Computational aspects of class group actions and applications to post-quantum cryptography

Houben, M.R.

Chapter 6

Generalized class polynomials

This chapter consists of a paper written together with Marco Streng. It has been published as

Marc Houben and Marco Streng. Generalized class polynomials. *Res. Number Theory*, 8(4):Paper No. 103, 26, 2022. `https://doi.org/10.1007/s40993-022-00400-2`.

Both authors of this paper contributed equally to the work.

Compared to the published version, we corrected minor typos. The numbering (of e.g. theorems and definitions) in the published version is different.

Acknowledgements

We thank Karim Belabas, Peter Bruin, Andreas Enge, David Kohel, Filip Najman, and Andrew Sutherland for helpful discussions and Edgar Costa for helping us use the LMFDB. We would further like to thank the organisers Samuele Anni, Valentijn Karemaker, and Elisa Lorenzo Garcia of AGC^{2}T 2021 for the inspiring setting in which the first ideas for this project came to be. Finally, we would like to thank the anonymous referees for their many comments that led to improvements of the exposition.

Abstract

The Hilbert class polynomial has as roots the j-invariants of elliptic curves whose endomorphism ring is a given imaginary quadratic order. It can be used to compute elliptic curves over finite fields with a prescribed number of points. Since its coefficients are typically rather large, there has been continued interest in finding alternative modular functions whose corresponding class polynomials are smaller. Best known are Weber's functions, which reduce the size by a factor of 72 for a positive density subset of imaginary quadratic discriminants. On the other hand, Bröker and Stevenhagen showed that no modular function will ever do better than a factor of 100.83. We introduce a generalization of class polynomials, with reduction factors that are not limited by the Bröker-Stevenhagen bound. We provide examples matching Weber's reduction factor. For an infinite family of discriminants, their reduction factors surpass those of all previously known modular functions by a factor at least 2.

6.1 Introduction

The *Hilbert class polynomial* $H_D[j]$ of the imaginary quadratic order $\mathcal{O}$ of discriminant D is the minimal polynomial of the j-invariant of an elliptic curve with endomorphism ring $\mathcal{O}$. It is a defining polynomial of the ring class field of $\mathcal{O}$ and can be used for constructing elliptic curves over a finite field with a given number of points. Its coefficients are however rather large, which limits its practical usefulness. Already in 1908, Weber [37] therefore introduced alternative *class invariants* to be used instead of j, which resulted in *class polynomials* with coefficients that have roughly 1/72 of the digits of the coefficients of the Hilbert class polynomial for certain discriminants.

There has been continued interest in alternative class invariants ever since (e.g. [2, 30, 18, 17, 31, 8, 10, 11, 4, 14, 12, 9]). None however matched, let alone surpassed, the factor 72 of Weber's functions. Moreover, Bröker and Stevenhagen [4] showed that no class invariant will ever do better than a factor 100.83. Under Selberg's eigenvalue conjecture [32, Conjecture 1], this bound reduces to 96.

We introduce *generalized (multivariate) class polynomials*, define an appropriate notion of their *reduction factor*, and show that this notion indeed gives a measure of their "size" compared to the Hilbert class polynomial (Section 6.3). Contrary to classical class polynomials, the reduction factors of generalized class polynomials are not limited by the Bröker-Stevenhagen bound.

We give a family of generalized class polynomials for which we prove that the reduction factor matches Weber's 72 for a large range of values of D, including infinitely many values of D where no reduction of 36 or better was previously known (Section 6.4). We also give an example that possibly achieves the factor 120 (Remark 6.7.6).

Though the focus of this paper is on introducing the generalized class invariants and studying their height, we also give a preliminary analysis indicating that the height reduction leads to a speed-up in their computation (Section 6.6), and we show how to use them for constructing elliptic curves over finite fields (Section 6.5).

6.2 Generalized class polynomials

Definition 6.2.1 By a *modular curve over* $\mathbf{Q}$ we mean a smooth, projective, geometrically irreducible curve C over $\mathbf{Q}$ together with a map $\psi : \mathbf{H} \to C(\mathbf{C})$ from the upper half space $\mathbf{H} \subset \mathbf{C}$ with the following property. There exists a positive integer N such that for every function $f \in \mathbf{Q}(C)$, the function $f \circ \psi$ is a modular function for $\Gamma(N)$ with all q-expansion coefficients in $\mathbf{Q}^{\mathrm{ab}}$.

We identify f with $f \circ \psi$ and we identify ψ with the induced morphism of curves $X(N) \to C$. △

For an order $\mathcal{O}$ in an imaginary quadratic number field K, we denote by $K_{\mathcal{O}}$ the associated ring class field. Let f be a modular function and $\tau \in \mathbf{H}$ imaginary quadratic, say a root of $aX^2 + bX + c$ for coprime integers a, b, c. The pair (f, τ) is called a *class invariant* for the imaginary quadratic order $\mathcal{O} = \mathbf{Z}[a\tau]$ if $f(\tau)$ lies in the ring class field $K_{\mathcal{O}}$. The *discriminant* D of the class invariant is the discriminant of

$\mathcal{O}$. The Galois group G of $K(f(\tau))/K$ is isomorphic via the Artin map to a quotient of the Picard group $\mathrm{Cl}(\mathcal{O})$. Associated to a class invariant is its minimal polynomial over K, also known as the *class polynomial*,

$$H_\tau[f] := \prod_{\sigma \in G} \big(X - \sigma(f(\tau))\big) \quad \in K[X].$$

Under additional restrictions, class polynomials can sometimes be shown to have coefficients in $\mathbf{Q}$ (cf. [9, Thm. 4.4], [13, Thm. 5.4]); in that case we call the class polynomials *real*. Oftentimes, a modular function admits class invariants for an infinite family of discriminants, determined by a certain congruence condition ([31], [9, Thm. 4.3]). Sometimes the discriminant uniquely determines the class polynomial for a given modular function.

Example 6.2.2 The modular j-function admits a unique class polynomial for any discriminant $D < 0$, called the *Hilbert class polynomial* $H_D[j] := H_\tau[j]$. It can be seen as a function on $\mathbf{P}^1$ whose zeros are the j-invariants of elliptic curves with CM by the imaginary quadratic order of discriminant D and whose poles are restricted to the point at infinity. ☆

We propose a generalization of class polynomials, seen as functions on modular curves of higher genus, for which the classical class polynomials can be viewed as the genus zero case. We will mostly restrict ourselves to the case of genus one, as this will make notation considerably less complicated. We discuss the arbitrary genus case in Section 6.7. Let C be a modular curve over $\mathbf{Q}$ with a smooth Weierstrass model $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, and suppose that $(x, \tau), (y, \tau)$ are class invariants for some imaginary quadratic $\tau \in \mathbf{H}$. Consider $G = \mathrm{Gal}(K(x(\tau), y(\tau))/K)$ and $m = \#G$. If we denote by $\mathcal{D}$ the divisor of the unique point at infinity of C, then $\mathcal{L}(\infty\mathcal{D})$ has a basis $b_0 = 1, b_1 = x, b_2 = y, b_3 = x^2, b_4 = xy, b_5 = x^3, b_6 = x^2y, \ldots$ (ordered by ascending degree). There exist $a_i \in K$, not all zero, such that

$$\sum_{i=0}^{m} a_i b_i(\tau) = 0. \tag{6.1}$$

In fact, up to scaling by an element of $K^\times$, there exists a unique function $F_\tau[C] = \sum_{i=0}^{m} a_i b_i \in K(C)$ such that

$$\mathrm{div} F_\tau[C] = \left[\sum_{\sigma \in G} (\sigma(\psi(\tau)))\right] + \left(-\sum_{\sigma \in G} \sigma(\psi(\tau))\right) - (m+1)\mathcal{D}. \tag{6.2}$$

Definition 6.2.3 We call $F_\tau[C]$ as in (6.2) a *generalized class function* for τ. The associated *generalized class polynomial* is the unique $H_\tau[C] \in K[X, Y]$ of degree ≤ 1 in Y such that $H_\tau[C](x, y) = F_\tau[C]$. △

We note that the polynomial $H_\tau[C]$ depends on the choice of x and y, but we leave this out of the notation. In Section 6.7 (and in particular Definition 6.7.3) we will allow more general divisors $\mathcal{D}$ and bases $\mathcal{B}$, leading to more general functions $F_\tau[C,\mathcal{B}]$ and polynomials $H_\tau[C,\mathcal{B}]$.

Definition 6.2.4 We call the point $P = \sum_{\sigma\in G}\sigma(\psi(\tau)) \in C(K)$ the *Heegner point* of the class function F. △

If the Heegner point P is the point at infinity, then $a_m = 0$. Otherwise, the point $-P$ is a zero of F. In particular, if $P = -(0,0)$, then $a_0 = 0$.

For $N \in \mathbf{Z}_{>0}$, we denote by $X^0(N)$ the smooth, projective, geometrically irreducible curve over $\mathbf{Q}$ with function field consisting of the modular functions for the modular group $\Gamma^0(N) = \{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbf{Z}) \mid b \equiv 0 \pmod N\}$ that have rational q-expansion. We denote by $X^0_+(N)$ the quotient of $X^0(N)$ by the Fricke-Atkin-Lehner involution $z \mapsto -N/z$, and write $\eta(z)$ for the Dedekind η-function

$$\eta(z) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n), \quad \text{where} \quad q = \exp(2\pi i z).$$

Example 6.2.5 Consider the genus one modular curve $C := X^0_+(119)$. Its conductor as an elliptic curve is 17 (Cremona label 17a4)[1]. A Weierstrass model for E is given by[2]

$$y^2 + 3xy - y = x^3 - 3x^2 + x, \tag{6.3}$$

where $x, y \in \mathbf{Q}(C)$ have respective q-expansions

$$\begin{aligned} x &= q^{-2} + q^{-1} + 1 + q + 2q^2 + 2q^3 + 3q^4 + 3q^5 + 4q^6 + 5q^7 + \ldots, \\ y &= q^{-3} + 1 + 2q + 2q^2 + 4q^3 + 4q^4 + 7q^5 + 9q^6 + 12q^7 + \ldots, \\ & \quad \text{where this time } q = \exp(2\pi i z/119). \end{aligned}$$

The "double eta quotient" $\mathfrak{w}_{7,17}$ given by

$$\mathfrak{w}_{7,17}(z) = \frac{\eta(z/7)\eta(z/17)}{\eta(z)\eta(z/119)} \tag{6.4}$$

[1] One way to deduce this is as follows. Using the command `J0(119).decomposition()` in SageMath [36] one finds that C has conductor 17. For each of the Weierstrass models of the now finitely many possible curves [23], there are finitely many options for the divisor of the function $\mathfrak{w}_{7,17}$ given by (6.4). The curve C has two rational CM points (both of discriminant -19), so given a possible Weierstrass model together with a possible divisor for $\mathfrak{w}_{7,17}$, one can first determine $\mathfrak{w}_{7,17}$ as a function of the Weierstrass coordinates x, y by evaluating in one CM point, and then determine whether it has the expected value in the other CM point. This process excludes all but one of the options, and we at once in fact deduce both the Weierstrass model (6.3) and the relation between $\mathfrak{w}_{7,17}$ and x and y (6.5).

[2] We note that a slightly "simpler" Weierstrass model $v^2 + uv + v = u^3 - u^2 - u$ exists by taking $u = x$ and $v = -y - 2x$, but the given model (6.3) turns out to yield slightly better practical reduction factors (see Section 6.4.5).

is invariant under the action of $\Gamma^0(N)$ [27, Thm. 1] and the Fricke-Atkin-Lehner involution [11, Thm. 2], hence also forms an element of the (rational) function field of C. It is related to x and y by

$$\mathfrak{w}_{7,17} = -y + x^2 - x. \tag{6.5}$$

The curve $X_+^0(119)$ has two cusps, and they are both rational. In the given Weierstrass model, these correspond to the point $(0,0)$ and the point at infinity. Numerical examples of generalized class polynomials specifically for $X_+^0(119)$ are given in Section 6.4.5. We will treat this curve as our main test case in the rest of the paper. ☆

6.3 Estimates and reduction factors

6.3.1 Reduction factors

We define the *reduction factor* of a modular curve C to be

$$r(C) = \frac{\deg(j : X(N) \to \mathbf{P}^1)}{\deg(\psi : X(N) \to C)}. \tag{6.6}$$

In the case $C = \mathbf{P}^1$, we denote this number also by $r(\psi)$ and our notation and terminology coincide with that of [4]. The number $r(\psi)^{-1}$ is denoted by $\widehat{c}(\psi)$ in [8] and by $c(\psi)$ in [9]. Bröker and Stevenhagen [4, Theorem 4.1][3] show $r(\psi) \leq 32768/325 \leq 100.83$. Under Selberg's eigenvalue conjecture, one can even prove $r(\psi) \leq 96$. The best known ψ achieves $r(\psi) = 72$. This result does not however apply directly to $r(C)$. For example, we have

$$r(X^0(N)) = N \prod_{p|N} (1 + \frac{1}{p}) \quad \text{and} \quad r(X_+^0(N)) = \frac{1}{2} r(X^0(N)) \quad \text{if} \quad N > 1. \tag{6.7}$$

Our main example $C = X_+^0(119)$ therefore achieves $r(C) = \frac{1}{2}(7+1)(17+1) = 72$. For (hyper)elliptic modular curves C we get $r(C) \leq 201.65$ (or $r(C) \leq 192$ under Selberg's eigenvalue conjecture), by applying the bounds to the x-function. Surprisingly, all elliptic curve quotients of $X^0(N)$ we found so far have $r \leq 72$ (Section 6.4.7). In Section 6.7 we will discuss higher-genus curves, which allow for unbounded $r(C)$.

Remark 6.3.1 In the applications we have in mind, the reduction factor is the main source of improvement in computational efficiency. It is important to note, however, that this number $r(C)$ does not tell the complete story, even in the "classical" setting ($C \cong \mathbf{P}^1$), for example for the following reasons.

1. There are many challenges when computing class polynomials, and even more with generalized class polynomials. See Section 6.6.

[3] The arXiv version v1 of [4] has weaker bounds than the final publication and needs to be combined with [21, Appendix 2] to get the same result.

2. In the CM method (Section 6.5), we will want to find a j-invariant in $\mathbf{F}_p$ from a point in $C(\mathbf{F}_p)$. This is done using the minimal polynomial of the j-function over $\mathbf{Q}(C)$, known as the *modular polynomial* (Lemma 6.5.1). This works best if the degree of j over $\mathbf{Q}(C)$ is small. For example, this degree is 1 for $C = X^0(N)$, is 2 for $C = X^0_+(N)$, and ranges from 1 to 20 in [9, Table 7.1], making $X^0_+(119)$ a good choice in this respect.

3. If the (generalized) class polynomial is not real, then its coefficients lie in an imaginary quadratic extension of $\mathbf{Q}$; roughly doubling its bit size. This issue can be avoided by imposing additional restrictions on C or τ, see Sections 6.4.2 and 6.4.3.

On the other hand, there are two important tricks that may be used in complementary directions, providing computational improvements beyond the reduction factor $r(C)$:

1. Under some constraints, typically when all primes dividing the level of the modular curve ramify in the CM field, both the degree and height of the class polynomial are cut in half. This happens for example in the record-computation of [14] for the Atkin invariant A_{71} when 71 divides the discriminant, leading to class polynomials that are $2^2 \cdot 36 = 144$ times smaller than the Hilbert class polynomial (note that the reduction factor $r(A_{71})$ is 36 in this case). The same trick also applies to generalized class polynomials, see Section 6.4.4, which in the case of $X^0_+(119)$ leads to a factor $2^2 \cdot 72 = 288$ in size reduction.

2. When the class number is composite, one can decompose the ring class field into a tower of fields whose defining polynomials have smaller degrees, also leading to a significant speed-up in the CM method [35].

These last two tricks only work when the class number is composite. We expect both of them to work well for generalized class polynomials, so will mainly restrict to the case of prime class number in our examples, as this more clearly illustrates the role of the parameter $r(C)$. ♢

The goal of the rest of this section is to show under some hypotheses that the reduction factor $r(C)$ is indeed an asymptotic reduction factor of the size of the polynomials involved. For that, we will first introduce the appropriate notions of “size”.

6.3.2 Measures of polynomials and heights of their roots

For a polynomial $A \in \mathbf{C}[X]$, let $|A|_1$ (resp. $|A|_\infty$) be the sum (resp. maximum) of the absolute values of the coefficients of A. The *Mahler measure* of a polynomial $A = a\prod_{i=1}^n (X - \alpha_i) \in \mathbf{C}[X]$ is

$$\mathcal{M}(A) = |a| \prod_i \max\{1, |\alpha_i|\}.$$

Lemma 6.3.2 *We have*

$$\begin{array}{rcccl} |A|_\infty & \leq & |A|_1 & \leq & (n+1)|A|_\infty, \\ \mathcal{M}(A) & \leq & |A|_1 & \leq & 2^n \mathcal{M}(A), \end{array}$$

$$\begin{array}{rcl} |\log |A|_1 - \log |A|_\infty| & \leq & \log(n+1), \\ |\log |A|_\infty - \log(\mathcal{M}(A))| & \leq & n \log(2). \end{array}$$

Proof. The first two inequalities are by definition and the third is Equation (6) of [24]. For its converse, observe that we have $|AB|_1 \leq |A|_1 |B|_1$, and hence also $|A|_1 \leq |a| \prod_i \max\{2, 2|\alpha_i|\} \leq 2^n \mathcal{M}(A)$. Then take logarithms. □

For an element α in a number field L of degree n, we define its *(absolute logarithmic) height* to be

$$h(\alpha) = \frac{1}{n} \sum_v \max\{0, \log |\alpha|_v\},$$

where the sum ranges over the Archimedean and non-Archimedean absolute values, suitably normalized (that is, those denoted $|| \cdot ||_v$ in [19, §B.1]). If α is a root of an irreducible $A \in \mathbf{Z}[X]$ of degree n, then we have

$$\log(\mathcal{M}(A)) = nh(\alpha). \tag{6.8}$$

Remark 6.3.3 Another measure for the complicatedness of A would be its total bit size, or the sum s of the logarithms of the absolute values of the nonzero coefficients. We will instead focus on $|A|_\infty$ for the following reasons.

First of all, for computational purposes, it is more useful to look at $p = \deg(A) \cdot \log |A|_\infty$, as the required precision (or number of primes with the CRT approach) is proportional to $\log |A|_\infty$ and the number of computations to do with that precision is proportional to $\deg(A)$.

Secondly, we get the impression from numerical computations that s is close to p. For example, the value of s/p is spread out over the interval $(0.75, 0.9)$ for the larger discriminants in both Section 6.4.5 and Example 6.7.4.

Finally, it is hard to prove lower bounds on s other than $s \geq \log |A|_\infty$, as it seems to already be hard to show that a sufficient proportion of coefficients is nonzero. ♢

6.3.3 Proof of the height reduction

Theorem 6.3.4 *Let C be a modular curve over $\mathbf{Q}$ and suppose that C is an elliptic curve of rank 0 with Weierstrass coordinates x and y. Suppose that $\tau \in \mathbf{H}$ ranges over a sequence of imaginary quadratic points for which C yields real generalized class polynomials $H_\tau[C]$, and with*

$$\frac{h(j(\tau))}{\log(\log(\# \operatorname{Cl}(\mathcal{O})))} \to \infty. \tag{6.9}$$

Scale each $H_\tau[C]$ such that it has coprime coefficients in $\mathbf{Z}$. *Then*

$$d \cdot \frac{\log |\mathbf{F}|_\infty}{\log |H_\tau[j]|_\infty} \to \frac{1}{r(C)},$$

where d is the degree of $K_\mathcal{O}$ over $K(\psi(\tau))$.

Remark 6.3.5 We argue that the hypothesis (6.9) is very reasonable. Under GRH, we have

$$\#\,\mathrm{Cl}(\mathcal{O}) = O(\sqrt{|D|}\log(\log|D|)), \tag{6.10}$$

where D is the discriminant of $\mathcal{O}$ (see [22, 9.Theorem 1 and 11. on page 371], suitably extended to arbitrary D.) Moreover, [8, §6.2] gives the approximation $\log|H_\tau[j]|_\infty \approx \pi\sqrt{|D|}S(D)$, with $S(D) = \sum_Q a^{-1}$, where the sum ranges over reduced primitive quadratic forms $Q = ax^2 + bxz + cz^2$ of discriminant D. We now give a heuristic lower bound of this sum on average over all $|D| \leq X$. We have $\sum_D S(D) \approx \sum_Q a^{-1}$, where this time the sum is taken over all reduced quadratic forms of negative discriminant $> -X$ (using the heuristic that imprimitive forms have a negligible contribution). As we are only computing a lower bound, we may restrict to $a \leq \sqrt{X/8}$. Then b ranges from $-a$ to a, and c ranges from a or $a+1$ to $\lfloor (X+b^2)/(4a)\rfloor$; a range that contains at least $\lfloor X/(8a)\rfloor$ integers. This yields at least roughly $X/4$ values of b and c for each a, hence $\sum_D S(D)$ is roughly at least $(X/4)\sum_{a^2\leq X/8} a^{-1} \geq \frac{1}{8}X\log(X)$. It follows that the average $S(D)$ is at least proportional to $\log|D|$. Thus, for "average" $S(D)$, we have that $\log|H_\tau[j]|_\infty$ is at least proportional to $\sqrt{|D|}\log|D|$. Combined with (6.10), (6.8), and Lemma 6.3.2, we find for such D that $h(j(\tau))/\log(\log(\#\,\mathrm{Cl}(\mathcal{O})))$ is at least proportional to $\log|D|/(\log(\log|D|))^2$. We thus see that (6.9) indeed holds for "average" $S(D)$. ♢

Theorem 6.3.4 is the analogue of the following result.

Theorem 6.3.6 (cf. Enge-Morain [8]) *Let f be a modular function and suppose that $\tau \in \mathbf{H}$ ranges over a sequence of imaginary quadratic points for which (f, τ) is a class invariant with $h(j(\tau)) \to \infty$. Then $d \cdot \frac{\log|H_\tau[f]|_\infty}{\log|H_\tau[j]|_\infty} \to \frac{1}{r(f)}$, where d is the degree of $K_\mathcal{O}$ over $K(f(\tau))$.*

The goal of the remainder of Section 6.3 is to prove Theorem 6.3.4. We start with a proof of Theorem 6.3.6.

Proof. Let m be the degree of $K(f(\tau))$ over K and let $n = dm$ be the degree of $K_\mathcal{O}$ over K. By Lemma 6.3.2 and (6.8), we get $|\frac{1}{n}\log|H_\tau[j]|_\infty - h(j(\tau))| \leq \log(2)$ and $|\frac{d}{n}\log|H_\tau[f]|_\infty - h(f(\tau))| \leq \log(2)$.

As $h(j(\tau)) \to \infty$, we also get

$$\frac{h(f(\tau))}{h(j(\tau))} \to \frac{1}{r(f)} \tag{6.11}$$

by [19, Proposition B.3.5(b)]. Altogether, this gives the result. □

Proposition 6.3.7 *Let C be a modular curve over $\mathbf{Q}$ and suppose that C is an elliptic curve of rank 0 with Weierstrass coordinates x and y. For every imaginary quadratic $\tau \in \mathbf{H}$ for which C yields a real generalized class polynomial $H_\tau[C]$, let m be the degree of $K(\psi(\tau))$ over K and let $d' \in \{1,2\}$ be the degree of $K(\psi(\tau))/K(x(\tau))$. Scale each $H_\tau[C]$ such that it has coprime coefficients in $\mathbf{Z}$. Then we have*

$$\left|\log|H_\tau[C]|_\infty - \frac{d'}{2}\log|H_\tau[x]|_\infty\right| < B\max\{1, m\log(\log(m))\},$$

for some constant B that only depends on C and the choice of Weierstrass model.

Proof. **We first put the equation for C in a nice form.** We have $C : y^2 + g(x)y = f(x)$. Without loss of generality we have $g = 0$ and $f \in \mathbf{Z}[X]$ monic of odd degree such that $f(z) \leq -1$ for all real $z \leq 0$. Indeed, we obtain $g = 0$ by the substitution $y' = y + \frac{1}{2}g(x)$, then do scalings $x' = vx$ and $y' = wy$ to make f integral and (thanks to its odd degree) monic, and then do a substitution $x' = x + c$ to make $f(z) \leq -1$ for all $z \leq 0$. This affects $H_\tau[C] = A + BY$ and $H_\tau[x]$ as follows. The first substitution changes A into $A + \frac{1}{2}g(X)B$, the second changes A into $A(vX)$ and B into $wB(vX)$, and the third changes A into $A(X + c)$. Each of these substitutions change $\log(\max\{|A|_1, |B|_1\})$ at most by $O(m)$, as does clearing the denominators afterwards.

Next, we relate a norm of $H_\tau[C]$ to $H_\tau[x]$. The extra elliptic curve point $(a/b^2, c/b^3) := \sum_{\sigma\in G}\sigma(\psi(\tau)) \in C(\mathbf{Q})$ from (6.2) (which is minus the Heegner point) is torsion by our assumption that C has rank 0. There are finitely many torsion points in $C(\mathbf{Q})$, hence finitely many possibilities for the polynomial $T = b^2X - a$. Writing $H_\tau[C] = A(X) + B(X)Y$, we get that $N(H_\tau[C]) = A(X)^2 + (-f(X))B(X)^2$ has the same divisor as the primitive polynomial $H_\tau[x]^{d'} \cdot T$, hence there is a constant $s \in \mathbf{Z} \setminus \{0\}$ with $N(H_\tau[C]) = sH_\tau[x]^{d'} \cdot T$.

We claim that $s = \pm 1$. If not, take a prime $p \mid s$ and consider the highest-weight term of $(H_\tau[C] \bmod p)$, where X has weight 2 and Y has weight $\deg(f)$. This gives rise to the highest-degree term of $(N(H_\tau[C]) \bmod p)$, which is therefore nonzero, a contradiction.

Now we use interpolation to bound $H_\tau[C]$ in terms of $H_\tau[x]$. We will choose interpolation points $z = g(i) \leq 0$. Note that for $z \leq 0$ we have

$$A(z)^2, B(z)^2 \leq A(z)^2 + (-f(z))B(z)^2 = N(H_\tau[C]) \leq \max\{1, |z|\}^m |H_\tau[x]|_1^e |T|_1,$$

and since there are finitely many polynomials T, we get

$$\log|A(z)|, \log|B(z)| \leq \frac{m}{2}\max\{0, \log|z|\} + \frac{d'}{2}\log|H_\tau[x]|_1 + O(1).$$

Interpolation then gives, for $P \in \{A, B\}$:

$$P(X) = \sum_{i=1}^{k} P(g(i)) \prod_{j \neq i} \frac{X - g(j)}{g(i) - g(j)}, \tag{6.12}$$

where $k = \deg(P) + 1 = O(m)$.

Taking $g(u) = -\log(eu)^2$, we find $|g(i) - g(j)| \geq |i - j| \min_{z \in [1,k]} |g'(u)| = |i - j| \min_{u \in [1,k]} 2\frac{\log(eu)}{u} = 2|i - j| \frac{\log(ek)}{k}$. So for each i there are at most $k/\log(k)$ values of $j \neq i$ with $|g(i) - g(j)| < 1$ and each of them has $|g(i) - g(j)| \geq 1/k$. We get

$$\log \prod_{j \neq i} \frac{1}{|g(i) - g(j)|} \leq (k/\log(k)) \log(k) = k = O(m).$$

For the other factors in (6.12), we have that $\log |X - g(j)|_1 \leq \log(1 + \log(em)^2) = O(\log(\log(m)))$, so $\log \prod_j |X - g(j)|_1 = O(m \log(\log(m)))$, as well as $\log |P(g(i))| \leq \frac{d'}{2} \log |H_\tau[x]|_1 + O(m \log(\log(m)))$. Taking the sum in (6.12) gives another $+\log(k)$, so that the end result is $\log |P(X)|_1 \leq \frac{d'}{2} \log |H_\tau[x]|_1 + O(m \log(\log(m)))$. By Lemma 6.3.2, this also holds with $|\cdot|_\infty$, which proves the upper bound on $\log |H_\tau[C]|_\infty$.

For the lower bound, note that $H_\tau[x]^{d'}$ is a factor of $Q = A^2 - f(X) \cdot B^2$, and we have $|Q|_1 \leq |A|_1^2 + |f|_1 |B|_1^2 \leq |f|_1 (m+1)^2 |H_\tau[C]|_\infty^2$. Using the fact that $\mathcal{M}$ is multiplicative by definition and is related to $|\cdot|_1$ and $|\cdot|_\infty$ by Lemma 6.3.2, we get exactly what we need: $d' \log |H_\tau[x]|_\infty \leq d' \log \mathcal{M}(H_\tau[x]) + O(m) \leq \log \mathcal{M}(Q) + O(m) \leq \log |Q|_1 + O(m) \leq 2 \log(|H_\tau[C]|_\infty) + O(m)$. □

Proof of Theorem 6.3.4. Denote again by $n = \#\mathrm{Cl}(\mathcal{O})$ the degree of $K_\mathcal{O}$ over K. First we apply Theorem 6.3.6 to x and get $dd' \frac{\log |H_\tau[x]|_\infty}{\log |H_\tau[j]|_\infty} \to \frac{2}{r(C)}$. Proposition 6.3.7, together with the hypothesis $h(j(\tau))/(n \log(\log(n))) \to \infty$, gives $\frac{1}{d'} \frac{\log |H_\tau[C]|_\infty}{\log |H_\tau[x]|} \to \frac{1}{2}$ (as in the proof of Theorem 6.3.6). The product of these two limits gives the result. □

Remark 6.3.8 Theorem 6.3.4 states that asymptotically the effect of the choice of a model of the curve C is negligible, as is the effect of replacing f by $2f$ or $f + 1$ or any other element of $\mathbf{Q}(f)$ in Theorem 6.3.6.

However, in practice the error terms can be quite large and depend on these choices. For example, if f is integral over $\mathbf{Z}[j]$ then $H_\tau[f]$ is monic, and if f^{-1} is integral over $\mathbf{Z}[j]$, then f has zero constant coefficient. This can make a difference in practical examples as it forces the coefficients at the beginning and end to be small, though this improvement is negligible asymptotically by the theorems. See also Remark 6.3.3. ◊

6.4 Class invariants for $X^0(N)$ and $X_+^0(N)$

In this section we assume that C is a quotient over $\mathbf{Q}$ of $X^0(N)$; in other words, C is a smooth, projective, geometrically irreducible curve over $\mathbf{Q}$ with function field consisting only of modular functions for $\Gamma^0(N)$ that have rational q-expansion. We will show how to obtain generalized class functions for every discriminant $D < 0$ that is square modulo $4N$ (Section 6.4.1).

In some cases we get further reductions from class invariants generating subfields of $K_{\mathcal{O}}$ or from real class polynomials (Sections 6.4.2–6.4.4).

In Sections 6.4.5–6.4.6 we study what this means for $X^0_+(119)$ and in Section 6.4.7 we look for more examples of elliptic curve quotients of $X^0(N)$.

6.4.1 Class invariants for $X^0(N)$

The following result does not require C to be an elliptic curve, except that (unless C is an elliptic curve) one needs to read the definitions in Section 6.7 for the parts about generalized class polynomials.

Proposition 6.4.1 (based on Schertz [31]) *Let $C = (C, \psi)$ be a quotient over $\mathbf{Q}$ of $X^0(N)$ and let $D < 0$ be a square modulo $4N$.*

There exist $a, b, c \in \mathbf{Z}$ with $a, c > 0$, $b^2 - 4ac = D$, $N \mid c$, and $\gcd(a, N) = \gcd(a, b, c) = 1$. Choose such a, b, c, let $\tau \in \mathbf{H}$ be a root of $aX^2 + bX + c$, with order $\mathcal{O} = \mathbf{Z}[a\tau]$, which has discriminant D. Then we have

$$\psi(\tau) \in C(K_{\mathcal{O}}),$$

thus giving rise to a generalized class polynomial $H_\tau[C]$.

*The Galois orbit of $\psi(\tau)$ can be computed as follows. There exists an N-*system*, that is, there exist $\tau_1, \ldots, \tau_n \in \mathbf{H}$ such that $(\tau_i\mathbf{Z} + \mathbf{Z})_i$ is a system of representatives of $\mathrm{Cl}(\mathcal{O})$ and such that τ_i is a root of $a_iX^2 + b_iX + c_i$ with $\gcd(a_i, N) = \gcd(a_i, b_i, c_i) = 1$ and $b_i \equiv b \bmod 2N$. Moreover, for any such choice, we have*

$$\mathrm{Gal}(K_{\mathcal{O}}/K) \cdot \psi(\tau) = \{\psi(\tau_i) : i = 1, \ldots, n\}.$$

Proof. For the existence of a, b, c, take an arbitrary square root b of D modulo $4N$, let $a = 1$, and $c = (b^2 - D)/4$. Then the existence of an N-system is [31, Proposition 3].

For any $f \in \mathbf{Q}(C)$, Theorem 4 of Schertz [31] states $f(\tau) \in K_{\mathcal{O}} \cup \{\infty\}$ and gives the $\mathrm{Gal}(K_{\mathcal{O}}/K)$-orbit as $\{g(N\tau_i) : i\}$, under an additional condition on the function $f(1/z)$. However, the condition on $f(1/z)$ is not needed, as stated in Theorems 3.9 and 4.4 of [13]. This proves the result. □

6.4.2 Real class polynomials from ramification

There are some situations in which we can actually get real class polynomials, cutting the total required bit size in half. The first such situation is when all primes dividing N ramify.

Proposition 6.4.2 (based on Enge-Morain [9]) *Let $C = (C, \psi)$ be a quotient over $\mathbf{Q}$ of $X^0(N)$ and let $D < 0$ be a discriminant divisible by N if N is odd and by $4N$ if N is even.*

There exist $a, b, c \in \mathbf{Z}$ with $a, c > 0$, $N \mid b, c$, $\gcd(a, N) = 1$, and $b^2 - 4ac = D$. Choose such a, b, c, let $\tau \in \mathbf{H}$ be a root of $aX^2 + bX + c$, with order $\mathcal{O} = \mathbf{Z}[a\tau]$, which has discriminant D.

Then the $\operatorname{Gal}(K_{\mathcal{O}}/K)$-orbit of $\psi(\tau)$ is stable under complex conjugation, and hence we may take $H_\tau[C] \in \mathbf{Q}[X, Y]$.

Proof. If D is odd, take $b = N$, and if D is even, take $b = 0$. If N is even, then we find $4N \mid b^2 - D$. If N is odd, then we find both $4 \mid b^2 - D$ and $N \mid b^2 - D$, hence also $4N \mid b^2 - D$. Let $a = 1$ and $c = (b^2 - D)/4$.

The complex conjugate of $\psi(\tau)$ is $\psi(-\overline{\tau})$ by the fact that the q-expansion coefficients are real. Here $-\overline{\tau}$ is a root of $aX^2 - bX + c$, and as $N \mid b$, we can choose the N-system in Proposition 6.4.1 in such a way that $-\overline{\tau} = \tau_i$ for some i. This proves the result. □

6.4.3 Real class polynomials from $X_+^0(N)$

The second situation in which we get real class polynomials is when working with quotients of $X_+^0(N)$.

Proposition 6.4.3 (based on Theorem 3.4 of Enge-Schertz [10]) *In the situation of Proposition 6.4.1, suppose furthermore that C is a quotient of $X_+^0(N)$, and that $\gcd(c/N, N) = 1$.*

Then the $\operatorname{Gal}(K_{\mathcal{O}}/K)$-orbit of $\psi(\tau)$ is stable under complex conjugation, and hence we may take $H_\tau[C] \in \mathbf{Q}[X, Y]$.

Proof. The complex conjugate of $\psi(\tau)$ is $\psi(-\overline{\tau})$ by the fact that the q-expansion coefficients are real. As ψ is invariant under the Fricke-Atkin-Lehner involution, this in turn is $\psi(\tau')$ with $\tau' = N/\overline{\tau}$, a root of $(c/N)X^2 + bX + Na$. As c/N is coprime to N, we can choose the N-system in Proposition 6.4.1 in such a way that $\tau' = \tau_i$ for some i. This proves the result. □

To use this result, we will need $\gcd(c/N, N) = 1$, which can be achieved most of the time, as follows.

Lemma 6.4.4 *If D is a square modulo $4N$ and $D = F^2 D_0$ for a negative fundamental discriminant D_0 and a positive integer F coprime to N, then there exist a, b, c as in Proposition 6.4.1 with $\gcd(c/N, N) = 1$.*

More generally, let $D < 0$ be a square modulo $4N$. Then there exist a, b, c as in Proposition 6.4.1 with $\gcd(c/N, N) = 1$ if and only if all of the following do not hold.

1. *there exists a prime $p \mid N$ with $\operatorname{ord}_p(N)$ odd and $\operatorname{ord}_p(D) > \operatorname{ord}_p(4N)$,*
2. *$m := \operatorname{ord}_2(N) > 0$ and D is of the form $2^{m+1}d$ with $d \equiv 1 \pmod 4$,*
3. *$m := \operatorname{ord}_2(N) > 0$ and D is of the form $2^m d$ with $d \equiv 1 \pmod 8$.*

Proof. The triple (a,b,c) exists if and only if there exists $b \in \mathbf{Z}$ such that for all $p \mid N$: $\mathrm{ord}_p(b^2 - D) = \mathrm{ord}_p(4N)$.

Suppose that we are not in case (1), (2), or (3). By the Chinese remainder theorem, it suffices to find one $b \in \mathbf{Z}$ for each $p \mid N$. So let $p \mid N$ be prime and let $k = \mathrm{ord}_p(4N)$ and $l = \mathrm{ord}_p(D)$. If $k < l$, then as we are not in case (2), we find that k is even, and we can take $b = p^{(k/2)}$. If $k = l$, then we can take $b = p^e$ with $e > k/2$. Now the case $k > l$ remains. As D is a square modulo $4N$, there exists $b_0 \in \mathbf{Z}$ be such that $D \equiv b_0^2 \pmod{4N}$. If $\mathrm{ord}_p(b_0^2 - D) = \mathrm{ord}_p(4N)$, then we are done, so suppose $\mathrm{ord}_p(b_0^2 - D) > k$.

Note that $2\,\mathrm{ord}_p(b_0) = l$, hence l is even. Let $b = b_0 + p^e$ with e to be determined later. We get $b^2 - D = (b_0^2 - D) + 2p^e b_0 + p^{2e}$, and the terms have valuation $> k$, $e + (l/2) + \mathrm{ord}_p(2)$, $2e$ respectively.

If $p \neq 2$, then we choose $e = k-(l/2)$, so $2e = k+(k-l) > k$, hence $\mathrm{ord}_p(b^2-D) = k$. If $p = 2$ and $k > l + 2$, then we choose $e = k - (l/2) - 1$, so $2e = k + (k - l - 2) > k$, hence $\mathrm{ord}_p(b^2 - D) = k$.

Now only the case $p = 2$ with $k - l \in \{1, 2\}$ remains. Write $d = 2^{-l}D$ and $b_1 = 2^{-(l/2)}b_0$, so b_1 is odd and $b_1^2 - d$ is divisible by 2^{k-l}.

In the case $k - l = 1$, we get $b_1^2 - d \equiv 0 \pmod 2$, and we claim that this is nonzero modulo 4. Indeed, b_1^2 is 1 modulo 4 and d is not (as we are not in case (2)). Therefore $\mathrm{ord}_2(b_1^2 - d) = 1$ and $\mathrm{ord}_2(b_0^2 - D) = 1 + l = k$, so we take $b = b_0$.

In the case $k - l = 2$, we get $b_1^2 - d \equiv 0 \pmod 4$, and we claim that this is nonzero modulo 8. Indeed, b_1^2 is 1 modulo 8, and d is not (as we are not in case (3)). Therefore $\mathrm{ord}_2(b_1^2 - d) = 2$ and $\mathrm{ord}_2(b_0^2 - D) = 2 + l = k$, so we take $b = b_0$.

Conversely, suppose that b exists.

In case (1), we have $\mathrm{ord}_p(D) > \mathrm{ord}_p(4N)$, hence $2\,\mathrm{ord}_p(b) = \mathrm{ord}_p(4N)$ is odd, contradiction.

In case (2), we have $\mathrm{ord}_2(b^2 - 2^{m+1}d) = m + 2$, hence $m + 1 = 2\,\mathrm{ord}_2(b) =: 2e$. Write $b = 2^e b_1$ and note $\mathrm{ord}_2(b_1^2 - d) = 1$, but $b_1^2 - d$ is 0 modulo 4.

In case (3), we have $\mathrm{ord}_2(b^2 - 2^m d) = m + 2$, hence $m = 2\,\mathrm{ord}_2(b) =: 2e$. Write $b = 2^e b_1$ and note $\mathrm{ord}_2(b_1^2 - d) = 2$, but $b_1^2 - d$ is 0 modulo 8.

It remains only to prove the first statement, for which it suffices to show that the exceptions (1), (2), and (3) all imply $\gcd(N, F) > 1$. In case (1), we see that $p^2 \mid D$ and if $p = 2$, then $p^4 \mid D$, hence $p \mid F$. In cases (2) and (3), write $D = 2^v d$ with $v \in \{m, m + 1\}$. As D is a square modulo 2^{m+2}, we find that v is even, and hence $D = (2^{v/2})^2 d$ for a discriminant d, so $2 \mid F$. □

Lemma 6.4.5 *Let N be the product of distinct odd primes $p_1, \ldots, p_k$. The negative discriminants that are a square modulo $4N$ and not in one of the exceptions of Lemma 6.4.4 have density*

$$\prod_{i=1}^{k} \frac{p_i^2 + p_i - 2}{2p_i^2}$$

in the set of all negative discriminants.

The negative fundamental discriminants that are a square modulo $4N$ (which are *not in one of the exceptions of Lemma 6.4.4) have density*

$$\prod_{i=1}^{k} \frac{p_i^2 + p_i - 2}{2(p_i^2 - 1)}$$

in the set of all fundamental negative discriminants.

Proof. Being a discriminant is the condition of being 0 or 1 modulo 4. It is equivalent to being a square modulo 4. This is independent of being a square modulo p_i that does not suffer from (1), which is happens for the $(p_i - 1)/2$ residue classes modulo p_i that are nonzero squares modulo p_i, and the $p_i - 1$ nonzero residue classes modulo p_i^2 that are zero modulo p_i. As $p_i(p_i - 1)/2 + p_i - 1 = (p_i^2 + p_i - 2)/2$, we get the first statement.

Being a fundamental discriminant means being nonzero modulo the squares of all odd primes and being $1, 5, 8, 9, 12, 13$ modulo 16. This happens for $\zeta(2)^{-1}(1-1/4)^{-1}\frac{6}{16}$ of all negative integers. In order to restrict this to products that satisfy the conditions of Lemma 6.4.4, we have to adjust the Euler product exactly by the given factor. □

For example, if $N = 119 = 7 \cdot 17$, then the numbers in Lemma 6.4.5 are > 0.2898 and $19/64 > 0.2968$.

6.4.4 Lower-degree class polynomials from ramification

In the case where all primes dividing N ramify, we get an even greater size reduction. The point $\psi(\tau)$ will then be defined over a subfield, cutting the degree of its minimal polynomial in half. This in turn also cuts the height of the coefficients of this polynomial in half, as we get $d \geq 2$ in Theorem 6.3.4. The amount of work required for computing the class polynomial, as well as the bit size of the polynomial (Remark 6.3.3), is related to the degree times the logarithm of the largest coefficient, and this product is reduced by a factor $\geq 2 \times 2 \times r(C) = 4r(C)$.

Proposition 6.4.6 (based on Enge-Schertz [12]) *Let $C = (C, \psi)$ be a quotient over $\mathbf{Q}$ of $X^0(N)$ and let $D = F^2 D_0 < 0$ be such that $N \mid D$, $\gcd(F, N) = 1$, and $D \notin \{N, 4N\}$.*

There exist $a, b, c \in \mathbf{Z}$ with $a > 0$, $N \mid b$, $c = N$, $b^2 - 4ac = D$, and $\gcd(a, b, c) = 1$. Choose such a, b, c, let $\tau \in \mathbf{H}$ be a root of $aX^2 + bX + c$, with order $\mathcal{O} = \mathbf{Z}[a\tau]$, which has discriminant D.

Let $\mathfrak{n} = ((-b + \sqrt{D})/2, a)$, and let $K_{\mathcal{O}}^{[\mathfrak{n}]}$ be the subfield of $K_{\mathcal{O}}$ fixed by the image of $\mathfrak{n}$ under the Artin map. Then $[\mathfrak{n}]$ has order 2 in $\mathrm{Cl}(\mathcal{O})$ and $\psi(\tau) \in C(K_{\mathcal{O}}^{[\mathfrak{n}]})$, where $K_{\mathcal{O}}$ has degree 2 over $K_{\mathcal{O}}^{[\mathfrak{n}]}$.

We get $m \leq \#\mathrm{Cl}(\mathcal{O})/2$ in the definition of $H_\tau[C]$, we get $H_\tau[C] \in \mathbf{Q}[X, Y]$, and we get and $d \geq 2$ in Theorem 6.3.4.

If $\mathfrak{a}_i$ are the ideals $\tau_i\mathbf{Z} + \mathbf{Z}$ of an N-system, then $\mathfrak{a}_i$ and $\mathfrak{a}_i\mathfrak{n}$ yield the same point $\psi(\tau_i)$, while $\mathfrak{a}_i^{-1}$ and $\mathfrak{a}_i^{-1}\mathfrak{n}$ yield $\overline{\psi(\tau_i)}$.

Proof. This is exactly what we get when applying [12, Theorem 9] to the coordinate functions f of C. □

6.4.5 Numerical results for $X^0_+(119)$

For the rest of this section we will return to our main Example 6.2.5, so set $N = 119 = 7 \cdot 17$. For any τ as in Proposition 6.4.3, we have $H_\tau[C] \in \mathbf{Q}[X, Y]$. By scaling, we may assume that the coefficients of $H_\tau[C]$ are integral and coprime, and that the leading coefficient (i.e. the coefficient of the monomial of highest degree as a function on C) is positive, and this uniquely determines $H_\tau[C] \in \mathbf{Z}[X, Y]$.

For any discriminant $D < 0$ coprime to N such that D is a square modulo N, there are two generalized class polynomials (depending on the choice of τ). We experimentally computed both of these for all fundamental discriminants of prime class number < 100. The main reason for restricting to prime class number is to exclude the two tricks of Remark 6.3.1; for these discriminants, the reduction factor thus provides a fair comparison with the Hilbert class polynomial. The method we employ numerically evaluates class invariants by their q-expansions, and finds a minimal polynomial relation (6.1) using lattice basis reduction (LLL). We leave faster methods for future research, but see Section 6.6 for the first ideas. Since the q-expansions can only be evaluated up to finite precision, this does not result in provably correct polynomials, although – based on heuristic estimates – they are highly unlikely to be incorrect.

A few examples of computed polynomials are listed in Table 6.1. Here, for the given discriminant D, we consistently chose τ such that its primitive equation is $X^2 + bX + (b^2 - D)/4$ with $b \in \mathbf{Z}_{>0}$ *minimal* satisfying $b^2 \equiv D \pmod{4N}$ and $\gcd((b^2 - D)/(4N), N) = 1$.

D	n	$F_\tau[C]$
-52	2	$y + 1$
-523	5	$x^3 + x^2 - 2xy - 3x - 2y$
-5347	13	$x^7 + 58x^6 - 13x^5y - 39x^5 - 143x^4y - 85x^4 - 135x^3y$ $-19x^3 - 51x^2y + 47x^2 + 7xy - 12x - y + 1$
-15139	29	$x^{15} + 1028x^{14} - 40x^{13}y + 37342x^{13} - 10557x^{12}y + 79865x^{12}$ $-167759x^{11}y - 385199x^{11} - 474165x^{10}y - 425857x^{10} - 69261x^9y$ $+345059x^9 + 493309x^8y + 309689x^8 + 168403x^7y - 132377x^7$ $-145439x^6y - 22165x^6 - 16029x^5y + 16139x^5 + 15225x^4y - 4867x^4$ $-7127x^3y - 456x^3 + 623x^2y + 423x^2 + 337xy - 65x - 64y$

Table 6.1: Some conjecturally correct generalized class functions for $C = X^0_+(119)$. The second column lists the class number n of the discriminant D.

Still assuming that $H_\tau[C]$ is scaled such that it has coprime coefficients in $\mathbf{Z}$, we denote by

$$r_A(\tau) := \frac{\log |H_\tau[j]|_\infty}{\log |H_\tau[C]|_\infty}$$

the *practical reduction factor* of τ. Under the assumption $h(j(\tau))/\log(\log(n)) \to \infty$ for $n = \#\operatorname{Cl}(\mathcal{O})$ (cf. Theorem 6.3.4) we have $d^{-1} r_A(\tau) \to r(C)$. Experimentally obtained practical reduction factors, plotted against both the class number n and $\log(|H_\tau[j]|_\infty)/\log(\log(n))$, can be seen in Figure 6.1. To visualize the role of the class number and the hypothesis $h(j(\tau))/\log(\log(n)) \to \infty$, the points of higher class number are given a darker color in the second figure.

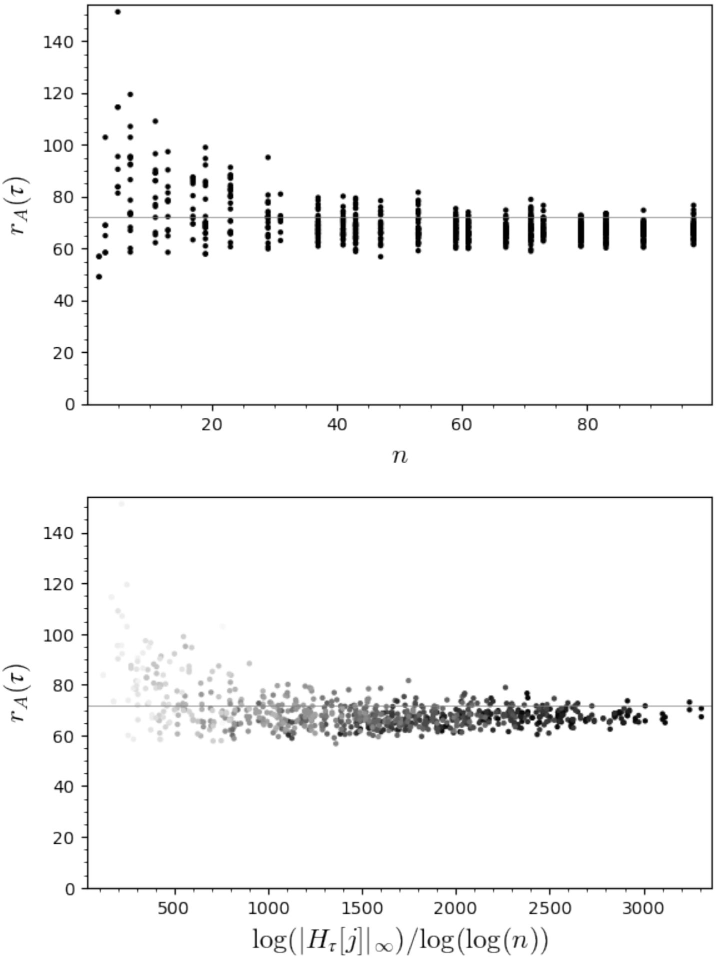


Figure 6.1: Practical reduction factors for $H_\tau[X_+^0(119)]$ for fundamental discriminants D with $\gcd(D, N) = 1$ and prime class number $n < 100$.

The values of the practical reduction factor $r_A(\tau)$ seem to be around their expected asymptotic value $r(C) = 72$ (represented by the horizontal grey line), though the convergence is not apparent; especially compared to, e.g. some classical class polynomials [8, Fig. 1]. However, in practical applications (see Section 6.5), the class numbers employed are typically several orders of magnitude higher (cf. e.g. [35]), so here we expect the speed of convergence not to cause major deviations in expected running times (cf. Section 6.6). For small class numbers, one can in practice even take advantage of this phenomenon by constructing generalized class polynomial with surprisingly good practical reduction factors by selecting a basis of $\mathcal{L}(\infty\mathcal{D})$ different from $1, x, y, x^2, xy, \dots$ (see Example 6.7.4).

6.4.6 Comparison with existing class invariants

Real class invariants typically arise subject to congruence conditions on the discriminant. For example, Weber's functions with reduction factor 72 are not known to give class invariants for discriminants $\equiv 5 \pmod 8$. The reduction factors obtained by class invariants coming from the family of (double) eta quotients $\mathfrak{w}_n$ and $\mathfrak{w}_{p,q}$ (such as the Weber function $\mathfrak{w}_2$, as well as the function $\mathfrak{w}_{7,17}$ of Example 6.2.5) have been extensively studied; cf. most notably [9]. These modular functions are not known to yield class invariants if D is not a square modulo $4n$ or $4pq$. Hence, to the best of our knowledge, they also are not applicable to discriminants $\equiv 5 \pmod 8$ as soon as n, p or q is even. Excluding these cases, the (double) eta quotient with highest known reduction factor is $\mathfrak{w}_9$, with a reduction factor of 36 [9, Table 7.1].

A less-studied generalization are *multiple eta quotients* [12], which are quotients of products of 2^k eta functions. As far as we know these do not yield reduction factors better than 36 for $k > 1$.

The only other known family of "good" class invariants (in the sense that they have large reduction factors) are the Atkin functions A_p for prime numbers p, defined to be the smallest-degree functions in $\mathcal{L}(\infty\mathcal{D})$, where $\mathcal{D}$ is the unique cusp of $X^0_+(p)$. The "best" known one here is A_{71}, again with a reduction factor of 36, owing to the fact that $X^0_+(71)$ has genus zero [14, §3]).

The curve $C = X^0_+(119)$ has a reduction factor $r(C) = 72$ and yields real class invariants whenever D is a square modulo $4 \cdot 7 \cdot 17$ and not divisible by 7^2 or 17^2. The set of such D has density $> 28.98\%$ among the set of all negative discriminants (by Lemma 6.4.5). Out of these discriminants, one-fourth are $\equiv 5 \pmod 8$. Hence, for at least $28.98\% \cdot \frac{1}{4} > 7.24\%$ of imaginary quadratic discriminants, the reduction factor exceeds the previously best known reduction factors by a factor of at least two.

Remark 6.4.7 One should note that the above comparison does not take into account the discussion of Remark 6.3.1. Most importantly, the reduction factor is *not* synonymous with the true size reduction of the class polynomials. Indeed, as noted in that remark, the record-breaking CM construction [35] uses the Atkin invariant A_{71} of reduction factor 36, because the effective size reduction of class polynomials is by a factor of roughly $2^2 \cdot 36 = 144$ for certain discriminants. However, by Section 6.4.4, the same trick applies to generalized class polynomials, leading for $X^0_+(119)$ to a size reduction of $2^2 \cdot 72 = 288$, again for a positive density subset of discriminants. In Figure 6.2 we plot the practical reductions in bit size we found compared to the Hilbert class polynomial using this trick. ◊

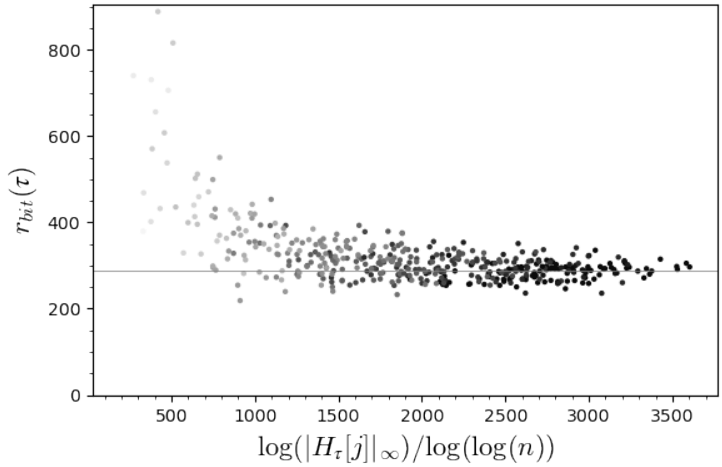


Figure 6.2: Bit-length reduction for $H_\tau[X_+^0(119)]$ for discriminants $D \equiv 0 \pmod{119}$ of class number $n < 100$.

Remark 6.4.8 Note that the "classical" class polynomial $H_\tau[x]$, arising from the function x on $X_+^0(119)$ by itself attains a reduction factor of 36 for the same 28.98% of discriminants. This beats all previously-known class invariants for a smaller subset ($\approx$ 1.2%) of discriminants: those that additionally are non-square modulo both 3 and 71. This x can be viewed as a generalisation of the Atkin functions to non-prime levels: it is the function of minimal degree in $\mathcal{L}(\infty\mathcal{D})$ for one of the cusps $\mathcal{D}$ of $X_+^0(119)$.

Similarly, the degree-two map of the hyperelliptic curve $X_+^0(191)$ (not to be confused with 119) has reduction factor 48, as observed by David Kohel in the AGC^{2}T 2021 Zulip group chat. This beats the reduction factor 32 of the Atkin function A_{191} of degree 3 on the same curve (see Example 6.7.2).

This shows that the search for generalized class invariants can even uncover new "classical" class invariants. ◊

6.4.7 More modular curves of genus one

We searched for more elliptic curves that could be used, and the results are in Tables 6.2, 6.3, and 6.4. In our search, we used the fact that $X_0(N)$ is well-studied and that there is an isomorphism $X_0(N) \to X^0(N) : z \mapsto Nz$. Surpisingly, we found lots of elliptic curves with reduction factor 72 and no elliptic curves with a greater reduction factor.

N		$g(X)$	$r(X)$	$\deg(\phi)$	$g(C)$	$r(C)$
119	$= 7 \cdot 17$	1	72	1	1	72
120	$= 2^3 \cdot 3 \cdot 5$	7	144	2	1	72
144	$= 2^4 \cdot 3^2$	5	144	2	1	72
176	$= 2^4 \cdot 11$	7	144	2	1	72
188	$= 2^2 \cdot 47$	9	144	2	1	72
131	$= 131$	1	66	1	1	66
75	$= 3 \cdot 5^2$	1	60	1	1	60
95	$= 5 \cdot 19$	1	60	1	1	60
171	$= 3^2 \cdot 19$	5	120	2	1	60
54	$= 2 \cdot 3^3$	1	54	1	1	54
81	$= 3^4$	1	54	1	1	54
90	$= 2 \cdot 3^2 \cdot 5$	4	108	2	1	54
108	$= 2^2 \cdot 3^3$	4	108	2	1	54
110	$= 2 \cdot 5 \cdot 11$	5	108	2	1	54
135	$= 3^3 \cdot 5$	4	108	2	1	54
136	$= 2^3 \cdot 17$	6	108	2	1	54
142	$= 2 \cdot 71$	8	108	2	1	54
159	$= 3 \cdot 53$	4	108	2	1	54
101	$= 101$	1	51	1	1	51
48	$= 2^4 \cdot 3$	1	48	1	1	48
56	$= 2^3 \cdot 7$	1	48	1	1	48
63	$= 3^2 \cdot 7$	1	48	1	1	48
64	$= 2^6$	1	48	1	1	48
84	$= 2^2 \cdot 3 \cdot 7$	4	96	2	1	48
96	$= 2^5 \cdot 3$	3	96	2	1	48
105	$= 3 \cdot 5 \cdot 7$	5	96	2	1	48
124	$= 2^2 \cdot 31$	6	96	2	1	48
128	$= 2^7$	3	96	2	1	48
141	$= 3 \cdot 47$	6	96	2	1	48
155	$= 5 \cdot 31$	4	96	2	1	48
191	$= 191$	2	96	2	0	48

Table 6.2: The curves $X = X^0_+(N)$ for which there exists a map $\phi : X \to C$ of degree ≤ 2 with $g(C) \leq 1$ and $r(C) \geq 48$. We used Furumoto-Hasegawa [15] and Jeon [20] to get a complete list.

In Section 6.7, we will allow curves of higher genus, which do achieve arbitrarily high values of $r(C)$. Moreover, our search is by no means exhaustive, as Tables 6.2 and 6.3 restrict to maps $\phi : X \to C$ of degree ≤ 2 and Table 6.4 only looks at one curve $X = X^0(N)$ per isomorphism class of curves C. For example, the curve $C = X^0_+(119)$ has $r(C) = 72$. However, in the Cremona database, it is listed as 17a4, and comes with a modular parametrization $\phi_{17} : X_0(17) \to C$ of degree 1, which has $r(\phi_{17}) = 18$. This is why C does not appear in Table 6.4.

N		$g(X)$	$r(X)$	$\deg(\phi)$	$g(C)$	$r(C)$
36	$= 2^2 \cdot 3^2$	1	72	1	1	72
60	$= 2^2 \cdot 3 \cdot 5$	7	144	2	1	72
72	$= 2^3 \cdot 3^2$	5	144	2	1	72
92	$= 2^2 \cdot 23$	10	144	2	1	72
94	$= 2 \cdot 47$	11	144	2	1	72
49	$= 7^2$	1	56	1	1	56
24	$= 2^3 \cdot 3$	1	48	1	1	48
32	$= 2^5$	1	48	1	1	48
42	$= 2 \cdot 3 \cdot 7$	5	96	2	1	48
48	$= 2^4 \cdot 3$	3	96	2	0	48
62	$= 2 \cdot 31$	7	96	2	1	48
69	$= 3 \cdot 23$	7	96	2	1	48

Table 6.3: The curves $X = X^0(N)$ for which there exists a map $\phi : X \to C$ of degree ≤ 2 with $g(C) \leq 1$ and $r(C) \geq 48$ and N is not already in Table 6.2. We used Ogg [28] and Bars [1] to get a complete list.

Finally, the tables are restricted to quotients of $X^0(N)$. Letting go of $X^0(N)$, we find that the genus-one modular curves $7C^1$, $8K^1$, $9H^1$, $12V^1$, $15I^1 = X_1(15)$, $16M^1$, $24J^1$, $27C^1$, $32E^1$ in the Pauli-Cummins database [6] all achieve $r(C) \in \{84, 96, 108\}$. We have not pursued these curves yet, as Proposition 6.4.1 does not apply to them.

6.5 Application: the CM method

Class polynomials are used in the *CM method* for constructing elliptic curves over finite fields with a specified characteristic polynomial of Frobenius.

The input to the CM method is a monic quadratic polynomial $P = x^2 - tx + q \in \mathbf{Z}[x]$, where q is a prime power coprime to t, and the discriminant $d = t^2 - 4q$ is negative. The output is an elliptic curve $E/\mathbf{F}_q$ with $q + 1 - t$ rational points, which has P as its characteristic polynomial of Frobenius.

The algorithm of the classical CM method (without using class invariants for now) is as follows. Let $K = \mathbf{Q}(\sqrt{d})$.

1. Compute the Hilbert class polynomial H_K of $\mathcal{O}_K$.
2. Find a root $j_0 \in \mathbf{F}_q$ of H_K (which is known to split into linear factors in $\mathbf{F}_q$).
3. Construct an elliptic curve $E/\mathbf{F}_q$ with $j(E) = j_0$. Compute all twists of E and return the one with $q + 1 - t$ rational points.

In practice, one can discard the curves for which $(q+1-t)Q \neq O$ for some random point Q, although there are also more straightforward methods to select the correct twist [29].

E	N	$r(X)$	$\deg(\phi)$	$\mathrm{rank}(E)$	$r(C)$
36a1	$2^2 \cdot 3^2$	72	1	0	72
92a1	$2^2 \cdot 23$	144	2	0	72
94a1	$2 \cdot 47$	144	2	0	72
144a1	$2^4 \cdot 3^2$	288	4	0	72
368e1	$2^4 \cdot 23$	576	8	1	72
558a1	$2 \cdot 3^2 \cdot 31$	1152	16	1	72
704a1	$2^6 \cdot 11$	1152	16	1	72
704k1	$2^6 \cdot 11$	1152	16	1	72
1728a1	$2^6 \cdot 3^3$	3456	48	1	72
1728v1	$2^6 \cdot 3^3$	3456	48	1	72
3456a1	$2^7 \cdot 3^3$	6912	96	1	72
3456e1	$2^7 \cdot 3^3$	6912	96	0	72
131a1	131	132	2	1	66
575a1	$5^2 \cdot 23$	720	12	1	60
711a1	$3^2 \cdot 79$	960	16	1	60
755b1	$5 \cdot 151$	912	16	1	57
999b1	$3^3 \cdot 37$	1368	24	1	57
49a1	7^2	56	1	0	56
1323m1	$3^3 \cdot 7^2$	2016	36	1	56
243a1	3^5	324	6	1	54
405c1	$3^4 \cdot 5$	648	12	1	54
459a1	$3^3 \cdot 17$	648	12	1	54
101a1	101	102	2	1	51
335a1	$5 \cdot 67$	408	8	1	51
591a1	$3 \cdot 197$	792	16	1	99/2
485b1	$5 \cdot 97$	588	12	1	49
723b1	$3 \cdot 241$	968	20	1	242/5
69a1	$3 \cdot 23$	96	2	0	48
105a1	$3 \cdot 5 \cdot 7$	192	4	0	48
141d1	$3 \cdot 47$	192	4	1	48
155c1	$5 \cdot 31$	192	4	1	48
213a1	$3 \cdot 71$	288	6	0	48

Table 6.4: The elliptic curves $E/\mathbf{Q}$ of conductor $<$ 500.000 such that the modular parametrization $\phi : X \to E$ according to the LMFDB [23, 5, 36] gives $r(C) \geq 66$ or gives $r(C) \geq 48$ and odd N.

As the degree and height of the Hilbert class polynomial grow quickly with the absolute value of the discriminant Δ_K of K, the CM method is only feasible for small values of $|\Delta_K|$. The record computation of [35] uses class invariants, specifically arising from the Atkin function A_{71}. Combined with the tricks listed in Remark 6.3.1 this allows to handle a case where $|\Delta_K| > 10^{16}$.

We will now describe how to apply the CM method using generalized class polynomials. Hence let C be an elliptic modular curve. Since we are working with alternative class invariants instead of the usual j-invariant, we will relate the two using *modular polynomials* as follows.

Lemma 6.5.1 *Let $d_j := [\mathbf{Q}(C,j) : \mathbf{Q}(C)]$. Then there exists a polynomial $\Psi_C = \sum_{i=0}^{d_j} f_i Z^i \in \mathbf{Z}[X,Y][Z]$ of degree d_j in Z such that*

(i) $\Psi_C(j) = 0$;

(ii) $\deg_Y(f_i) \leq 1$ for each i;

(iii) the coefficients (in $\mathbf{Z}$) of Ψ_C viewed as an element of $\mathbf{Z}[X,Y,Z]$ are coprime;

(iv) viewed as elements of $\mathbf{Q}(C)$, the f_i have at most one common zero in $C(\overline{\mathbf{Q}})$.

Furthermore, Ψ_C is unique up to sign.

Proof. Consider the minimal polynomial $\Psi_C^0 = \sum_{i=0}^{d_j} g_i Z^i \in \mathbf{Q}(C)[Z]$ of j over $\mathbf{Q}(C)$. Let

$$\mathcal{E} := \sum_{P \in C \setminus \{O\}} \min_i(\mathrm{ord}_P(g_i))(P).$$

Then $\mathcal{E} - \left(\sum_{P\in C} \mathrm{ord}_P(\mathcal{E})P\right) - (\deg(\mathcal{E})-1)(O)$ is a $\mathbf{Q}$-rational principal divisor. There is a unique function g up to $\mathbf{Q}^\times$-scaling such that $\mathrm{div}(g) = \mathcal{E}$. Dividing each g_i by g gives $g_i \in \mathcal{L}(\infty(O)) = \mathbf{Q}[x,y]$ satisfying ((iv)) and unique up to $\mathbf{Q}^\times$. Now take representatives f_i satisfying ((ii)) and scale to get ((iii)), which makes Ψ_C unique up to sign. □

For each curve C with which we would like to apply the generalized CM method, the polynomial $\Psi_C \in \mathbf{Z}[X,Y,Z]$ can be precomputed and stored. Next we need a criterion for which discriminants D yields class invariants. For example, if $C = X_+^0(N)$ then this is given by Proposition 6.4.1. Now, given a desired characteristic polynomial of Frobenius $x^2 - tx + q$ such that $D = t^2 - 4q$ satisfies this criterion, we have the following algorithm for sufficiently large $|D|$.

(1) Compute a generalized class function F of discriminant D as well as its Heegner point Q.

(2a) Find a zero $P = (x,y) \in C(\mathbf{F}_q)$ of F that is neither $-Q$ nor a common root of the polynomials $f_1, \ldots, f_{d_j}$ of Lemma 6.5.1.

(2b) Find all roots $j_0 \in \mathbf{F}_q$ of the polynomial $\Psi_C(x,y,Z) \in \mathbf{F}_q[Z]$.

(3) For each j_0, construct an elliptic curve $E/\mathbf{F}_q$ with $j(E) = j_0$ and all of its twists up to isomorphism over $\mathbf{F}_q$. Return one with $q + 1 - t$ rational points.

The main advantage compared to the classical CM method, both in terms of memory and speed, is expected to be in the (dominant) first step (1) (see Section 6.6). Out of the computationally non-dominant steps, only (2a) is less straightforward. One way to proceed would be as follows.

(i) Compute $F_x := N_{\mathbf{F}_q(C)/\mathbf{F}_q(x)}(F)$.

(ii) Find a root $x \in \mathbf{F}_q$ of F_x.

(iii) Solve for the corresponding value of y using the linear polynomial $H_\tau[C](x, Y)$, or continue with both solutions y coming from the Weierstrass equation.

Remark 6.5.2 The polynomial F_x is very close to the classical class polynomial $H_\tau[x]$; indeed, it has the same roots, together with one additional root at the x-coordinate of the Heegner point of F. The norm computation in step ((i)) is however computationally asymptotically dominated by the computation of F. ♢

6.6 The computational benefits of our invariants

6.6.1 Space complexity of the functions

The advantage of using generalized class functions lies in their size. This already gives a serious advantage when storing one or more class polynomials for later use, e.g. for various values of q in the CM method. Additionally, one would expect the smaller size to make the generalized class polynomials less expensive to compute. Again for C a modular elliptic curve with a given Weierstrass model, we present a preliminary analysis of the cost of computing a generalized class polynomial $H_\tau[C]$ when compared to the "classical" class polynomial $H_\tau[x]$ (though recall that the latter already dominates all previously-known class invariants along a positive density subset of discriminants for $C = X_+^0(119)$, cf. Section 6.4.6).

6.6.2 Speed of complex analytic computation

We now explain how to adapt the complex analytic approximation algorithm to generalized class polynomials.

To compute the classical class polynomial $H_\tau[x]$ one first evaluates $x(\tau)$ and all its conjugates, which are of the form $x_i(\tau_i)$, where x_i and τ_i can be obtained using Shimura's reciprocity law [18] or N-systems [31]. Then one multiplies the linear polynomials $X - x_i(\tau_i)$ together in a binary tree using fast multiplication algorithms.

As $H_\tau[C]$ has roughly half the height, we only need half the precision at each step. This gives a great speed-up when evaluating $x_i(\tau_i)$, but then we also need to compute $y_i(\tau_i)$. Fortunately that should only take a fraction of the time required for computing

$x_i(\tau_i)$, as we can first compute it to low precision and then obtain as many digits as desired quickly using

$$y = \frac{-g(x) + \sqrt{g(x)^2 + 4f(x)}}{2}$$

for $C : y^2 + g(x)y = f(x)$.

The **binary tree** step is harder to analyze. Instead of having polynomials $A(X) = \prod_{i \in S}(X - x_i(\tau))$ to multiply for various subsets $S \subset \{1, 2, \dots, n\}$, we will have pairs (F, Q) with $F = A(X) + B(X)Y$ and

$$\mathrm{div}(F) = \sum_{i \in S}(P_i) + (Q) - (\#S + 1)\mathcal{D}.$$

Instead of a single multiplication A_1A_2 to go from S_1 and S_2 to $S_3 = S_1 \sqcup S_2$, we now need to compute the point $Q_3 = Q_1 + Q_2$ (with the elliptic curve group law) and a function F_3 with

$$\mathrm{div}(F_3) = \sum_{i \in S}(P_i) + (Q_3) - (\#S_3 + 1)\mathcal{D} = \mathrm{div}(F_1) + \mathrm{div}(F_2) + (Q_3) + (O) - (Q_1) - (Q_2).$$

The following formula can be used:

$$F_3 = \frac{F_1\ F_2\ R \mod (Y^2 - f(X))}{(X - x(Q_1))(X - x(Q_2))}, \quad \text{where} \tag{6.13}$$

$$R = (x(Q_1) - x(Q_2))\ Y\ +\ (y(-Q_2) - y(-Q_1))\ X \tag{6.14}$$

$$+\ x(Q_2)y(-Q_1) - x(Q_1)y(-Q_2), \tag{6.15}$$

and where the reduction modulo $Y^2 - f(X)$ keeps the outcome of degree ≤ 1 in Y.

We can multiply F_1 with F_2 using three multiplications of half the degree, by the same trick that is used in Karatsuba multiplication. Indeed, let

$$C = A_1A_2, \quad D = B_1B_2, \quad \text{and} \quad E = (A_1 + B_2)(A_2 + B_1)$$

to get $F_1F_2 = (C + Df) + (E - C - D)Y$. So computing F_3 involves three polynomial multiplications of half the degree of F_1 and F_2, as well as various multiplications and long divisions by fixed-degree polynomials and various additions and subtractions. The most serious computations in the binary tree are now done with half the degree *and* half the number of digits, but three times as often, which takes 3/16th of the time with naive multiplication and still less than 3/4 of the time with quasi-linear-time multiplication. The impact of the extra additions and subtractions, as well as the extra multiplications by a linear polynomial in X and Y and long division by the denominator of (6.13) requires further analysis, but we expect this to be minor. Regardless, for large discriminants, the main bottleneck is in memory complexity (as noted in [7, Section 7]), and here we obtain an improvement of a factor of 1/2 when passing from $H_\tau[x]$ to $H_\tau[C]$.

6.6.3 Adapting the CRT method

Overview of CRT class polynomial computation

We now heuristically estimate the expected speed-up when computing $H_\tau[C]$ instead of $H_\tau[x]$ using the (currently state-of-the-art) CRT method for class polynomial computation [14, 34, 35]. We restrict to the case of C such that all q-expansion coefficients of x and y are rational, and will analyse some steps only in the main case where C is a quotient of $X_+^0(N)$. To keep our exposition simple, we will not treat the main improvement of [35], even though we do expect it to combine well with our generalized class invariants. We plan to give a more detailed account and an implementation in future work.

For the CM method, it is more efficient to directly compute class polynomials modulo q using the *online CRT* as in [34, Section 2]. In other words, we never write down $H_\tau[C] \in \mathbf{Z}[X, Y]$, but instead compute $(H_\tau[C] \bmod q) \in \mathbf{F}_q[X, Y]$ directly from $(H_\tau[C] \bmod p)$ for p in a set S of small primes. The space complexity of the CM method is then $n \log(q)$, which is independent of our choice of class function. The set S must be chosen in such a way that $\prod_{p \in S} p$ is larger than 4 times the largest coefficient.

By cutting the number of digits in half when switching from x to C, we essentially cut $\#S$ in half. If the amount of work that we do for each prime p does not grow too much, then our class function $H_\tau[C]$ yields a speed-up over the classical class polynomial $H_\tau[x]$.

What needs to be done for each p is the following.

1. Enumerate all E'' with endomorphism ring $\mathcal{O}$ and compute the appropriate points in $C(\mathbf{F}_p)$.
2. Compute $(F \bmod p)$ by putting together the information from Step 1.

In practice, for “typical” discriminants D with 9 to 14 digits, Sutherland [34, Sections 8.3 and 8.4] finds that performing Steps 1 and 2 together $\#S$ times is the dominant part of the CRT method.

We will now argue why we expect each of these steps to take (much) less than twice as long with the generalized class polynomial for suitable C. Together with the fact that our set S is only half the original size due to the reduction factor, this means that computing $H_\tau[C]$ takes less time than computing $H_\tau[x]$.

Enumerating via the Fricke involution

Step 1 is already very subtle in the case of a single class invariant f. Indeed, there could be multiple Galois orbits of values $f(\tau)$ for the same order $\mathcal{O}$, and hence multiple irreducible class polynomials $H_{\tau_i}[f] \in K[X]$. In the CRT method, one has to make sure to compute the polynomials $(H_{\tau_i}[f] \bmod p)_p$ for the same value of i, and only for τ_i for which this is a class invariant. This issue is addressed in detail in [14, Section 4].

We will first explain how to adapt one solution to our main case of quotients C of $X_+^0(N)$ where N is coprime to the conductor of $\mathcal{O}$ and $D = \mathrm{disc}(\mathcal{O})$ is a square modulo $4N$.

We adapt the method of Section 4.3 of [14] as follows. We have $\mathbf{Q}(X^0(N)) = \mathbf{Q}(j, j_N)$, where $j_N(z) = j(z/N) = j(W_N z)$ for the Fricke-Atkin-Lehner involution $W_N : z \mapsto -N/z$ (this follows for example from [33, Proposition 6.9]). In particular, every function $f \in \mathbf{Q}(C)$ for a quotient C of $X^0(N)$ can be expressed as a rational function in j and j_N. In practice, these expressions can be quite large, but (analogously to [14, Lemma 2]) we can also obtain the value $f(z)$ as a root of $\gcd(\Psi_f(X, j(z)), \Psi_{f\circ W_N}(X, j_N(z)))$ instead.

In the particular case where C is a quotient of $X^0_+(N)$, we even have $\mathbf{Q}(C) \subset \mathbf{Q}(X^0_+(N)) = \mathbf{Q}(j + j_N, j \cdot j_N)$, and we can use Ψ_f instead of $\Psi_{f\circ W_N}$.

So instead of enumerating just the j-values, we wish to link them with the corresponding j_N-values, and we do that as follows.

Suppose that N is coprime to the conductor of $\mathcal{O}$ and that D is a square modulo $4N$. Then by Lemma 6.4.4 we get $a, b, c \in \mathbf{Z}$ with $a, c > 0$, $b^2 - 4ac = D$, $N \mid c$, and $\gcd(ac/N, N) = \gcd(a, b, c) = 1$. In line with Lemma 2 of [14] we could even take $c = N$ by replacing a by ac/N. We take $z = \frac{-b+\sqrt{D}}{2a}$, $\mathfrak{n} = a\bar{z}\mathbf{Z} + N\mathbf{Z}$, and $\mathfrak{a} = z\mathbf{Z} + \mathbf{Z}$. Then we have $\mathcal{O} = az\mathbf{Z} + \mathbf{Z}$, and we find that $\mathfrak{n}$ is an invertible $\mathcal{O}$-ideal with $\mathcal{O}/\mathfrak{n} \cong \mathbf{Z}/N\mathbf{Z}$. In fact, we find $\overline{\mathfrak{n}}\mathfrak{a} = z\mathbf{Z} + N\mathbf{Z}$ and hence

$$\sigma_{[\mathfrak{n}]} j(z) = j(\mathfrak{n}^{-1}\mathfrak{a}) = j(\overline{\mathfrak{n}}\mathfrak{a}) = j_N(z).$$

Exactly as in Section 4.3 of [14], we list the j-values of elliptic curves over $\mathbf{F}_p$ with endomorphism ring $\mathcal{O}$, and arrange them into unoriented $[\mathfrak{n}]$-isogeny cycles. If C is a quotient of $X^0_+(N)$ over $\mathbf{Q}$, then for each edge of this graph, we find the f-value from the two j-values of the end points. (In the case where the $[\mathfrak{n}]$-isogeny cycles are 2-cycles, we only get one f-value per 2-cycle and we get a lower-degree class polynomial $H_\mathcal{T}[f]$.)

In practice, we could do this for $f = x$ exactly as in [14], and then solve for y using $\Psi_C(x, y, j) = 0$, which is linear in y. The only additional work compared to what is done in [14] is computing and solving the linear equation to get y, which is much faster than all the other steps.

In particular, Step 1 takes much less than twice as long with C than with x, while we need to do it only half as often, which leads to a speed-up. Further research into these modular polynomials is needed in order to determine the exact gain.

To also make this work for quotients of $X^0(N)$ that are not quotients of $X^0_+(N)$, one would need to compute oriented $[\mathfrak{n}]$-isogeny cycles.

Other tricks for enumerating

The methods from [14, Sections 4.1 and 4.2] also seem amenable.

The main computational tool at the beginning of Section 4.1 is the modular polynomial $\Phi_{\ell,f}$, which we generalize from f to C as follows.

Let $\Phi_{\ell,C}$ be a Gröbner basis of the ideal in $\mathbf{Q}[X_1, Y_1, X_2, Y_2]$ of polynomials that vanish on $\{(\psi(z), \psi(\ell z)) : z \in \mathbf{H}\}$, with respect to the lexicographic ordering with $X_1 > Y_1 > X_2 > Y_2$. To get from $\psi(z)$ to all possible values of $\psi(\ell z)$, one substitutes $\psi(z)$ for (X_1, Y_1), and then solves first for X_2 and then for Y_2. For each C and ℓ this works for all but a finite set of primes p. Such multivariate modular polynomials would need to be precomputed. One possible starting point for computing these would be [25, 26], which compute multivariate (Hilbert) modular polynomials, each with a different method. For yet another approach to computing modular polynomials, see [3].

We expect the reduction factor to also give a reduction of the size of these multivariate modular polynomials, but on the other hand, we need two of them: one to solve for x of an isogenous curve, and one to evaluate in x and get y. As evaluating is faster than solving, we expect the use of these modular polynomials to take much less than twice as long (and we need to do it only half as often, because we have half as many primes).

The 'Trace Trick' of [14, Section 4.2] enables the use of the Weber function $\mathfrak{f}$ in the CRT method. In case we would also need this trick, for some more exotic curves C, we could consider applying it with arbitrary functions $f \in \mathbf{Q}(C)$ such as $f = ax + by$ for small integers a and b. In loc. cit. the relevant trace is computed with much fewer primes, so it is ok to apply this with the lower reduction factor of f.

We did not yet consider the general algorithm of [14, Section 4.4]. It is the method that works for all class invariants, but is only practical under additional restrictions. We do not have examples of generalized class invariants where this trick is needed. The challenging step to generalize is factoring a large-degree function in $\mathbf{Q}(C)$ in order to obtain the small class functions.

Constructing a function from its roots

In the CRT setting the multiplications and long-divisions by small-degree polynomials of Section 6.6.2 only take time $O(nM(\log(p)))$ per level, which is asymptotically dominated by the $O(M(n\log(p))$ time of the multiplications of large-degree polynomials. Therefore, Step 2 seems to take about 1.5 times as long per prime p for $H_\mathcal{T}[C]$ when compared to $H_\mathcal{T}[x]$.

The total running time

Concluding this preliminary analysis, we estimate the cost of computing $H_\mathcal{T}[C]$ to be significantly lower compared to $H_\mathcal{T}[x]$, though further research, in particular into (the implementation of) modular polynomials for C is required to determine the exact gain. This is beyond the scope of the current paper, which focuses on introducing the generalized class functions and their height reduction. We plan to give a more detailed account and an implementation in future work.

6.7 General curves and bases

Now suppose that our modular curve C is not necessarily an elliptic curve. Let $\mathcal{D}$ be an effective divisor over $\mathbf{Q}$ on C and let $\mathcal{B} = \{b_0, b_1, \ldots\}$ be a $\mathbf{Q}$-basis of $\mathcal{L}(\infty\mathcal{D})$ ordered by ascending degree.

The classical case is the case where we have one modular function f and we take $C = \mathbf{P}^1$, $\psi = f = (f : 1)$, $\mathcal{D} = ((1 : 0)) = (\infty)$, and $\mathcal{B} = \{1, f, f^2, \ldots\}$. The case of all previous sections of this paper is the case where C is an elliptic curve given by a Weierstrass equation, $\mathcal{D} = ((0:1:0))$, and $\mathcal{B} = \{1, x, y, x^2, xy, x^3, x^2y, \ldots\}$.

Example 6.7.1 One systematic way to choose a $\mathbf{Q}$-basis of $\mathcal{L}(\infty\mathcal{D})$ is as follows. First choose $x \in \mathcal{L}(\infty\mathcal{D}) \setminus \mathbf{Q}$ of some degree d. (For example, one can take $x = f$ with $d = 1$ in the classical case, and $x = x$ with $d = 2$ in the elliptic case.) Now, let $y_0 = 1$ and choose y_j for $j = 1, 2, \ldots, d-1$ in such a way that

$$y_j \in \mathcal{L}(m_j\mathcal{D}) \setminus \langle y_k x^e : k < j, e \in \mathbf{Z}\rangle,$$

where m_j is minimal such that this set is non-empty. This way we obtain a vector $\vec{y} = (y_0, \ldots, y_{d-1})$ of d functions. (For example, in the classical case we have $\vec{y} = 1$, and in the elliptic case we chose $\vec{y} = (1, y)$.) Then $\mathcal{B} = \{x^e y_j : e \in \mathbf{Z}_{\geq 0}, j \in \{0, 1, 2, \ldots, d-1\}\}$ is a basis of $\mathcal{L}(\infty\mathcal{D})$. We order this basis by ascending degree $de + m_j$, and if two elements have the same degree, then we put the one with lowest j first. ☆

Example 6.7.2 Consider the modular curve $X_+^0(191)$ (not to be confused with 119), which is hyperelliptic with model $t^2 = s^6 + 2s^4 + 2s^3 + 5s^2 - 6s + 1$ [16, Table 3], and the unique cusp is at $\mathcal{D} = ((1:1:0))$. One of the possible bases of $\mathcal{L}(\infty\mathcal{D})$ obtained by the recipe above is $\mathcal{B} = \{1, x, y_1, y_2, x^2, x^2y_1, x^2y_2, \ldots\}$, where $x = (t + s^3 + s + 1)/2$, $y_1 = sx$, and $y_2 = s(y_1 + 1)$. The degrees of these functions are respectively 3, 5, and 7.

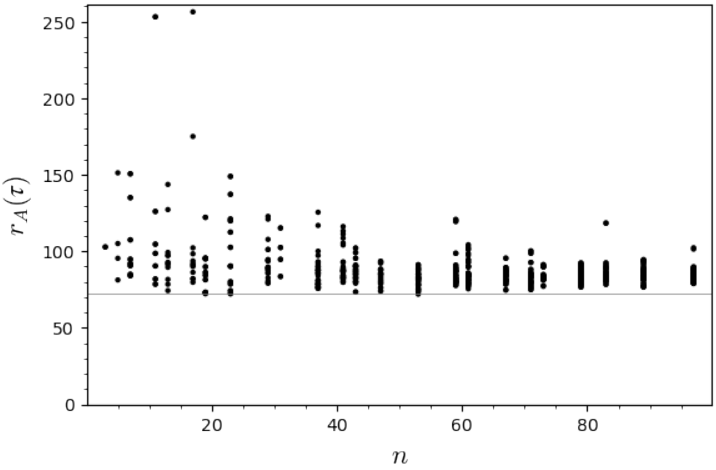
The function x is, up to multiplicative and additive constants, equal to the Atkin function A_{191}. The reduction factors are $r(C) = 96$, $r(s) = 48$, and $r(A_{191}) = 32$. ☆

As in Section 6.2, let $\tau \in \mathbf{H}$ imaginary quadratic and assume that (b_i, τ) is a class invariant for every $b_i \in \mathcal{B}$. Then, again unique up to scaling, we obtain a non-zero function $F_\tau[C, \mathcal{B}] = \sum_{i=0}^k a_i b_i \in K(C)$ $(a_i \in K)$ with k minimal such that $\sum_{i=0}^k a_i b_i(\tau) = 0$.

Definition 6.7.3 We call this $F_\tau[C, \mathcal{B}]$ the *generalized class function* for the triple $C, \mathcal{B}, \tau$. If $\mathcal{B}$ is as in Example 6.7.1 then we again refer to the associated polynomial $H_\tau[C, \mathcal{B}] \in K[X, Y_1, \ldots, Y_d]$ (of total degree ≤ 1 in $Y_1, \ldots, Y_d$ and such that $H_\tau[C, \mathcal{B}](x, y_1, \ldots, y_d) = F_\tau[C, \mathcal{B}]$) as the *generalized class polynomial.* △

Example 6.7.4 It turns out that, already for the case of elliptic curves, allowing the freedom of the choice of basis of may in reality lead to potentially better practical reduction factors. Revisiting our main example $C := X_+^0(119)$, denote by $w := \mathfrak{w}_{7,17}$ the function (6.4) and by $z := x + y$ the sum of the Weierstrass coordinates for the model (6.3). Now consider the basis $\mathcal{B} := \{1, x, z, w, xz, wx, wz, w^2, wxz, w^2x, \ldots\}$ of $\mathcal{L}(\infty\mathcal{D})$. The resulting generalized class polynomials corresponding to the discriminants of Table 6.1 are listed in Table 6.5. We get practical reduction factors in Figure 6.3 that are better than those in Figure 6.1.

A likely explanation for this improvement is that now not only the poles, but also the zeroes are as much restricted to the cusps of $X_+^0(119)$ as possible. Indeed, the points $O = (0 : 1 : 0)$ and $P = (0, 0)$ are the cusps, while $2P$ and $3P$ are rational CM points. Now $\operatorname{div}(w) = 4(P) - 4(O)$, $\operatorname{div}(x) = (P) + (3P) - 2(O)$, and $\operatorname{div}(y) = 2(P) + (2P) - 3(O)$. In particular, the function w is a modular unit. As explained in Remark 6.3.8, modular units in the classical setting give better practical reduction factors than non-units, even though the reduction factors are asymptotically the same. ☆



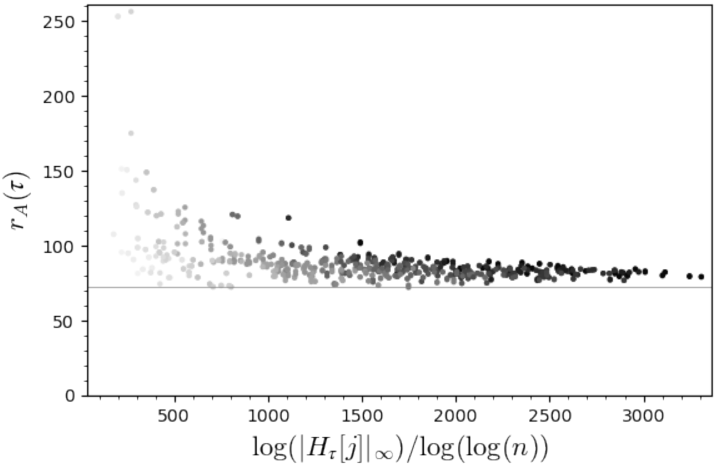


Figure 6.3: Practical reduction factors for $H_\tau[X_+^0(119), \mathcal{B}]$ for fundamental discriminants D with $\gcd(D, N) = 1$ and prime class number $n < 100$.

D	n	$F_\tau[C, \mathcal{B}]$
-52	2	$z - x + 1$
-523	5	$xw - xz - x + 3w + z$
-5347	13	$xw^3 - 10xw^2z + 42xw^2 + 48w^3 + 13xwz + 35w^2z + 62xw$ $+104w^2 + 39xz + 90wz - 11x + 41w + 39z + 1$
-15139	29	$xw^7 - 33xw^6z + 5874xw^6 + 849w^7 - 2119xw^5z - 3865w^6z$ $+31183xw^5 - 4249w^6 + 2200xw^4z - 15449w^5z + 36423xw^4$ $-29399w^5+6066xw^3z-46282w^4z+46223xw^3-27578w^4+6207xw^2z$ $-30128w^3z + 31320xw^2 - 47581w^3 + 6757xwz - 35595w^2z$ $+8017xw - 17181w^2 - 742xz - 10159wz - x - 2797w + 22z$

Table 6.5: Some conjecturally correct generalized class functions for the curve $C = X_+^0(119)$ using the $\mathcal{L}(\infty\mathcal{D})$-basis $\mathcal{B} := \{1, x, z, w, xz, wx, wz, w^2, wxz, w^2x, \ldots\}$.

Theorem 6.7.5 *Let $C : y^2 + g(x)y = f(x)$ with $f, g \in \mathbf{Q}[X]$ be a hyperelliptic curve such that $4f(x) + g(x)^2$ has odd degree and $\mathrm{Jac}(C)(\mathbf{Q})$ is finite. Set $\mathcal{D}$ to be the unique point at infinity and choose the basis $\mathcal{B} = \{1, x, x^2, y, xy, x^2y, \ldots\}$ of $\mathcal{L}(\infty\mathcal{D})$. Then Theorem 6.3.4 and Proposition 6.3.7 also hold for C and $H_\tau[C, \mathcal{B}]$.*

Proof. The original proof now goes through with only the following change. There are finitely many possibilities for the class c of the divisor $-\sum_\sigma((\sigma(\psi(\tau))) - \mathcal{D})$ by our assumption that $\mathrm{Jac}(C)(\mathbf{Q})$ is finite. For every c, choose a representative $\sum_{i=1}^m((P_i) - \mathcal{D})$ with m minimal and consider a primitive polynomial $T \in \mathbf{Z}[X]$ with roots $x(P_i)$ for $i = 1, \ldots, m$. □

Remark 6.7.6 Our proofs of Theorems 6.3.4 and 6.7.5 heavily rely on the fact that Heegner points are torsion. To completely remove the assumption on ranks, one would therefore need to bound the Heegner points, even in the rank-one case. Moreover, the proofs rely on the hyperelliptic equation where we use that $|a| \leq |a + bi|$ for real numbers a and b. Though we expect an analogue of these results to hold for general modular curves, this would require additional ideas. Do note that such an analogue would yield arbitrarily high reduction factors for generalized class polynomials by (6.7). For example, for $C = X_+^0(239)$ of genus 3 we already obtain $r(C) = 120$, exceeding the Bröker-Stevenhagen bound. ♢

6.8 Bibliography

[1] Francesc Bars. Bielliptic modular curves. *J. Number Theory*, 76(1):154–165, 1999.

[2] Brian J. Birch. Weber's class invariants. *Mathematika*, 16:283–294, 1969.

[3] Reinier Bröker, Kristin Lauter, and Andrew V. Sutherland. Modular polynomials via isogeny volcanoes. *Math. Comp.*, 81(278):1201–1231, 2012.

[4] Reinier Bröker and Peter Stevenhagen. Constructing elliptic curves of prime order. In *Computational arithmetic geometry*, volume 463 of *Contemp. Math.*, pages 17–28. Amer. Math. Soc., Providence, RI, 2008.

[5] John Cremona. *Algorithms for modular elliptic curves.* Cambridge University Press, Cambridge, 1992.

[6] Chris J. Cummins and Sebastian Pauli. Congruence subgroups of $\mathrm{PSL}(2, \mathbb{Z})$ of genus less than or equal to 24. *Experiment. Math.*, 12(2):243–255, 2003. Interactive database at `https://mathstats.uncg.edu/sites/pauli/congruence/csg1.html`.

[7] Andreas Enge. The complexity of class polynomial computation via floating point approximations. *Math. Comp.*, 78(266):1089–1107, 2009.

[8] Andreas Enge and François Morain. Comparing invariants for class fields of imaginary quadratric fields. In *Algorithmic number theory (Sydney, 2002)*, volume 2369 of *Lecture Notes in Comput. Sci.*, pages 252–266. Springer, Berlin, 2002.

[9] Andreas Enge and François Morain. Generalised Weber functions. *Acta Arith.*, 164(4):309–342, 2014.

[10] Andreas Enge and Reinhard Schertz. Constructing elliptic curves over finite fields using double eta-quotients. *Journal de Théorie des Nombres de Bordeaux*, 16:555–568, 2004.

[11] Andreas Enge and Reinhard Schertz. Modular curves of composite level. *Acta Arith.*, 118(2):129–141, 2005.

[12] Andreas Enge and Reinhard Schertz. Singular values of multiple eta-quotients for ramified primes. *LMS Journal of Computation and Mathematics*, 16:407–418, 2013.

[13] Andreas Enge and Marco Streng. Schertz style class invariants for genus two, 2016. preprint, arXiv:1610.04505.

[14] Andreas Enge and Andrew V. Sutherland. Class invariants by the CRT method. In *Algorithmic number theory*, volume 6197 of *Lecture Notes in Comput. Sci.*, pages 142–156. Springer, Berlin, 2010.

[15] Masahiro Furumoto and Yuji Hasegawa. Hyperelliptic quotients of modular curves $X_0(N)$. *Tokyo J. Math.*, 22(1):105–125, 1999.

[16] Steven Galbraith. *Equations for modular curves.* PhD thesis, St. Cross College, 1996. `https://www.math.auckland.ac.nz/~sgal018/thesis.pdf`.

[17] Alice Gee. Class invariants by Shimura's reciprocity law. volume 11, pages 45–72. 1999. Les XXèmes Journées Arithmétiques (Limoges, 1997).

[18] Alice Gee and Peter Stevenhagen. Generating class fields using Shimura reciprocity. In *Algorithmic number theory (Portland, OR, 1998)*, volume 1423 of *Lecture Notes in Comput. Sci.*, pages 441–453. Springer, Berlin, 1998.

[19] Marc Hindry and Joseph H. Silverman. *Diophantine geometry*, volume 201 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 2000. An introduction.

[20] Daeyeol Jeon. Bielliptic modular curves $X_0^+(N)$. *J. Number Theory*, 185:319–338, 2018.

[21] Henry H. Kim. Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2. *J. Amer. Math. Soc.*, 16(1):139–183, 2003. With appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak.

[22] John E. Littlewood. On the class-number of the corpus $p(\sqrt{-k})$. *Proc. Lond. Math. Soc.*, 27:358–372, 1928.

[23] The LMFDB Collaboration. The L-functions and modular forms database. `http://www.lmfdb.org`, 2022. [Online; accessed March and August 2022].

[24] Kurt Mahler. An application of Jensen's formula to polynomials. *Mathematika*, 7:98–100, 1960.

[25] Chloe Martindale. Hilbert modular polynomials. *J. Number Theory*, 213:464–498, 2020.

[26] Enea Milio and Damien Robert. Modular polynomials on Hilbert surfaces. *J. Number Theory*, 216:403–459, 2020.

[27] Morris Newman. Construction and application of a class of modular functions. *Proc. London Math. Soc. (3)*, 7:334–350, 1957.

[28] Andrew P. Ogg. Hyperelliptic modular curves. *Bull. Soc. Math. France*, 102:449–462, 1974.

[29] Karl Rubin and Alice Silverberg. Choosing the correct elliptic curve in the CM method. *Math. Comp.*, 79(269):545–561, 2010.

[30] Reinhard Schertz. Die singulären Werte der Weberschen Funktionen $\mathfrak{f}$, $\mathfrak{f}_1$, $\mathfrak{f}_2$, γ_2, γ_3. *J. Reine Angew. Math.*, 286/287:46–74, 1976.

[31] Reinhard Schertz. Weber's class invariants revisited. *J. Théor. Nombres Bordeaux*, 14(1):325–343, 2002.

[32] Atle Selberg. On the estimation of Fourier coefficients of modular forms. In *Proc. Sympos. Pure Math., Vol. VIII*, pages 1–15. Amer. Math. Soc., Providence, R.I., 1965.

[33] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions.* Kanô Memorial Lectures, No. 1. Iwanami Shoten Publishers, Tokyo; Princeton University Press, Princeton, N.J., 1971. Publications of the Mathematical Society of Japan, No. 11.

[34] Andrew V. Sutherland. Computing Hilbert class polynomials with the Chinese remainder theorem. *Math. Comp.*, 80(273):501–538, 2011.

[35] Andrew V. Sutherland. Accelerating the CM method. *LMS J. Comput. Math.*, 15:172–204, 2012.

[36] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.5)*, 2022. `https://www.sagemath.org`.

[37] Heinrich Weber. *Algebraische Zahlen*, volume 3 of *Lehrbuch der Algebra.* Friedrich Vieweg, 1908.