain the principal result of this paper.

Theorem 8.4. *Let F be a Klein form of index $n \in \{2, 3, 4, 5\}$ and S_F denote the set of primes dividing $n\,\Delta_F$, together with the archimedean prime. Suppose that F is irreducible (in case $n = 2$ or 5), or that F contains no linear factors over $\mathbb{Q}[x, y]$ (if $n = 3$), or no linear or quadratic factors over $\mathbb{Q}[x, y]$ (if $n = 4$). If $n = 4$, further assume that δ_4, as given in Section 2, is not an integral square. To F, we associate a corresponding family of Frey-Hellegouarch curves $E_{x,y}$, as well as a companion form $\tilde{F}$, family of curves $\tilde{E}_{x,y}$, and set of primes $\tilde{S}_F$. Let $u_0 \in \mathbb{Z}_{S_F}^*$ and suppose that (35) has a solution in integers x_0, y_0 and (necessarily nonzero) z_0, and prime l. Let N_0 be the conductor of $E_0 = E^{(t)}_{x_0,y_0}$, where t is chosen such that E_0 is semistable outside S_F, and let N_1 be given by (36). Then one of the following occurs:*

(i) there exists an effectively computable absolute constant c such that inequality (41) holds, or

(ii) there exist integers x_1 and y_1 for which

$$\rho_l^{E^{(t)}_{x_1,y_1}} \simeq \rho_l^{E_0}, \;\; N(E^{(t)}_{x_1,y_1}) = N_1 \;\; \text{and} \;\; F(x_1, y_1) \in \mathbb{Z}_{S_F}^*, \;\; \text{or}$$

(iii) there exist integers x_1 and y_1 for which

$$\rho_l^{\tilde{E}^{(t)}_{x_1,y_1}} \simeq \rho_l^{E_0}, \;\; N(\tilde{E}^{(t)}_{x_1,y_1}) = N_1 \;\; \text{and} \;\; \tilde{F}(x_1, y_1) \in \mathbb{Z}_{\tilde{S}_F}^*.$$

Proof. Let x_0, y_0 and z_0 be integers and l a prime, satisfying equation (2). Without loss of generality, we may suppose that $l > 163$. Appealing to Proposition 8.1, we deduce the existence of a weight 2 newform f, of trivial character and level $N_1 \in \mathbb{Z}^*_{S_F}$ given by (36), satisfying $\rho_l^{E_0} \simeq \rho_l^f$. In case f is not rational (i.e. $[K_f : \mathbb{Q}] > 1$), inequality (41) is immediate from Proposition 8.3. If, on the other hand, f is rational, let us denote by E the elliptic curve corresponding to f via the Eichler-Shimura relation.

If $E_0[n] \not\simeq E[n]$ (as Galois modules), we may apply Proposition 7.4 (where we appeal, for the irreducibility conditions, to Propositions 6.8 and 6.10, and, for the ramification condition, to Proposition 6.11). Together with the Hasse-Weil bounds, this result immediately implies (a stronger version of) inequality (41).

If $E_0[n] \simeq E[n]$, then, from Theorem 6.6, E is necessarily isomorphic over $\mathbb{Q}$ to either $E^{(t)}_{x_1,y_1}$ or $\tilde{E}^{(t)}_{x_1,y_1}$, for some integers x_1 and y_1. The remaining parts of (ii) and (iii) therefore follow from Lemma 6.7. □

We note that sometimes additional methods may be applied to show that we cannot have $\rho_l^{E^{(t)}_{x_1,y_1}} \simeq \rho_l^{E_0}$ or $\rho_l^{\tilde{E}^{(t)}_{x_1,y_1}} \simeq \rho_l^{E_0}$ and thereby treat the corresponding superelliptic equations (for large enough prime exponents). An example where we can carry this out through a simple comparison of images of inertia is given in Remark 10.3. Other arguments involving elliptic curves with complex multipication can be found in Section 13.

An immediate corollary of Theorem 8.4, from which Theorem 1.1 is a direct consequence (in case $n = 2$), is the following

Corollary 8.5. *Let $n \in \{2, 3, 4, 5\}$ and suppose that $F(x, y) \in \mathbb{Z}[x, y]$ is an irreducible Klein form of index n with companion form $\tilde{F}$. If $n = 4$, suppose also that δ_4 is not the square of an integer. If the Thue-Mahler equations*

$$F(x, y) \in \mathbb{Z}^*_{S_F} \quad \text{and} \quad \tilde{F}(x, y) \in \mathbb{Z}^*_{\tilde{S}_F}$$

each have no solutions (in coprime integers x, y), then the generalized superelliptic equation

$$F(x, y) = z^l, \quad \gcd(x, y) = 1$$

has at most finitely many solutions in integers x, y, z and integer $l \geq \max\left\{2, \frac{30-7n}{6-n}\right\}$.

We should remind the reader that $F = \tilde{F}$ in case $n = 2$ or 4. The remainder of this paper is devoted to exploring applications of Theorem 8.4 and Corollary 8.5. In particular, we will attempt to demonstrate that their attendant hypotheses are frequently, perhaps usually, met.

9. A cubic family

In Section 12, we will sketch a heuristic indicating that Theorem 1.1 is applicable to "almost all" cubic forms. Our goal in this section is to provide an explicit infinite family of cubic forms F which satisfy the hypotheses of Theorem 1.1. To achieve this, we will exhibit a family of forms for which the corresponding Thue-Mahler equations (3) possess no solutions whatsoever. To our knowledge, no infinite family of Thue-Mahler equations with corresponding set of (noninert) primes S of unbounded cardinality, has hitherto been completely solved (though treating families of Thue equations or inequalities has become relatively commonplace; see e.g. [28] and [55]).

Let us define

$$F_{a,b}(x,y) = bx^3 - ax^2y - (a+3b)\,xy^2 - by^3, \tag{43}$$

where a and b are coprime nonzero integers, so that

$$\Delta_{F_{a,b}} = (a^2 + 3ab + 9b^2)^2. \tag{44}$$

It follows that a cubic field corresponding to a root of $F_{a,b}(x,1) = 0$ is Galois and cyclic; indeed all cyclic cubic fields arise in this fashion (see e.g. [57]). The existence of a nontrivial automorphism for these forms will prove helpful in the sequel; to be specific, we have

$$F_{a,b}(x,y) = F_{a,b}(-x-y,x) = F_{a,b}(y,-x-y). \tag{45}$$

An advantage of working with forms of this shape is that the method of Thue-Siegel may be used to deduce good "irrationality measures" for the roots of the equation $F_{a,b}(x,1) = 0$. Such an approach to solving corresponding Thue inequalities has been carried out in [28], [55] and [56]. In case $b = 1$, the forms $F_{a,1}(x,y)$ have been termed the *simplest cubic forms* (with corresponding *simplest cubic fields*). For our purposes, since we wish to find forms F for which equation (3) is insoluble, it is clear that we cannot take $b = 2^k$ for any non-negative integer k (or else $F(1,0) \in \mathbb{Z}_{S_F}^*$). Choosing $b = 3$ (we can obtain a similar result for any fixed $b \neq 2^k$), we prove the following

Theorem 9.1. *Let a be an integer such that $a^2 + 9a + 81$ is squarefree and S the set of primes dividing $2\,(a^2 + 9a + 81)$, together with an Archimedean prime. If neither a nor $-a - 9$ is in the set $\{-4, 8, 22, 31\}$, then the Thue-Mahler equation*

$$3x^3 - ax^2y - (a+9)xy^2 - 3y^3 \in \mathbb{Z}_S^* \tag{46}$$

has no solutions in integers. In these exceptional cases, the solutions to (46) are as follows

a	$\pm(x,y)$	$F_{a,3}(x,y)$
-40	$(-25,2),(2,23),(23,-25)$	± 1
-31	$(-19,2),(2,17),(17,-19)$	± 109
-17	$(-5,1),(1,4),(4,-5)$	± 7
-5	$(-2,1),(1,-2),(1,1)$	± 1
-4	$(-1,-1),(-1,2),(2,-1)$	± 1
8	$(-4,-1),(-1,5),(5,-4)$	± 7
22	$(-17,-2),(-2,19),(19,-17)$	± 109
31	$(-23,-2),(-2,25),(25,-23)$	± 1

Applying Theorem 1.1 (since, as we shall see, $F_{a,3}(x,y)$ is irreducible in $\mathbb{Q}[x,y]$) immediately yields Theorem 1.2

It is worth noting that the arguments employed in [28], [55] and [56] do not lead to Theorem 9.1 directly. Indeed, in the course of proving this result, one encounters Thue inequalities of the shape

$$|F_{a,b}(x,y)| \leq k, \tag{47}$$

with k larger than can be handled by direct application of the techniques of, say, [56] (i.e of size $\Delta_{F_{a,b}}^{1/2}$ rather than $\Delta_{F_{a,b}}^{1/4}$).

9.1. **From Thue-Mahler equations to Thue equations.** One property of the forms $F_{a,b}(x,y)$, though a simple observation, is key to the proof of Theorem 9.1. From (44), $\Delta_{F_{a,b}} = \delta^2$, where $\delta = a^2 + 3ab + 9b^2$. We have

Lemma 9.2. *Let a and b be coprime integers and p prime with $p \nmid 3b$. If 3 does not divide $\nu_p(\delta)$, then, for every $x, y \in \mathbb{Z}$ with $F_{a,b}(x,y) \equiv 0 \pmod{p^{\nu_p(\delta)+1}}$, we have $(x,y) \equiv (0,0) \pmod{p}$.*

Proof. Using the fact that $p \nmid b$ and the homogeneity of $F_{a,b}$, it obviously suffices to prove that $f(x) = F_{a,b}(x,1) \equiv 0 \pmod{p^{\nu_p(\delta)+1}}$ has no solution with $x \in \mathbb{Z}$. To see this, define $h(x) = f''(x)/2 = -a + 3bx$, so that we have the relation

$$27b^2 f(x) + (2a + 3b + 3h(x))\delta = h(x)^3. \tag{48}$$

Together with the fact that $\mathrm{Res}(2a + 3b, \delta) = 27$, if there exists an integer x with $p^{\nu_p(\delta)+1} \mid f(x)$, then necessarily $p^{\nu_p(\delta)+1} \mid \delta$. This proves the lemma. □

Remark 9.3. If $3 \mid \nu_p(\delta)$, then it can easily happen that $f(x)$ has roots in $\mathbb{Z}_p$. Take e.g. $b = 3$ and $a = 848$. Then $\delta = 7^3 \cdot 13 \cdot 163$ and $f(x)$ had three roots in $\mathbb{Z}_7$ (in the cubic number field defined by $f(x)$, the prime 7 splits completely).

Remark 9.4. If $\nu_p(\delta) = 1$, then $f(x)$ is a translate of an Eisenstein polynomial at p (and hence irreducible over $\mathbb{Q}$). Indeed, let $t \in \mathbb{Z}$ satisfy $f(t) \equiv f'(t) \equiv 0 \pmod{p}$ and write

$$f(y + t) = by^3 + h(t)y^2 + f'(t)y + f(t).$$

By (48), we have $p \mid h(t)$ whereby, from Lemma 9.2, $p^2 \nmid f(t)$ and hence $f(y+t)$ is Eisenstein in y at p.

We will suppose, here and henceforth, that $b = 3$ and, since

$$F_{a,3}(x,y) = F_{-a-9,3}(-y,-x)$$

that, without loss of generality, $a \geq -4$. We will also assume $a^2 + 9a + 81$ to be squarefree. Lemma 9.2 thus enables us to conclude, for coprime x, y, and p a divisor of $a^2 + 9a + 81$, that we have

$$F_{a,3}(x,y) \not\equiv 0 \pmod{p^2}.$$

Since, further, $F_{a,3}(x,y) \equiv 1 \pmod 2$, it follows that $F_{a,3}(x,y) \in \mathbb{Z}_S^*$ is equivalent to $F_{a,3}(x,y)$ dividing $a^2 + 9a + 81$. In other words, our family of Thue-Mahler equations has been reduced to a number of families of Thue equations. Theorem 9.1 will now follow from resolving the equations

$$3x^3 - ax^2y - (a+9)xy^2 - 3y^3 = k \tag{49}$$

where $k \mid a^2 + 9a + 81$.

Before we proceed, it is worth noting that the restriction to values of a for which $a^2 + 9a + 81$ is squarefree is not too severe. Indeed, as a simple application of sieving, together with the observation that $a^2 + 9a + 81 \equiv 0 \pmod{p}$ is solvable precisely when the Legendre symbol $\left(\frac{-3}{p}\right) = 1$, we have

$$\lim_{n\to\infty} \frac{1}{n} \left| 1 \leq a \leq n \ : \ a^2 + 9a + 81 \text{ is squarefree } \right| = \frac{2}{3} \prod_{p \equiv 1 \pmod 3} \left(1 - \frac{2}{p^2}\right),$$

where the product is over primes p. In particular, $a^2 + 9a + 81$ is squarefree for rather more than 90% of $a \equiv \pm 1 \pmod 3$.

9.2. **Solutions as convergents.** In this subsection, we will show that solutions to (49) correspond to convergents in the infinite simple continued fraction expansions to the roots of $F_{a,3}(x,1)=0$. To simplify our subsequent arguments somewhat, we first apply standard computer algebra packages (either Pari/GP or Magma have routines for solving Thue equations) to treat small values of a:

Proposition 9.5. *If $-4 \le a \le 5000$ and there exist integers x, y and k satisfying (49) with $k \ge 1$ and $k \mid a^2+9a+81$, then we have a, x, y and k as follows*

a	k	(x,y)
-4	1	$(-1,-1),(-1,2),(2,-1)$
8	7	$(-4,-1),(-1,5),(5,-4)$
22	109	$(-17,-2),(-2,19),(19,-17)$
31	1	$(-23,-2),(-2,25),(25,-23)$
56	3721	$(-13,-1),(-1,14),(14,-13)$

Remark 9.6. The computation indicated here takes a surprisingly long time in Magma. A (much) faster way to carry out such a calculation, at least for a of moderate size, would be to appeal to Corollary 9.10 of subsection 9.3.

We will assume, from now on, that $a > 5000$ and that x and y are coprime nonzero integers with

$$y > 0, \;\; -y/2 < x < y \;\text{ and }\; xy(x+y) \not\equiv 0 \pmod 3. \tag{50}$$

Equivalently, (x,y) are of the form

$$(x,y) = (y-3j, y), \;\; j = 1, \ldots, \left[\frac{y-1}{2}\right], \;\; y \equiv \pm 1 \pmod 3, \; \gcd(y,j) = 1.$$

To see that this is without loss of generality, note that, for a cubic form F, we have $F(-x,-y) = -F(x,y)$, whereby we may assume $y > 0$. Appealing to (45) and noting that $F_{a,3}(1,1) = -2a-9$ fails to divide $a^2+9a+81$ for a coprime to 3, leads to the desired inequalities (this essentially follows along the lines of Lemma 10 (b) of [28]). Since we suppose that x, y and a satisfy (49) where $k \mid a^2+9a+81$, necessarily $xy(x+y)$ is coprime to 3 and thus

$$\left|\theta_1 - \frac{x}{y}\right| \left|\theta_2 - \frac{x}{y}\right| \left|\theta_3 - \frac{x}{y}\right| = \frac{|k|}{3y^3}, \tag{51}$$

where $\theta_1 < \theta_2 < \theta_3$ are roots of $F_{a,3}(x,1) = 0$. Via Newton's method, we have, for $a \ge 11$,

$$-1 - 3/a < \theta_1 < -1,$$

$$-3/a < \theta_2 < -3/a + 18/a^2,$$

and

$$a/3 + 1 < \theta_3 < a/3 + 1 + 4/a.$$

We will actually study θ_2 much more carefully later. For our purposes, however, since (50) implies that $-1/2 < x/y < 1$, the above inequalities are enough to yield

$$\left|\theta_2 - \frac{x}{y}\right| < \frac{2\,|k|}{a\,y^3}. \tag{52}$$

Our goal at this stage is to show that x/y is a convergent in the continued fraction expansion to θ_2. From classical theory, this is certainly the case if we have

$$\left|\theta_2 - \frac{x}{y}\right| < \frac{1}{2y^2}.$$

If we have, say, $|k| \leq 10a$, then it follows that x/y is a convergent to θ_2 if $y \geq 40$. For the values of $y < 40$, there are precisely 171 corresponding pairs (x, y) satisfying (50). For each such pair, $F_{a,3}(x, y)$ is a linear polynomial in a with integer coefficients; computing its resultant with $a^2 + 9a + 81$, we find that

$$\mathrm{Res}_a(F_{a,3}(x, y), a^2 + 9a + 81) = 3^5\, R^3,$$

for some positive integer $R = R(x, y)$. We thus have that $F_{a,3}(x, y)$ divides R (again using that $a^2 + 9a + 81$ is squarefree). For the cases of x, y under consideration, since we assume $a \geq -4$, this can occur only for

(x, y)	$F_{a,3}(x, y)$	R
$(-1, 5)$	$20a - 153$	7
$(-2, 19)$	$646a - 14103$	109
$(-2, 25)$	$1150a - 35649$	193

where $F_{a,3}(x, y)$ divides R for $a = 8, 22$ and 31, respectively.

Let us now suppose that $|k| > 10a$, with $a > 5000$ and $y \geq 40$. To handle these larger values of k requires a new approach. We appeal to (a special case of) work of Hoshi and Miyake [24] (see also [35]):

Theorem 9.7. *(Theorem 5.4 of* [24]*) If m and n are rational numbers, then the splitting fields over $\mathbb{Q}$ of the polynomials $x^3 - mx^2 - (m+3)x - 1$ and $x^3 - nx^2 - (n+3)x - 1$ coincide precisely when there exists $z \in \mathbb{Q}$ such that either*

$$n = \frac{m(z^3 - 3z - 1) - 9z(z+1)}{mz(z+1) + z^3 + 3z^2 - 1} \quad \text{or} \quad n = -\frac{m(z^3 + 3z^2 - 1) + 3(z^3 - 3z - 1)}{mz(z+1) + z^3 + 3z^2 - 1}.$$

Writing $m = a/3$ and $n = a_1/3$ for integers a and a_1, and putting $z = y/x$ for x and y coprime integers, we have

Corollary 9.8. *If a and a_1 are integers, then the splitting fields over $\mathbb{Q}$ of the polynomials $F_{a,3}(x, 1)$ and $F_{a_1,3}(x, 1)$ coincide precisely when there exist coprime integers x and y such that*

$$a_1 = a + \frac{(a^2 + 9a + 81)xy(x+y)}{F_{a,3}(x, y)} \quad \text{or} \quad -a_1 - 9 = a + \frac{(a^2 + 9a + 81)xy(x+y)}{F_{a,3}(x, y)}.$$

Since, for a putative solution to equation (46), $F_{a,3}(x, y)$ divides $a^2 + 9a + 81$, it follows that

$$a_1 = a + \frac{(a^2 + 9a + 81)\, x\, y\, (x+y)}{F_{a,3}(x, y)}$$

is an integer. From the fact that $a^2 + 9a + 81$ is assumed to be squarefree, we have that the discriminant of the associated cubic field is also $(a^2 + 9a + 81)^2$. This is a consequence of the fact that every prime dividing a squarefree $a^2 + 9a + 81$ necessarily ramifies in this field. We may thus conclude that

$$a^2 + 9a + 81 \mid a_1^2 + 9a_1 + 81,$$

whereby

$$a^2 + 9a + 81 \mid (a - a_1)(a + a_1 + 9).$$

Writing $a^2+9a+81 = F_{a,3}(x,y)\cdot M$ and noting that $\gcd(F_{a,3}(x,y), xy(x+y))$ divides 3 (and hence is equal to 1), we thus have

$$2a+9+Mxy(x+y) \equiv 0 \pmod{F_{a,3}(x,y)}. \tag{53}$$

Since $xy(x+y)$ is even, the left hand side of (53) is necessarily nonzero, and so

$$|2a+9+Mxy(x+y)| \geq |F_{a,3}(x,y)|\,. \tag{54}$$

We will use this inequality to show that, in this case as well, x/y is a convergent to θ_2.

To begin, notice that if $x>0$, then

$$|F_{a,3}(x,y)| = ax^2y+(a+9)xy^2+3(y^3-x^3) > axy(x+y),$$

while

$$|Mxy(x+y)| < \frac{a^2+9a+81}{10a}xy(x+y),$$

contradicting (54) and the assumption $a>5000$. We may thus assume the inequality $-y/2<x<0$, so that

$$|F_{a,3}(x,y)| \geq a|x|y^2-ax^2y-3(y^3-x^3) > a|x|y^2-ax^2y-27y^3/8. \tag{55}$$

We will treat small and large values of $|x|$ separately. If $|x|\leq 54$, then either $y\leq 20|x|$ (which leads to a finite set of pairs (x,y) which we treat as previously; we find no new solutions to (49)) or we have $y>20|x|$ and hence, with the first inequality in (55),

$$|F_{a,3}(x,y)| > \frac{19}{20}\,a|x|y^2-3(y^3-x^3).$$

If we further suppose that $y<a|x|/4$, then $-1/2<x/y<-4/a$ and so, via our inequalities upon the θ_i, $F_{a,3}(x,y)$ is positive. We have

$$F_{a,3}(x,y) > \frac{1}{5}\,a|x|y^2-3|x|^3 > \frac{3}{16}\,ay^2,$$

where we have used that $y\geq 40$, $-54<x<0$ and $a>5000$. It follows from (54) that

$$\max\{2a+9, |M||x|y(x+y)\} \geq \frac{3}{16}\,ay^2$$

and hence

$$\frac{3}{16}\,ay^2 \leq |M||x|y(x+y) \leq 54\,|M|\,y^2,$$

i.e. $|M|\geq a/288$. Recalling that $F_{a,3}(x,y)M$ divides $a^2+9a+81$, we thus have

$$\frac{a}{288} \leq |M| < \frac{16(a^2+9a+81)}{3a\,y^2},$$

contradicting $y\geq 40$ and $a>5000$.

For the values of x with $-54<x<0$, we may thus suppose that $y\geq a|x|/4$. For $x<0$ fixed, it is a routine exercise in calculus to show that $F_{a,3}(x,y)$ is monotone decreasing as a function of y in the interval $[a|x|/4,\infty)$. Writing $x=-x_0$ where $x_0\in\{1,2,4,\ldots,53\}$ and setting $a=3a_0+i$ for $i\in\{1,2\}$,

$$F_{a,3}(-x_0, a_0x_0+s) = c_2a_0^2+c_1a_0+c_0,$$

where

$$c_2=(6+i)x_0^3-3x_0^2s,\ c_1=-ix_0^3+(15s+2is)x_0^2-6s^2x_0$$

and

$$c_0=-3x_0^3-isx_0^2+(is^2+9s^2)x_0-3s^3.$$

If $x_0 \geq 4$, it is easy to check that

$$F_{a,3}\left(-x_0, a_0x_0 + 2x_0 + [ix_0/3]\right) > a^2 + 9a + 81$$

and

$$F_{a,3}\left(-x_0, a_0x_0 + 2x_0 + [ix_0/3] + 1\right) < -(a^2 + 9a + 81),$$

and hence we may assume that $x_0 \in \{1, 2\}$. In case $x_0 = 2$, we have that

$$F_{a,3}\left(-2, 2a_0 + i + 2\right) = (24 - 4i)a_0^2 + (-4i^2 + 20i + 72)a_0 + (-i^3 + 4i^2 + 36i + 24),$$

and

$$F_{a,3}\left(-2, 2a_0 + i + 5\right) = (-4i - 12)a_0^2 + (-4i^2 - 28i)a_0 + (-i^3 - 11i^2 - 15i + 51),$$

and so

$$F_{a,3}\left(-2, 2a_0 + i + 2\right) > a^2 + 9a + 81$$

and

$$F_{a,3}\left(-2, 2a_0 + i + 5\right) < -(a^2 + 9a + 81).$$

We have

$$F_{a,3}\left(-2, 2a_0 + i + 3\right) = (12 - 4i)a_0^2 + (-4i^2 + 4i + 72)a_0 + (-i^3 - i^2 + 33i + 57)$$

and

$$F_{a,3}\left(-2, 2a_0 + i + 4\right) = -4ia_0^2 + (-4i^2 - 12i + 48)a_0 + (-i^3 - 6i^2 + 16i + 72),$$

where, since x and y are coprime, we necessarily have $i = 2$ and $i = 1$, respectively. Computing

$$\mathrm{Res}_{a_0}(4a_0^2 + 64a_0 + 111, 9a_0^2 + 39a_0 + 103) = 109^3$$

and

$$\mathrm{Res}_{a_0}(4a_0^2 - 32a_0 - 81, 9a_0^2 + 33a_0 + 91) = 109^3,$$

it follows that $4a_0^2 + 64a_0 + 111$ or $4a_0^2 - 32a_0 - 81$ divides 109, a contradiction.

In case $x_0 = 1$, arguing as before we find that

$$|F_{a,3}\left(-1, a_0 + j\right)| > a^2 + 9a + 81,$$

provided $j \leq -1$ or $j \geq 6$. We compute, for each $j \in \{0, 1, 2, 3, 4, 5\}$,

$$\mathrm{Res}_{a_0}((6 + i - 3j)a_0^2 - (j^2 - (15 + 2i)j + i)a_0 + (9 + i)j^2 - ij - 3, (3a_0 + i)^2).$$

In each case, we find this to be of the form T^3 where $T \leq 91$, again contradicting $a = 3a_0 + i > 5000$.

We may thus assume $|x| \geq 54$. It follows from (53) that

$$|Mxy(x + y)| \geq |F_{a,3}(x, y)| - 2a - 9 > a|x|y^2 - ax^2y - 27y^3/8 - 2a - 9$$

and so

$$|Mxy(x + y)| \geq a|x|y^2/2 - 27y^3/8 - 2a - 9.$$

If $y < 2|x|a/27$ then $a|x|y^2/2 - 27y^3/8 > a|x|y^2/4$ and so

$$|Mxy(x + y)| \geq a|x|y^2/4 - 2a - 9 > a|x|y^2/5,$$

since $y > 40$. On the other hand, $|M| < a/10 + 9 + 81/a$, whereby

$$|Mxy(x + y)| < a|x|y^2/9.$$

We thus have $y \geq 2|x|a/27 \geq 4a$ and so x/y is a convergent to θ_2.

9.3. **Applying the method of Thue-Siegel.** Our work in the preceding subsection led to the conclusion that x/y is a convergent in the continued fraction expansion to θ_2. To reduce this to a finite problem, we appeal to an irrationality measure for θ_2, derived from the method of Thue-Siegel:

Theorem 9.9. *(Theorem 2.9 of* [56]*) Let a and b be positive integers with $a \geq 31b^4$ and suppose that $F_{a,b}(\theta, 1) = 0$ with $-1 < \theta < 0$. Then if p and q are integers with $q \geq \frac{a+\frac{3}{2}b}{9.04}$, we have*

$$\left|\theta - \frac{p}{q}\right| > \frac{1}{c(a,b)q^{1+\lambda}},$$

where

$$\lambda = \frac{\log\left(\sqrt{a^2+3ab+9b^2}\right) + 0.83}{\log\left(a+\frac{3}{2}b\right) - 2\log b - 1.3} < 2$$

and

$$c(a,b) = 17.43\sqrt{a^2+3ab+9b^2}\,\left(\frac{2.47}{b^2}\right)^{\lambda}.$$

With care, we can improve this result slightly, but it is adequate for our purposes. It implies the following.

Corollary 9.10. *(Theorem 2.10 of* [56]*) If a and b are positive integers with $a \geq 31\,b^4$, k is a positive integer, and if x and y are nonzero integers satisfying (47), with $-1/2 < x/y \leq 1$ and $y \geq \max\left\{\frac{a+\frac{3}{2}b}{9.04}, \sqrt[3]{\frac{ak}{1.99b^2}}\right\}$, then*

$$y^{2-\lambda} < 18.34\,\left(\frac{2.47}{b^2}\right)^{\lambda} k,$$

where λ is as in Theorem 9.9.

Applying Corollary 9.10 with $k \leq a^2 + 9a + 81$, together with the fact that λ in Theorem 9.9 is decreasing monotonically for suitably large values of a, we may readily compute that

$$\begin{array}{ll} y < a^{15} & \text{if } a > 5000, \\ y < a^{8.6} & \text{if } a > 10^4, \\ y < a^{4.6} & \text{if } a > 10^5, \\ y < a^{3.6} & \text{if } a > 10^6. \end{array}$$

9.4. **Continued fraction expansions to θ_2.** From our preceding arguments, $x/y = p_n/q_n$ for some convergent in the simple continued fraction expansion to θ_2, where q_n is bounded above by the upper bound for y at the end of the previous subsection. For $5000 < a \leq 10^6$, we compute the continued fraction expansion for θ_2 using Pari/GP and verify that $F_{a,3}(p_n, q_n)$ fails to divide $a^2 + 9a + 81$ in each case. To treat the remaining values of $a > 10^6$ (and hence $q_n < a^{3.6}$), we begin by noting that we can explicitly compute the first few terms, in the simple continued fraction expansion

$$\theta_2 = [a_0; a_1, a_2, \ldots].$$

Specifically, applying Newton's method, we find that, at least for $a \geq 86$,

$$(a_0, a_1, a_2, a_3, a_4, a_5) = \begin{cases} \left(-1, 1, \left[\frac{a+2}{3}\right], 2, 1, \left[\frac{a-34}{54}\right]\right) & \text{if } a \equiv 1 \pmod 3, \\ \left(-1, 1, \left[\frac{a+2}{3}\right], 1, 2, \left[\frac{a-14}{54}\right]\right) & \text{if } a \equiv 2 \pmod 3. \end{cases}$$

Here $[x]$ denotes the greatest integer $\leq x$. It follows that the first few convergents p_n/q_n to θ_2 are given by

$$-\frac{1}{1}, \frac{0}{1}, -\frac{1}{1+\left[\frac{a+2}{3}\right]}, -\frac{2}{3+2\left[\frac{a+2}{3}\right]}, -\frac{3}{4+3\left[\frac{a+2}{3}\right]}$$

and

$$-\frac{1}{1}, \frac{0}{1}, -\frac{1}{1+\left[\frac{a+2}{3}\right]}, -\frac{1}{2+\left[\frac{a+2}{3}\right]}, -\frac{3}{5+3\left[\frac{a+2}{3}\right]},$$

for $a \equiv 1 \pmod 3$ and $a \equiv 2 \pmod 3$, respectively.

If we suppose that x/y is a convergent to θ_2 for which $F_{a,3}(x,y)$ divides $a^2 + 9a + 81$, then, since this latter quantity is squarefree and coprime to 3, we have that both $(x,y) = \pm(p_n, q_n)$ for some n, whereby $F_{a,3}(p_n, q_n)$ divides $a^2 + 9a + 81$, and that

$$p_n q_n (p_n + q_n) \not\equiv 0 \pmod 3. \tag{56}$$

We note that condition (56) is not satisfied for $n \in \{0, 1, 4\}$. Writing $a = 3j + s$ for $s \in \{1, 2\}$, we have that

$$F_{3j+s,3}(p_2, q_2) = F_{3j+s,3}(-1, j+2) = sj^2 + (3s+6)j + 2s + 9$$

and hence $F_{a,3}(p_2, q_2) \mid a^2 + 9a + 81$ precisely when

$$sj^2 + (3s+6)j + 2s + 9 \quad \text{divides} \quad 9j^2 + (6s+27)j + s^2 + 9s + 81.$$

A short calculation shows that this does not occur. Specifically, we have that

$$(13 - 4s)\, F_{3j+1,3}(p_2, q_2) > a^2 + 9a + 81$$

and hence necessarily $a^2 + 9a + 81 = M\, F_{3j+1,3}(p_2, q_2)$ for some integer M with $1 \leq M \leq 12 - 4s$, whereby

$$(9 - Ms)\, j^2 + (6s + 27 - (3s+6)M)\, j + s^2 + 9s + 81 - (2s+9)\, M = 0.$$

This equation has no rational roots for the given choices of M and s. Since

$$F_{3j+1,3}(p_3, q_3) = F_{3j+1,3}(-2, 2j+5) = -4j^2 + 32j + 81$$

and

$$F_{3j+2,3}(p_3, q_3) = F_{3j+1,3}(-1, j+3) = -j^2 + j + 9,$$

we can argue similarly to conclude that $x/y = p_n/q_n$ with $n \geq 5$. To show that, in fact, $n \geq 6$, note that

$$\frac{p_5}{q_5} = \begin{cases} \dfrac{-3\left[\frac{a-34}{54}\right]-2}{3\left[\frac{a-34}{54}\right]\left[\frac{a+2}{3}\right]+2\left[\frac{a+2}{3}\right]+4\left[\frac{a-34}{54}\right]+3} & \text{if } a \equiv 1 \pmod 3, \\ \dfrac{-3\left[\frac{a-14}{54}\right]-1}{3\left[\frac{a-14}{54}\right]\left[\frac{a+2}{3}\right]+\left[\frac{a+2}{3}\right]+5\left[\frac{a-14}{54}\right]+2} & \text{if } a \equiv 2 \pmod 3. \end{cases}$$

Writing $a = 54k + t$ for $k \in \mathbb{Z}$, $1 \leq t \leq 53$ and $\gcd(t, 3) = 1$, and arguing as before, we find that $F_{a,3}(p_5, q_5)$ fails to divide $a^2 + 9a + 81$, in every case. It follows that $y > q_5$ and hence, since we assume $a > 10^6$, $y > a^2/55$.

To finish our proof, we need to handle values of y satisfying

$$\frac{1}{55}\, a^2 < y < a^{3.6} \quad \text{for} \quad a > 10^6. \tag{57}$$

If we continue further with our examination of the infinite simple continued fraction for θ_2, perhaps unsurprisingly, complications arise. Indeed, the next few partial quotients, for suitably large a, depend on the value of a modulo 54. Specifically,

we have that if $a \equiv t \pmod{54}$ with $1 \leq t \leq 53$ and $\gcd(t,3) = 1$, then the sequence $a_6, a_7, \ldots$ begins with a sequence α_t of terms $a_6, a_7, \ldots, a_{k(t)}$, followed by $a_{k(t)+1} = \left[\frac{a-\beta_t}{3402}\right]$, where

t	α_t	β_t	t	α_t	β_t
1	$2,3,2,1,3,1$	2701	28	$1,14,2,3$	1000
2	$1,3,1,2,3,2$	1514	29	$3,2,14,1$	3215
4	$2,26,1,1$	1732	31	$1,107$	31
5	$1,5,2,1,4,1$	2813	32	$2,1,11,3$	1112
7	$1,1,4,1,8,1$	3085	34	$21,1,1,2$	1384
8	$1,8,1,4,1,1$	1898	35	$2,1,1,21$	197
10	$1,1,1,1,20,1$	3250	37	$9,1,4,2$	1549
11	$1,20,1,1,1,1$	2063	38	$2,4,1,9$	362
13	$1,1,1,11,2,1$	2335	40	$6,2,1,5$	634
14	$107,1$	3416	41	$1,1,26,2$	1715
16	$1,2,2,15$	232	43	$4,1,2,3,1,1$	1933
17	$15,2,2,1$	2447	44	$1,1,3,2,1,4$	746
19	$1,3,3,8$	451	46	$3,1,2,1,1,1,1,1$	2152
20	$8,3,3,1$	2612	47	$1,1,1,1,1,2,1,3$	911
22	$1,4,1,2,5,1$	2884	49	$3,11,1,2$	1183
23	$5,1,2,6$	563	50	$1,2,11,1,1,1$	2264
25	$1,7,3,4$	835	52	$2,1,1,1,2,1,2,1$	2536
26	$4,3,7,1$	2996	53	$1,2,1,2,1,1,1,2$	1295

These are valid for $a \geq 6818$. It is worth noting here that if $t \equiv 1, 7 \pmod 9$ then α_t is just the reverse of α_{t+1}, while if $t \equiv \pm 4 \pmod 9$, the same is true of α_t and $\alpha_{t\mp 17}$.

Even this knowledge leads us only, after much work, to the conclusion that $y \gg a^3$. Indeed, we find ourselves confronted with rather dramatic combinatorial explosion. One way to overcome these technical difficulties is to appeal to an argument of Wakabayashi, introduced in [54]. Instead of relying upon a continued fraction expansion to θ_2 with integer partial quotients, following Wakabayashi (see section 8 of [56] for details), we compute one with rational partial quotients. Let us suppose that, for a given real number ψ, we choose a rational numbers k_0 such that $k_0 < \psi < k_0 + 1$, and define recursively for $i \geq 1$, $\psi_i = 1/(\psi_{i-1} - k_{i-1})$, where $\psi_0 = \psi$ and, in each case, rational k_i are chosen with $k_i < \psi_i < k_i + 1$. We call

$$\psi = [k_0; k_1, k_2, \ldots]$$

a *continued fraction expansion with rational partial quotients for* ψ. If we further define convergents p_n/q_n via

$$\begin{cases} p_0 = 1, & p_1 = k_0, & p_{n+1} = k_n p_n + p_{n-1} \ \ (n \geq 1), \\ q_0 = 0, & q_1 = 1, & q_{n+1} = k_n q_n + q_{n-1} \ \ (n \geq 1), \end{cases}$$

then if $k_n \geq 1$ for $n \geq 1$, necessarily $q_n \to \infty$ and p_n/q_n converges to ψ. We have the following

Proposition 9.11. *Let ψ be a real number and p_n/q_n $(n = 1, 2, \ldots)$ be the convergents defined by a continued fraction expansion with rational partial quotients for ψ. Suppose, for $n \geq 0$, that d_n is a positive rational number with the property that*

$d_n p_n$ and $d_n q_n$ are integers. If, for a given positive integer n, there exist integers p and q satisfying

$$\frac{q_n}{d_{n-1}} \le q < \frac{q_{n+1}}{d_n} \tag{58}$$

and

$$\left|\psi - \frac{p}{q}\right| < \frac{1}{d_n\left(d_{n-1}+d_{n+1}\right)q^2}, \tag{59}$$

then we may conclude that $p/q = p_n/q_n$.

Proof. This is Theorem 5 of [54], together with the observation that we may choose positive rational numbers d_n (rather than positive integers, as Wakabayashi does) with the property that $d_n p_n$ and $d_n q_n$ are integers, and reach an identical conclusion. □

In our case, we may take $\theta_2 = [k_0; k_1, k_2, \ldots]$, where

$$k_0 = -1,\ k_1 = 1,\ k_2 = \tfrac{a}{3}+1,\ k_3 = \tfrac{a}{6}+\tfrac{3}{4},\ k_4 = \tfrac{8a}{21}+\tfrac{12}{7},\ k_5 = \tfrac{49a}{288}+\tfrac{49}{64}$$

and $k_6 = \frac{384a}{1001}+\frac{1728}{1001}$. Corresponding convergents are

$$p_0 = 1,\ p_1 = -1,\ p_2 = 0,\ p_3 = -1,\ p_4 = -\tfrac{a}{6}-\tfrac{3}{4},\ p_5 = -\tfrac{4a^2}{63}-\tfrac{4a}{7}-\tfrac{16}{7},$$

$$p_6 = -\tfrac{7a^3}{648}-\tfrac{7a^2}{48}-\tfrac{143a}{144}-\tfrac{5}{2},\ p_7 = -\tfrac{16a^4}{3861}-\tfrac{32a^3}{429}-\tfrac{896a^2}{1287}-\tfrac{464a}{143}-\tfrac{944}{143},$$

$$q_0 = 0,\ q_1 = 1,\ q_2 = 1,\ q_3 = \tfrac{a}{3}+2,\ q_4 = \tfrac{a^2}{18}+\tfrac{7a}{12}+\tfrac{5}{2},$$

$$q_5 = \tfrac{4a^3}{189}+\tfrac{20a^2}{63}+\tfrac{16a}{7}+\tfrac{44}{7},\ q_6 = \tfrac{7a^4}{1944}+\tfrac{91a^3}{1296}+\tfrac{11a^2}{16}+\tfrac{245a}{72}+\tfrac{117}{16},$$

$$q_7 = \tfrac{16a^5}{11583}+\tfrac{128a^4}{3861}+\tfrac{1568a^3}{3861}+\tfrac{3616a^2}{1287}+\tfrac{1568a}{143}+\tfrac{208}{11}.$$

We may choose

$$d_0 = 1,\ d_1 = 1,\ d_2 = 1,\ d_3 = 3,\ d_4 = 36,\ d_5 = \frac{189}{4},\ d_6 = 3888,\ d_7 = \frac{11583}{16}.$$

Note that from (57), we have

$$\frac{q_3}{d_2} < y < \frac{q_7}{d_6}.$$

We will appeal to Proposition 9.11 with $\psi = \theta_2$ and $n = 3, 4, 5$ and 6. Let us begin by observing that the inequality

$$\left|\theta_2 - \frac{x}{y}\right| < \frac{1}{d_n\left(d_{n-1}+d_{n+1}\right)y^2} \tag{60}$$

is a consequence of (52), (57) and the fact that $a > 10^6$, at least for $n = 3$ and 4. We may thus apply Proposition 9.11 to conclude that either $x/y = p_3/q_3$, $x/y = p_4/q_4$, or $y \ge q_5/d_4$. In the first case, we have that $y = a + 6$, contradicting (57). In the second,

$$\frac{x}{y} = \frac{-6a-27}{2a^2+21a+90}$$

contrary to $xy(x+y) \not\equiv 0 \pmod 3$. We thus have

$$y \ge \frac{a^3}{1701}+\frac{5a^2}{567}+\frac{4a}{63}+\frac{11}{63}$$

which, with (52), implies inequality (60) for $n = 5$ and 6. From Proposition 9.11, we conclude that $x/y = p_5/q_5$ or $x/y = p_6/q_6$, i.e. that

$$\frac{x}{y} = \frac{-3a^2 - 27a - 108}{a^3 + 15a^2 + 108a + 297} \quad \text{or} \quad \frac{-42a^3 - 567a^2 - 3861a - 9720}{14a^4 + 273a^3 + 2673a^2 + 13230a + 28431}.$$

Again, this contradicts $xy(x + y) \not\equiv 0 \pmod 3$, completing the proof of Theorem 9.1.

10. A quartic family

It appears to be somewhat harder to find a convenient family of Klein forms of index 3, since families for which corresponding Thue inequalities have been treated in the literature fail to satisfy (7). We will consider instead (Klein) forms of the shape

$$F_a(x, y) = 2x^4 + 3x^3y - 3c(a)\, x^2y^2 + 3d(a)\, xy^3 - 3e(a)\, y^4,$$

where

$$2\, c(a) = 18a^4 + 204a^3 + 867a^2 + 1595a + 1038,$$
$$4\, d(a) = 72a^6 + 1224a^5 + 8643a^4 + 32106a^3 + 65399a^2 + 68268a + 28332$$

and

$$\begin{aligned} 8\, e(a) = \;& 81a^8 + 1836a^7 + 18153a^6 + 101871a^5 + 353472a^4 \\ & + 773229a^3 + 1036930a^2 + 776604a + 248112. \end{aligned}$$

Via Eisenstein's criterion at the prime 3, $F_a(x, y)$ is irreducible in $\mathbb{Q}[x, y]$, at least provided $a \equiv \pm 1 \pmod 3$. Writing $\kappa_a = (2a + 3)(12a^2 + 84a + 163)$, we prove

Theorem 10.1. *Let $a \equiv 17, 37 \pmod{60}$ be an integer and suppose that κ_a is squarefree. If, further, κ_a has no prime divisors congruent to $3, 17, 27$ or 33 modulo 40, then the equation*

$$F_a(x, y) = z^n$$

has at most finitely many solutions in coprime integers x and y, and integers z and $n \geq 3$.

Proof. Assume henceforth that κ_a is squarefree (an old result of Erdős [17] ensures that this occurs for infinitely many values of a). We begin by noting that the family of forms $F_a(x, y)$ has a number of properties reminiscent of our cubic family. The discriminant of $F_a(x, y)$, for instance, satisfies $\Delta_{F_a} = -3^3\kappa_a^6$, while we have corresponding Hessian $H_a(x, y) = \kappa_a\, H_a^*(x, y)$, where

$$H_a^*(x, y) = -a_1(a)x^4 + 6b_1(a)x^3y - 3c_1(a)x^2y^2 + 3d_1(a)xy^3 - 3e_1(a)y^4,$$

with

$$a_1(a) = 6a + 17, \;\; b_1(a) = 6a^3 + 51a^2 + 143a + 118,$$
$$2c_1(a) = 54a^5 + 765a^4 + 4308a^3 + 11925a^2 + 16028a + 8292,$$
$$2d_1(a) = 54a^7 + 1071a^6 + 9063a^5 + 42159a^4 + 115659a^3 + 185702a^2 + 160308a + 57096$$

and

$$\begin{aligned} 16e_1(a) = \;& 162a^9 + 4131a^8 + 46656a^7 + 304002a^6 + 1249038a^5 \\ & + 3322347a^4 + 5648376a^3 + 5820520a^2 + 3229248a + 711216. \end{aligned}$$

We note that, for every $a \in \mathbb{Z}$, the coefficients of $F_a(x, y)$ and $H_a^*(x, y)$ actually lie in $\mathbb{Z}$.

A straightforward calculation yields, if $H_a^*(x, y)$ is even, that necessarily $H_a^*(x, y)$ is divisible by 2^5. Since, in such a case, we have both δ_3 and $F_a(x, y)$ odd, we

may conclude via (30) that $\nu_2(N(\tilde{E}_{x,y})) > 0$. Theorem 8.4 thus implies the desired result, unless there exist integers x and y for which either $F(x,y) \in \mathbb{Z}^*_{S_{F_a}}$ or $H_a^*(x,y) \in \mathbb{Z}^*_{S_{F_a}}$.

To treat these equations, analogous to Lemma 9.2, we require information about the lifting of roots of $F_a(x,1)$ modulo p.

Lemma 10.2. *Let $a \in \mathbb{Z}$, $p > 3$ be a prime and suppose that $\nu_p(\kappa_a) = 1$. Then, for every pair of coprime integers x and y, we have*

$$\nu_p(F_a(x,y)) \text{ and } \nu_p(H_a^*(x,y)) \in \{0,2\}.$$

Proof. Let $F(x,y)$ be any quartic Klein form as in (6), with $\alpha_i \in \mathbb{Z}$ and $p \nmid \alpha_0$. Assume further that $\nu_p(\delta_3) = 3$. Then the discriminant of F is divisible by p and so F has a repeated factor over $\mathbb{F}_p$. From (7), we may readily conclude that this factor is not an irreducible quadratic over $\mathbb{F}_p$. It follows that F has a root modulo p. After a suitable linear transformation, we may assume that p divides α_3 and α_4. Repeatedly appealing to (7) and our assumption that

$$p^3 \mid 27\delta_3 = 72\alpha_0\alpha_2\alpha_4 + 9\alpha_2\alpha_3\alpha_4 - 27\alpha_0\alpha_3^2 - 27\alpha_4\alpha_1^2 - 2\alpha_2^3,$$

we find that either $p \nmid \alpha_1$, in which case $p^3 \mid \alpha_4$, $p^2 \mid \alpha_3$, $p \mid \alpha_2$ and F has a simple root modulo p, or $p \mid \alpha_1$, whence $p^2 \mid \alpha_4$, $p^2 \mid \alpha_3$ and $p \mid \alpha_2$. In the latter case, any root mod p is automatically a root mod p^2. Such a root cannot lift to a root modulo p^3, since this would imply $p^3 \mid \alpha_4$ and $p^2 \mid \alpha_2$, and hence $p^4 \mid \delta_3$, contradicting $\nu_p(\delta_3) = 3$.

To apply this to the situation at hand, first of all note that $p \nmid F_a(1,0)$ and, by a resultant computation, also $p \nmid H_a^*(1,0)$. Next, it is a relatively easy matter to check that all primes lying above p in the field defined by $F_a(x,1)$ (and hence the field defined by $H_a^*(x,1)$) ramify (but not necessarily completely). This implies that $F_a(x,1)$ and $H_a^*(x,1)$ have no root in $\mathbb{Z}_p$, and hence no simple roots modulo p. We are thus in the second case of the preceding paragraph, whereby the lemma follows. □

Suppose, from now on, that $a \equiv \pm 1 \pmod 3$ (so that δ_3 is coprime to 3). Then

$$\nu_p(F_a(x,y)) \in \{0,2\}$$

for every coprime x, y, and each $p \mid \kappa_a$. Since $\nu_3(F_a(x,y)) \in \{0,1\}$, it follows that if $F_a(x,y)$ is an S_{F_a} unit, necessarily

$$F_a(x,y) = \pm 3^\delta z^2$$

for $\delta \in \{0,1\}$ and z an odd integer, coprime to 3. Assuming, further, that $a \equiv 1 \pmod 4$, then $F_a(x,y)$ is either even or $F_a(x,y) \equiv 1 \pmod 8$, whereby $F_a(x,y) = z^2$ for some z, coprime to 6. In particular, it follows that $F_a(x,y) \equiv 1 \pmod 3$, contradicting $F_a(x,y) \equiv 2x^4 \pmod 3$.

Let us next suppose that $H_a(x,y)$ is an S_{F_a} unit, whereby the same is necessarily the case for $H_a^*(x,y)$. Again, we have that $\nu_p(H_a^*(x,y)) \in \{0,2\}$ for every coprime x, y, and each $p \mid \kappa_a$, and that $\nu_3(H_a^*(x,y)) \in \{0,1\}$. Since a short calculation implies $H_a^*(x,y) \equiv 1 \pmod 8$, it follows that $H_a^*(x,y) = z^2$ for some z dividing κ_a.

Now suppose that p is a prime dividing κ_a. The proof of Lemma 10.2 shows that we have

$$H_a^*(x,y) \equiv (-6a - 17)(x + my)^4 \pmod p,$$

for some integer m. It follows from $H_a^*(x,y) = z^2$ that either $p \mid z$ or $-6a-17$ is a quadratic residue modulo p. If $p \mid 2a+3$ then

$$\left(\frac{-6a-17}{p}\right) = \left(\frac{-8}{p}\right) = \left(\frac{-2}{p}\right),$$

and hence $-6a-17$ is a quadratic residue modulo p precisely when $p \equiv 1, 3 \pmod 8$. If $p \mid 12a^2+84a+163$, then, since the discriminant of this quadratic is $-2^8 \cdot 3$, there exists an integer k such that

$$k^2 \equiv -3 \pmod p \quad \text{and} \quad a \equiv -\frac{7}{2} \pm \frac{2k}{3} \pmod p.$$

Thus $-6a-17 \equiv 4 \pm 4k \pmod p$, whereby, again,

$$\left(\frac{-6a-17}{p}\right) = \left(\frac{1 \pm k}{p}\right) = \left(\frac{1 \pm k}{p}\right)^3 = \left(\frac{(1 \pm k)^3}{p}\right) = \left(\frac{-8}{p}\right) = \left(\frac{-2}{p}\right).$$

We conclude that, if $p \mid \kappa_a$ and $p \equiv 5$ or 7 modulo 8, then necessarily $p \mid z$.

Restricting attention to $a \equiv 2 \pmod 5$, it is easy to see that $H_a^*(x,y)$ cannot be congruent to -1 modulo 5 and that $\kappa_a \equiv 3 \pmod 5$. Since, by hypothesis, every prime divisor of κ_a congruent to 1 or $3 \pmod 8$ is also $\equiv \pm 1 \pmod 5$, we thus conclude that $z \equiv \pm 2 \pmod 5$, contradicting $H_a^*(x,y) \equiv 1 \pmod 5$. □

It is likely that the restrictions we impose here upon the prime divisors of κ_a are unnecessary. Indeed, provided only that κ_a squarefree and a is coprime to 3, we would expect the Thue equations associated to $F_a(x,y)$ and $H_a^*(x,y)$ to have no solutions, with perhaps finitely many exceptions. We may check with Pari GP or Magma that this is the case for all $a \in \mathbb{Z}$ with $3 \nmid a$ and $|a| \leq 50$, except if $a = -2$ (where the associated form satisfies $F_{-2}(0,1) = 3$).

Remark 10.3. It is actually possible to handle this anomalous case $a = -2$ as follows. We have

$$F_{-2}(x,y) = 2x^4 + 3x^3y + 42x^2y^2 + 204xy^3 + 3y^4$$

and, in the notation of Theorem 8.4, we can find a suitable twist by t such that $N_1 = 3^\alpha \cdot 43^2$, for $\alpha \in \{2,3\}$. For all elliptic curves E with conductor N_1, the equation $\tilde{j}(x,y) = j_E$ has no rational roots, and hence the equality $N(\tilde{E}^{(t)}_{x_1,y_1}) = N_1$ cannot be satisfied (the Thue-Mahler equation $\tilde{F}(x_1,y_1) \in \mathbb{Z}^*_{\tilde{S}_F}$ has no solutions).

On the other hand, the equality $j(x,y) = j_E$ actually does have solutions, whereby there exist coprime integers x_1 and y_1 for which $E^{(t)}_{x_1,y_1}$ has conductor N_1. These elliptic curves correspond to the isogeny class $16641e1$ in Cremona's notation, with j-invariant $2^{18} \cdot 3^3 \cdot 5^3$. Since these curves have complex multiplication, we may argue as we will do in Section 13 to show that the equation $F_{-2}(x,y) = z^l$ has no solutions for suitably large prime $l \equiv 1 \pmod 3$. We can, however, do rather better by examining images of inertia. Let E be an elliptic curve in isogeny class $16641e1$, whereby a twist of E by -3 has good reduction at 3. This means that, for $l > 3$ we have $\#\rho_l^E(I_3) = 2$, where $I_3 \subset \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ denotes an inertia subgroup at 3. Conversely, from the explicit formulae for $\Delta(x,y)$ and $c_4(x,y)$, it is a simple matter to check that E_0 has no quadratic twist with good reduction at 3, whereby $\#\rho_l^{E_0}(I_3) > 2$. Theorem 8.4 now tells us that the generalized superelliptic equation in question has no solutions for any suitably large prime exponent l.

Remark 10.4. Though it is very likely that the hypotheses of Theorem 10.1 are satisfied for infinitely many a, this does not appear to follow in a straightforward fashion from, say, application of the half-dimensional sieve to the polynomial κ_a. It appears that, with a certain amount of work, it should be possible to generalize the family $F_a(x, y)$ to one with two parameters and analogous properties, which would, in all likelihood, be provably infinite.

11. Higher degree families of Klein forms

It is a relatively easy matter to find infinite families of Klein forms of degrees 6 and 12 for which we can guarantee that equation (2) has finitely many solutions in integers x, y, z and $l \geq 2$. We appeal to the following result.

Proposition 11.1. *Let F be a Klein form of index n and suppose that $p \notin S_F$ is prime. If p has the additional property that $F(x, y) \equiv 0 \pmod{p}$ for all integers x and y, then there are at most finitely many solutions to equation (2) in integers x, y, z and $l \geq 2$. In particular, there are no such solutions with prime l satisfying*

$$\log l > 2^7 \cdot 3^4 \left(\prod_{q \in S_F \setminus \{2,3\}} q^2 \right) \log \left(1 + \sqrt{p}\right).$$

Proof. Suppose that F is a Klein form of index n and that p is a prime, coprime to $n\,\Delta_F$, with the property that $F(x, y) \equiv 0 \pmod{p}$ for all integers x and y. It follows that if x_0, y_0, z_0 is a solutions in integers to (2), where l is prime, then the Frey-Hellegouarch curve $E^{(t)}_{x_0,y_0}$ (with t chosen as usual) has (multiplicative) bad reduction at p. From (38), either $l \leq 163$ or there exists a weight 2 cuspidal newform f, of trivial character and level coprime to p, for which, provided $p \neq l$, we have

$$a_p(f) \equiv \pm(p + 1) \pmod{\mathfrak{L}},$$

where $\mathfrak{L}$ is a prime lying above l in K_f, the field of definition for the Fourier coefficients of the form f. From the Weil bounds, we thus have

$$l \leq (p + 1 + 2\sqrt{p})^{[K_f:\mathbb{Q}]},$$

whereby arguing as in Section 8, we conclude as desired. □

This result is much more specialized than it first appears. In fact, the only pairs (n, p) for which the hypotheses of Proposition 11.1 are satisfied are $(n, p) = (4, 3), (4, 5)$, and $(5, 11)$. To see this, note first that, since a primitive form $F(x, y)$ of degree k can have at most k zeros in $\mathbb{P}_1(\mathbb{F}_p)$, we necessarily have $p \leq k - 1$. To be precise, we require that

$$F(x, y) \equiv xy \left(\prod_{i=1}^{p-1} (x - iy) \right) G(x, y) \pmod{p}, \tag{61}$$

where $G(x, y)$ is a form of degree $k - p - 1$, with no linear factors over $\mathbb{F}_p[x, y]$ (the latter condition to ensure that p does not divide Δ_F). Since we assume that $p \neq n$, this immediately contradicts the existence of primes p satisfying the hypotheses of Proposition 11.1 for $n \in \{2, 3\}$. We may thus suppose that F has index 4 or 5 (which justifies the assumption $l \geq 2$ in the statement of Proposition 11.1). It follows that the only candidates for primes p are $p \in \{3, 5\}$ (if $n = 4$) and $p \in \{2, 3, 7, 11\}$ (if $n = 5$). To rule out $(n, p) = (5, 2), (5, 3)$ and $(5, 7)$, we appeal to (9), in conjunction

with (61); in each case, we may assume, without loss of generality, that $G(x, y)$ is monic in x. Suppose that F is a Klein form with coefficients as given in (6). From (9), we thus have

$$\alpha_5 \equiv \alpha_7 \equiv \alpha_1\alpha_3 + \alpha_2 \equiv \alpha_1\alpha_{11} + \alpha_3\alpha_9 + \alpha_6 \equiv 0 \pmod 2,$$

while (61) and our assumptions upon G imply that α_0 and α_{12} are even, while α_1 is odd, and hence, from the second equation in (9), $\alpha_2\alpha_3 + \alpha_4 \equiv 0 \pmod 2$. A short computation using the fact that G has no linear factors in $\mathbb{F}_2[x, y]$, leads to the conclusion that

$$G(x, y) \equiv \sum_{i=0}^{9} \alpha_i x^{9-i} y^i \pmod 2$$

where $(\alpha_0, \ldots, \alpha_9)$ is one of

$$(1, 0, 1, 0, 0, 1, 1, 0, 0, 1), (1, 0, 1, 0, 0, 1, 1, 1, 1, 1), (1, 1, 1, 1, 1, 0, 0, 0, 1, 1)$$

or $(1, 1, 1, 1, 1, 0, 0, 1, 0, 1)$. In each case, the combination of the third, fifth and seventh equations in (9) leads to a contradiction, modulo 4.

If $(n, p) = (5, 3)$, since (9) implies that $\alpha_i \equiv 0 \pmod 3$ for each $i \in \{4, 5, 7, 8\}$, whence, via the last equation of (9),

$$\alpha_1\alpha_{11} - \alpha_2\alpha_{10} + \alpha_6^2 \equiv 0 \pmod 3,$$

we find that every octic form $G(x, y)$ for which $xy\,(x^2 - y^2)\,G(x, y)$ satisfies (9) modulo 3, has a linear factor in $\mathbb{F}_3[x, y]$. In case $(n, p) = (5, 7)$, if we write $G(x, y) = x^4 + ax^3y + bx^2y^2 + cxy^3 + dy^3$, then the first 3 equations in (9) imply that

$$2b + a^2 \equiv c + ab \equiv d + ac + 4b^2 \pmod 7.$$

A routine check with Magma verifies that $G(x, y)$ satisfying these congruences all have linear factors modulo 7.

For the remaining pairs $(n, p) = (4, 3), (4, 5)$, and $(5, 11)$, in each case there exist infinitely many inequivalent forms satisfying the hypotheses of Proposition 11.1. To see this, for $(n, p) = (4, 3)$, note that we necessarily have

$$G(x, y) \in \left\{x^2 + y^2, x^2 + xy + 2y^2, x^2 + 2xy + 2y^2\right\}.$$

For each of these we may check, via (8), that $xy\,(x^2 - y^2)\,G(x, y)$ is a Klein form modulo 3, which leads to 3 families of sextic forms satisfying the hypotheses of Proposition 11.1 with $(n, p) = (4, 3)$. By way of example, fixing for simplicity $F(1, 0) = 3$, we find the families

$$\begin{array}{ll} (3, b, 15c, 90d)_6 & \text{with } b \equiv 1 \pmod 3,\ c^2 - d \equiv 1 \pmod 3, \\ (3, b, 5c, 10d)_6 & \text{with } b \equiv -c \equiv d \equiv 1 \pmod 3 \text{ and } bd - c^2 \equiv 3 \pmod 9, \\ (3, b, 5c, 10d)_6 & \text{with } b \equiv c \equiv d \equiv 1 \pmod 3 \text{ and } bd - c^2 \equiv -3 \pmod 9. \end{array}$$

Similarly, $xy\,(x^4 - y^4)$, satisfies (8), modulo 5, which provides us with infinitely many forms with $(n, p) = (4, 5)$, for instance

$$(5, b, 25c, 250d)_6 \text{ with } b \equiv 1 \pmod 5,\ c^2 - d \equiv 1 \pmod 5.$$

Finally, for $(n, p) = (5, 11)$, the fact that $xy\,(x^{10} - y^{10})$ is a Klein form modulo 11 leads to the desired conclusion. An example of a family of forms in this case is provided by

$$(11, 12, 22c, 55d)_{12} \text{ with } c \equiv 16 \pmod{11^4},\ \ d \equiv 1 \pmod{11^4}.$$

One readily shows that each of these families contain infinitely many $\mathrm{GL}_2(\mathbb{Q})$ inequivalent forms. We thus have

Corollary 11.2. *If $n \in \{4, 5\}$, there are infinitely many $\mathrm{GL}_2(\mathbb{Q})$-inequivalent Klein forms of index n for which equation (2) has at most finitely many solutions in integers x, y, z and $l \geq 2$.*

12. Heuristics for cubic forms

As we have seen, there exist infinitely many cubic forms with the property that equation (3) is insoluble in integers. In this section, we will sketch a heuristic to indicate that this is the usual state of affairs, that, in fact, a "typical" binary cubic form $F(x, y) \in \mathbb{Z}[x, y]$ has this property (and hence that Theorem 1.1 applies to almost all cubic forms, excepting a set of density zero).

For a cubic form $F(x, y)$, we have $F(-x, -y) = -F(x, y)$ and hence if F represents an integer m, it also necessarily represents $-m$. We may thus restrict attention to the representation of positive integers by F. Let us quantify what we mean by a "typical" form. We begin by noting a result of Davenport [10], [11] (we can sharpen the error term here by appealing to work of Shintani [45], [46], but this is unnecessary for our argument):

Theorem 12.1. *(Davenport, 1951) Let $H_3(A, B)$ denote the number of $\mathrm{GL}_2(\mathbb{Z})$ equivalence classes of primitive, irreducible binary cubic forms $F \in \mathbb{Z}[x, y]$, with $A < \Delta_F \leq B$. Then, as $X \to \infty$,*

$$H_3(0, X) = \frac{5}{4\pi^2} X + O(X^{15/16})$$

and

$$H_3(-X, 0) = \frac{15}{4\pi^2} X + O(X^{15/16}).$$

Note that the seeming discrepancy between this result and that stated in [10] and [11] derives from a missing factor of 3 in the statement of the main theorem of [10], together with the fact that the estimates in [10] and [11] are for classes of *properly equivalent*, rather than *equivalent*, forms; i.e. for $SL_2(\mathbb{Z})$ instead of $\mathrm{GL}_2(\mathbb{Z})$ equivalence classes, with, additionally, no restriction to primitive forms. The assumption here that our forms be irreducible can be relaxed, via application of Lemma 3 of [10].

With this result in mind, our goal is to derive a heuristic to suggest that the number of $\mathrm{GL}_2(\mathbb{Z})$ equivalence classes of primitive, irreducible binary cubic forms $F \in \mathbb{Z}[x, y]$, with $-X < \Delta_F \leq X$, say, for which (3) has integer solutions, which we will denote $H_3^{(3)}(-X, X)$, satisfies

$$H_3^{(3)}(-X, X) = o(X) \ \text{ as } \ X \to \infty,$$

i.e. such forms are "atypical" in the sense that they have zero density in the set of all classes of forms.

We begin by noting that we can unconditionally bound the number of integers up to a given X which are represented by a fixed irreducible binary cubic form. The following pair of results are the main theorem of Thunder [52] and a special case of Theorem 1 of Bean [1], respectively.

Proposition 12.2. *Let $F(x,y) \in \mathbb{Z}[x,y]$ be a cubic form of discriminant Δ_F which is irreducible over $\mathbb{Q}$ and let $X \geq 1$. Let $N_F(X)$ and A_F denote the number of integral solutions to the inequality $|F(x,y)| \leq X$, and the area of the region*

$$\left\{(x,y) \in \mathbb{R}^2 \ : \ |F(x,y) \leq 1\right\},$$

respectively. Then

$$\left|N_F(X) - X^{2/3} A_F\right| < 9 + \frac{2008\, X^{1/2}}{|\Delta_F|^{1/2}} + 3156\, X^{1/3}.$$

Proposition 12.3. *If $F(x,y) \in \mathbb{Z}[x,y]$ is a cubic form of discriminant $\Delta_F \neq 0$, then*

$$|\Delta_F|^{1/6}\, A_F \leq \frac{3\,\Gamma(1/3)^2}{\Gamma(2/3)} < 16.$$

Taken together, for suitably large $|\Delta_F|$ and X, these results imply that

$$N_F(X) < X^{2/3},$$

and, of particular interest for our purposes, that

$$\#\{1 \leq n \leq X \ : \ F(x,y) = n \text{ for some } (x,y) \in \mathbb{Z} \times \mathbb{Z}\} < X^{2/3}. \tag{62}$$

Suppose next that S_F is the set of primes p dividing $2\,\Delta_F$. We would like to find an upper bound for

$$\#\{1 \leq n \leq X \ : \ n \in \mathbb{Z}^*_{S_F}\}.$$

Write $\psi(X,Y)$ for the number of Y-smooth integers $\leq X$, and denote by p_t, the t-th prime. It follows that

$$\#\{1 \leq n \leq X \ : \ n \in \mathbb{Z}^*_{S_F}\} \leq \psi(X, p_{\omega(\Delta_F)+1}),$$

where $\omega(m)$ denotes the number of distinct prime divisors of a positive integer m. If $X \geq |\Delta_F|^{1/16}$, say, then

$$\omega(\Delta_F) \ll \frac{\log X}{\log\log X},$$

where the implied constant is absolute. It follows that $p_{\omega(\Delta_F)+1} \ll \log X$ and hence, for large enough $|\Delta_F|$ and $X \geq |\Delta_F|^{1/16}$, we have

$$\#\{1 \leq n \leq X \ : \ n \in \mathbb{Z}^*_{S_F}\} \leq \psi(X, \log^{7/6} X) = X^{1/7+o(1)} \ll X^{1/6}. \tag{63}$$

Readers interested in bounds for $\psi(X,Y)$ (and much more besides) are directed to the survey article of Granville [22].

We will make the following heuristic assumption. Let us write m_F for the smallest positive integer represented by the form F, i.e.

$$m_F = \min\left\{m \in \mathbb{N} : \text{ there exist } x,y \in \mathbb{Z} \text{ with } F(x,y) = m\right\}.$$

We will assume that the number of $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of binary cubic forms with $|\Delta_F| \leq X$ and $m_F < |\Delta_F|^{1/16}$ is $o(X)$ as $X \to \infty$. In fact, for our purposes, all we need is a like statement with $1/16$ replaced by any fixed $\delta > 0$. We actually believe that the same conclusion holds with the exponent $1/16$ replaced by any $\delta < 1/4$. By an old theorem of Mordell [34], this would be the strongest possible statement of this nature, as $m_F \leq (|\Delta_F|/23)^{1/4}$, for every cubic form F. Lemma 4 of Davenport [10] ensures that for all but $O(X^{15/16})$ classes of cubic forms, as $X \to \infty$, the Hermite reduced form for F satisfies $F(1,0) > |\Delta_F|^{1/16}$. In many cases, but not all, $m_F = F(1,0)$ for such forms.

With our heuristic assumption in hand, we appeal to a straightforward density argument. Let X be large. Then for all but $o(X)$ classes of primitive cubic forms $F(x,y)$ with $|\Delta_F| \leq X$, we have $m_F \geq |\Delta_F|^{1/16}$, whereby, from (62) and (63), the expected number of positive integers that are in

$$\mathbb{Z}_{S_F}^* \cap \{F(x,y) : (x,y) \in \mathbb{Z} \times \mathbb{Z}\}$$

is, for a given form outside this exceptional set,

$$\int_{|\Delta_F|^{1/16}}^{\infty} \frac{X^{2/3}}{X} \cdot \frac{X^{1/6}}{X} dX \ll |\Delta_F|^{-1/96}.$$

It follows (where we are of course relying upon the rather speculative independence of the representation of an integer by a cubic form F from that of said integer being an S_F-unit) that the "probability" that a given non-exceptional form F represents an S_F-unit tends to 0 as $|\Delta_F| \to \infty$.

We can derive similar heuristics (of equal plausibility) for higher degree Klein forms (whereby our expectation is that there are at most finitely many solutions to equation (2) in integers x, y, z and $l \geq \max\{2, 7-k\}$, for "almost all" Klein forms of degrees $k = 4, 6$ and 12). For $k > 3$, however, Klein forms constitute a density zero subset of the set of all forms of degree k.

13. Diagonal forms

In this section, we will turn our attention to what are, in some sense, the simplest binary forms. We call a binary form *diagonal* if it is $\mathrm{GL}_2(\mathbb{Z})$ equivalent to a form of the shape $F(x,y) = ax^k + by^k$, for integers a and b. It is clear from equations (7), (8) and (9) that the only diagonal Klein forms are of degree $k = 3$. Since such forms have discriminant

$$\Delta_F = -3^3 \ (ab)^2 \,,$$

the conditions of Theorem 1.1 are never satisfied ($F(1,0)$ divides Δ_F, for instance). Despite this, we can still appeal to Theorem 8.4 to deduce some useful Diophantine information.

We begin by noting that our conductor calculations become rather more straightforward for diagonal cubic forms; for simplicity, we will initially restrict our attention to those forms with odd coefficients. Let U denote the set of integers congruent to ± 1 modulo 9, and V consist of those $\equiv \pm 2, \pm 4 \pmod 9$.

Proposition 13.1. *Suppose that a and b are cubefree coprime odd integers and let $F(x,y) = ax^3 + by^3$. Define the family of associated elliptic curves $E_{x,y}$ as in (20). Then, if x and y are coprime integers, the conductor $N(E_{x,y}^{(t)})$ of $E_{x,y}^{(t)}$ satisfies, for t equal to one of $\pm 1, \pm 2, \pm 3, \pm 6$,*

$$N(E_{x,y}^{(t)}) = 2^{\alpha} \cdot 3^{\beta} \prod_{p \mid ab} p^2 \prod_{\substack{p \mid ax^3 + by^3 \\ p \nmid 6ab}} p,$$

where

α	*Conditions*
0	*if* $\nu_2(ax^3 + by^3) = 4$
1	*if* $\nu_2(ax^3 + by^3) \geq 5$
2	*if* $\nu_2(xy) \geq 2$
3	*if* $\nu_2(xy) = 1$ *or* $\nu_2(ax^3 + by^3) \in \{2, 3\}$
5	*if* $\nu_2(ax^3 + by^3) = 1$

and

β	*Conditions*
1	*if* $3 \mid ax^3+by^3$ *and either* $a,b,ab \in V$ *or* $a,b \in U$,
2	*if* $3 \nmid ax^3+by^3$ *and either* $a,b,ab \in V$ *or* $a,b \in U$,
3	*if either* $a,b \in U$, *or* $a,b \in V, ab \in U$, *or* $3 \mid ab$.

Proof. The computations of $\nu_p(N(E_{x,y}))$ is straightforward for $p > 3$, but requires some more work (and twisting to minimize) for $p = 2$ and 3. We use Tate's algorithm; in practice we appeal to Papadopolous [36] for most (but not all) cases. □

As an example of the kind of result one may obtain from Theorem 8.4 in this situation, we have the following.

Theorem 13.2. *Let a and b be odd integers such that ax^3+by^3 is irreducible in $\mathbb{Q}[x,y]$ and suppose that all solutions to the equation*

$$ax^3+by^3 = \prod_{p \mid 6ab} p^{\alpha_p} \tag{64}$$

in coprime nonzero integers x and y, and nonnegative integers α_p, satisfy $\alpha_2 \in \{1,4\}$. Then there exists an effectively computable constant l_0 such that the equation

$$ax^3+by^3 = z^l \tag{65}$$

has no solution in nonzero integers x, y and z (with $\gcd(x,y)=1$), and prime $l \equiv 1 \pmod 3$ with $l \geq l_0$. If all solutions to (64) with x and y coprime integers have $\alpha_2 \leq 4$, then there exists an effectively computable constant l_1 such that equation (65) has no solutions in odd integers x, y and prime $l \geq l_1$.

Proof. Suppose that we have $ax_0^3+by_0^3 = z_0^l$ where x_0, y_0 and z_0 are nonzero integers with $\gcd(x_0,y_0)=1$, and l is prime. Then, via Theorem 8.4, we find that either l is bounded as in (41), or we deduce the existence of an integer solution (x_1,y_1) to equation (64), for which we have, writing E_i for $E^{(t)}_{x_i,y_i}$ (with t as in Proposition 13.1),

$$\nu_2(N(E_1)) = \nu_2(N(E_0)).$$

By Proposition 13.1, it follows that

$$\nu_2(N(E_0)) \in \{1,2,3\}$$

(and that $\nu_2(N(E_0))=1$, if $x_0\,y_0$ is odd). From our assumptions about solutions to equation (64), we thus have $x_1y_1 = 0$, whereby E_1 has j-invariant 0 and hence complex multiplication by $\mathbb{Z}[(1+\sqrt{-3})/2]$. If $l \equiv 1 \pmod 3$, then l splits in $\mathbb{Z}[(1+\sqrt{-3})/2]$, and so $\rho_l^{E_1}(G_{\mathbb{Q}})$ and hence $\rho_l^{E_0}(G_{\mathbb{Q}})$ are contained in the normalizer of a split Cartan subgroup of $\mathrm{GL}_2(\mathbb{F}_l)$. By work of Momose [33] (and since we may suppose $l > 13$), we conclude that

$$j_{E_0} = \frac{2^8 \cdot 3^3 ab(x_0y_0)^3}{(ax_0^3+by_0^3)^2} = \frac{2^8 \cdot 3^3 ab(x_0y_0)^3}{z_0^{2l}} \in \mathbb{Z}[1/2].$$

In fact, Merel [32] showed that $j_{E_0} \in \mathbb{Z}$, but the weaker result is adequate for our purposes. Indeed, since we suppose x_0, y_0, z_0 to be nonzero (with $\gcd(x_0,y_0)=1$), we reach the desired contradiction, provided l is suitably large in terms of a and b, unless $z_0 = \pm 1$ (which itself contradicts our assumptions upon possible solutions to (64)). □

Remark 13.3. In case $l \equiv -1 \pmod 3$ in the preceding proof, the image of $\rho_l^{E_0}$ is contained in the normalizer of a non-split Cartan subgroup of $\mathrm{GL}_2(\mathbb{F}_l)$. A conjecture of Serre (stated as a question in Section 4.3 of [43]) then implies, for suitably large l, that E_0 has complex multiplication. Assuming this conjecture (or something rather weaker, such as integrality of the corresponding j-invariant j_{E_0}), the restriction to $l \equiv 1 \pmod 3$ in Theorem 13.2 can be removed.

As a rough indication of the applicability of Theorem 13.2, we note that there are precisely three pairs of coprime odd positive integers (a, b) with $ab < 100$ and $a < b$, for which all solutions to equation (64)) in coprime nonzero integers x and y, and nonnegative integers α_p, satisfy $\alpha_2 \in \{1, 4\}$:

$$(a, b) \in \{(1, 57), (1, 83), (3, 19)\}.$$

In addition to these, if we require only $\alpha_2 \leq 4$, we find the following pairs (a, b) :

$$(1, 15), (1, 23), (1, 43), (1, 47), (1, 51), (1, 99), (3, 11), (3, 13), (3, 25), (5, 9).$$

It is a simple matter to find diagonal cubic forms $ax^3 + by^3$ for which the corresponding Thue-Mahler equations immediately reduce to Thue equations, in a similar fashion as that of our cubic family in Section 9. For example, if we take a and b to be cubefree nonzero coprime integers with ab even and $ab^2 \not\equiv \pm 1 \pmod 9$, then if x and y are coprime integers satisfying (64), it follows that $ax^3 + by^3 = 3^\delta k$, where either $\delta \in \{0, 1\}$ and $k \mid ab$ (more specifically, $\nu_p(k) \in \{\nu_p(ab), 0\}$ for each $p \mid ab$), if ab is coprime to 3, or $\delta = 0$ and $k \mid ab$, otherwise. We thus reduce (64) to solving the family of Thue equations $a_0x^3 + b_0y^3 = c$, where a_0, b_0 range over all positive cubefree integers for which $\mathbb{Q}(\sqrt[3]{a_0b_0^2}) = \mathbb{Q}(\sqrt[3]{ab^2})$ and $c = 1$ or 3 (where a_0b_0 is coprime to 3 in the latter case). From work of Stender [50], the number of such equations with a solution in nonzero integers is, with a pair of exceptions, at most one. Specifically, combining Satz 1 and Satz 2 of [50], we have

Theorem 13.4. *(Stender) Let $D > 1$ be a cubefree integer. Then, if a and b are cubefree positive integers, there is at most one equation of the shape $ax^3 + by^3 = c$ with $c \in \{1, 3\}$, $\gcd(ab, c) = 1$ and $\mathbb{Q}(\sqrt[3]{a/b}) = \mathbb{Q}(\sqrt[3]{D})$, with as many as a single solution in nonzero integers x and y, unless $D = 2$ (or 4), or $D = 20$ (or 50). For these values of D, we have solutions corresponding to*

$$(a, b, c, x, y) = (2, 1, 1, 1, -1), (2, 1, 3, 1, 1), (2, 1, 3, 4, -5), (4, 1, 3, 1, -1)$$

and

$$(a, b, c, x, y) = (20, 1, 1, 7, -19), (5, 2, 3, 1, -1),$$

respectively. If, for a given D, we have a solution to such an equation $ax^3 + by^3 = c$ with $xy \neq 0$, then either

- *(i)* $\min\{a, b\} = 1$, $c = 1$, *and* $(a, b, c) \notin \{(19, 1, 1), (20, 1, 1), (28, 1, 1)\}$, *in which case* $\eta = x\sqrt[3]{a} + y\sqrt[3]{b}$ *is the fundamental unit in the field* $\mathbb{Q}(\sqrt[3]{ab^2})$, *or*
- *(ii)* $(a, b, c) \in \{(19, 1, 1), (20, 1, 1), (28, 1, 1)\}$, *where* $\eta = x\sqrt[3]{a} + y\sqrt[3]{b}$ *is the square of the fundamental unit in the field* $\mathbb{Q}(\sqrt[3]{ab^2})$, *or*
- *(iii) either* $\min\{a, b\} > 1$, *or* $\min\{a, b\} = 1$ *and* $c = 3$, *and* $\eta = \frac{1}{c}\left(x\sqrt[3]{a} + y\sqrt[3]{b}\right)^3$ *is the fundamental unit in* $\mathbb{Q}(\sqrt[3]{ab^2})$, *or its square (with the sole exception of* $(a, b, c) = (2, 1, 3)$*).*

Arguing as in the proof of Theorem 13.2 (only without appeal to Proposition 13.1), we thus have

Theorem 13.5. *Let D be an even positive cubefree integer with $D \not\equiv \pm 1 \pmod 9$ and let $0 < \epsilon < 1$ be the fundamental unit in $\mathbb{Q}(\sqrt[3]{D})$. Suppose that ϵ is not of the shape $x\sqrt[3]{a} + y\sqrt[3]{b}$ and that neither ϵ nor ϵ^2 is of the shape $\frac{1}{c}\left(x\sqrt[3]{a} + y\sqrt[3]{b}\right)^3$, with $c \in \{1, 3\}$, x and y nonzero integers, and a and b coprime positive cubefree integers for which $\mathbb{Q}(\sqrt[3]{ab^2}) = \mathbb{Q}(\sqrt[3]{D})$. Then there exists an effectively computable constant l_0, depending only on D, such that, for each such pair a and b, the equation*

$$ax^3 + by^3 = z^l$$

has no solutions in nonzero integers x, y and z, with $\gcd(x, y) = 1$, and prime $l \equiv 1 \pmod 3$ with $l \geq l_0$.

A short computation with Pari GP indicates that this theorem is applicable to the following $D < 150$:

$$D \in \{34, 38, 74, 78, 84, 86, 92, 94, 102, 106, 114, 132, 138, 142, 146\}.$$

We are unaware of simple criteria for determining whether, given D, the fundamental unit of $\mathbb{Q}(\sqrt[3]{D})$ takes a "binomial" form as in the hypotheses of Theorem 13.5. Our calculations indicate that this occurs infrequently (and hence that this theorem is applicable to "most" cubefree D satisfying the given congruences modulo 18).

14. Examples and computations

Given a Klein form F, we know essentially three distinct ways to determine, in an effective manner, whether the corresponding Thue-Mahler equations have solutions or not. The first is to solve the equations, as in Tzanakis and de Weger [53], through a combination of lower bounds in linear forms in complex and p-adic logarithms, with techniques from computational Diophantine approximation. A second method is to appeal to the syzygy (12), which shifts the problem to one of determining the S-integral points on a collection of (Mordell) elliptic curves. Since the number of such curves is exponential in the cardinality of S_F, in many situations this method appears impractical.

A third approach is to actually compute models for all elliptic curves $E/\mathbb{Q}$ at the corresponding levels, say via modular symbols, and to check whether or not there exists one with $E[n]$ isomorphic to that coming from the Klein form. One has the feeling that, in all but the simplest cases, this last approach is the least computationally efficient (though we are unaware of a complexity analysis to confirm this).

In the remainder of this section, we will illustrate the first and third of these methods, providing the results of somewhat extensive computations.

14.1. **Solving Thue-Mahler equations.** In the case of cubic forms, both as a "reality check" on our heuristics, and to illustrate the utility of these methods, we implemented the algorithm for solving Thue-Mahler equations detailed in Tzanakis and de Weger [53]. After computing (Hermite) reduced representatives for each class of irreducible, primitive cubic forms with $|\Delta_F| \leq 10^6$, in each case we solved the corresponding equation (3). We summarize our results in the following table. More details, including lists of forms satisfying the hypotheses of Theorem 1.1, are available from the authors on request. Let us define, as before, $H_3(A, B)$ to be the number of $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of primitive irreducible binary cubic forms $F \in \mathbb{Z}[x, y]$ with discriminant $A < \Delta_F \leq B$ and let $H_3^*(A, B)$ to be the analogous

counting function for those forms for which equation (3) has no integral solutions. Then we have

X	$H_3(0,X)$	$H_3^*(0,X)$	$H_3(-X,0)$	$H_3^*(-X,0)$
10^2	2	0	7	0
10^3	30	0	160	0
10^4	484	4	2201	20
10^5	6765	69	26875	741
10^6	83636	2174	303136	19311

Recall that our heuristic predicts

$$\lim_{X\to\infty} \frac{H_3^*(0,X)}{H_3(0,X)} = \lim_{X\to\infty} \frac{H_3^*(-X,0)}{H_3(-X,0)} = 1.$$

Dan Shanks once noted that $\log\log\log x$ tends to infinity "with great dignity." (Math. Comp. 13 (1959), page 272); we believe something similar to be occurring here.

14.2. **Studying elliptic curves at candidate levels.** Given a Klein form F, if we have access to a full set of isomorphism classes of elliptic curves $E/\mathbb{Q}$ at the various levels N that can arise from possible solutions to equation (2), we can, in principle, do without most of the theory we have developed so far. Indeed, we may simply appeal to the congruences from Proposition 8.1 to eliminate the possibility of a given elliptic curve giving rise to our newform f, for large enough prime exponent l. It is possible, however, to carry out this check in a more elegant fashion, with no need for explicit calculation of Fourier coefficients for the various elliptic curves (Frey-Hellegouarch and otherwise) encountered.

Let F be an irreducible Klein form of index n (with δ_n nonsquare, in case $n = 4$) and, as previously, denote by $j(x, y)$ the j-invariant associated to $E_{x,y}$. By a direct application of Theorem 8.4, for an elliptic curve E at a candidate level N, with j-invariant j_0, we need only check that j_0 is not of the form $j(x, y)$ for (coprime) integers x and y and, if $n = 3$ or 5, that j_0' is not of the form $j(x, y)$. Here $j_0' = 1728^2/j_0$, if $n = 3$, and $j_0' = J(j_0)$ (as in (31)), if $n = 5$. If $j_0' = \infty$, this value can be ignored, since then $j_0' = j(x, y)$ corresponds to a root of F, which is irreducible by assumption. Note that solving the equation $j_0 = j(x, y)$ (or $j_0' = j(x, y)$) amounts to finding rational roots of a univariate polynomial over the rationals. If $j_0 = 0$ or $j_0 = 1728$, and this check fails to eliminate E, then it could still happen that K_n^E is reducible or that the fields determined by $F(x, 1)$ and $K_n^E(x, 1)$ are not isomorphic. If this is the case, one can still eliminate E.

Note that what we are doing here amounts to considering only the values of the Fourier coefficients a_p of a candidate curve E, up to sign. Of course, there may occur instances where such an approach fails to discard an elliptic curve E which we can actually eliminate through careful examination of the a_p (taking their signs into account). In practice, it appears that such a situation does not arise particularly often, especially if care is taken to identify the levels N which can genuinely correspond to solutions of our superelliptic equations.

We already know that our Frey-Hellegouarch curve attached to (2), after level lowering, corresponds to a modular form at some level in $\mathbb{Z}_{S_F}^*$. In fact, we can typically restrict the possible levels occuring somewhat further. Although tedious in many cases, especially when in comes down to determining possible $\nu_2(N)$ and

$\nu_3(N)$, this is a straightforward task. For more information we refer to the appendix. The remainder of this section contains a number of examples, for each index $n \in \{2,3,4,5\}$, where we explicitly determine "minimal" such N.

14.2.1. *Cubic forms.* We list reduced forms $F(x,y) = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)_3$ with $|\Delta_F| \leq 10^4$ and the property that equation (3) has no integral solutions. Our first table contains forms of negative discriminant:

α_0	α_1	α_2	α_3	Δ_F	α_0	α_1	α_2	α_3	Δ_F
3	2	5	3	-2063	3	1	6	5	$-79 \cdot 89$
5	6	7	3	-2423	3	4	2	6	$-2^2 \cdot 1931$
3	8	9	7	-2591	3	4	10	6	$-2^2 \cdot 1931$
3	4	9	5	-5087	3	7	8	9	-7823
3	7	10	9	$-19 \cdot 269$	3	7	12	11	$-17 \cdot 487$
3	1	7	3	$-2^2 \cdot 1283$	3	5	8	9	$-37 \cdot 251$
3	5	6	7	$-13 \cdot 443$	3	4	3	7	-9343
5	1	4	3	-6271	3	1	8	-3	-9551
6	8	10	3	$-2^2 \cdot 31 \cdot 53$	3	2	8	6	$-2^2 \cdot 2411$
3	1	8	3	-6983	3	5	3	7	$-2^2 \cdot 2459$

The corresponding levels we need to consider for these examples are, after suitable twisting, $N = 2^{\beta_2} \prod_{p|\Delta_F} p$, where the product is over odd prime p and $\beta_2 \in \{0,5\}$ or $\{2,3\}$, if 2 divides, or fails to divide Δ_F, respectively.

If we consider forms of positive discriminant, there are precisely four classes with $\Delta_F < 10^4$:

α_0	α_1	α_2	α_3	Δ_F	Level(s)
3	2	-7	-3	67^2	$2^\delta \cdot 67^2$, $\delta \in \{2,3\}$
3	1	-8	-3	73^2	$2^\delta \cdot 73^2$, $\delta \in \{2,3\}$
3	4	-7	-3	8017	$2^\delta \cdot 8017$, $\delta \in \{2,3\}$
3	-1	-10	-3	$7^2 \cdot 13^2$	$2^\delta \cdot 7^2 \cdot 13^2$, $\delta \in \{2,3\}$

All the levels involved in these 24 examples are smaller than 200000, which means that all elliptic curves at these levels (i.e. with these conductors) are available in Cremona's database. For each of these forms, therefore, there was no need to solve the corresponding Thue-Mahler equations. Instead, one could simply have checked that the equation $j_0 = j(x,y)$ has no rational roots, for each j-invariant j_0 of an elliptic curve at the given levels.

14.2.2. *Quartic forms.* Binary forms of degree 4 have, as noted in Section 2, exactly two independent invariants, traditionally denoted I and J – in our notation, $I = \tau_4(F) = 0$ and $J = 27\,\delta_3$, where $\Delta_F = -27\,\delta_3^2$. In Cremona [6], one finds an algorithm (based on Julia's theory of reduction) for finding an $SL_2(\mathbb{Z})$-reduced representative for each class of quartic forms with given (I,J). Implementing this in the case $I = 0$, we may list quartic Klein forms with $|\Delta_F|$ below a given bound. We assume, replacing F by $-F$ if need be, that $\delta_3 > 0$. Applying Proposition 14 of [6], we may thus conclude that there exists a (not necessarily Julia-reduced) representative $(\alpha_0, \alpha_1, \alpha_2, \alpha_3)_4$ for each form with invariants $I = 0$ and $J = 27\,\delta_3$, satisfying

$$-\frac{\sqrt{3}}{2}\,\delta_3^{1/3} \leq \alpha_0 \leq \frac{1}{\sqrt{3}}\,\delta_3^{1/3}, \quad -2|\alpha_0| < \alpha_1 \leq 2|\alpha_0|,$$

and

$$H_1 \le 8\alpha_0\alpha_2 - 3\alpha_1^2 \le H_2,$$

where

$$H_1 = 6\,\delta_3^{1/3}\,\max\left\{-2(\alpha_0+\delta_3^{1/3}), \alpha_0 - \sqrt{4\delta_3^{2/3}-3\alpha_0^2}\right\}$$

and

$$H_2 = 6\,\delta_3^{1/3}\,\min\left\{-2\,\alpha_0, \alpha_0 + \sqrt{4\delta_3^{2/3}-3\alpha_0^2}\right\}.$$

In the following, we list (Julia) reduced representatives $F = (\alpha_0, \alpha_1, \alpha_2, \alpha_3)_4$ for classes of Klein forms with $|\delta_3| \le 4000$, and for which the corresponding Thue-Mahler equation has no small solutions (i.e. solutions with, say, $|x|, |y| \le 100$). We suspect that the superelliptic equations corresponding to these forms are, for suitably large exponents, insoluble.

α_0	α_1	α_2	α_3	$\lvert\delta_3\rvert$	α_0	α_1	α_2	α_3	$\lvert\delta_3\rvert$
6	13	−9	−9	$3\cdot 599$	5	−8	−18	19	3491
4	1	−18	28	1907	2	1	−21	51	3517
2	−13	−9	−7	2053	7	−13	−24	−4	3581
6	3	−21	−15	2411	5	16	−12	−7	3659
4	5	3	−25	2683	7	−11	−18	8	$3\cdot 1237$
2	−5	24	−8	2789	4	7	6	−28	3739
6	−3	9	−25	$3\cdot 1103$	2	7	−18	28	3907
6	15	−3	13	3391	2	3	21	−15	3923
2	7	−12	24	3391					

To prove our suspicions correct (without actually solving the Thue-Mahler equations in the traditional manner), we observe that the corresponding levels we are required to consider for these examples are, after suitable twisting,

$$N = 3^\beta \cdot \mathrm{rad}_3(\delta_3),\ \beta \in \{2,3\},$$

for those forms with δ_3 coprime to 3, and

$$N = 3^\beta \cdot \mathrm{rad}_3(\delta_3),\ \beta \in \{1,4\},$$

otherwise. Here, we denote by $\mathrm{rad}_3(\delta_3)$ the product of primes $\neq 3$ dividing δ_3. We note that for some of these examples where $\nu_3(N) = 2$, there exists a quadratic twist with $\nu_3(N) = 1$, but for such a form, we always still need to consider the case $\nu_3(N) = 2$.

Again, for each of these forms, data for all corresponding elliptic curves are available. It is again an easy matter to check that for all these quartic forms F and all j-invariants j_0 of the elliptic curves at the levels corresponding to F, we have that the associated equations $j_0 = j(x,y)$ and $1728^2/j_0 = j(x,y)$ both have no rational roots.

We conclude this subsection with a number of illustrative examples of quartic Klein forms.

Example 14.1. Consider the Klein form $F = (2, 55, 429, 85)_4$, which has $\delta_3 = 163^2$ and corresponding levels $N = 3^\beta \cdot 163$, $\beta \in \{2,3\}$. The equations $j_0 = j(x,y)$ have no solutions, but the equations $1728^2/j_0 = j(x,y)$ do have a solution in one case, namely for the elliptic curve

$$E \ : \ y^2 + y = x^3 - 18x - 34,$$

which has conductor $3^2 \cdot 163$ and j-invariant $j_0 = -884736/163$. We have $1728^2/j_0 = j(-85, 8)$. This is also illustrated by the fact that the 3-division polynomial of E, given by $3x^4 - 108x^2 - 405x - 324$, defines the same field as

$$F(x, 1) = 2x^4 + 55x^3 + 429x^2 + 85x - 7084.$$

On further inspection, one can actually discard this elliptic curve E by using an image of inertia argument (this is basically possible because its twist over $\sqrt{-3}$ has conductor 163, but a similar twist of the Frey-Hellegouarch curve still has potentially good additive reduction at 3).

Our final two examples of this subsection are chosen to illustrate the necessity of considering the companion form H and of the inclusion of the prime 2 in $\tilde{S}_F$. In both cases, one may check that straightforward image of inertia arguments cannot be used to eliminate the elliptic curves in question.

Example 14.2. Consider $F = (2, 1, -18, 4)_4$ with $\delta_3 = 1637$. We can show that the equation $j_0 = j(x, y)$ is insoluble, for each j-invariant j_0 of an elliptic curve $E/\mathbb{Q}$ of conductor $3^2 \cdot 1637$ and $3^3 \cdot 1637$. On the other hand, we have $H(23, 27) = 1637^2$, corresponding to an $E/\mathbb{Q}$ of conductor $N = 3^3 \cdot 1637$ (the conductor of $E_{1,0}$ is $2 \cdot 3^3 \cdot 1637$).

Example 14.3. Consider $F = (5, 4, -120, 85)_4$ with $\delta_3 = 43 \cdot 101^2$. Again, the equation $j_0 = j(x, y)$ has no solutions. We have $H(7, -6) = 2^4 \cdot 3 \cdot 43^2 \cdot 101$, corresponding to an $E/\mathbb{Q}$ of conductor $N = 3^3 \cdot 43 \cdot 101^2$. Here, the conductor of $E_{1,0}$ is precisely $5N$.

14.2.3. *Sextic forms.* For sextic forms (and the same remark applies to forms of higher degree), there are a number of viable ways to identify distinguished forms in a given $\mathrm{GL}_2(\mathbb{Z})$-equivalence class (see e.g. Cremona and Stoll [51], and Edwards [13]). These reduction theories are quite involved, however, and we will content ourselves with a simplistic search for examples, by considering only forms with small naive height (i.e. $\max |\alpha_i|$). Such a search reveals a number of examples which potentially satisfy the hypotheses of Corollary 8.5 (in that the corresponding Thue-Mahler equations have no small solutions). We have not actually attempted to run the algorithm of Tzanakis and de Weger for such forms; the dependence upon the degree of the form (or more specifically, upon the number of fundamental units in the field corresponding to $F(x, 1)$) is a severe one.

Remark 14.4. Our search revealed many sextics which could be obtained from a sextic in the list below by a matrix transformation over $\mathbb{Z}$ (not necessarily with unit determinant). We also found cases where the matrix transformation could not be defined over $\mathbb{Z}$, but the resulting fields are isomorphic. In our list, we have suppressed such examples.

Form	$\lvert\delta_4\rvert$	Level(s)
$(3,2,25,-40)_6$	$2^2 \cdot 331$	$2^2 \cdot 331$
$(3,8,20,50)_6$,	227	$2^\delta \cdot 227,\ \delta \in \{2,3\}$
$(3,4,-105,-540)_6$	$2^2 \cdot 239$	$2^3 \cdot 239$
$(5,4,-40,-140)_6$	$2^2 \cdot 491$	$2^2 \cdot 491$
$(3,4,20,10)_6$,	251	$2^\delta \cdot 251,\ \delta \in \{2,3\}$
$(3,1,-10,-80)_6$	$2^2 \cdot 397$	$2^\delta \cdot 397,\ \delta \in \{1,3\}$
$(5,9,5,40)_6$	419	$2^3 \cdot 419$
$(5,2,-70,180)_6$	$2^2 \cdot 439$	$2^2 \cdot 439$
$(3,8,20,140)_6$	$2^2 \cdot 907$	$2^2 \cdot 907$
$(3,5,-140,530)_6$	$2^2 \cdot 461$	$2^\delta \cdot 461,\ \delta \in \{1,3\}$
$(7,6,20,50)_6$	523	$2^\delta \cdot 523,\ \delta \in \{2,3\}$
$(7,6,-15,120)_6$	$2^2 \cdot 3^3 \cdot 41$	$2^2 \cdot 3^3 \cdot 41$
$(3,4,-25,10)_6$	557	$2^\delta \cdot 557,\ \delta \in \{2,3\}$
$(3,2,-20,50)_6$	571	$2^\delta \cdot 571,\ \delta \in \{2,3\}$

In all these cases, the levels we encounter are sufficiently small that data for all elliptic curves are available. It is again an easy matter to check that for each of these sextic forms F and all j-invariants j_0 of the elliptic curves at the levels corresponding to F, we have that the associated equation $j_0 = j(x,y)$ has no rational roots. As before, we conclude that the corresponding superelliptic equations have no solutions for suitably large prime exponents.

Remark 14.5. As noted previously, squares of sextic Klein forms arise as covers of cubic Klein forms. By way of example, the first sextic form in the above table

$$F(x,y) = 3x^6 + 2x^5y + 25x^4y^2 - 40x^3y^3 - 55x^2y^4 + 6xy^5 - 29y^6$$

satisfies $(3F(x,y))^2 = G(p(x,y), q(x,y))$ with

$$\begin{aligned} G(u,v) &= 3u^3 + u^2v + 5uv^2 - 2v^3 \\ p(x,y) &= 3x^4 - 10x^2y^2 + 16xy^3 + 11y^4 \\ q(x,y) &= 4y(3x^3 + x^2y + 5xy^2 - 2y^3). \end{aligned}$$

The cubic form G has discriminant $-3^2 \cdot 331$, while the discriminant of F is not divisible by 3. In fact, there are no cubic Klein forms such that one of the corresponding level is $2^2 \cdot 331$. Similar remarks hold for the other sextic forms in the table.

14.2.4. *Duodecic forms.* A routine search, similar to that for sextic forms, reveals some examples of Klein forms of index 5 and moderate corresponding levels, for

which the Thue-Mahler equations have no small solutions.

Form	$\lvert\delta_5\rvert$	Level(s)
$(2,3,-385,4125)_{12}$	$5^3\cdot 61$	$5\cdot 61$
$(2,1,-165,1375)_{12}$	$5^2\cdot 101$	$5\cdot 101$
$(2,3,66,-1100)_{12}$	$3^3\cdot 5\cdot 7^2$	$3^3\cdot 5\cdot 7$
$(2,5,-473,5115)_{12}$	$5\cdot 331$	$5\cdot 331$
$(2,3,-264,1760)_{12}$	$3^3\cdot 5^2\cdot 13$	$3^3\cdot 5\cdot 13$
$(2,5,-308,2640)_{12}$	$5\cdot 379$	$5\cdot 379$
$(2,1,-198,-1540)_{12}$	$5^2\cdot 89$	$5^2\cdot 89$
$(2,3,143,-1155)_{12}$	$5^3\cdot 97$	$5^2\cdot 97$
$(2,1,-88,1760)_{12}$	$5^3\cdot 107$	$5^2\cdot 107$
$(2,1,-66,220)_{12}$	163	$5^\delta\cdot 163,\ \delta\in\{0,1\}$
$(2,3,-66,-1100)_{12}$	$3^5\cdot 17$	$3^5\cdot 17$

In each case, all corresponding elliptic curve data are available. Again, one checks that, for each of these duodecic forms F and all j-invariants j_0 of the elliptic curves at the levels corresponding to F, the associated equations $j_0 = j(x,y)$ and $J(j_0) = j(x,y)$ have no rational roots.

14.3. **Small bounds for the exponents.** If not only all elliptic curves but actually all newforms at the appropriate levels are known, one can compute explicit bounds for the prime exponent l which, in practice, turn out to be much smaller than that provided by our main theorem. Again, there is no need to invoke all the machinery we have constructed – we can simply apply standard congruences provided by the modular method. In this subsection, we highlight some examples of forms from Section 14.2 for which we are able to compute all newforms at the appropriate levels.

In order to apply the modular machinery, we require the mod l Galois representation in question to be irreducible. Thanks to Mazur [31], we know that this is automatically the case if $l > 163$. In fact, assuming only $l > 13$, we obtain a like result, provided only that we exclude finitely many possibilities for the corresponding j-values. As is well-known, these are given by j in the following set :

$$\{-17\cdot 373^3/2^{17}, -17^2\cdot 101^3/2, -2^{15}\cdot 3^3, -7\cdot 11^3, -7\cdot 137^3\cdot 2083^3,\\ -2^{18}\cdot 3^3\cdot 5^3, -2^{15}\cdot 3^3\cdot 5^3\cdot 11^3, -2^{18}\cdot 3^3\cdot 5^3\cdot 23^3\cdot 29^3\}.$$

It is straightforward to check if the j-invariant of a given Frey-Hellegouarch curve can be equal to one of these values. In each of the examples we consider here, this is not the case.

In what follows, we tabulate various Klein forms, together with a set $\mathfrak{S}$ of prime exponents $l > 13$ for which we have not ruled out the possibilities of (2) having a solution using standard congruences coming from the modular method (i.e. (37) and (38) for small primes p). For $l > 13$, $l \notin \mathfrak{S}$, our conclusion is, in each case, that (2) has no (nonzero) integer solutions.

Form	$\mathfrak{S}$
$(3,2,-7,-3)_3$	$\{23,67\}$
$(3,4,2,6)_3$	$\{19\}$
$(3,4,10,6)_3$	$\emptyset$
$(2,51,369,73)_4$	$\{19,37\}$
$(3,2,25,-40)_6$	$\emptyset$
$(5,4,-40,-140)_6$	$\emptyset$
$(5,9,5,40)_6$	$\emptyset$
$(2,3,66,-1100)_{12}$	$\emptyset$
$(2,3,-264,1760)_{12}$	$\emptyset$
$(2,3,-66,-1100)_{12}$	$\{17\}$

With work, we can likely eliminate possible solutions corresponding to (some) primes $l \leq 13$ or $l \in \mathfrak{S}$. Our aim here, however, is merely to demonstrate that the bounds on l provided by our main theorem are, in specific applications, overly pessimistic.

15. Ackowledgments

The authors would like to thank Ryotaro Okazaki, Peter Olver and the anonymous referee for numerous valuable suggestions.

Appendix A. Conductor calculations

A.1. **Quartic Klein forms at $p = 2$.** To find minimal (integral) models for our Frey-Hellegouarch curves in case F is a Klein form of index 3 with Δ_F odd, we consider quadratic twists of (20) over $\mathbb{Q}((-1)^{\delta/2})$ (where δ is as defined in (27)), translated as follows (we write H for $H(x,y)$ and G for $G(x,y)$ for concision):

$$E_{x,y} \; : \; Y^2 + Y = X^3 + \frac{3\,H}{16}X + \frac{(-1)^\delta G - 16}{64}, \tag{66}$$

if $H(x,y)$ is even,

$$E_{x,y} \; : \; Y^2 + XY + Y = X^3 - X^2 + \frac{3\,H - 5}{16}X + \frac{(-1)^\delta G - 3H - 17}{64}, \tag{67}$$

if $H(x,y) \equiv 7 \pmod{16}$, and

$$E_{x,y} \; : \; Y^2 + XY = X^3 - X^2 + \frac{3\,H + 3}{16}X + \frac{(-1)^\delta G - 3H - 1}{64}, \tag{68}$$

if $H(x,y) \equiv 15 \pmod{16}$.

Corresponding quantities associated to the models $E_{x,y}$ in (66) – (68) are

$$\begin{aligned} \Delta(x,y) &= 3^3\,\delta_3\,F(x,y)^3 \\ c_4(x,y) &= -3^2 H(x,y) \\ c_6(x,y) &= (-1)^{\delta+1}\,3^3\,G(x,y)/2 \\ j(x,y) &= \frac{-3^3\,H(x,y)^3}{\delta_3 F(x,y)^3}. \end{aligned}$$

It is by no means obvious that these curves have integral coefficients. To see that they do, note that

$$\begin{aligned} H(x,y) \equiv \quad & (8\alpha_0\alpha_2 - \alpha_1^2)x^4 + (8\alpha_0\alpha_3 + 4\alpha_1\alpha_2)x^3y + (2\alpha_1\alpha_3 + 4\alpha_2^2)x^2y^2 \\ & + (8\alpha_1\alpha_4 + 4\alpha_2\alpha_3)xy^3 + (8\alpha_2\alpha_4 - \alpha_3^2)y^4 \pmod{16}. \end{aligned}$$

If α_1 is even, then, from the assumption that δ_3 is odd, necessarily α_0 and α_3 are odd, whereby from (7), $4 \mid \alpha_1$ and $2 \mid \alpha_2$. It follows that

$$H(x,y) \equiv 8x^3y + 2\alpha_1\alpha_3x^2y^2 + 4\alpha_2\alpha_3xy^3 - \alpha_3^2y^4 \pmod{16},$$

and hence either $H(x,y) \equiv -1 \pmod 8$ or $H(x,y) \equiv 0 \pmod{16}$. Via symmetry, we reach a like conclusion if α_3 is even. On the other hand, if α_1 and α_3 are both odd, then (7) implies that $\alpha_1\alpha_3 \equiv -1 \pmod 4$ and $\alpha_2 \equiv 1 \pmod 2$, whereby, from (28), precisely one of α_0, α_4, say α_0, is even. We thus have that either $H(x,y) \equiv -1 \pmod 8$ or that $H(x,y)$ is even whereby xy is necessarily odd, and

$$H(x,y) \equiv -(\alpha_1 - \alpha_3)^2 + 4 \equiv 0 \pmod{16}.$$

If $H(x,y)$ is even, then necessarily, from Proposition 2.1, $F(x,y)$ is odd, whereby, from (12), $G(x,y) \equiv 16 \pmod{32}$. We claim that $(-1)^\delta G(x,y) \equiv 16 \pmod{64}$.

If α_2 is odd, then (7) implies that $\alpha_1\alpha_3 \equiv -1 \pmod 4$ and so, from (28), α_0 and α_4 have opposite parity. Since $F(x,y)$ is odd, it follows that xy is even, contradicting the fact that

$$H(x,y) \equiv -\left(\alpha_1x^2 + \alpha_3y^2\right)^2 \equiv 0 \pmod 4. \tag{69}$$

We may thus assume that α_2 is even and hence, from (7), (28) and (69), that either $4 \mid \alpha_1$, $2 \mid y$ and $2 \nmid \alpha_0$, or $4 \mid \alpha_3$, $2 \mid x$ and $2 \nmid \alpha_4$. In the first case,

$$G(x,y) \equiv 16\,\alpha_3 \pmod{64},$$

while, in the second,

$$G(x,y) \equiv -16\,\alpha_1 \pmod{64}.$$

In each case, $(-1)^\delta G(x,y) \equiv 16 \pmod{64}$.

Next, suppose that $H(x,y)$ is odd (so that $H(x,y) \equiv -1 \pmod 8$). From (7), it is straightforward to check that

$$(-1)^\delta G - 3H \equiv 1 \text{ or } 17 \pmod{64},$$

for $H \equiv -1$ or $7 \pmod{16}$, respectively (where we note that the choice of δ ensures that $(-1)^\delta G \equiv -2 \pmod 8$).

A.2. **Conductor calculations at $p \,\|\, \delta_n$.** In many cases where a prime $p \,\|\, \delta_n$, for a given Klein form F, we are able to be precise about the reduction at p of our Frey-Hellegouarch curves E_{x_0,y_0}. Let, as usual, $F(x,y) = \sum_{i=0}^k \alpha_i x^{k-i}y^i$ denote a Klein form of index n.

Lemma A.1. *Let p be a prime with $p \nmid \alpha_0$ and suppose $p \,\|\, \delta_n$. Furthermore, assume that $p \neq 2$ (if $n = 4$), and that $p > 5$ (if $n = 5$). Then*

$$F \equiv \alpha_0\,(x - ay)^{k-1}(x - by) \pmod p$$

for some $a, b \in \mathbb{Z}$ with $a \not\equiv b \pmod p$. In particular, for the Hessian of F, we have

$$H \equiv -\alpha_0^2\,(a-b)^2(x - ay)^{2k-4} \pmod p.$$

Proof. The formula for H follows immediately from that for F. Since $p \mid \delta_n$ (and hence Δ_F), F necessarily has a repeated factor modulo p. If no such factor is linear, this places severe restrictions upon the coefficients of F (modulo p). Using the relations (7), (8) and (9), we see that this is in fact impossible, so F must have a double root over $\mathbb{F}_p$. After a linear translation, we can assume that $\alpha_k \equiv \alpha_{k-1} \equiv 0 \pmod{p}$. Again appealing to (7), (8) and (9), we find that also $\alpha_{k-2} \equiv \ldots \equiv \alpha_2 \equiv 0 \pmod{p}$. It remains, then, to show that $p \nmid \alpha_1$. Assume to the contrary that $p \mid \alpha_1$. In case the index $n = 2$ or 3, the formulae for δ_n immediately imply that $p^2 \mid \delta_n$, a contradiction. In case the index $n = 4$ or 5, (8) and (9) yield $p^2 \mid a_6$ (if $(n, p) = (4, 5)$ or $(n, p) = (5, 11)$, this is less immediate and one must take care). Except when $(n, p) = (5, 11)$, we may thus conclude that $p^2 \mid \delta_n$, again a contradiction. In the remaining case, one obtains $11 \mid a_1, 11^2 \mid a_2$, and $11^3 \mid a_i$, for $i = 3, 4, 5, 6$, which is enough to imply that $11^2 \mid \delta_5$. This proves the lemma. □

Proposition A.2. *Let p be a prime with $p \nmid \alpha_0 n$ and suppose $p \| \delta_n$. If the index $n = 2$ or 3, then the Frey-Hellegouarch curve E_{x_0,y_0} associated to a solution of equation (2) has multiplicative reduction at p.*

Proof. In the model for our Frey-Hellegouarch curve we have $p \mid \Delta(x_0, y_0)$, so it suffices to prove that $p \nmid c_4(x_0, y_0)$. This would follow if $p \nmid H(x_0, y_0)$. If we suppose $p \mid H(x_0, y_0)$, then Lemma A.1 implies that $p \mid x_0 - ay_0$, and so $p \mid F(x_0, y_0)$. Since $l > 1$, it follows that $p^2 \mid F(x_0, y_0)$. Using a linear translation, we can assume, without loss of generality, that $a \equiv 0 \pmod{p}$ and so $F \equiv \alpha_0 x^{k-1}(x - by) \pmod{p}$, i.e. $\alpha_2 \equiv \ldots \equiv \alpha_k \equiv 0 \pmod{p}$. The fact that $p \mid x_0$, together with $p^2 \mid F(x_0, y_0)$, therefore implies that $p^2 \mid \alpha_k$. From the expressions for δ_n given in Section 2, we may conclude, in case $n = 2$ or 3, that $p^2 \mid \delta_n$, a contradiction. □

For index $n = 4$ and 5, we do not obtain quite such a strong result. Indeed it can happen that the root $x_0 - ay_0$ modulo p lifts to a root in $\mathbb{Z}_p$. Furthermore, there are actually cases where (with similar assumptions to those of the lemma and proposition above) the Frey-Hellegouarch curve actually has additive reduction at p. Even in such instances, however, we are, in a certain sense, not too far from multiplicative reduction.

Proposition A.3. *Let p be a prime with $p \nmid \alpha_0 n$ and suppose $p \| \delta_n$, for $n = 4$ or 5. Furthermore, assume that $p > 3$ if $n = 5$. Then the Frey-Hellegouarch curve E_{x_0,y_0} associated to a solution of (2) with $l > 7$, or its quadratic twist over $\mathbb{Q}(\sqrt{\pm p})$, has multiplicative reduction at p.*

Proof. Since $p \neq 2$, it suffices to prove that E_{x_0,y_0} has potentially multiplicative reduction, i.e. $\nu_p(j(x_0, y_0)) < 0$. We calculate

$$\nu_p(j(x_0, y_0)) = 3\,\nu_p(H(x_0, y_0)) - n\,\nu_p(F(x_0, y_0)) - 1.$$

From Lemma A.1, if $p \nmid F(x_0, y_0)$, then $p \nmid H(x_0, y_0)$, in which case $\nu_p(j(x_0, y_0)) = -1 < 0$. It remains only to consider when $p \mid F(x_0, y_0)$. If $n = 4$ (so that $k = 6$), then, from Proposition 2.1, $\nu_p(H(x_0, y_0)) \leq k(k-1)/3{=}10$. Together with $\nu_p(F(x_0, y_0)) \geq l$, we obtain $\nu_p(j(x_0, y_0)) < 0$. A similar argument, in case $n = 5$, leads to a like conclusion, at least for $l \geq 29$. To treat smaller values of l, note that there certainly exist (nondegenerate) $x, y \in \mathbb{Z}$ with $p \nmid F(x, y)$, so that $\nu_p(j(x, y)) = -1$. It follows from the theory of Tate curves that $5 \mid \#\rho_5^{E_{x,y}}(I_p)$, where $I_p \subset \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ denotes an inertia subgroup of p. Since $\rho_5^{E_{x,y}} \simeq \rho_5^{E_{x_0,y_0}}$,

we also have $5 \mid \#\rho_5^{E_{x_0,y_0}}(I_p)$. On the other hand, if E_{x_0,y_0} has potentially good reduction at p, then no primes ≥ 5 can divide $\#\rho_5^{E_{x_0,y_0}}(I_p)$. This contradiction concludes the proof. □

References

[1] M. A. Bean, An isoperimetric inequality for the area of plane regions defined by binary forms, *Compositio Math.* 92 (1994), 115–131.

[2] M. Bennett, J. Ellenberg and N. Ng, The Diophantine equation $A^4 + 2^\delta B^2 = C^n$, *Int. J. Number Theory* 6 (2010), 1–27.

[3] F. Beukers, The generalized Fermat equation, notes available at `www.math.uu.nl/people/beukers/Fermatlectures.pdf`

[4] C. Breuil, B. Conrad, F. Diamond and R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises, *J. Amer. Math. Soc.* 14 (2001), 843–939.

[5] H. Carayol, Formes modulaires et représentations galoisiennes à valeurs dans un anneau local complet, *Contemp. Math.* 165 (1994), 213–237.

[6] J. E. Cremona, Reduction of binary cubic and quartic forms, *LMS J. Comput. Math.* 2 (1999), 62–92.

[7] S. Dahmen, Classical and modular methods applied to Diophantine equations, PhD thesis, University of Utrecht, 2008. Permanently available at `igitur-archive.library.uu.nl/dissertations/2008-0820-200949/UUindex.html`

[8] H. Darmon and A. Granville, On the equations $z^m = F(x, y)$ and $Ax^p + By^q = Cz^r$, *Bull. London Math. Soc.* 27 (1995), 513–543.

[9] H. Darmon and L. Merel, Winding quotients and some variants of Fermat's last theorem, *J. Reine Angew. Math.* 490 (1997), 81–100.

[10] H. Davenport, On the class-number of binary cubic forms. I. *J. London Math. Soc.* 26 (1951), 183–192; ibid, 27 (1952), 512.

[11] ———, On the class-number of binary cubic forms. II. *J. London Math. Soc.* 26 (1951), 192–198.

[12] L.V. Dieulefait and J. Jimenez, Solving Fermat-type equations $x^4 + dy^2 = z^p$ via modular Q-curves over polyquadratic fields, *J. Reine Angew. Math.* 633 (2009), 183–196.

[13] J. Edwards, Platonic Solids and Solutions to $X^2 + Y^3 = dZ^r$, PhD thesis, University of Utrecht, 2005. Permanently available at `igitur-archive.library.uu.nl/dissertations/2006-0208-200155/UUindex.html`

[14] ———, A complete solution to $X^2 + Y^3 + Z^5 = 0$, *J. Reine Angew. Math.*, 571 (2004), 213–236.

[15] J. Ellenberg, Galois representations attached to $\mathbb{Q}$-curves and the generalized Fermat equation $A^4 + B^2 = C^p$, *Amer. J. Math.*, 126 (2004), 763–787.

[16] N. D. Elkies, ABC implies Mordell, *Internat. Math. Res. Notices*, 7 (1991), 99–109.

[17] P. Erdős, Arithmetical properties of polynomials, *J. London Math. Soc.* 28 (1953), 416–425.

[18] G. Faltings, Endlichkeitssätze für abelsche Varietäten über Zahlkörpern, *Invent. Math.* 73 (1983), 349–366.

[19] T. Fisher, The Hessian of a genus one curve, *Proc. London Math. Soc.*, to appear.

[20] P. Gordan, Beweis, dass jede Covariante und Invariante einer binären Form eine ganze Function mit numerischen Coefficienten einer endlichen Anzahl solcher Formen ist, *J. Reine Angew. Math.* 69 (1868), 323–354.

[21] ———, Vorlesungen über Invariantentheorie, Teubner, Leipzig, 1887.

[22] A. Granville, Smooth numbers: computational number theory and beyond, Algorithmic Number Theory, MSRI Publications, Volume 44, 2008, 267–323.

[23] D. Hilbert, Theory of Algebraic Invariants, Cambridge Mathematical Library, 1993.

[24] A. Hoshi and K. Miyake, A geometric framework for the subfield problem of generic polynomials via Tschirnhausen transformation, Number theory and applications, 65–104, Hindustan Book Agency, New Delhi, 2009.
[25] F. Klein, Lectures on the icosahedron and the solution of equations of the fifth degree, Dover publications, New York, 1956 (translation of original 1884 edition)
[26] A. Kraus, Majorations effectives pour l'équation de Fermat généralisée, *Canad. J. Math.* 49 (1997), no. 6, 1139–1161.
[27] J. C. Lagarias, H. L. Montgomery and A. M. Odlyzko, A bound for the least ideal in the Chebotarev density theorem, *Invent. Math.* 54 (1979), no. 3, 271–296.
[28] G. Lettl, A, Pethő and P. Voutier, Simple families of Thue inequalties, *Trans. Amer. Math. Soc.* 351 (1999), 1871–1894.
[29] W. J. Leveque, On the equation $y^m = f(x)$, *Acta Arith.* 9 (1964), 209–219.
[30] G. Martin, Dimensions of the space of cusp forms and newforms on $\Gamma_0(N)$ and $\Gamma_1(N)$, *J. Number Theory* 112 (2005), no. 2, 298–331.
[31] B. Mazur, Rational isogenies of prime degree, *Invent. Math.* 44 (1978), 129–162.
[32] L. Merel, Normalizers of split Cartan subgroups and supersingular elliptic curves, Diophantine geometry, 237–255, CRM Series, 4, Ed. Norm., Pisa, 2007.
[33] F. Momose, Rational points on the modular curves $X_{split}(p)$, *Compositio Math.* 52 (1984), 115–137.
[34] L. J. Mordell, On numbers represented by binary cubic forms, *Proc. London Math. Soc.* 48 (1943), 198–228.
[35] P. Morton, Characterizing cyclic cubic extensions by automorphism polynomials, *J. Number Theory* 49 (1994), 183–208.
[36] I. Papadopoulos, Sur la classification de Néron des courbes elliptiques en caractéristique résiduelle 2 and 3, *J. Number Theory* 44 (1993), 119–152.
[37] P. Olver, Classical Invariant Theory, Cambridge University Press, 1999.
[38] K. A. Ribet, On modular representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms, *Invent. Math.* 100 (1990), 431–476.
[39] ———, Report on mod l representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, in Motives, Proc. Symp. Pure Math. **55**:2 (1994), 639–676.
[40] K. Rubin and A. Silverberg, Families of elliptic curves with constant mod p representations, Elliptic curves, modular forms, & Fermat's last theorem (Hong Kong, 1993), 148–161, Ser. Number Theory, I, Int. Press, Cambridge, MA, 1995.
[41] ———, Mod 2 representations of elliptic curves, *Proc. Amer. Math. Soc.* 129 (2001), no. 1, 53–57.
[42] A. Schinzel and R. Tijdeman, On the equation $y^m = P(x)$, *Acta Arith.* 31 (1976), 199–204.
[43] J.-P. Serre, Propriété galoisiennes des points d'ordre fini des courbes elliptiques, *Invent. Math.* **15** (1972), 259–331.
[44] ———, Quelques applications du théorème de densité de Chebotarev, *Publications mathématique de l'I.H.É.S.*, tome 54 (1981), 123–201.
[45] T. Shintani, On Dirichlet series whose coefficients are class numbers of integral binary cubic forms. *J. Math. Soc. Japan* **24** (1972), 132–188.
[46] ———, On zeta-functions associated with the vector space of quadratic forms. *J. fac. Sci. Univ. Tokyo, Sec. Ia* **22** (1975), 25–66.
[47] X (C. L. Siegel), The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + \ldots + k$ *J. London Math. Soc.* 1 (1926), 66–68.
[48] C. L. Siegel, Einige Anwendungen diophantischer Approximationen, *Abh. Preuss. Akaad. Wiss. Phys. Math. Kl.* (1929), 41–69.
[49] A. Silverberg, Explicit families of elliptic curves with prescribed mod N representations, Modular forms and Fermat's last theorem (Boston, MA, 1995), 447–461, Springer, New York, 1997.

[50] H. Stender, Lösbare Gleichungen $ax^n - by^n = c$ und Grundeinheiten für einige algebraische Zahlkörper vom Grade $n = 3, 4, 6$, *J. Reine Angew. Math.* **290** (1977), 24–62.

[51] M. Stoll and J.E. Cremona, On the reduction theory of binary forms, *J. Reine Angew. Math.* **565** (2003), 79–99.

[52] J. Thunder, On cubic Thue inequalities and a result of Mahler, *Acta Arith.* 83 (1998), 31–44.

[53] N. Tzanakis and B.M.M. de Weger, How to explicitly solve a Thue-Mahler equation, *Compositio Math.* 84 (1992), 223-288.

[54] I. Wakabayashi, Cubic Thue inequalities with negative discriminant, *J. Number Theory* 97 (2002), 222–251.

[55] ———, Simple families of Thue inequalities, *Ann. Sci. Math. Québec* 31 (2007), 211–232.

[56] ———, Number of solutions for cubic Thue equations with automorphisms, *Ramanujan J.* 14 (2007), 131–154.

[57] L.C. Washington, A family of cyclic quartic fields arising from modular curves, *Math. Comp.* 57 (1991), 763–775.

[58] A. Wiles, Modular elliptic curves and Fermat's Last Theorem, *Ann. Math* 141 (1995), 443–551.