SHIMURA CURVES AND THE ABC CONJECTURE

HECTOR PASTEN

Abstract. In this work we introduce new methods to study the abc conjecture and related problems. These methods are based on the theory of Shimura curves. The three main applications that we obtain are the following:

1. we show new and stronger unconditional effective bounds for the minimal discriminant and Faltings height of elliptic curves over \( \mathbb{Q} \) in terms of the conductor,
2. we prove that for any given \( \epsilon > 0 \), if \( a, b, c \) are coprime positive integers with \( a + b = c \) then \( d(abc) \ll \epsilon \operatorname{rad}(abc)^{8/3+\epsilon} \), where \( d \) is the number of divisors function, and
3. we show that appropriate bounds for modular degrees imply Szpiro’s conjecture in the context of modular elliptic curves over totally real number fields, thus extending beyond \( \mathbb{Q} \) a classical modular approach of Frey to Szpiro’s conjecture.

Contents

1. Introduction 2
2. Basic notation 10
3. Review of the classical modular approach 11
4. Preliminaries on Shimura curves 13
5. The degree of Shimura curve parameterizations 17
6. A refinement of the Ribet-Takahashi formula 21
7. Bounding the modular degree 27
8. Infinite part: Norm comparisons 31
9. Finite part: Differentials on integral models 34
10. Bounds for the Manin constant 37
11. Counting imaginary quadratic extensions of \( \mathbb{Q} \) 45
12. Arakelov degrees 49
13. Arakelov height of Heegner points 50
14. Integrality and lower bounds for the \( L_2 \)-norm 54
15. Linear forms in logarithms 59
16. Bounds for the product of exponents of the minimal discriminant 62
17. Counting quadratic extensions of a totally real number field 65
18. A modular approach to Szpiro’s conjecture over number fields 67
19. Acknowledgments 76
References 77

Date: August 29, 2017.
2010 Mathematics Subject Classification. Primary 11G18; Secondary 11G05, 14G40.
Key words and phrases. Shimura curves, elliptic curves, abc conjecture.
1. Introduction

This work is concerned with Szpiro’s conjecture, the \(abc\) conjecture, and related problems. Our purpose is twofold: On the one hand we obtain new unconditional results towards these open problems, while on the other hand we introduce new techniques that we expect to be useful not just in the applications discussed here.

Our main innovation is the introduction of Shimura curve parametrizations of elliptic curves to study these problems.

First, we extend Frey’s classical modular approach to Szpiro’s conjecture and the \(abc\) conjecture (cf. [31]; see also Section 3), from the case of modular curves \(X_0(N)\) to Shimura curves. That is to say, we show that appropriate bounds for the degree of Shimura curve parametrizations of elliptic curves over \(\mathbb{Q}\) imply Szpiro’s conjecture. Serious complications arise due to the lack of \(q\)-expansions for modular forms on Shimura curves, as \(q\)-expansions are crucially used in Frey’s argument. We also show how to bound such degrees in terms of congruences of systems of Hecke eigenvalues. The reward for extending the modular approach to the Shimura curve case is that one gets unconditional effective bounds for Szpiro’s conjecture over \(\mathbb{Q}\) that are sharper than the unconditional bounds available in the literature. See Paragraph 1.2 for details.

Secondly, we introduce a new “modular method” to obtain progress on Szpiro’s conjecture and the \(abc\) conjecture, which is fundamentally different to Frey’s modular degree approach. For an elliptic curve \(E\) over \(\mathbb{Q}\), instead of bounding the degree of morphisms of the form \(X \to E\) for a Shimura curve \(X\), our method requires one to bound the variation of this degree as we vary \(X\) among the several available Shimura curves \(X\) for the same \(E\). At this point we can state in a precise way one of the key applications of this new approach without having to introduce technical notation. As usual, for a positive integer \(n\) we write \(\text{rad}(n)\) for the radical of \(n\) and \(d(n)\) for the number of positive integer divisors of \(n\).

**Theorem 1.1 (d(abc) Theorem).** Let \(\epsilon > 0\). There is a constant \(K_\epsilon\) depending only on \(\epsilon\) such that for all triples of coprime positive integers \(a, b, c\) with \(a + b = c\), one has

\[
d(abc) \leq K_\epsilon \cdot \text{rad}(abc)^{8/3 + \epsilon}.
\]

Theorem 1.1 is qualitatively stronger than the results that one can obtain from the theory of linear forms in \(p\)-adic logarithms, because it gives a bound polynomial on \(\text{rad}(abc)\) for the product of the exponents in the prime factorization of an \(abc\) triple, while transcendental methods give such a bound only for the maximum of the exponents. See Paragraph 1.3 for details and for other applications of these techniques.

We remark that several intermediate results are obtained along the way (cf. Paragraph 1.5), including bounds for the \(p\)-adic valuation of the Manin constant of elliptic curves in the case of additive reduction.

The last part of this work is focussed on developing a Shimura curve approach to Szpiro’s conjecture over number fields, see Paragraph 1.4. We are able to address the case of modular semi-stable elliptic curves over totally real number fields, and we prove that appropriate conjectural bounds for the degree of modular parametrizations in this case also imply Szpiro’s conjecture. One of the key tools used in our argument is recent work of Yuan and Zhang [93] on Colmez’s conjecture.

The rest of this introduction is devoted to a more precise description of the problems that we consider and the results that we obtain.
1.1. **The problems.** For a number field \( F \) and an elliptic curve \( E \) over \( F \), we write \( \Delta_E \) for the norm of the minimal discriminant ideal \( \mathcal{D}_E \) of \( E \) over \( F \), and \( N_E \) for the norm of the conductor ideal \( \mathcal{A}_E \) of \( E \) over \( F \).

Early in the eighties, Lucien Szpiro proposed the following:

**Conjecture 1.2** (Szpiro). Let \( F \) be a number field. There are constants \( \alpha, \beta > 0 \) depending only on \( F \) such that for every elliptic curve \( E \) over \( F \) we have \( \Delta_E \leq \alpha N_E^\beta \).

As usual, we write \( \text{rad}(n) \) for the radical of a positive integer \( n \). In the case \( F = \mathbb{Q} \), Szpiro’s conjecture applied to elliptic curves arising from Frey’s construction implies:

**Conjecture 1.3.** There is a constant \( \kappa > 1 \) such that for every triple of coprime positive integers \( a, b, c \) satisfying \( a + b = c \), we have \( c \leq \text{rad}(abc)^\kappa \).

The latter conjecture has been proposed in a refined form by Masser and Oesterlé in 1985. The refined version is known as the Masser-Oesterlé conjecture, or the \( abc \) conjecture.

**Conjecture 1.4** (\( abc \) Conjecture). Let \( \epsilon > 0 \). There is a constant \( K_\epsilon \) depending only on \( \epsilon \), such that for all triples of coprime positive integers \( a, b, c \) satisfying \( a + b = c \) we have

\[
c \leq K_\epsilon \cdot \text{rad}(abc)^{1+\epsilon}.
\]

See [65] for more details on the previous open problems.

Also related to Szpiro’s conjecture, there is the height conjecture of Frey [31]. If \( E \) is an elliptic curve over a number field \( F \), we write \( h(E) \) for the Faltings height of \( E \) over \( F \) (whose definition will be recalled in Sections 3 and 18). A basic property of \( h(E) \) is the estimate

\[
\frac{1}{[F : \mathbb{Q}]} \log \Delta_E \leq 12h(E) + 16
\]

(cf. Lemma 18.1 below) which makes it clear that Szpiro’s conjecture is a consequence of the following

**Conjecture 1.5** (Height conjecture). Let \( F \) be a number field. There are constants \( \alpha, \beta > 0 \) depending only on \( F \) such that for every elliptic curve \( E \) over \( F \) we have \( h(E) \leq \alpha \log N_E + \beta \).

As of today, all unconditional bounds for these conjectures are of exponential nature. The strongest unconditional bound for Szpiro’s conjecture and the height conjecture over \( \mathbb{Q} \) are obtained by the methods of Murty and Pasten [63]. For an elliptic curve \( E \) over \( \mathbb{Q} \), the bounds in loc. cit. take the form

\[
\log \Delta_E = O(N_E \log N_E) \quad \text{and} \quad h(E) = O(N_E \log N_E)
\]

with an effective implicit constant, see the discussion in Paragraph 1.2 below for more details. We also mention [28] where the set of elliptic curve failing Szpiro’s conjecture (over \( \mathbb{Q} \), for a given exponent) is shown to be small in an asymptotic sense.

On the other hand, the strongest unconditional bounds for the \( abc \) conjecture are due to Stewart and Yu [79], after previous work by Stewart and Tijdeman [77] and Stewart and Yu [78]. For \( \epsilon > 0 \) and an \( abc \) triple \( a, b, c \), these bounds take the form

\[
\log(c) = O_\epsilon (\text{rad}(abc)^{\alpha+\epsilon})
\]

with an effective implicit constant, where \( \alpha = 15 \) is obtained in [77], \( \alpha = 2/3 \) is obtained in [78], and \( \alpha = 1/3 \) is obtained in [79].
In this paper we don’t discuss the interesting subject of applications of these conjectures, but instead, our goal is to make progress on them and to introduce new tools. Our main results on these conjectures fall into three categories:

(i) improved effective unconditional exponential bounds for the height conjecture and Szpiro’s conjecture over \( \mathbb{Q} \);
(ii) unconditional upper bounds (polynomial on the conductor) for the product of the exponents of the minimal discriminant of elliptic curves, and similarly for the number of divisors of \( abc \) when \( a, b, c \) are coprime positive integers with \( a + b = c \);
(iii) a new “modular approach” to the height conjecture and Szpiro’s conjecture over totally real number fields.

Our methods involve a number of tools, such as classical analytic number theory, Arakelov geometry, Galois representations, height formulas related to Colmez’s conjecture, linear forms in logarithms, automorphic forms, etc. However, as mentioned before, the main innovation in this work is the introduction of new techniques based on Shimura curves to approach Szpiro’s conjecture and the \( abc \) conjecture.

Below we give a more detailed description of some of our main results in the three aspects mentioned above, with references to more technical theorems stated later in the text and some discussion on the methods. We also present additional results of independent interest that will be obtained along the way.

We use Vinogradov’s notation \( X \ll Y \), as well as the equivalent notation of Landau \( X = O(Y) \). When this notation is used, if the implicit constant depends on any additional parameters, this will be indicated as a subscript.

1.2. New bounds for Szpiro’s conjecture and the height conjecture over \( \mathbb{Q} \). We prove:

**Theorem 1.6** (Improved effective bounds). Let \( \epsilon > 0 \). There is an effectively computable constant \( c(\epsilon) \) depending only on \( \epsilon \) such that every elliptic curve \( E \) over \( \mathbb{Q} \) satisfies

\[
h(E) \leq \left( \frac{1}{48} + \epsilon \right) N_E \log N_E + c(\epsilon) \quad \text{and} \quad \log |\Delta_E| \leq \left( \frac{1}{4} + \epsilon \right) N_E \log N_E + c(\epsilon).
\]

Furthermore, if the Generalized Riemann Hypothesis holds (for the Rankin-Selberg convolution \( L \)-function of two Hecke eigenforms of weight 2), then

\[
h(E) \leq \left( \frac{1}{24} + \epsilon \right) N_E \log \log N_E + c(\epsilon) \quad \text{and} \quad \log |\Delta_E| \leq \left( \frac{1}{2} + \epsilon \right) N_E \log \log N_E + c(\epsilon).
\]

See Theorem 7.5 for a more precise statement.

Previously, the best available results for the height conjecture and Szpiro’s conjecture for all elliptic curves over \( \mathbb{Q} \) were obtained by the methods of Murty and Pasten [63], with an approach different to the one used here. For comparison purposes, let us recall them. The bounds in Theorem 7.1 [63] are

\[
h(E) \leq \left( \frac{1}{12} + \epsilon \right) N_E \log N_E + c(\epsilon), \quad \log |\Delta_E| < (1 + \epsilon) N_E \log N_E + c(\epsilon)
\]
while the bounds in Proposition 10.8 \([50]\) (obtained by numerical improvements of the argument in \([63]\)) are

\[
h(E) < \left( \frac{1}{16} + \epsilon \right) N_E \log N_E + c(\epsilon), \quad \log |\Delta_E| < \left( \frac{3}{4} + \epsilon \right) N_E \log N_E + c(\epsilon).
\]

So, Theorem 1.6 gives an improvement on the existing results. Furthermore, under GRH our bounds are of lower order of magnitude.

The fact that we obtain better bounds and that we are able to relate the problem to GRH is due to a conceptual difference of the arguments: previously, only congruences among Fourier expansions (with Fourier coefficients \(\mathbb{Z}\)) of modular forms that are not necessarily eigenforms were considered. The \(L\)-function of such a modular form does not need to satisfy GRH in general, and the Rankin-Selberg convolution \(L\)-function is not defined for all of them. Here, instead, we prove our bounds by a different method based on congruences between Hecke eigenforms, whose (Rankin-Selberg) \(L\)-functions are in fact expected to satisfy GRH.

Let \(\varphi(n)\) denote Euler’s totient function. For semi-stable elliptic curves we obtain stronger estimates.

**Theorem 1.7** (Improved effective bounds; almost semi-stable case). Let \(\epsilon > 0\), let \(S\) be a (possibly empty) finite set of primes and let \(P\) be the product of the primes in \(S\), with \(P = 1\) if \(S\) is empty. There is an effectively computable constant \(c(\epsilon, S)\) depending only on \(\epsilon\) and \(S\) such that every elliptic curve \(E\) over \(\mathbb{Q}\) with semi-stable reduction away from \(S\) satisfies

\[
h(E) \leq \frac{P}{\varphi(P)} \left( \frac{1}{48} + \epsilon \right) \varphi(N_E) \log N_E + c(\epsilon, S), \quad \text{and}
\]

\[
\log |\Delta_E| \leq \frac{P}{\varphi(P)} \left( \frac{1}{4} + \epsilon \right) \varphi(N_E) \log N_E + c(\epsilon, S).
\]

Furthermore, if we assume GRH then

\[
h(E) \leq \frac{P}{\varphi(P)} \left( \frac{1}{24} + \epsilon \right) \varphi(N_E) \log \log N_E + c(\epsilon, S), \quad \text{and}
\]

\[
\log |\Delta_E| \leq \frac{P}{\varphi(P)} \left( \frac{1}{2} + \epsilon \right) \varphi(N_E) \log \log N_E + c(\epsilon, S).
\]

See Corollary 7.8. Also, we refer the reader to Section 7 for additional effective bounds and related results.

Accordingly, the applications regarding effective bounds for the height of \(S\)-integral solutions to the unit equation, to Mordell’s equation, and to other Diophantine equations discussed in \([49, 50, 63]\) follow in a sharper form from our work (unconditionally, or in an even stronger form under GRH) by using Theorem 1.6, Theorem 1.7, or the more precise results from Section 7, instead of the results and methods from \([63]\).

It is worth pointing out that the proof of the effective results in Section 7 does not involve the theory of linear forms in logarithms.

1.3. Product of exponents and number of divisors. The \(d(abc)\) Theorem stated before (Theorem 1.1; see Theorem 16.3 below) is one of the main applications of our work in the context of the \(abc\) Conjecture. To put this theorem in context, let us consider more generally bounds of the form

\[
d(abc) \ll_{\epsilon} \text{rad}(abc)^{\beta + \epsilon}
\]
for $\beta$ fixed. Expressing $d(n)$ in terms of the prime factorization of $n$, we see that this bound implies (and in fact, due to the presence of $\epsilon > 0$, it is equivalent to)

$$\prod_{p|abc} v_p(abc) \ll \epsilon \rad(abc)^{\beta + \epsilon}$$

for the same $\beta$, where $p$ varies over primes and $v_p$ is the $p$-adic valuation. On the other hand, the known bounds of the form (1.1) are equivalent to

$$\max_{p|abc} v_p(abc) \ll \epsilon \rad(abc)^{\alpha + \epsilon}$$

with the same $\alpha$. Although the value $\beta = 8/3 = 2.66...$ that we obtain is not small enough to fully recover (1.1) with the best current exponent $\alpha = 1/3$, qualitatively our result is a considerable improvement in the sense that Theorem 1.1 gives a bound for the product of the exponents in the prime factorization of $abc$ which is polynomial on $\rad(abc)$, while the existing results only bound the maximum of these exponents polynomially on $\rad(abc)$.

It might also be useful to compare the quantities $\log c$ and $d(abc)$ not just on theoretical grounds, but on actual examples instead. The next table records this comparison for the top five highest quality (in the standard terminology) known $abc$ triples [76]:

|   |   |   |   |   |
|---|---|---|---|---|
|   | 2  | $3^{11}109$ | 23$^9$ | 15.677... | 264 |
| 11$^2$ | $3^{12}5^67^3$ | $2^{21}23$ | 17.691... | 11088 |
| 19·1307 | 7·29$^2$31$^8$ | $2^83^{22}5^4$ | 36.152... | 223560 |
| 283 | $5^{11}13^2$ | $2^83^817^3$ | 22.833... | 23328 |
| 1 | 2·3$^7$ | 5$^4$7 | 8.383... | 160 |

One can also ask how far is Theorem 1.1 from the optimal expected bound for $d(abc)$ by a function of $\rad(abc)$. After taking logarithms, the $d(abc)$ Theorem gives a bound of the form

$$\log d(abc) \ll \log \rad(abc)$$

for $abc$ triples, while the $abc$ conjecture readily implies the conjectural bound

$$\log d(abc) \ll (\gamma) \frac{\log \rad(abc)}{\log \log \rad(abc)}.$$

Except for the implicit constant, this last conjectural bound for $d(abc)$ in terms of $\rad(abc)$ cannot be improved.

Our method also gives other results of interest, see Section 16 for details. For the convenience of the reader, let us mention the following (see Theorem 16.2 below):

**Theorem 1.8** (Product of valuations of the minimal discriminant). Let $\epsilon > 0$. There is a constant $K_\epsilon$ depending only on $\epsilon$ such that for all semi-stable elliptic curves $E$ over $\mathbb{Q}$ we have

$$\prod_{p|\Delta_E} v_p(\Delta_E) \leq K_\epsilon \cdot N_E^{\frac{2}{3} + \epsilon}.$$ 

In particular, the product of the Tamagawa numbers of $E$ for finite primes is bounded polynomially in terms of the conductor of $E$.

6
Let \( \text{Tam}_p(E) \) be the Tamagawa factor at \( p \) for an elliptic curve \( E \). Of course, for any \( \eta > 0 \) one expects

\[
\prod_{p < \infty} \text{Tam}_p(E) \ll \eta \, N_E^\eta
\]

but to the best of our knowledge, this was not known unconditionally for any fixed exponent \( \eta > 0 \). Theorem 1.8 establishes such a bound unconditionally in the semi-stable case, for any \( \eta > 9/2 \).

Also, note that since \( \prod_{p | \Delta_E} v_p(\Delta_E) \) controls level-lowering congruences (in the sense of Ribet) for the Hecke eigenform associated to \( E \), Theorem 1.8 also establishes an unconditional bound for such congruences, in a precise sense.

A word about the proofs. Theorems 1.1 and 1.8, as well as related results obtained in Sections 14 and 16, are not proved by bounding the degree of modular parametrizations — unlike earlier results in the literature such as in [63]. Our methods here are completely different to Frey’s modular degree approach, and instead of bounding the degree of modular parametrizations of elliptic curves, one needs to bound the \textit{variation} of the degree across various modular parametrizations coming from different Shimura curves to the same elliptic curve. This variation is quantified in Theorem 6.1 below, which is an extension of earlier results by Ribet and Takahashi [71, 81], and can be seen as an arithmetic analogue of partial derivatives (cf. the discussion after Theorem 6.1). After this point, the key estimates are then proved by means of Arakelov geometry; as usual, this requires technical work on integral models as well as on the complex analytic side, which in our case takes the form of certain norm comparisons. Furthermore, results related to Colmez’s conjecture translate part of the problem into a question of analytic number theory regarding the existence of suitable auxiliary imaginary quadratic number fields. Unsurprisingly, the prime \( p = 2 \) is problematic (at least for Theorem 1.1) and its contribution to the bounds needs to be dealt with by other means, namely, linear forms in logarithms.

Another technical point that requires attention is the Manin constant of classical modular parametrizations since we need to establish bounds for its valuation at primes where the relevant elliptic curve has additive reduction.

One can ask whether the \( d(abc) \) Theorem can be proved by directly using the theory of linear forms in \( p \)-adic logarithms instead of the approach developed in this paper. We suspect that this is unlikely, and in Paragraph 15.2 we present a heuristic justification for this guess.

1.4. **Towards Szpiro’s conjecture over number fields.** In precise terms, Frey’s classical modular approach to Szpiro’s conjecture shows that the height conjecture and Szpiro’s conjecture over \( \mathbb{Q} \) would follow from the following:

**Conjecture 1.9** (Modular degree conjecture). \textit{There is a constant \( \alpha \) such that for every elliptic curve \( E \) over \( \mathbb{Q} \) there is a modular parametrization \( \phi : X_0(N_E) \to E \) satisfying \( \deg \phi \leq N_E^\alpha \).}

See Section 8 for an overview of Frey’s modular approach over \( \mathbb{Q} \). Over number fields other than \( \mathbb{Q} \), no analogue of this modular approach has been developed in the literature. Of course it is reasonable to first focus on the case of elliptic curves over a totally real number field \( F \), since in this case one expects the existence of modular parametrizations by Shimura curves associated to quaternion \( F \)-algebras (provided that the conductor ideal \( N_E \) satisfies some mild necessary conditions). It is possible to formulate a conjectural bound for the
degree of such a modular parametrization, and we do so in Conjecture 18.11 which can be considered a number field version of the modular degree conjecture stated above.

To be precise, Conjecture 18.11 actually concerns a more intrinsic quantity $\delta_E$ canonically attached to a modular elliptic curve $E$ over a totally real number field $F$, and Theorem 18.4 relates $\delta_E$ to the degree of a modular parametrization in this context.

However, the important problem in this setting is, of course, to show that the Faltings height of an elliptic curve can be bounded in terms of the degree of a modular parametrization coming from a Shimura curve. The key technical complication is that the relevant Shimura curves do not have cuspidal points, and therefore, $q$-expansions are not available. We overcome this difficulty by using CM points instead, and at this point the recent deep results of Yuan and Zhang [93] are essential. This fix for the lack of $q$-expansions, however, raises a number of additional complications of a more technical nature.

In Section 18 we carry out this program, and we show that the modular approach is also available over totally real fields under certain technical hypothesis. For the sake of simplicity, let us discuss here just the case when $F$ is a real quadratic field; more general results are included in Section 18 especially, see Theorem 18.13. The real quadratic case is particularly convenient because all elliptic curves over such fields are modular (cf. [29]) so that we don’t need to include modularity as an assumption.

**Theorem 1.10** (Modular approach for real quadratic fields). *If the bound for the modular degree proposed in Conjecture 18.11 holds for real quadratic fields and for their totally real quadratic extensions, then there are constants $\alpha, \beta$ such that the following holds:

Let $F$ be a real quadratic number field and let $d_F$ be the absolute value of its discriminant. Let $E$ be a semi-stable elliptic curve over $F$ not having everywhere good reduction. Then

\[ h(E) < \alpha \log(d_F N_E) \quad \text{and} \quad \Delta_E \leq (d_F N_E)^{\beta}. \]

This theorem realizes the modular approach to the height conjecture and Szpiro’s conjecture over all real quadratic fields $F$ simultaneously. We remark that the proof naturally yields a dependence on $F$ through the quantity $d_F$, in accordance with Vojta’s conjectures for algebraic points of bounded degree (cf. Conjecture 5.2.6 in [86]).

In Paragraph 18.7 we present an application of our modular approach, giving unconditional exponential bounds for Szpiro’s conjecture over a totally real number field $F$, at least for elliptic curves $E$ over $F$ admitting a Shimura curve parametrization. The bounds will be of the form

\[ h(E) \ll_{F, \epsilon} N_E^{1+\epsilon} \quad \text{and} \quad \log \Delta_E \ll_{F, \epsilon} N_E^{1+\epsilon} \]

when applicable; see Theorem 18.14 for a precise statement.

Let us recall that the aim in [93] is to establish an averaged version of Colmez’s conjecture, which is done by first developing a suitable theory of heights and integral models of Shimura curves and then proving a height formula for CM points. We will not use the Yuan-Zhang results on averaged Colmez’s conjecture per se, but instead, we will need their theory of integral models and their height formula for CM points.

1.5. Additional results. In order to establish the theorems discussed so far, we need to prove a number of intermediate results, some of which are listed here as we believe that they are of independent interest. Precise statements of the results outlined in this paragraph will be given after the appropriate notation is introduced.
Theorem 6.1 proves an extension of the Ribet-Takahashi formula [71, 81] with an estimate for the contribution at every prime, even in the presence of reducible residual Galois representations.

Theorem 7.6 provides a version of the modular approach over $\mathbb{Q}$ using Shimura curves rather than modular curves. Unlike our modular approach to Szpiro’s conjecture over totally real fields (cf. Section 18), here we don’t use Heegner points and instead we deduce the result from the classical case. The bound is sharper than in the general case of totally real number fields.

Theorem 8.1 yields a norm comparison for the $L_2$-norm and the $L_{\infty}$-norm of holomorphic weight 2 modular forms associated to quaternion algebras over totally real fields, in the compact case. Note that in the non-compact case the literature is abundant and the bounds are strong (cf. [7] and the references therein), while in the compact case bounds have been obtained as the weight grows (cf. [22]).

Theorem 10.1 proves a bound for the $p$-adic valuation of the Manin constant of classical modular parametrizations of elliptic curves over $\mathbb{Q}$ at primes $p$ which are not necessarily of semi-stable reduction. As the Frey elliptic curves arising from $abc$ triples are not always semi-stable at the prime $p = 2$, we expect that this bound for the Manin constant will have applications in related projects.

Finally, we mention Theorem 14.1. This result gives a lower bound for the Petersson norm of a non-zero, weight 2, holomorphic modular form $f$ on a compact Shimura curve over $\mathbb{Q}$ assuming that $f$ is $\mathbb{Z}$-integral in the following sense: Seen as a differential on the Shimura curve, $f$ extends to a regular differential on a certain standard integral model of the Shimura curve over $\mathbb{Z}$. We don’t assume that $f$ is a Hecke eigenform. This can be seen as a lower bound for the first successive minimum of the space of global sections of the (suitably) metrized canonical sheaf of Shimura curves.

1.6. Other references. We conclude this introduction by briefly recalling some literature around the $abc$ Conjecture, although only tenuously related to our work. We are not attempting to make a survey on the $abc$ Conjecture and the list is by no means complete, but we feel that it is close in taste to the topics in this article.

In [85], Ullmo gave estimates for the Faltings height of $J_0(N)$ when $N$ is squarefree coprime to 6. He conjectured that the Faltings height of $J_0(N)$ is somewhat evenly distributed among the simple factors of $J_0(N)$ according to their dimension, which would imply a form of Szpiro’s conjecture thanks to his height estimates.

In [67], Prasanna established integrality results in the context of the Jacquet-Langlands correspondence for Shimura curves, which allowed him to compute local contributions of the Faltings height of jacobians of Shimura curves over $\mathbb{Q}$ away from an explicit set of primes.

In [30], Freixas i Montplet used the Jacquet-Langlands correspondence to compute Arakelov-theoretic invariants of certain Shimura curves over $\mathbb{Q}$, but only after discarding the contribution of finitely many primes.

Despite all this progress, the computation (or asymptotic evaluation) in terms of the level of the Faltings height of the jacobian of quaternionic Shimura curves remains open. This is not used in our work.

From a completely different angle, S.-W. Zhang [25] showed that appropriate bounds on the height of the Gross-Schoen cycle in triple products of curves, would imply the $abc$ conjecture.
Regarding the abc Conjecture for number fields other than \( \mathbb{Q} \), the method of linear forms in logarithms is available and it has been worked out by Surroca [80] obtaining effective unconditional exponential bounds. Sharper exponential bounds have been proved by Györy [38]. However, as in the case of \( \mathbb{Q} \), this approach has well understood technical limitations; see Baker’s article [6] for a discussion on conjectural strengthenings to the theory of linear forms in logarithms that would be needed in order to approach the abc Conjecture.

2. Basic notation

In this brief section we introduce the basic notation used throughout the paper. Further notation and preliminary material will be introduced as needed.

As said in the introduction, we will be using the standard asymptotic notation \( X = O(Y) \) of Landau, or equivalently, Vinogradov’s notation \( X \ll Y \). Any dependence on parameters will be indicated as sub-index. The notation \( X \asymp Y \) means that both \( X \ll Y \) and \( Y \ll X \) hold. The symbol \( \epsilon \) will always refer to a given small positive number.

We will be using the following arithmetic functions: \( \varphi(n) \) is the Euler totient function, \( d(n) \) is the number of divisors of \( n \), \( \sigma_1(n) \) is the sum of the divisors of \( n \), \( \omega(n) \) is the number of distinct prime divisors of \( n \), and \( \mu(n) \) is the Möbius function.

Recall from the introduction that for a number field \( F \) the absolute value of its discriminant is \( d_F \). Also, if \( E \) is an elliptic curve over \( F \), we write \( \mathcal{D}_E \) for its minimal discriminant ideal, and \( \mathcal{N}_E \) for its conductor ideal. We write \( \Delta_E \) for the norm of \( \mathcal{D}_E \), and \( N_E \) for the norm of \( \mathcal{N}_E \).

Note that when \( F = \mathbb{Q} \), the absolute value of the minimal discriminant of \( E \) is precisely \( \Delta_E \), and the conductor of \( E \) equals \( N_E \). We will often write simply \( N \) instead of \( N_E \) if there is no risk of confusion. We will be willing to discard finitely many (isomorphism classes of) elliptic curves in several of our arguments, which is the same as letting \( N \) be large enough, by Shafarevich’s theorem. Such a reduction is of course effective.

Let \( N \) be a positive integer. We will say that a factorization \( N = DM \) is admissible if \( D, M \) are coprime positive integers with \( D \) being the product of an even number of distinct prime factors, possibly \( D = 1 \).

Given a positive integer \( N \) with an admissible factorization \( N = DM \), we denote by \( X^0_D(M) \) the canonical model over \( \mathbb{Q} \) (determined by special points) of the compactified Shimura curve associated to an Eichler order of index \( M \) for the quaternion \( \mathbb{Q} \)-algebra of discriminant \( D \). The particular case \( D = 1, M = N \) is simply denoted by \( X_0(N) \). The Jacobian of \( X^0_D(M) \) (resp. \( X_0(N) \)) is denoted by \( J^0_D(M) \) (resp. \( J_0(N) \)). Later, in Section 4, we will review in closer detail some relevant facts about Shimura curves that will be needed in the arguments.

Write \( S_2(N) \) for the complex vector space of weight 2 holomorphic cuspidal modular forms for the congruence subgroup \( \Gamma_0(N) \). Given an elliptic curve \( E \) over \( \mathbb{Q} \), the modularity theorem associates to \( E \) a Hecke eigenform \( f \in S_2(N) \), the space of cuspidal modular forms for \( \Gamma_0(N) \). Here, \( N = N_E \). The form \( f \) is new of level \( N \) and we assume that it is normalized so that its first Fourier coefficient at \( i \infty \) is 1. Associated to \( f \), the Shimura construction gives an optimal quotient \( q_{1,N} : J_0(N) \to A_{1,N} \) defined over \( \mathbb{Q} \), with \( A_{1,N} \) isogenous to \( E \). More generally, for each admissible factorization \( N = DM \), the Jacquet-Langlands correspondence gives an optimal quotient \( q_{D,M} : J^0_D(M) \to A_{D,M} \) defined over \( \mathbb{Q} \), with \( A_{D,M} \) isogenous to \( E \). (As usual, the word “optimal” means “with connected kernel”.)
In this setting, the \((D, M)\)-modular degree of \(E\) (or simply modular degree if the pair \((D, M)\) is understood), denoted by \(\delta_{D,M}\), is defined as follows: The dual map \(q_{D,M}^* : A_{D,M} \to J_0(D)\) satisfies that \(q_{D,M}q_{D,M}^* \in \text{End}(E)\) is multiplication by a positive integer, and \(\delta_{D,M}\) is defined to be this integer. For \(D = 1\), we simply call \(\delta_{1,N}\) the modular degree of \(E\). Although the notation \(\delta_{D,M}\) does not indicate the elliptic curve \(E\), these numbers do depend on \(E\), and sometimes it will be convenient to make this dependence explicit by writing \(\delta_{D,M}(E)\) instead.

The integers \(\delta_{D,M}\) are a central object in this paper. We will later see that they are “almost” equal to the degree of suitably chosen maps \(\phi_{D,M} : X_0(D) \to E_{D,M}\), but a priori they cannot be interpreted in this way. In fact, for \(D \neq 1\) the curve \(X_0(D)\) does not have \(\mathbb{Q}\)-rational points (not even cuspidal points) so it is not a priori clear how to map \(X_0(D)\) to \(J_0(D)\).

In the particular case \(D = 1\), however, the cusp \(\infty\) is \(\mathbb{Q}\)-rational and we can use it to define an embedding \(j_N : X_0(N) \to J_0(N)\). The composition \(\phi = q_{1,N}j_N : X_0(N) \to A_{1,N}\) is well-known to have degree \(\delta_{1,N}\).

We will consider the pull-back of relative differentials in several contexts, so let us introduce the notation once and for all. Given morphisms of \(B\)-schemes \(f : X \to Y\) for a base scheme \(B\), the sheaf \(f^*\Omega^1_{X/B}\) is not in general a sub-sheaf of \(\Omega^1_{X/B}\), and for sections \(s \in H^0(V, \Omega^1_{X/B})\) \((V \text{ open in } Y)\) the section \(f^*s\) belongs to \(H^0(f^{-1}V, f^*\Omega^1_{X/B})\). Nevertheless, there is a canonical map \(f_{X/Y/B}^* : f^*\Omega^1_{X/B} \to \Omega^1_{X/Y/B}\) sitting in the fundamental exact sequence \(f^*\Omega^1_{Y/B} \to \Omega^1_{X/B} \to \Omega^1_{X/Y} \to 0\), and we define \(f^*s \in H^0(f^{-1}V, \Omega^1_{X/B})\) to be the image of \(f^*s\) under \(f_{X/Y/B}^*\). A similar construction applies in the analytic setting.

### 3. Review of the classical modular approach

In this section we recall the classical modular approach to the \(abc\) conjecture and Szpiro’s conjecture. Everything in this section is well-known and we include it for later reference, and to clarify the similarities and differences with our approach in other parts of the paper. In fact, none of our results stated in the introduction is obtained directly from the classical approach reviewed in this section, although some of the ideas will be useful.

An \(abc\) triple is a triple of positive integers \((a, b, c)\) such that \(a, b, c\) are coprime and \(a + b = c\). For each \(abc\) triple \((a, b, c)\) the Frey curve \(E_{a,b,c}\) is defined by the affine equation
\[
y^2 = x(x - a)(x + b).
\]
The absolute value of the discriminant and the conductor of \(E_{a,b,c}\) satisfy \(\Delta_{E_{a,b,c}} \sim (abc)^2\) and \(N_{E_{a,b,c}} \sim \text{rad}(abc)\). Furthermore, \(E_{a,b,c}\) is semistable away from 2, see [65]. From here, it is clear that Conjecture [1.2] implies Conjecture [1.3] and that any partial result for Conjecture [1.2] which applies to Frey elliptic curves yields a partial result for the \(abc\) conjecture.

For an elliptic curve \(E\) over \(\mathbb{Q}\), let \(h(E)\) be the Faltings height of \(E\) over \(\mathbb{Q}\) (this is not the semi-stable Faltings height). For \(\omega_E\) a global Néron differential of \(E\) we have
\[
h(E) = -\frac{1}{2} \log \left( \frac{i}{2} \int_{E(\mathbb{C})} \omega_E \wedge \overline{\omega_E} \right).
\]
Furthermore, by a formula of Silverman [75] we get
\[
\log \Delta_E \leq 12h(E) + 16.
\]
We have the standard modular parameterization \( \phi = q_{1,N}j_N : X_0(N) \to A_{1,N} \) whose degree is \( \delta_{1,N} = \delta_{1,N}(E) \). The degree of a minimal isogeny between \( A_{1,N} \) and \( E \) is uniformly bounded by 163 thanks to Mazur \cite{60} and Kenku \cite{51}, so

\[
|h(A_{1,N}) - h(E)| \leq \frac{1}{2} \log 163 < 3.
\]

Let \( f \in S_2(N) \) be the associated normalized Hecke eigenform and let \( c_f \) denote the (positive) Manin constant of the optimal quotient \( A_{1,N} \), which is defined as the absolute value of the scalar \( c \) satisfying that the pull-back of the Néron differential \( \omega_{A_{1,N}} \) to \( h \) via the composition \( h \to X_0(N) \to A_{1,N} \) is \( 2\pi i c f(z) dz \). The Manin constant is a positive integer (cf. \cite{26}); further details on the Manin constant will be discussed in Section 10.

Let us fix the notation

\[
\mu_h(z) = \frac{dx \wedge dy}{y^2} \quad (z = x + yi \in h)
\]

for the usual hyperbolic measure on \( h \).

Frey \cite{31} observed that by pulling-back the \((1,1)\)-form \( \omega_{A_{1,N}} \wedge \omega_{A_{1,N}} \) from \( A_{1,N} \) to \( X_0(N) \), one obtains

\[
(3.2) \quad \log \delta_{1,N}(E) = 2 \log(2\pi c_f) + 2 \log(\|f\|_{2,\Gamma_0(N)}) + 2h(A_{1,N})
\]

where the Petersson norm of \( f \) is given by

\[
\|f\|^2_{2,\Gamma_0(N)} := \int_{\Gamma_0(N) \setminus h} |f(z)|^2 y^2 d\mu_h(z).
\]

Since the Manin constant \( c_f \) is a positive integer and since \( f \) has Fourier coefficients in \( \mathbb{Z} \), one finds by integrating \( f \cdot \bar{f} \) on \( \{ z \in h : |x| < 1/2 \text{ and } y > 1 \} \) that

\[
\log(2\pi c_f) + \log(\|f\|_{2,\Gamma_0(N)}) > -6.
\]

(Actually, as pointed out in \cite{62}, one has \( 2 \log \|f\|_{2,\Gamma_0(N)} \sim \log N \) with an effective error term by \cite{42}, but the previous lower bound is trivial and sufficient for our current purposes.) Therefore by (3.2) we obtain

\[
(3.3) \quad h(E) \leq h(A_{1,N}) + 3 \leq \frac{1}{2} \log \delta_{1,N} + 9.
\]

The Modular Degree Conjecture \cite{1.9} can be restated in a more intrinsic way as follows

**Conjecture 3.1** (Modular degree conjecture). As \( E \) varies over elliptic curves over \( \mathbb{Q} \), we have

\[
\log \delta_{1,N}(E) \ll \log N_E.
\]

In view of (3.1) and (3.3), it is immediate that the modular degree conjecture implies Szpiro’s conjecture, and that any partial result for the modular degree conjecture gives a partial result for Szpiro’s conjecture.
4. Preliminaries on Shimura curves

In this section we recall some facts about the theory of Shimura curves. We work over \( \mathbb{Q} \) for simplicity, as this is the case that will be used most of the time in this work. Just in Section \[15\] we will need algebraic and arithmetic results about Shimura curves over totally real number fields, and the necessary facts will be recalled there.

The results in this section are by now standard. They can be found in [36], [73], and [94], among other references. We include them to fix the notation and to simplify the exposition.

4.1. Shimura curves. Let \( D \) be a squarefree positive integer with an even number of prime divisors, possibly \( D = 1 \). Let \( B \) the unique (up to isomorphism) quaternion \( \mathbb{Q} \)-algebra ramified exactly at the primes dividing \( D \). For each prime \( p \nmid D \) fix an isomorphism \( B_p = M_2(\mathbb{Q}_p) \), and under this identification we take the maximal order \( O_p = M_2(\mathbb{Z}_p) \). For \( v \mid D \) we also let \( O_p \) denote a maximal order in the completion \( B_p \). These choices determine a maximal order \( O_B \subseteq B \) with \( O_B \otimes \mathbb{Z}_p = O_p \) in \( B_p \), for each prime \( p \). We also fix an identification at the infinite place \( B_\infty = M_2(\mathbb{R}) \), which determines an action of \( B^\times \) on \( \mathfrak{h}^\pm = \mathbb{C} \setminus \mathbb{R} \) by fractional linear transformations.

Let \( A^\infty = \mathbb{Q} \otimes \hat{\mathbb{Z}} \) be the ring of finite adeles of \( \mathbb{Q} \). Let \( \mathbb{B} := B \otimes A^\infty \) and write \( O_B = O_B \otimes \hat{\mathbb{Z}} \), the maximal order in \( \mathbb{B} \) induced by \( O_B \). For each compact open subgroup \( U \subseteq \mathbb{B}^\times \) let \( X_U \) be the compactified Shimura curve over \( \mathbb{Q} \) associated to \( U \). More precisely, the set of complex points of \( X_U \) is the complex curve

\[
X_U^{an} = B^\times \backslash \mathfrak{h}^\pm \times \mathbb{B}^\times /U \cup \{ \text{cusps} \}
\]

and the model over \( \mathbb{Q} \) is the canonical one defined in terms of special points. The cuspidal points are necessary only when \( D = 1 \). The curve \( X_U \) is irreducible over \( \mathbb{Q} \).

4.2. Projective system. For \( U \subseteq V \) compact open subgroups of \( \mathbb{B}^\times \), the natural map \( X_U \to X_V \) is defined over \( \mathbb{Q} \), and they define a projective system \( \{ X_U \} \). The limit is a \( \mathbb{Q} \)-scheme \( X \) whose complex points are given by

\[
X^{an} = B^\times \backslash \mathfrak{h}^\pm \times \mathbb{B}^\times \cup \{ \text{cusps} \}.
\]

The scheme \( X \) comes with a right \( \mathbb{B}^\times \) action by \( \mathbb{Q} \)-automorphisms, under which \( X/U \simeq X_U \). This right \( \mathbb{B}^\times \) action is compatible with the right action on the projective system \( \{ X_U \} \) given by the \( F \)-maps \( g: X_U \to X_{g^{-1}Ug} \) for \( g \in \mathbb{B}^\times \).

4.3. Components. Let \( \mathbb{Q}^+ \) be the set of strictly positive rational numbers. Let \( B^\times_+ \) be the subgroup of \( B^\times \) consisting of elements with reduced norm in \( \mathbb{Q}^+ \). Then \( B^\times_+ \) acts on the upper half plane \( \mathfrak{h} \), and the natural map

\[
B^\times_+ \backslash \mathfrak{h} \times \mathbb{B}^\times /U \to B^\times \backslash \mathfrak{h}^\pm \times \mathbb{B}^\times /U
\]

is an isomorphism. So, we can identify \( X_U^{an} \) with the compactification of the former. Let \( \text{rn} : \mathbb{B} \to A^\infty \) be the reduced norm, then the connected components of \( X_U^{an} \) are indexed by the class group \( C(U) := \mathbb{Q}^+_x \backslash A^\infty,x /\text{rn}(U) \) via the natural map induced by \( \text{rn} \). The number of connected components of \( X_U^{an} \) is \( h_U := \#C(U) \). The component associated to \( a \in C(U) \) is denoted by \( X_{U,a}^{an} \).

Given \( g \in \mathbb{B}^\times \) define

\[
\Gamma_{U,g} := gUg^{-1} \cap B^\times_+.
\]
We see that $\Gamma_{U,g} \subseteq GL_2(\mathbb{R})^+$, and its image $\tilde{\Gamma}_{U,g}$ in $PSL_2(\mathbb{R})$ is a discrete subgroup acting on $\mathfrak{h}$ on the left. This gives rise to the (compactified) complex curve

$$X_{U,g}^{an} := \tilde{\Gamma}_{U,g}\mathfrak{h} \cup \{\text{cusps}\}$$

where the cusps are only needed if $D = 1$. It comes with the obvious complex uniformization

$$\xi_{U,g} : \mathfrak{h} \to X_{U,g}^{an}.$$ 

For each $a \in C(U)$ choose $g_a \in \mathbb{B}^\times$ with $[rn(g_a)] = a$ in $C(U)$. One has the bi-holomorphic bijection

$$\bigcup_{a \in C(U)} \tilde{\Gamma}_{U,g_a}\mathfrak{h} \cup \{\text{cusps}\} \to X_{U,a}^{an}, \quad \tilde{\Gamma}_{U,g_a} : z \mapsto [z, g_a]$$

respecting the projections onto $C(U)$. We can identify the component $X_{U,a}^{an}$ with $X_{U,g_a}^{an}$. Thus, after the choice of $g_a$ for each $a \in C(U)$ is made, there is no harm in simplifying the notation as follows: $\Gamma_{U,a} = \Gamma_{U,g_a}$, $\tilde{\Gamma}_{U,a} = \tilde{\Gamma}_{U,g_a}$, and $X_{U,a}^{an} = X_{U,g_a}^{an}$.

Let $H_U$ be the field extension of $\mathbb{Q}$ associated to $C(U)$ by class field theory. Each component $X_{U,a}^{an}$ has a model $X_{U,a}$ over $H_U$, so that $X_U \otimes H_U = \coprod_{a \in C(U)} X_{U,a}$.

### 4.4. Heegner points

We say that an imaginary quadratic field $K$ satisfies the Heegner hypothesis for $D$ if every prime $p|D$ is inert in $K$. In particular, $K/\mathbb{Q}$ is unramified at primes dividing $D$.

Let $K$ satisfy the Heegner hypothesis for $D$. Then there is a $\mathbb{Q}$-algebra embedding $\psi : K \to B$ which is optimal in the sense that $\psi^{-1}O_B = O_K$; we fix such an optimal embedding. We have that $\psi(K^\times) \subseteq B_+^\times$ and there is a unique $\tau_K \in \mathfrak{h}$ which is fixed by all elements of $\psi(K^\times)$ (a more accurate notation would be $\tau_{K,\psi}$, but any other choice of optimal embedding leads to an equivalent theory). These choices determine the point

$$P_K = [\tau_K, 1] \in B_+^\times \times \mathbb{B}^\times \subseteq X^{an}$$

which maps to a point $P_{K,U} \in X_{U,a}^{an}$ for each compact open subgroup $U$. The point $P_K$ (hence, all the $P_{K,U}$) is algebraic and defined over $K^{ab}$, the maximal abelian extension of $K$. Since $\psi$ was chosen as an optimal embedding, it follows from Shimura reciprocity that the point $P_{K,O_3^\times}$ of $X_{O_3^\times}$ has residue field $H_K$, the Hilbert class field of $K$.

We will refer to these points as $D$-Heegner points.

### 4.5. Jacobians

We adopt the convention that the Jacobian of an irreducible curve is the Albanese variety, so that its dual is the identity component of the Picard variety —this clarification is relevant in terms of functoriality. The Jacobian of $X_U$ is denoted by $J_U$, and the torus of its complex points is $J_U^{an}$. Similarly, for $X_{U,a}$ its Jacobian is $J_{U,a}$, having $J_{U,a}^{an}$ as set of complex points. Note that $J_U$ is defined over $\mathbb{Q}$ and $J_{U,a}$ is defined over $H_U$. Furthermore $J_{U,a}^{an}$ is the Jacobian of the irreducible complex curve $X_{U,a}^{an}$, while the points of $J_{U,a}^{an}$ correspond to divisor classes on $X_{U,a}^{an}$ having degree 0 on each component $X_{U,a}^{an}$, hence $J_{U,a}^{an} = \coprod_{a \in C(U)} J_{U,a}^{an}$. This decomposition is defined over $H_U$, namely $J_U \otimes H_U = \coprod_{a \in C(U)} J_{U,a}$. 

\[14\]
4.6. Modular forms and differentials. The space of cuspidal holomorphic weight 2 modular forms for the discrete subgroup \( \Gamma_{U,a} \) acting on \( \mathfrak{h} \) is denoted by \( S_{U,a} \) (the cuspidal condition is only needed when \( D = 1 \)). All the curves \( X_{U,a}^{an} \) have the same genus \( g_U \), and the uniformization \( \xi_{U,a} : \mathfrak{h} \to \tilde{\Gamma}_{U,a} \backslash \mathfrak{h} \) induces by pull-back an isomorphism \( \Psi_{U,a} : H^0(X_{U,a}^{an}, \Omega^1) \to S_{U,a} \) given by the condition that the image of a differential \( \omega \) is the modular form \( \Psi_{U,a}(\alpha) \in S_{U,a} \) satisfying

\[
\xi_{U,a}^* \omega = \Psi_{U,a}(\alpha)dz
\]

with \( z \) the complex variable on \( \mathfrak{h} \). In particular \( \dim S_{U,a} = g_U \). Thus we get isomorphisms

\[
H^0(X_U, \Omega^1) \otimes \mathbb{C} \cong H^0(X_{U,a}^{an}, \Omega^1) = \bigoplus_{a \in C(U)} H^0(X_{U,a}^{an}, \Omega^1) \cong \bigoplus_{a \in C(U)} S_{U,a}.
\]

The inner product \((-,-)_U\) on \( H^0(X_{U,a}^{an}, \Omega^1) \) is defined so that the direct summands \( H^0(X_{U,a}^{an}, \Omega^1) \) are orthogonal, and for \( \omega_1, \omega_2 \in H^0(X_{U,a}^{an}, \Omega^1) \)

\[
(\omega_1, \omega_2)_U := \frac{i}{2} \int_{X_{U,a}^{an}} \omega_1 \wedge \overline{\omega_2}.
\]

The (non-normalized) Petersson inner product \((-,-)_{U,a}\) on \( S_{U,a} \) is defined by

\[
\langle h_1, h_2 \rangle_{U,a} := \int_{U \backslash \mathfrak{h}} h_1(z) \overline{h_2(z)} \Im(z)^2 d\mu_{\mathfrak{h}}(z)
\]

for \( h_1, h_2 \in S_{U,a} \). We note that for \( \omega_1, \omega_2 \in H^0(X_{U,a}^{an}, \Omega^1) \) we get

\[
(\omega_1, \omega_2)_U = \langle \Psi_{U,a}(\omega_1), \Psi_{U,a}(\omega_2) \rangle_{U,a}.
\]

Thus, \((-,-)_U\) will also be called the Petersson inner product.

4.7. Classical subgroups. Using the maximal order \( O_B \) of \( B \) we recall some standard choices for \( U \). For a positive integer \( n \), define \( U^D(n) = (1 + nO_B)\times \). These open compact subgroups \( U^D(n) \) determine a cofinal system in the projective system \( \{U^D \}_{U^D} \). We say that a compact open \( U \) has level \( n \) if \( U^D(n) \subseteq U \) and for every other positive integer \( n' \) with \( U^D(n') \subseteq U \) we have \( n|n' \).

For \( M \) coprime to \( D \) we define \( U_0^D(M) \) prime by prime as follows. For \( p \nmid M \) finite, we let \( U_0^D(M)_p = O_p \), i.e. maximal at \( p \). For \( p|M \) the identification \( B_p = M_2(\mathbb{Q}_p) \) allows us to define

\[
U_0^D(M)_p = \left\{ g \in GL_2(\mathbb{Z}_p) : g \equiv \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \mod M \right\}.
\]

The subgroup \( U_1^D(M) \) is defined similarly, except that the last condition is replaced by

\[
U_1^D(M)_p = \left\{ g \in GL_2(\mathbb{Z}_p) : g \equiv \begin{bmatrix} * & * \\ 0 & 1 \end{bmatrix} \mod M \right\}.
\]

Both \( U_0^D(M) \) and \( U_1^D(M) \) have level \( M \).

For the groups \( U^D(n), U_0^D(M), \) and \( U_1^D(M) \), the notation \( X_U \) is replaced by \( X^D(n), X_0^D(M), \) and \( X_1^D(M) \) respectively (the upper script \( D \) keeps track of the quaternion algebra, up to isomorphism — all other relevant choices are implicit). Similarly for their Jacobians, sets of complex points, components, spaces of modular forms (always cuspidal of weight 2), etc. To simplify notation, we may omit the upper script \( D \) in the case \( D = 1 \).
The reduced norm maps $U = U_0^D(M_1) \cap U_1^D(M_2)$ surjectively onto the maximal open compact subgroup $\mathbb{Z}^\times \subseteq \mathbb{A}^{\times\times}$, so $C(U)$ is trivial in this case, as the narrow class number of $\mathbb{Q}$ is 1. Thus, the Shimura curve associated to a compact open subgroup of the form $U_0^D(M_1) \cap U_1^D(M_2)$ is geometrically connected.

4.8. **Hecke action.** (cf. [94] Sec. 1.4, 3.2) Let $U$ be a compact open subgroup of $\mathbb{B}^\times$ of level $m$. Recall that for each positive integer $n$ coprime to $Dm$ one has a Hecke correspondence $\mathcal{T}_{U,n}$ on $X_U$. It is defined over $\mathbb{Q}$ and has degree $\sigma_1(n)$, the sum of divisors of $n$, in the sense that the induced push-forward map $\text{Div}(X_U) \to \text{Div}(X_U)$ multiplies degrees by $\sigma_1(n)$.

Hecke correspondences generate a commutative ring $\mathbb{T}_U$. This ring acts (contravariantly) by endomorphisms on $H^0(X_U, \Omega^1)$ via pull-back and trace. Let $\mathbb{T}_U$ be the image of $\mathbb{T}^c_U$ in $\text{End}_\mathbb{Q} H^0(X_U, \Omega^1)$, and denote by $\mathbb{T}_{U,n}$ the image of $\mathcal{T}^c_{U,n}$. The ring $\mathbb{T}_U$ also acts (covariantly) by endomorphisms on $J_U$, and the image of $\mathbb{T}^c_U$ in $\text{End}_\mathbb{Q} J_U$ is isomorphic to $\mathbb{T}_U$. In fact, the action of $\mathbb{T}_U$ on $J_U$ is compatible with that on $H^0(X_U, \Omega^1)$ in the following sense: The action of $\mathbb{T}_U$ on $J_U$ induces by pull-back an action on $H^0(J_U, \Omega^1)$, and the canonical isomorphism $H^0(J_U, \Omega^1) \simeq H^0(X_U, \Omega^1)$ is $\mathbb{T}_U$-equivariant for this action.

4.9. **Systems of Hecke eigenvalues.** For $M$ coprime to $D$ and for $U = U_0^D(M)$ we write $\mathbb{T}_{D,M}$ instead of $\mathbb{T}_U$, similarly for the Hecke operators $\mathbb{T}_{D,M,n}$. The action of $\mathbb{T}_{D,M}$ on $V_{D,M} := H^0(X_0^D(M), \Omega^1) \otimes \mathbb{C}$ is simultaneously diagonalizable, which gives rise to systems of Hecke eigenvalues $\chi : \mathbb{T}_{D,M} \to \mathbb{C}$. Such a $\chi$ takes values in the ring of integers of a totally real number field. The associated isotypical subspaces $V_{D,M}^\chi$ are orthogonal to each other for the Petersson product.

If $m|M$ the map $X_0^D(M) \to X_0^D(m)$ induces by pull-back an inclusion $V_{D,m} \subseteq V_{D,M}$, which satisfies that for $n$ coprime to $DM$ one has $T_{D,M,n}|_{V_{D,m}} = T_{D,m,n}$. Hence we can lift systems of Hecke eigenvalues from $\mathbb{T}_{D,m}$ to $\mathbb{T}_{D,M}$. Those $\chi : \mathbb{T}_{D,M} \to \mathbb{C}$ arising in this way for $m|M$ with $m \neq M$ are called old, and the remaining $\chi$ are called new. The minimal $m|M$ from which $\chi$ arises is called the level of $\chi$. Multiplicity one holds for the new eigenspaces, that is, if $\chi$ is new then $\dim V_{D,M}^\chi = 1$ (cf. [94]).

In the previous discussion, one can of course replace $V_{D,M}$ by $S_{D}^{0}(M) := S_{U_0^D(M)}$.

4.10. **Jacquet-Langlands.** Write $N = DM$ with $M$ coprime to $D$ and observe that our notation gives $S_{1}^{0}(N) = S_{2}(N)$ when $D = 1$, where as usual $S_{2}(N)$ denotes the space of weight 2 holomorphic cuspidal modular forms for the congruence subgroup $\Gamma_{0}(N)$.

For $d$ a divisor of $N$, write $S_{2}(N)^d = \oplus_{d|\chi} S_{2}(N)^{\chi}$ where the notation “$d|\chi$ ” means that $\chi$ varies over the systems of Hecke eigenvalues of $\mathbb{T}_{1,N}$ whose level is divisible by $d$. The Jacquet-Langlands correspondence gives an isomorphism

$$JL : S_{0}^{D}(M) \to S_{2}(N)^{D}$$

such that for every $n$ coprime to $N$ we have

$$JL \circ T_{D,M,n} = T_{1,N,n}|_{S_{2}(N)^{D}} \circ JL.$$ 

Hence, $JL$ induces a quotient map $\mathbb{T}_{1,N} \to \mathbb{T}_{D,M}$ which realizes the systems of Hecke eigenvalues of $\mathbb{T}_{D,M}$ precisely as the systems of Hecke eigenvalues of $\mathbb{T}_{1,N}$ with level divisible by $D$. 

16
4.11. The Shimura construction. We say that two systems of Hecke eigenvalues \( \chi_1, \chi_2 : \mathbb{T}_{D,M} \to \mathbb{Q} \) are equivalent if \( \chi_1 = \sigma \chi_2 \) for some automorphism \( \sigma \in G_\mathbb{Q} \). The equivalence class of \( \chi \) is denoted by \([\chi]\). Note that the degree of the field generated by the values of \( \chi \) is \( \#([\chi]) \), and the property that \([\chi]\) be new is well-defined. All the elements of \([\chi]\) have the same kernel, which we denote by \( \mathbb{I}_{[\chi]} \).

Given a class \([\chi]\) of systems of Hecke eigenvalues, define the (connected) abelian subvariety

\[ K_{[\chi]} := \mathbb{I}_{[\chi]} \cdot J_0^D(M) \leq J_0^D(M). \]

Then \( K_{[\chi]} \) is defined over \( \mathbb{Q} \) and one obtains the quotient \( q_{[\chi]} : J_0^D(M) \to A_{[\chi]} \) with kernel \( K_{[\chi]} \), also defined over \( \mathbb{Q} \). If \([\chi]\) is new, then the abelian variety \( A_{[\chi]} \) is simple over \( \mathbb{Q} \) and has dimension \( \#([\chi]) \). Since \( K_{[\chi]} \) is \( \mathbb{T}_{D,M} \)-stable, \( \mathbb{T}_{D,M} \) also acts on \( A_{[\chi]} \) making \( q_{[\chi]} \) Hecke equivariant.

Furthermore, \( \text{End}(A_{[\chi]}) \) contains a ring \( O_{[\chi]} \) isomorphic to \( \chi(\mathbb{T}_{D,M}) \) (for any \( \chi \) in \([\chi]\)) which is an order in a totally real field of degree \( \#([\chi]) \). Fixing any choice of \( \chi \) and isomorphism \( O_{[\chi]} \simeq \chi(\mathbb{T}_{D,M}) \), we have that \( \mathbb{T}_{D,M} \) acts on \( A_{[\chi]} \) via \( \chi \). Thus, the kernel of \( \mathbb{T}_{D,M} \to \text{End}(A_{[\chi]}) \) is precisely \( \mathbb{I}_{[\chi]} \).

The maps \( q_{[\chi]} \) induce an isogeny \( J_0^D(M) \to \prod_{[\chi]} A_{[\chi]} \). Given any \([\chi]\) define \( B_{[\chi]} \) as the identity component of the kernel of \( J_0^D(M) \to \prod_{[\chi] \neq [\chi]} A_{[\chi]} \). That is, \( B_{[\chi]} \) is the identity component of \( \bigcap_{[\chi] \neq [\chi]} K_{[\chi]} \). The abelian variety \( B_{[\chi]} \) is defined over \( \mathbb{Q} \), it is stable under the action of \( \mathbb{T}_{D,M} \), and it is the largest abelian subvariety of \( J_0^D(M) \) on which \( \mathbb{I}_{[\chi]} \) acts as 0. The composition \( B_{[\chi]} \to J_0^D(M) \to A_{[\chi]} \) is an isogeny defined over \( \mathbb{Q} \), thus, \( J_0^D(M) = \sum_{[\chi]} B_{[\chi]} \).

5. The degree of Shimura curve parameterizations

5.1. The modular degree. In this section we fix an elliptic curve \( A \) over \( \mathbb{Q} \) of conductor \( N \) which is an optimal quotient \( q : J_0^D(M) \to A \) for an admissible factorization \( N = DM \). The associated modular degree \( \delta_{D,M} \) will be simply denoted by \( \delta \). The kernel of \( q \) is denoted by \( K \), the image of \( q^\vee : A \to J_0^D(M) \) is denoted by \( B \), and the system of Hecke eigenvalues attached to \( A \) is denoted by \( \chi_0 \). Note that \( \chi_0 \) is \( \mathbb{Z} \)-valued. Also, since \( q^\vee \) is an injective, we get an isomorphism \( A[\delta] \to B[\delta] = M \cap B \) which respects both the Galois action and the Hecke action.

The goal of this section is to develop some tools to handle the quantity \( \delta \).

5.2. Maps to the Jacobian. Since \( X_0^D(M) \) does not have a canonical embedding into \( J_0^D(M) \) when \( D > 1 \), we cannot a priori interpret \( \delta \) as the degree of a map \( X_0^D(M) \to A \), unlike the case of classical modular parameterizations.

Let us momentarily address the general case. Let \( U \subseteq \mathbb{B}^\times \) be a compact open subgroup of level \( m \). Let \( t : X_U \to X_U \) be an algebraic correspondence defined over \( \mathbb{Q} \) of degree 0 (meaning that it multiplies degrees of divisors by 0) with the following additional property (*) : \( t \) maps complex points \( x \in X_U^{an} \) to degree zero divisors \( t(x) \) supported on the same complex component containing \( x \). Then for each \( a \in C(U) \) one gets a morphism \( j_{t,a} : X_{U,a}^{an} \to J_{U,a}^{an} \), and via the canonical isomorphism \( J_{U,a}^{an} = \prod_a J_{U,a}^{an} \) they descend to a morphism \( j_t : X_U \to J_U \) defined over \( \mathbb{Q} \). By construction, for each \( a \in C_U^{an} \) the following diagram

\[ \begin{array}{ccc} X_{U,a}^{an} & \to & X_{U,a}^{an} \\ \downarrow & \downarrow & \downarrow \\ J_{U,a}^{an} & \to & J_{U,a}^{an} \\ \end{array} \]

would be valid.
commutes:

\[
\begin{array}{ccc}
X_{U,n}^{an} & \longrightarrow & X_{U}^{an} \\
\downarrow & & \downarrow \\
\text{Jac}(X_{U,n}^{an}) & \longrightarrow & J_{U}^{an}.
\end{array}
\]

We obtain a morphism \( j_{t}^{*} : H^{0}(J_{U}, \Omega^{1}) \rightarrow H^{0}(X_{U}, \Omega^{1}) \) which has no reason to be an isomorphism (e.g. take \( t = 0 \)). Since the ring \( \mathbb{T}_{U} \) of Hecke correspondences is commutative, we get that if in addition \( t \in \mathbb{T}_{U} \), then for every \( n \) coprime to \( Dm \) we have \( j_{t}^{*} \circ T_{U,n} = T_{U,n} \circ j_{t}^{*} \). In particular, the map \( j_{t}^{*} \) is \( \mathbb{T}_{U} \)-equivariant.

A particularly convenient choice of \( t \in \mathbb{T}_{U} \) of degree 0 is given by the Eisenstein correspondences

\[
E_{U,n}^{c} := T_{U,n} - \sigma_{1}(n) \cdot \Delta_{U}
\]

where \( \Delta_{U} \) is the diagonal correspondence of \( X_{U} \), and \( n \) is coprime to \( Dm \).

In the particular case \( U = U_{0}^{D}(M) \), taking \( n \) coprime to \( N = DM \) and recalling that \( C(U_{0}^{D}(M)) \) is trivial (thus, condition (*) trivially holds), we can define \( j_{D,M,n} : X_{0}^{D}(M) \rightarrow J_{0}^{D}(M) \) as the map \( j_{t} \) with \( t = E_{D,M,n}^{c} := E_{U_{0}^{D}(M),n}^{c} \). This need not be an embedding.

5.3. Shimura curve parameterizations. Returning to our setting \( U = U_{0}^{D}(M) \) and the elliptic curve \( A \) arising as an optimal quotient \( q : J_{0}^{D}(M) \rightarrow A \), for each \( n \) coprime to \( N \) we obtain a map

\[
\phi_{D,M,n} = q j_{D,M,n} : X_{0}^{D}(M) \rightarrow A
\]

defined over \( \mathbb{Q} \).

**Proposition 5.1.** The morphism \( \phi_{D,M,n} \) is non-constant of degree \((a_{n}(A) - \sigma_{1}(n))^{2} \cdot \delta \), where \( a_{n}(A) \) is the \( n \)-th coefficient of the L-function of the elliptic curve \( A \).

**Proof.** Choose any complex point \( p_{0} \in X_{0}^{D}(M)^{an} \) and consider the embedding \( j_{p_{0}} : X_{0}^{D}(M)^{an} \rightarrow J_{0}^{D}(M)^{an} \) over \( \mathbb{C} \), given by \( x \mapsto [x - p_{0}] \). One easily checks that the \( \mathbb{C} \)-map \( \phi_{p_{0}} = q j_{p_{0}} : X_{0}^{D}(M)^{an} \rightarrow A_{\mathbb{C}} \) has degree equal to the modular degree \( \delta \).

Let \( E_{D,M,n} \) be the image of \( E_{D,M,n}^{c} \) under \( \mathbb{T}_{D,M} \rightarrow \mathbb{T}_{D,M} \subseteq \text{End}(J_{0}^{D}(M)) \). A direct computation shows that for each complex point \( x \in X_{0}^{D}(M)^{an} \)

\[
j_{D,M,n}(x) = (E_{D,M,n} \circ j_{p_{0}})(x) + j_{D,M,n}(p_{0}).
\]

By the Eichler-Shimura congruence relation we have that \( E_{D,M,n} \) acts on \( A \) as multiplication by \( a_{n}(A) - \sigma_{1}(n) \). Composing with \( q \) we get that for every complex point \( x \in X_{0}^{D}(M)^{an} \)

\[
\phi_{D,M,n}(x) = (a_{n}(A) - \sigma_{1}(n)) \cdot \phi_{p_{0}}(x) + \phi_{D,M,n}(p_{0}).
\]

Note that \( a_{n}(A) - \sigma_{1}(n) \neq 0 \) by Hasse’s bound. The result follows. \( \square \)

5.4. Spectral decomposition of the modular degree. Given \( [\chi] \) an equivalence class of systems of eigenvalues on \( \mathbb{T}_{D,M} \) different to \([\chi_{0}]\), let us define the congruence modulus

\[
\eta_{[\chi_{0}]}([\chi]) := [\mathbb{Z} \cdot f : \mathbb{I}_{[\chi]} \cdot f]
\]

where \( f \) is any non-zero element of \( V_{D,M}^{\chi_{0}} \) (unique up to scalar). It is easy to see that the primes dividing \( \eta_{[\chi_{0}]}([\chi]) \) correspond to congruence primes in a suitable sense, but more is true. This positive integer measures the congruences between the systems of Hecke eigenvalues \( \chi \) and \( \chi_{0} \) in the following precise sense.
Proposition 5.2. The number \( \eta_{[\chi_0]}([\chi]) \) is the largest positive integer \( m \) with the following property:

Given any polynomial \( P \in \mathbb{Z}[x_1, \ldots, x_\ell] \) and positive integers \( n_1, \ldots, n_\ell \) coprime to \( N \), if

\[
P(\chi(T_{D,M,n_1}), \ldots, \chi(T_{D,M,n_\ell})) = 0,
\]

then \( m \) divides the integer \( P(\chi_0(T_{D,M,n_1}), \ldots, \chi_0(T_{D,M,n_\ell})) \), which equals \( P(a_{n_1}(A), \ldots, a_{n_\ell}(A)) \).

Proof. Clear from the definitions and the relation between the values of \( \chi_0 \) and the \( L \)-function of \( A \) given by the Eichler-Shimura congruence. \( \square \)

The previous proposition not only justifies the name of the congruence modulus, but also, it will be useful for estimating the size of \( \eta_{[\chi_0]}([\chi]) \). In the classical setting of the modular curve \( X_0(N) \), the numbers \( \eta_{[\chi_0]}([\chi]) \) measure congruences between eigenforms and can be computed using Fourier expansions (of course, in the more general setting the Fourier expansions are not available).

The relation with the \((D, M)\)-modular degree is given by

Theorem 5.3. The \((D, M)\)-modular degree \( \delta \) divides

\[
\prod_{[\chi] \neq [\chi_0]} \eta_{[\chi_0]}([\chi]).
\]

The product runs over all classes \([\chi]\) of systems of Hecke eigenvalues on \( \mathbb{T}_{D,M} \) different to \([\chi_0]\).

We give two different proofs of this fact, as they can be of independent interest.

5.5. First proof: torsion of abelian varieties. Enumerate the classes of systems of Hecke eigenvalues on \( \mathbb{T}_{D,M} \) as \( c_0, c_1, \ldots, c_s \) with \( c_0 = [\chi_0] \). Write \( A_j \) and \( B_j \) instead of \( A_{c_j} \) and \( B_{c_j} \), in particular \( A_0 = A \) and \( B_0 = B \). Define the following abelian sub-varieties of \( J_0^D(M) \):

\[
C_i := \sum_{j=1}^s B_j, \quad i = -1, 0, 1, \ldots, s.
\]

For each \( 0 \leq j \leq s \) define \( G_j = B_j \cap C_j \). Note that \( B_j/G_j \) is an abelian variety defined over \( \mathbb{Q} \) with a natural \( \mathbb{T}_{D,M} \)-action, such that the quotient \( B_j \to B_j/G_j \) is a \( \mathbb{T}_{D,M} \)-equivariant isogeny.

Consider the addition maps \( \sigma_j : B_j \times C_j \to C_{j-1} \). Geometrically, their kernels are

\[
\ker(\sigma_j) = \{(P, -P) : P \in G_j(F)\}.
\]

If we compose the induced isomorphism \( C_{j-1} \cong (B_j \times C_j)/\ker(\sigma_j) \) with the projection onto \( B_j/G_j \) and restrict to \( B_0 \cap C_{j-1} \), then for each \( 1 \leq j \leq s \) we obtain a map

\[
u_j : B_0 \cap C_{j-1} \to B_j/G_j
\]

defined over \( \mathbb{Q} \). Note that \( u_j \) is \( \mathbb{T}_{D,M} \)-equivariant with respect to the \( \mathbb{T}_{D,M} \)-actions on \( B_0 \cap C_{j-1} \) and on \( B_j/G_j \).

First proof of Theorem 5.3. Let \( p \) be a prime number and let \( \alpha = v_p(\delta) \). We will show that \( p^\alpha \) divides \( \prod_{j=1}^s \eta_{c_0}(c_j) \).
We inductively define algebraic points \( Q_j \in J_0^D(M) \) and integers \( \gamma_j \geq 0 \) for \( 0 \leq j \leq s \). Let \( Q_0 \in B_0 \cap C_0 = B_0[\delta] \) be any algebraic point of order exactly \( p^\alpha \) and set \( \gamma_j = 0 \). For \( 0 \leq j < s \), let

\[
\gamma_{j+1} := \min\{ \gamma \geq 0 : p^\gamma Q_j \in C_{j+1} \} \quad \text{and} \quad Q_{j+1} := p^{\gamma_{j+1}} Q_j.
\]

Note that \( \gamma_{j+1} \) exists and it is at most \( \alpha \), and that for each \( j \) we have that \( Q_j \in C_j \). Also, \( Q_0 \) has order \( p^\alpha \) and \( Q_s = 0 \) because \( C_s = (0) \). All the points \( Q_j \) are multiples of \( Q_0 \) so that they have order a power of \( p \), and more precisely for \( 1 \leq j \leq s \) we have that

\[
\text{ord}(Q_j) = \text{ord}(Q_{j-1})/p^{\gamma_j}.
\]

Hence, for each \( 0 \leq j \leq s \) we have \( \text{ord}(Q_j) = p^\alpha - \sum_{i=j}^{s} \gamma_i \) and taking \( j = s \) we get \( \alpha = \sum_{i=0}^{s} \gamma_i \).

Thus, it suffices to show that \( p^{\gamma_j} \) divides \( \eta_0(c_j) = \lfloor Zf : \mathbb{I}_{c_j}f \rfloor \) for each \( 1 \leq j \leq s \).

Let us fix an index \( 1 \leq j \leq s \). We have \( Q_{j-1} \in B_0 \) as it is a multiple of \( Q_0 \). Hence \( Q_{j-1} \in B_0 \cap C_{j-1} \), thus we can evaluate \( u_j \) at \( Q_{j-1} \) to get the point \( u_j(Q_{j-1}) \in B_j/G_j \).

We claim that \( p^{\gamma_j} \) divides \( \text{ord}(u_j(Q_{j-1})) \). If \( \gamma_j = 0 \) then there is nothing to prove, so let us consider the case \( \gamma_j > 0 \). Observe that \( \text{ord}(u_j(Q_{j-1})) \) is a power of \( p \), so it suffices to show that given any integer \( 0 \leq \gamma < \gamma_j \) one has \( p^\gamma u_j(Q_{j-1}) \neq 0 \). For the sake of contradiction, suppose that \( p^\gamma u_j(Q_{j-1}) = 0 \) for some \( 0 \leq \gamma < \gamma_j \). Recall that \( Q_{j-1} \in C_{j-1} \), and choose some \( (Q'_{j-1}, Q''_{j-1}) \in \sigma_{j-1}^{-1}(Q_{j-1}) \). Then \( p^\gamma Q'_{j-1} \in G_j \) because \( p^\gamma u_j(Q_{j-1}) = 0 \). This means that \((-p^\gamma Q'_{j-1}, p^\gamma Q''_{j-1}) \in \ker(\sigma_j) \), and since \( (Q'_{j-1}, Q''_{j-1}) \in \sigma_{j-1}^{-1}(Q_{j-1}) \), we deduce that the following point is in \( \sigma_{j-1}^{-1}(p^\gamma Q_{j-1}) \subseteq B_j \times C_j \):

\[
(-p^\gamma Q'_{j-1}, p^\gamma Q''_{j-1}) + p^\gamma (Q'_{j-1}, Q''_{j-1}) = (0, p^\gamma Q_{j-1}).
\]

In particular, \( p^\gamma Q_{j-1} \in C_j \) which contradicts the minimality of \( \gamma_j \). Thus, \( p^{\gamma_j} \) divides \( \text{ord}(u_j(Q_{j-1})) \).

Finally, take any \( t \in \mathbb{I}_{c_j} \) and \( 0 \neq f \in V_{D,M}^{\chi_0} \). Then \( tf = \chi_0(t)f \). Since \( t \) annihilates \( B_j \), it also annihilates \( B_j/G_j \), thus \( t \cdot u_j(Q_{j-1}) = 0 \). As \( Q_{j-1} \in B_0 \), we get \( tQ_{j-1} = \chi_0(t)Q_{j-1} \) so that

\[
0 = t \cdot u_j(Q_{j-1}) = u_j(tQ_{j-1}) = u_j(\chi_0(t)Q_{j-1}) = \chi_0(t)u_j(Q_{j-1}).
\]

Thus \( \text{ord}(u_j(Q_{j-1})) \) divides \( \chi_0(t) \), which gives \( p^{\gamma_j}|\mu_t \). Therefore \( p^{\gamma_j} \) divides \( \lfloor Zf : \mathbb{I}_{c_j}f \rfloor \). \( \Box \)

5.6. **Second proof: projectors and the Hecke algebra.** Write \( \mathbb{E} = \text{End}(J_0^D(M)) \). Then \( \mathbb{E} \) has a right action via pull-back on \( H^0(J_0^D(M)^{an}, \Omega^1) \) which is canonically isomorphic to \( V_{D,M} \). This right action on \( V_{D,M} \) extends the action of the commutative subring \( \mathbb{T}_{D,M} \subseteq \mathbb{E} \). Since \( \mathbb{E} \) acts faithfully on \( V_{D,M} \) we can identify \( \mathbb{E}^{op} \) with its image in \( \text{End}_C(V_{D,M}) \).

Let \( \pi_0 \in \text{End}_C(V_{D,M}) \) be the orthogonal projection (with respect to the Petersson inner product) onto the subspace \( V_{D,M}^{\chi_0} \). Equivalently, \( \pi_0 \) is the projection onto the factor \( V_{D,M}^{\chi_0} \) in the direct sum decomposition \( V_{D,M} = \oplus_{\chi} V_{D,M}^{\chi} \).

**Lemma 5.4.** \( \pi_0 \in \mathbb{E}^{op} \otimes \mathbb{Q} \). Furthermore, the denominator of \( \pi_0 \) with respect to \( \mathbb{E}^{op} \) is the modular degree \( \delta \). That is, \( \delta \) is the least positive integer \( m \) with the property that \( m\pi_0 \in \mathbb{E}^{op} \).

**Proof.** In the case of the classical modular curve \( X_0(N) \) this is proved in \([\text{[17]}]\), see also \( \text{[3]} \). The general case is not harder, and we include a proof for the convenience of the reader.

Consider the map \( \varpi = q^\ast \circ q : J_0^D(M) \to J_0^D(M) \) where we recall that \( q : J_0^D(M) \to A \). Note that \( \varpi \in \mathbb{E} \) and \( \varpi|_B : B \to B \) is multiplication by \( \delta \), by definition of the modular
degree. Taking pull-back of holomorphic differentials and recalling the isogeny decomposition $J_0^D(M) \to \prod_{[\chi]} A_{[\chi]}$ we see that $\varpi = \delta \pi_0$ in $\text{End}_C(V_{D,M})$.

It only remains to prove that given any positive integer $m$ such that $\varpi_m := m \pi_0$ is in $\mathbb{E}$, on has that $\delta$ divides $m$. In fact, take such an $m$. From the isogeny decomposition of $J_0^D(M)$ one deduces that $\varpi_m \cdot J_0^D(M) = B \subseteq J_0^D(M)$. Hence, we obtain a map $J_0^D(M) \to B \cong A$ defined over $\mathbb{Q}$ whose restriction to $B$ is multiplication by $m$. Recalling the definition of $\delta$ in terms of the optimal quotient map $q$, we see that $\delta$ divides $m$. \hfill \square

**Second proof of Theorem 5.3.** For each system of Hecke eigenvalues $\chi : \mathbb{T}_{D,M} \to \mathbb{Q}$, let $L_\chi$ be the totally real number field generated by $\chi(\mathbb{T}_{D,M})$. The rule $t \mapsto (\chi(t))_\chi$ defines a ring morphism

$$\psi : \mathbb{T}_{D,M} \to \prod_\chi L_\chi$$

which is injective because $\mathbb{T}_{D,M}$ acts faithfully on $V_{D,M} = \bigoplus_\chi V_{D,M}^\chi$.

For each class $c$ of systems of Hecke eigenvalues on $\mathbb{T}_{D,M}$, define the ring $R_c := \prod_{\chi \in c} L_\chi$ so that we can re-write the previous ring morphism as

$$(5.1) \quad \psi : \mathbb{T}_{D,M} \to \prod_c R_c.$$

Fix a choice of non-zero $f \in V_{D,M}^{\chi_0}$. For each class $c' \neq [\chi_0]$, choose an element $t_{c'} \in \mathbb{I}_{c'}$ such that $t_{c'} \cdot f = \eta_{[\chi_0]}(c') \cdot f$. Then we have that the $[\chi_0]$-component of $\psi(t_{c'})$ in $(5.1)$ is $\eta_{[\chi_0]}(c')$, while the $c'$-component of $\psi(t_{c'})$ is 0.

Finally, let $\varpi = \prod_{c \neq [\chi_0]} t_c \in \mathbb{T}_{D,M} \subseteq \mathbb{E}$ and observe that $\psi(\varpi) \in \prod_c R_c$ has the integer $\prod_{c \neq [\chi_0]} \eta_{[\chi_0]}(c)$ in the $[\chi_0]$-component, and 0 in all other components. It follows that

$$\varpi = \left( \prod_{c \neq [\chi_0]} \eta_{[\chi_0]}(c) \right) \pi_0,$$

and by Lemma 5.3 we get that $\delta$ divides $\prod_{c \neq [\chi_0]} \eta_{[\chi_0]}(c)$. \hfill \square

6. A Refinement of the Ribet-Takahashi Formula

6.1. **Notation.** Consider an elliptic curve $E$ defined over $\mathbb{Q}$ with conductor $N = N_E$ and minimal discriminant $\Delta_E$. Recall that for each admissible factorization $N = DM$ we have the optimal quotients $q_{D,M} : J_0^D(M) \to A_{D,M}$ with modular degree $\delta_{D,M}$ respectively. The elliptic curves $A_{D,M}$ are isogenous over $\mathbb{Q}$ to $E$, although they need not be isomorphic.

In this section we consider $E$ (and $N$) as varying. So, the implicit constants in error terms will always be independent of $E$. In fact, as always in this paper, any dependence of the implicit constants on other parameters will be indicated explicitly as a subscript.

6.2. **The formula.** Ribet and Takahashi (cf. [71] and [81]) proved a formula for the fraction $\delta_{1,N}/\delta_{D,M}$, which in the case when $E$ has no non-trivial rational cyclic isogeny and $M$ is squarefree, reads

$$\frac{\delta_{1,N}}{\delta_{D,M}} = \prod_{p \mid D} v_p(\Delta_E).$$

The requirement that $M$ be squarefree can be relaxed, thanks to the argument in [81]. However, the above equality is known to be false in cases when $E$ has non-trivial isogenies,
and the formula of Ribet and Takahashi in the general case includes a correction factor which takes into account reducible residual Galois representations.

The correction factor is a rational number supported on primes $\ell$ for which the Galois representation $E[\ell]$ is reducible, although the exponents of those primes are not controlled in the literature.

We will need the following version of the Ribet-Takahashi formula with finer control on the correction factor, not just the primes of its support.

**Theorem 6.1.** Let $S$ be a finite set of primes. There is a constant $K_S$ depending only on $S$ (in particular, independent of $E$) with the following property:

If $M$ is squarefree away from $S$, then there is a rational number $\gamma_{D,M,E}$ supported on primes $\leq 163$, with numerator bounded from above by $163^{\omega(D)}$ and denominator bounded from above by $K_S^{\omega(D)}$, such that

$$\frac{\delta_{1,N}}{\delta_{D,M}} = \gamma_{D,M,E} \cdot \prod_{\ell \mid D} v_\ell(\Delta_E).$$

In particular, if $D \neq 1$ (thus, $D \geq 6$) we have

$$\log \delta_{1,N} \leq \log \delta_{D,M} + \log \left( \prod_{\ell \mid D} v_\ell(\Delta_E) \right) + 5.1 \cdot \omega(D)$$

and

$$\log \delta_{1,N} - \log \delta_{D,M} = \log \left( \prod_{\ell \mid D} v_\ell(\Delta_E) \right) + O_S \left( \frac{\log D}{\log \log D} \right)$$

where the implied constant depends at most on the set $S$.

At this point, let us make a heuristic (and admittedly naive) remark. The formula in the previous theorem can be seen as an arithmetic analogue of partial derivatives in the following sense:

Given a monomial $x_1^{e_1} \cdots x_n^{e_n}$ and a subset $J \subseteq \{1, \ldots, n\}$ we can recover the product of the exponents $e_j$ for $j \in J$ as a “rate of change in the direction of the selected variables $x_j$ for $j \in J$”

$$\prod_{j \in J} e_j = \frac{\partial^{\#J}}{\partial x_1^{e_1} \cdots x_n^{e_n} |_{x=(1, \ldots, 1)}}.$$

If we write the factorization $\Delta_E = \prod_{\ell \mid N} p_\ell^{v_\ell(\Delta_E)}$ and consider the expression on the right as a monomial, then “formal partial derivatives with respect to the primes $\ell \mid D$” would produce the factor $\prod_{\ell \mid D} v_\ell(\Delta_E)$. On the other hand, the fraction $\delta_{1,N}/\delta_{D,M}$ can be seen as a “rate of change in the direction of the selected primes $\ell \mid D$”, since it measures the variation of the degree of a modular parametrization of the elliptic curve $E$ under consideration when we change the modular curve $X_0(N)$ by the Shimura curve $X_0^D(M)$.

While this heuristic might be relevant in the context of the usual analogies between function fields and number fields, we remark that the arguments in this paper do not depend on it.
6.3. Lemmas on congruences.

**Lemma 6.2.** Let \( \ell > 163 \) be a prime. For every elliptic curve \( A \) over \( \mathbb{Q} \) of conductor \( N \) there are infinitely many primes \( r \nmid \ell N \) satisfying
\[
a_r(A) \not\equiv r + 1 \pmod{\ell}.
\]

*Proof.* As \( \ell > 163 \), from [60] we have that \( A[\ell] \) is an irreducible Galois module, and the claim follows from the proof of Theorem 5.2(c) in [69]. \( \square \)

If \( A \) is an elliptic curve over \( \mathbb{Q} \) and \( m \) is a positive integer, the Galois representation on \( A[m] \) is denoted by
\[
\rho_{A[m]} : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}/m\mathbb{Z}).
\]
This depends on a choice of isomorphism \( \mathbb{A} \), and note also that
\[
A,\ell \in \mathbb{A},...H \subseteq \mathbb{A},\ell \subseteq \mathbb{A},\ell + 1 \pmod{\ell}.
\]

**Lemma 6.3.** Let \( \ell \) be a prime. There is a positive integer \( c(\ell) \) depending only on \( \ell \) such that for any given \( n \geq c(\ell) \) there is a finite set \( \mathcal{F}(\ell, n) \subseteq \mathbb{Q} \) with the following property:

Let \( A \) be an elliptic curve defined over \( \mathbb{Q} \). If \( j_A \notin \mathcal{F}(\ell, n) \) (the \( j \)-invariant of \( A \)), then
\[
im(\rho_{A[\ell]}(I)) \supseteq \{ \gamma \in SL_2(\mathbb{Z}/\ell^n\mathbb{Z}) : \gamma \equiv I \pmod{\ell^c} \}.
\]

*Proof.* For each positive integer \( m \) and each subgroup \( H \leq GL_2(\mathbb{Z}/m\mathbb{Z}) \) such that \( \det(H) = (\mathbb{Z}/m\mathbb{Z})^* \), there is a congruence subgroup \( \Gamma_H \leq SL_2(\mathbb{Z}) \) and a geometrically connected, open congruence modular curve \( Y_H \) defined over \( \mathbb{Q} \) with a map \( \pi_H : Y_H \to Y(1) = \mathbb{A}^1 \) over \( \mathbb{Q} \). The set of complex points of \( Y_H \) is \( \Gamma_H \backslash \mathbb{I} \). Distinct \( m \) and \( H \) can give the same \( \Gamma_H \). It is a classical result that for congruence subgroups the genus grows with the level (cf. [21]).

If \( A \) is an elliptic curve over \( \mathbb{Q} \) with \( \rho_{A[m]}(G_{\mathbb{Q}}) \) conjugate to a subgroup of such an \( H \), then it gives rise to a \( \mathbb{Q} \)-rational point \( P_H(A, m) \in Y_H(\mathbb{Q}) \) satisfying \( \pi_H(P_H(A, m)) = j_A \).

Given a positive integer \( m \) and an elliptic curve \( A/\mathbb{Q} \) we define the subgroup
\[
H_{A,m} = \rho_{A[m]}(G_{\mathbb{Q}}) \leq GL_2(\mathbb{Z}/m\mathbb{Z}).
\]
Observe that \( \det(H_{A,m}) = (\mathbb{Z}/m\mathbb{Z})^* \) because \( \det(\rho_{A[m]}(Frob_p)) = p \pmod{m} \) for all but finitely many primes \( p \), and note also that \( P_{H_{A,m}}(A, m) \in Y_{H_{A,m}}(\mathbb{Q}) \).

Given a positive integer \( c \) and taking \( m = \ell^n \) for some \( n \geq c \), the failure of (6.1) can be expressed by saying that \( H_{A,\ell^n} \) does not contain the group
\[
S_{c,\ell^n} := \{ \gamma \in SL_2(\mathbb{Z}/\ell^n\mathbb{Z}) : \gamma \equiv I \pmod{\ell^c} \}.
\]
We note that the level of \( \Gamma_{H_{A,\ell^n}} \) is certain power of \( \ell \). If (6.1) fails, then the level of \( \Gamma_{H_{A,\ell^n}} \) is larger than \( \ell^c \); for otherwise \( H_{A,\ell^n} \) would contain \( S_{c,\ell^n} \). Thus, suitable choice of \( c \) will ensure that whenever (6.1) fails for \( A \) and \( n \geq c \), one has that \( Y_{H_{A,\ell^n}} \) has geometric genus at least 2. For \( \ell \) and \( n \geq c \) fixed, there are only finitely many groups \( H_{A,\ell^n} \leq GL_2(\mathbb{Z}/\ell^n\mathbb{Z}) \) as we vary \( A \), all of them giving modular curves of geometric genus at least 2. Let \( \Gamma_1, ... \Gamma_t \) be these finitely many subgroups, then we can take \( \mathcal{F}(\ell, n) = \bigcup_{i=1}^t \pi_{\Gamma_i}(Y_{\Gamma_i}(\mathbb{Q})) \), which is finite by Faltings’s theorem. We take \( c(\ell) = c \). \( \square \)

**Lemma 6.4.** Let \( \ell \) be a prime. There is a positive integer \( b = b_\ell \) and a finite set \( \mathcal{G}_\ell \subseteq \mathbb{Q} \), both depending only on \( \ell \), such that the following holds:

Let \( A \) be an elliptic curve over \( \mathbb{Q} \) with \( j_A \notin \mathcal{G}_\ell \). There are infinitely many primes \( r \nmid N\ell \) with \( a_r(A) \not\equiv r + 1 \pmod{\ell^b} \).

23
Proof. Let $c = c(\ell)$ be as in Lemma 6.3. We claim that $b = 4c$ and $\mathcal{G}_\ell = \mathcal{F}(\ell, b)$ have the desired property. For the sake of contradiction, let $A$ be an elliptic curve over $\mathbb{Q}$ with $j_A \not\in \mathcal{G}_\ell$ and suppose that all sufficiently large primes $r$ satisfy

$$X^2 - a_r(A)X + r \equiv (X - 1)(X - r) \mod \ell^b.$$ 

The Cayley-Hamilton theorem applied to the action on $A[\ell^b]$ of the Frobenius element $F_r = \rho_{A[\ell^b]}(Frob_r)$, gives that $(F_r - 1) \circ (F_r - r)$ is the zero endomorphism on $A[\ell^b]$ for all but finitely many primes $r$. Note that by Lemma 6.3

$$\gamma := \left[ 1 + \frac{\ell b c}{\ell c} \frac{\ell c}{1} \right] \in im(\rho_{A[\ell^b]})$$

so the Chebotarev density theorem gives infinitely many primes $r$ for which $F_r = \gamma$. We have

$$\ker(\gamma - 1) = \ker \left[ \frac{\ell 2c}{\ell c} \frac{\ell c}{0} \right] = \ell^{b-c}A[\ell^c]$$

so that $\# \ker(\gamma - 1) = \ell^{2c}$. On the other hand, we have that

$$\gamma - r = \left[ 1 - r + \frac{\ell 2c}{\ell c} \frac{\ell c}{1 - r} \right]$$

has image of size at least $\ell^{b-c}$ (by looking at the top-right entry, say), so that $\# \ker(\gamma - r) \leq \ell^{b+c}$. It follows that for each of the infinitely many primes $r$ with $F_r = \gamma$ we have

$$\ell^{2b} = \# \ker((F_r - 1) \circ (F_r - r)) \leq \# \ker(\gamma - 1) \cdot \# \ker(\gamma - r) \leq \ell^{b+3c}$$

which is not possible, because $b = 4c$. \qed

Lemma 6.5. Let $S$ be a finite set of primes. For all primes $\ell$ there is an integer $\beta_S(\ell)$ with the following properties:

(i) If $\ell > 163$, then we have $\beta_S(\ell) = 1$.

(ii) For every elliptic curve $A/\mathbb{Q}$ semi-stable away from $S$, there are infinitely many primes $r$ such that $a_r(A) \not\equiv r + 1 \mod \ell^{\beta_S(\ell)}$.

Proof. Given a finite set of primes $S$ and a finite set of rational numbers $J$, there are only finitely many elliptic curves over $\mathbb{Q}$ with $J$-invariant in $J$ and semi-stable reduction outside $S$. The result now follows from Lemmas 6.2 and 6.4. \qed

6.4. Component groups. Let $A$ be an abelian variety over $\mathbb{Q}$ and let $\mathcal{A}$ be its Néron model. For a prime $p$, we denote by $\Phi_p(A)$ the group of connected components of $\mathcal{A}_p$, the special fibre at $p$. Then $\Phi_p(A)$ is a finite abelian group with a $G_{\mathbb{F}_p}$-action, and the rule $A \mapsto \Phi_p(A)$ is functorial. Given $\theta : A \to B$ a morphism of abelian varieties over $\mathbb{Q}$, we write $\theta_{p,*} : \Phi_p(A) \to \Phi_p(B)$ for the induced map. In the particular case of multiplication by an integer $n$, that is $[n] : A \to A$, we have that $[n]_{p,*}$ is multiplication by $n$ on $\Phi_p(A)$.

Let $X_p(A)$ be the character group $Hom_{\mathbb{F}_p}(\tau_p(A), \mathbb{G}_m)$ where $\tau_p(A)$ is the toric part of $\mathcal{A}_p$. One has the monodromy pairing $\nu_{A,p} : X_p(A) \times X_p(A^\vee) \to \mathbb{Z}$ which, in the case when $A$ comes with an isomorphism $A \simeq A^\vee$ (e.g. when $A$ is the Jacobian of a curve, in particular when $A$ is an elliptic curve) becomes a $\mathbb{Z}$-valued bilinear pairing on $X_p(A)$.

Suppose now that $A$ is an elliptic curve over $\mathbb{Q}$ with minimal discriminant $\Delta_A$ and with multiplicative reduction at $p$. Then $\Phi_p(A)$ is a cyclic group of order $c_p(A) := \nu_p(\Delta_A)$. 

24
Lemma 6.6. Let $A$ and $B$ be elliptic curves which are isogenous over $\mathbb{Q}$ and suppose that $p$ is a prime of multiplicative reduction for one (hence both) of them. Then $c_p(A)/c_p(B)$ is a rational number whose multiplicative height is at most 163.

Proof. Let $\alpha : A \to B$ be an isogeny of minimal degree; be results of Mazur [60] and Kenku [51] we know that $n := \deg(\alpha) \leq 163$. Let $\beta : B \to A$ be the dual isogeny, so that $\beta \alpha = [n]$ on $A$. Since $A$ and $B$ have multiplicative reduction at $p$, we have isomorphisms of abstract groups $\Phi_p(A) = \mathbb{Z}/c_p(A)\mathbb{Z}$ and $\Phi_p(B) = \mathbb{Z}/c_p(B)\mathbb{Z}$, under which we obtain a commutative diagram

\[
\begin{array}{ccc}
\mathbb{Z}/c_p(A)\mathbb{Z} & \xrightarrow{\alpha_p \ast} & \mathbb{Z}/c_p(B)\mathbb{Z} \\
& \searrow \downarrow{p} \beta_p \ast & \\
& \mathbb{Z}/c_p(A)\mathbb{Z} & \\
\end{array}
\]

From this we get that $c_p(A)/(n, c_p(A)) = \# \text{im}(n \cdot)$ divides $\# \text{im}(\beta_p \ast)$, which divides $c_p(B)$. Thus, the numerator of $c_p(A)/c_p(B)$ divides $n$, and similarly for the denominator using $\alpha \beta$ instead. \hfill \square

6.5. Fixed curve. Let $N = DM$ be an admissible factorization of $N = N_E$ and let $\chi_{D,M}$ be the system of Hecke eigenvalues on $\mathbb{T}_{D,M}$ attached to $E$. Assume that the prime $p$ exactly divides $N$, then $A_{D,M}$ has multiplicative reduction at $p$. Let $x_{D,M} \in X_p(J_0^D(M))$ be a generator for the $\chi_{D,M}$-isotypical component of $X_p(J_0^D(M))$ under the Hecke action, and define the integer

$$h_p(J_0^D(M), \chi_{D,M}) := u_{J_0^D(M),p}(x_{D,M}, x_{D,M})$$

with $u_{J_0^D(M),p}$ the monodromy pairing. We also define the integers

$$i_p(J_0^D(M), \chi_{D,M}) = \# \text{im}(\xi_{p, \ast} : \Phi_p(J_0^D(M)) \to \Phi_p(A_{D,M}))$$

$$j_p(J_0^D(M), \chi_{D,M}) = \# \text{coker}(\xi_{p, \ast} : \Phi_p(J_0^D(M)) \to \Phi_p(A_{D,M}))$$

and observe that $c_p(A) = i_p(J_0^D(M), f) \cdot j_p(J_0^D(M), f)$. The following two results are proved in [81] (see also [71]) under the additional assumption that $M$ is squarefree, but the proofs only require that $p$ exactly divides $N$, i.e. that $E$ (hence $A_{D,M}$) has multiplicative reduction at $p$.

Lemma 6.7 (cf. Theorem 2.3 [81]). We have

$$\delta_{D,M} = \frac{h_p(J_0^D(M), \chi_{D,M}) \cdot j_p(J_0^D(M), \chi_{D,M})}{i_p(J_0^D(M), \chi_{D,M})}.$$ 

Lemma 6.8 (cf. Theorem 2.4 [81]). If $p | D$, then $j_p(J_0^D(M), \chi_{D,M}) = 1$, i.e. $\xi_{p, \ast}$ is surjective.

In [81], the case $p | M$ is only fully addressed under the requirement that the Galois module $E[\ell]$ be irreducible for all primes $\ell$, in which case one has $i_p(J_0^D(M), \chi_{D,M}) = 1$, i.e. $\xi_{p, \ast}$ has trivial image. However, if some of the Galois modules $E[\ell]$ are reducible (which is the case for Frey elliptic curves, for instance), then no information about $v_{\ell}(i_p(J_0^D(M), \chi_{D,M}))$ is given in [81]; the situation is similar in [71]. For our purposes, the following lemma will suffice.
Lemma 6.9. Let $S$ be a finite set of primes. If $N = DM$ is squarefree away from $S$ and $p$ exactly divides $M$, then for every prime $\ell$ we have

$$v_\ell(i_p(J_0^D(M), \chi_{D,M})) \leq \beta_S(\ell) - 1$$

with $\beta_S(\ell)$ as in Lemma 6.5. In particular, if $N = DM$ is squarefree away from $S$ and $p$ exactly divides $M$, then $i_p(J_0^D(M), \chi_{D,M})$ divides an integer $\kappa_S$ which only depends on the set $S$ and, moreover, $\kappa_S$ is supported on the primes $\leq 163$.

Proof. We follow the idea of the proof of Proposition 3 in [71]. The group $\Phi_p(J_0^D(M))$ is Eisenstein in the sense that for $r \nmid N$, the Hecke operator $T_r$ acts on it as multiplication by $r + 1$ (cf. [70]). On the other hand, since $A_{D,M}$ is the optimal quotient associated to $\chi_{D,M}$, the action of $T_r$ on $J_0^D(M)$ induces multiplication by $\chi_{D,M}(T_r) = a_r(A_{D,M})$ on $A_{D,M}$. Hence, $r + 1 - a_r(A_{D,M})$ acts as 0 on $im(\xi_{p,*})$, which is a cyclic group of order $i_p(J_0^D(M), \chi_{D,M})$ because it is a subgroup of the cyclic group $\Phi_p(A_{D,M})$.

It follows that $i_p(J_0^D(M), \chi_{D,M})$ divides $r + 1 - a_r(A_{D,M})$ for every prime $r \nmid N$. Let $\ell$ be a prime, then, under the assumption that $N$ is squarefree away from $S$, Lemma 6.5 affords infinitely many primes $r$ for which $v_\ell(r + 1 - a_r(A_{D,M})) < \beta_S(\ell)$. Taking any of these primes $r$ gives the result. \hfill \Box

6.6. Varying the factorization of $N$. Before proving Theorem 6.1 we need one further lemma regarding the quantity $h_p(J_0^D(M), \chi_{D,M})$.

Let $d$ be squarefree with an even number of prime factors, let $p$ and $q$ be distinct primes not dividing $d$, and let $m$ be a positive integer coprime to $dpq$. Suppose that the conductor of $E$ is $N = dpqm$. Consider the systems of Hecke eigenvalues $\chi = \chi_{dpqm}$ on $\mathbb{T}_{dpqm}$ and $\chi' = \chi_{dpq,m}$ on $\mathbb{T}_{dpq,m}$, both attached to $E$. Let $J = J_0^D(pqm)$ and $J' = J_0^{dpq}(m)$ and consider the optimal quotients $J \to A$, $J' \to A'$ associated to $\chi$ and $\chi'$ respectively; in particular $A$ and $A'$ are elliptic curves over $\mathbb{Q}$ with multiplicative reduction at $p$ and $q$. The following result appears in [81] under the implicit assumption that $m$ is squarefree, but this condition is not necessary because Proposition 3.1 [81] does not need it.

Lemma 6.10 (cf. Theorem 3.2 (a)). We have $h_q(J, \chi) = h_p(J', \chi')$.

Proof of Theorem 6.7. If $D = 1$ there is nothing to prove, so we assume that $D > 1$.

First, let $p, q$ be different prime divisors of $D$, let $d$ be a divisor of $D/(pq)$ with an even number of prime factors (possibly $d = 1$), and let $m = N/(dpq)$ so that $M|m$ and $N = dpqm$. Then $m$ is squarefree away from $S$. By Lemma 6.7 and Lemma 6.10 we can write

$$\frac{\delta_{d,pqm}}{\delta_{dpq,m}} = \frac{h_p(J, \chi) \cdot j_q(J, \chi)/i_q(J, \chi)}{h_p(J', \chi') \cdot j_q(J', \chi')/i_q(J', \chi')} = \frac{i_p(J', \chi') \cdot j_q(J', \chi')}{i_p(J', \chi') \cdot i_q(J', \chi')}$$

where $J = J_0^D(pqm)$, $J' = J_0^{dpq}(m)$ and with $\chi = \chi_{dpqm}$, $\chi' = \chi_{dpq,m}$ the Hecke eigenvalues associated to $E$. Using Lemma 6.5 we find

$$\frac{\delta_{d,pqm}}{\delta_{dpq,m}} = \frac{c_p(A_{dpq,m}) \cdot j_q(J, \chi)}{i_q(J, \chi)} = \frac{c_p(A_{dpq,m}) \cdot c_q(A_{dpqm})}{i_q(J, \chi)^2}.$$

By Lemma 6.6 there is a rational number $\gamma_{p,q}$ of multiplicative height $\leq 163^2$ supported on primes $\leq 163$ with

$$\frac{\delta_{d,pqm}}{\delta_{dpq,m}} = c_p(E) \cdot c_q(E) \cdot \frac{\gamma_{p,q}}{i_q(J, \chi)^2}.$$
By Lemma [6.9] we see that \( i_q(J, \chi) \) divides \( \kappa S \), where \( \kappa S \) is an integer that only depends on \( S \) and is supported on primes \( \leq 163 \). Thus, the factor \( \gamma_p/q(J, \chi)^2 \) is a rational number supported on primes \( \leq 163 \), whose numerator is \( \leq 163^2 \) and whose denominator is \( \leq (163\kappa S)^2 \).

Finally, write \( D = p_1q_1 \cdots p_qq_\nu \) where \( p_s, q_s \) are the distinct prime divisors of \( D \), with \( 2\nu = \omega(D) \). Then by repeated applications of the previous formula to the choices \( (d, p, q) = (1, p_1, q_1), (p_1q_1, p_2, q_2), \ldots \) (there are \( \nu = \omega(D)/2 \) of them) we get

\[
\delta_{1, N}/\delta_{D, M} = \gamma \cdot \prod_{p | D} c_p(E)
\]

where \( \gamma \) has the required form. The last bound stated in the theorem follows from the elementary estimate \( \omega(n) \ll (\log n)/\log \log n \). □

7. Bounding the modular degree

Our goal in this section is to give bounds for the modular degrees \( \delta_{D, M} \) associated to an elliptic curve \( E \) of conductor \( N \), with \( N = DM \) an admissible factorization (both unconditionally and under the Generalized Riemann Hypothesis) and to use them in the context of Szpiro’s conjecture. This last point needs some attention; the classical modular approach to Szpiro’s conjecture translates bound for \( \delta_{1, N} \) into bounds for the Faltings height of an elliptic curve, but the analogous transition is not available in the literature for the more general case of \( \delta_{D, M} \) due to the lack of a Fourier expansion for the modular forms in \( S^D_0(M) \) when \( D > 1 \).

7.1. Counting systems of Hecke eigenvalues. Let \( s(n) = \dim S_2(n)^{\text{new}} \) and let \( r_{D, M} \) be the number of systems of Hecke eigenvalues on \( T_{D, M} \). Here, \( N \) is a positive integer and \( N = DM \) is an admissible factorization. By multiplicity one in \( S_2(n)^{\text{new}} \) and by the Jacquet-Langlands correspondence we have

\[
r_{D, M} = \sum_{m | M} s(Dm).
\]

On the other hand, from Theorem 1 in [59] (see also Appendix B in [39]) together with Lemma 17 in [59], we have the following bound for \( s(n) \):

\[
s(n) \leq \frac{\varphi(n)}{12} + \frac{7}{12} \cdot 2^{\omega(n)} + \mu(n).
\]

The functions \( \varphi, 2^\omega, \mu \) are multiplicative, \( (D, M) = 1 \) and \( D \) is squarefree, so we find

\[
r_{D, M} \leq \frac{\varphi(D)}{12} \sum_{m | M} \varphi(m) + \frac{7}{12} \cdot 2^{\omega(D)} \sum_{m | M} 2^\omega(m) + \mu(D) \sum_{m | M} \mu(m)
\]

\[
\leq \frac{1}{12} \cdot \varphi(D)M + \frac{7}{12} \cdot d(DM^2) + 1.
\]

Writing \( N = DM \), we deduce:

Proposition 7.1. We have

\[
r_{D, M} \leq \frac{1}{12} \cdot \varphi(D)M + \frac{7}{12} \cdot d(DM^2) + 1.
\]
Thus, given \( \epsilon > 0 \), for \( N \gg_{\epsilon} 1 \) with an effective implicit constant, we have

\[
r_{D,M} < \left( \frac{1}{12} + \epsilon \right) \varphi(D)M.
\]

(The asymptotic bound follows by recalling that \( \varphi(n) \gg n/\log \log n \).) We remark that similar methods give \( r_{D,M} \gg \varphi(D)M \).

### 7.2. Unconditional bound.

**Theorem 7.2.** Given any elliptic curve \( E \) over \( \mathbb{Q} \) with conductor \( N \) and an admissible factorization \( N = DM \), we have

\[
\log \delta_{D,M}(E) \leq \left( \frac{1}{12} \cdot \varphi(D)M + \frac{7}{12} \cdot d(DM^2) \right) \left( \log N + \frac{4 \log N}{\log \log N} \right).
\]

Furthermore, given any \( \epsilon > 0 \), for \( N \gg_{\epsilon} 1 \) with an effective implicit constant, we have

\[
\log \delta_{D,M} < \left( \frac{1}{24} + \epsilon \right) \varphi(D)M \log N.
\]

**Proof.** First we bound \( \eta_{[\chi_{D,M}]}(c) \) for \( c \neq [\chi_{D,M}] \) a class of systems of Hecke eigenvalues on \( \mathbb{T}_{D,M} \). Let \( f \in S_2(N) \) be the normalized newform corresponding to \( \chi_{D,M} \) by Jacquet-Langlands. Take any \( \chi \in c \), let \( d|M \) be its level and let \( g \in S_2(Dd)_{\text{new}} \subseteq S_2(N) \) be the only normalized newform which corresponds to \( \chi \) by Jacquet-Langlands.

We have \( f \neq g \). Furthermore, the modular forms \( F, G \) obtained by deleting the Fourier coefficients of \( f, g \) (respectively) of index not coprime to \( N \), have level dividing \( N^2 \) (cf. the proof of Theorem 1 in [5]). Hence, the Fourier expansions of \( F \) and \( G \) differ at some index bounded by \( 2(\dim S_2(N^2) - 1) \). It follows that there is some integer \( n_c \) coprime to \( N = DM \) satisfying \( a_{n_c}(f) \neq a_{n_c}(g) \) and

\[
(7.1) \quad n_c \leq \frac{N^2}{6} \prod_{p|N} \left( 1 + \frac{1}{p} \right) \leq \frac{1}{6} \cdot N^2(1 + \log N) < N^3.
\]

Let \( P(x) \in \mathbb{Z}[x] \) be the (monic) minimal polynomial of \( a_{n_c}(g) = \chi(T_{D,M,n_c}) \). Then \( \deg(P) \leq \#c \) and

\[
0 \neq |P(a_{n_c}(f))| \leq (2d(n_c) \cdot n_c^{1/2})^\#c,
\]

where we have used the Hasse-Weil bound on the Fourier coefficients of a normalized eigenform of weight 2. From [61] we have the following explicit bound for the divisor function \( d(n) \), valid for \( n > 2 \):

\[
(7.2) \quad d(n) < \exp \left( 1.5379 \cdot (\log 2) \cdot \frac{\log n}{\log \log n} \right) < 3^{\log n/\log \log n}.
\]

By Proposition [5,2] the previous divisor bound, and the inequalities in (7.1) we obtain

\[
\eta_{[\chi_{D,M}]}(c) \leq (2d(n_c) \cdot n_c^{1/2})^\#c \cdot 3^{\log n/\log \log n} \cdot \frac{\log n}{\log \log n}.
\]

By the previous divisor bound, and the inequalities in (7.1) we obtain

\[
\eta_{[\chi_{D,M}]}(c) \leq (2d(n_c) \cdot n_c^{1/2})^\#c \cdot 3^{\log n/\log \log n} \cdot \frac{\log n}{\log \log n}.
\]

By the previous divisor bound, and the inequalities in (7.1) we obtain

\[
\eta_{[\chi_{D,M}]}(c) \leq (2d(n_c) \cdot n_c^{1/2})^\#c \cdot 3^{\log n/\log \log n} \cdot \frac{\log n}{\log \log n}.
\]
So we get
\[
\log \eta_{[\chi_{D,M}]}(c) < \#c \cdot \left( \log N + \frac{4 \log N}{\log \log N} \right).
\]
Varying \( c \neq [\chi_{D,M}] \) over the classes of systems of Hecke eigenvalues on \( T_{D,M} \), Theorem 5.3 gives
\[
\log \delta_{D,M} \leq \sum_{c \neq [\chi_{D,M}]} \log \eta_{[\chi_{D,M}]}(c)
< (r_{D,M} - 1) \cdot \left( \log N + \frac{4 \log N}{\log \log N} \right).
\]
Here we used the fact that \( r_{D,M} = \sum_{c} \#c \), summing over all classes \( c \) of systems of Hecke eigenvalues on \( T_{D,M} \). By Proposition \( \square \) we obtain the claimed explicit bound.

The proof of the asymptotic bound is similar, but using instead the (effective) estimate
\[
n_{c} \ll_{\epsilon} N^{1+\epsilon}
\]
from Lemma 11 in [62].

7.3. **Under GRH.** The following result follows from Proposition 5.22 in [15] specialized to classical modular forms — the necessary properties for Rankin-Selberg \( L \)-functions in this setting have been established in [55]. See also [34].

**Theorem 7.3.** There is an effective constant \( C \) such that the following holds:

Let \( f, g \) be normalized Hecke newforms of weight 2 and level dividing \( N \). Assume that the Generalized Riemann Hypothesis holds for the Rankin-Selberg \( L \)-functions \( L(s, f \otimes g) \). Then there is a prime number \( p_{f,g} < C \cdot (\log N)^2 \) not dividing \( N \), satisfying \( a_{p_{f,g}}(f) \neq a_{p_{f,g}}(g) \).

Using this, we get

**Theorem 7.4.** Suppose that the Generalized Riemann Hypothesis for Rankin-Selberg \( L \)-functions of modular forms holds. Let \( \epsilon > 0 \). Then for \( N \gg_{\epsilon} 1 \) with an effective implicit constant, we have
\[
\log \delta_{D,M} \leq \left( \frac{1}{12} + \epsilon \right) \varphi(D)M \log \log N.
\]

**Proof.** The proof is similar to that of Theorem 7.2 except that we replace the integer \( n_{c} \) by the prime \( p_{f,g} \), with \( c, f \) and \( g \) as in the cited proof, and \( p_{f,g} \) as in Theorem 7.3. \( \square \)

7.4. **Application to Szpiro’s conjecture: Classical modular parameterizations.** The classical modular approach to Szpiro’s conjecture (cf. Section 3, especially the estimates (3.1) and (3.3)) together with our bounds for \( \delta_{D,M} \) specialized to \( D = 1, M = N \), give:

**Theorem 7.5.** For all elliptic curves \( E \) over \( \mathbb{Q} \) of conductor \( N \) we have
\[
h(E) \leq \frac{1}{24} \left( N + 7d(N^2) \right) \left( \log N + \frac{4 \log N}{\log \log N} \right) + 9
\]
and
\[
\log |\Delta_E| \leq \frac{1}{2} \left( N + 7d(N^2) \right) \left( \log N + \frac{4 \log N}{\log \log N} \right) + 124.
\]
Furthermore, given $\epsilon > 0$, for $N \gg_\epsilon 1$ with an effective implicit constant we have

$$h(E) < \left(\frac{1}{48} + \epsilon\right) N \log N \quad \text{and} \quad \log |\Delta_E| < \left(\frac{1}{4} + \epsilon\right) N \log N.$$ 

Finally, if we assume GRH, for $N \gg_\epsilon 1$ with an effective implicit constant we have

$$h(E) < \left(\frac{1}{24} + \epsilon\right) N \log \log N \quad \text{and} \quad \log |\Delta_E| < \left(\frac{1}{2} + \epsilon\right) N \log \log N.$$ 

7.5. Application to Szpiro’s conjecture: Shimura curve parameterizations. Here is an extension of the modular approach to Szpiro’s conjecture, using Shimura curve parameterizations coming from $X_0^D(M)$ instead of the classical modular parameterization from $X_0(N)$.

**Theorem 7.6** (The Shimura curve approach to abc). Let $\epsilon > 0$. For all elliptic curves $E$ of conductor $N \gg_\epsilon 1$ (with an effective implicit constant), and for any admissible factorization $N = DM$ we have

$$\log |\Delta_E| < (6 + \epsilon) \log \delta_{D,M}(E) \quad \text{and} \quad h(E) < \left(\frac{1}{2} + \epsilon\right) \log \delta_{D,M}(E).$$

**Proof.** The case $D = 1$ is known, so we can assume $D > 1$. The classical modular approach gives upper bounds for $h(E)$ and $\log |\Delta_E|$ in terms of $\log \delta_{1,N}$ (cf. (3.1) and (3.3)). The result now follows from the first (effective) inequality in Theorem 6.1, together with the estimates

$$\log \prod_{p\mid D} v_p(\Delta_E) \leq \log d(\Delta_E)$$

$$< \frac{1.07 \log |\Delta_E|}{\log \log |\Delta_E|} \quad \text{(by the divisor bound (7.2))}$$

$$\leq \frac{1.07 \log |\Delta_E|}{\log \log N} \quad \text{(for the first bound)}$$

$$\leq \frac{1.07 (12h(E) + 16)}{\log \log N} \quad \text{(by (3.1); for the second bound).}$$

Of course, instead of the current asymptotic formulation, Theorem 7.6 can be given an exact formulation more amenable for computations with a completely explicit error term if desired.

Theorem 7.6 together with our bounds for the modular degree (cf. Theorems 7.2 and 7.4) give:

**Theorem 7.7.** For $\epsilon > 0$ and $N \gg_\epsilon 1$ (with an effective implicit constant), for each admissible factorization $N = DM$ we have the following bounds valid for all elliptic curves $E$ over $\mathbb{Q}$ with conductor $N$:

$$h(E) < \begin{cases} (\epsilon + 1/24) \varphi(D)M \log N & \text{with an accessible error term} \\
(\epsilon + 1/48) \varphi(D)M \log N & \text{unconditional} \\
(\epsilon + 1/24) \varphi(D)M \log \log N & \text{under GRH} \end{cases}$$
and similarly
\[
\log |\Delta_E| < \begin{cases} 
(\epsilon + 1/2) \varphi(D) M \log N & \text{with an accessible error term} \\
(\epsilon + 1/4) \varphi(D) M \log N & \text{unconditional} \\
(\epsilon + 1/2) \varphi(D) M \log N & \text{under GRH.}
\end{cases}
\]

The comment “with an accessible error term” refers to the fact that the bound can be obtained with completely explicit lower order terms if desired, thanks to the first estimate in Theorem 7.2.

Let us remark that in the almost semi-stable case one can replace \( \varphi(D) M \) by \( \varphi(N) \) by suitable choice of admissible factorization \( N = DM \). Let us record the result here.

**Corollary 7.8.** Let \( S \) be a finite set of primes and let \( P \) be the product of the elements of \( S \). For \( \epsilon > 0 \) and \( N \gg \epsilon, S \) with an effective implicit constant, if \( E \) is an elliptic curve over \( \mathbb{Q} \) semi-stable away from \( S \), then we have
\[
h(E) < \begin{cases} 
\frac{P}{\varphi(P)} (\epsilon + 1/48) \varphi(N) \log N & \text{unconditional} \\
\frac{P}{\varphi(P)} (\epsilon + 1/24) \varphi(N) \log \log N & \text{under GRH}
\end{cases}
\]
and similarly
\[
\log |\Delta_E| < \begin{cases} 
\frac{P}{\varphi(P)} (\epsilon + 1/4) \varphi(N) \log N & \text{unconditional} \\
\frac{P}{\varphi(P)} (\epsilon + 1/2) \varphi(N) \log \log N & \text{under GRH.}
\end{cases}
\]

Note that if \( S = \emptyset \) (i.e., for semi-stable elliptic curves) the factor \( P/\varphi(P) \) is 1, and that for \( P \gg 1 \) one has \( P/\varphi(P) < (\epsilon \gamma + \epsilon) \log \log N < 2 \log \log P \).

**Proof.** Let \( p_N \) be the largest prime factor of \( N \) away from \( S \). Then for all but finitely many \( E \) we have that \( p_N \) exists. Furthermore, \( p_N \rightarrow \infty \) as \( N \rightarrow \infty \) by Shafarevich’s theorem.

Let \( N \) be sufficiently large, so that \( p_N \) exists. If \( N \) has an even number of prime factors away from \( S \), take \( D \) to be the product of them. Otherwise, take \( D \) as the product of them except \( p_N \). Let \( M = N/D \), then
\[
\varphi(D) M = \varphi(N) \prod_{p \mid M} \left(1 - \frac{1}{p}\right)^{-1} \leq \varphi(N) \prod_{p \mid p_N \cdot P} \left(1 - \frac{1}{p}\right)^{-1} = \varphi(N) \cdot \frac{p_N}{p_N - 1} \cdot \frac{P}{\varphi(P)}.
\]

The result now follows from Theorem 7.7. Note that this argument is effective. \( \square \)

8. Infinite part: Norm comparisons

8.1. The result. This section is purely analytic and there is no additional difficulty in momentarily considering a more general case.

Let \( F \) be a totally real number field of degree \( n \) with real embeddings \( \tau_j \) (1 ≤ \( j \) ≤ \( n \)). Let \( B \) be a quaternion division \( F \)-algebra with exactly one split place at infinity, say \( \tau_1 \). Let \( B^\times = B \otimes \mathbb{Z} \) and let \( B^\times_1 \) be the elements of \( B^\times \) with totally positive reduced norm. For each compact open subgroup \( U \subseteq B^\times \) and for each \( g \in B^\times \) consider the group \( \Gamma_{U,g} = gUg^{-1} \cap B^\times_1 \) and its image \( \tilde{\Gamma}_{U,g} \) in \( PSL_2(\mathbb{R}) \) via \( \tau_1 \). The quotient \( X_{U,g}^an = \tilde{\Gamma}_{U,g} \backslash \mathfrak{h} \) is a compact, connected, complex curve, because \( B \) is a division algebra.
Let $S_{U,g}$ be the space of weight 2 holomorphic modular forms for the action of $\tilde{\Gamma}_{U,g}$ on $\mathfrak{h}$. Since we are assuming that $B$ is a division algebra, the cuspidality condition would be vacuous.

On $S_{U,g}$ we have the $L_2$-norm induced by the Petersson inner product
\[ \|h\|_{U,g,2} := \left( \int_{\tilde{\Gamma}_{U,g}\backslash \mathfrak{h}} |h(z)|^2 \Im(z)^2 d\mu_0(z) \right)^{1/2}. \]

We also have the supremum norm on $S_{U,g}$
\[ \|h\|_{U,g,\infty} := \sup_{z \in \mathfrak{h}} |h(z)|\Im(z). \]

(Observe that the $L_2$-norm is not normalized, so it is not invariant by shrinking $U$.) Our goal in this section is to compare these two norms.

**Theorem 8.1.** Keep the previous notation. There is a number $\nu_n > 0$ depending only on the degree $n = [F : \mathbb{Q}]$, such that if $\tilde{\Gamma}_{U,g}$ acts freely on $\mathfrak{h}$, then for all $h \in S_{U,g}$ we have
\[ \|h\|_{U,g,\infty} \leq \nu_n \cdot \|h\|_{U,g,2}. \]

Note that $\tilde{\Gamma}_{U,g}$ acts freely on $\mathfrak{h}$ if and only if the quotient map $\mathfrak{h} \to X_{U,g}^{an}$ is unramified. This can always be achieved by suitably shrinking $U$; however, for our purposes later, we will need to be precise about this point.

These two norms have been compared in the case of classical modular curves (non-compact case, see for instance [7]) or in the compact case assuming that the weight is large (cf. [22]). As in the non-compact case, one might expect that for every $\epsilon > 0$, the quantity $\nu_n$ should be replaced by a factor $\ll_n, \epsilon \mathrm{Vol}(X_{U,g})^{-1/2+\epsilon}$, where the volume is taken with respect to $d\mu_0$. Improvements in this direction are available in the non-compact case (see for instance [7]), but the techniques do not work in the absence of Fourier expansions. In any case, Theorem 8.1 suffices for our purposes.

8.2. **Injectivity radius.** Let $U$ be an open compact subgroup of $\mathbb{B}^\times$, let $g \in \mathbb{B}^\times$, and write $\Gamma := \Gamma_{U,g} = gUg^{-1} \cap B_+^\times$.

For $\gamma \in B_+^\times$ we write $\gamma^*$ and $\tilde{\gamma}$ for its image in $SL_2(\mathbb{R})$ and in $PSL_2(\mathbb{R})$ respectively (via $\tau_1$).

The systole of $\Gamma$ is defined as
\[ \sigma_\Gamma := \min \{ d_\mathfrak{h}(x, \tilde{\gamma} \cdot x) : x \in \mathfrak{h}, \gamma \in \Gamma, x \neq \tilde{\gamma} \cdot x \} \]
where $d_\mathfrak{h}$ is the hyperbolic distance in $\mathfrak{h}$. Note that $\sigma_\Gamma$ is twice the injectivity radius $\rho_\Gamma$ of $X_{U,g}^{an}$.

Recall that if $u \in SL_2(\mathbb{R})$ is hyperbolic (i.e. $|\text{tr}(u)| > 2$), then for every $x \in \mathfrak{h}$ we have
\[ d_\mathfrak{h}(x, \gamma^* x) = 2 \log |\lambda_u| \]
where $\lambda_u$ is the eigenvalue of $u$ with largest absolute value. We say that $\gamma \in B_+^\times$ is hyperbolic if $\gamma^*$ is, and we write $\lambda_{\gamma} = \lambda_{\gamma^*}$.

**Lemma 8.2.** Let $\gamma \in \Gamma$ be hyperbolic. Define $\beta := \text{rn}(\gamma)^{-1} \gamma^2$, where, as before, $\text{rn}$ denotes the reduced norm. Then we have the following:

1. $\beta \in B_+^\times$, $\text{rn}(\beta) = 1$ and it is in some maximal order of $B$. 

(ii) $\beta$ is hyperbolic and $d_h(x, \beta x) = 2d_h(x, \gamma x)$ for every $x \in h$

(iii) Let $K = F(\lambda_\beta)$. Then $[K : F] = 2$.

(iv) $\lambda_\beta \in O_K^\times$.

(v) $1/\lambda_\beta$ is a conjugate of $\lambda_\gamma$ over $F$.

(vi) all other conjugates of $\lambda_\beta$ over $\mathbb{Q}$ (if any) have modulus 1.

Proof. Since $\gamma \in \Gamma \subseteq B^\times$, we have the first two assertions in (i) Also, since $\gamma \in gUg^{-1} \cap B^\times$ we see that $\text{rn}(\gamma) \in O_F^\times$, so $\text{rn}(\gamma)^{-1}\gamma^2$ is in some maximal order of $B$, proving (i). Item (ii) follows because $\tilde{\beta} = \tilde{\gamma}^2$.

In view of (i), the properties (iii)-(vi) for $\beta$ are known, see for instance Section 12.3 of [57]. □

In particular, $\lambda_\beta$ as in the previous lemma is a Salem number of degree $2n$, where $n = [F : \mathbb{Q}]$. Using the available partial progress on Lehmer’s conjecture for the Mahler measure, one gets

Lemma 8.3. If $\tilde{\Gamma}_{U,g}$ acts freely on $h$, then

$$\rho_\Gamma \geq \frac{1}{2(\log(6n))^3}.$$ 

Proof. Since $B$ is a division algebra, $\tilde{\Gamma}_{U,g}$ contains no non-trivial parabolic elements, and since it acts freely on $h$, it contains no elliptic elements either. Thus, every non-hyperbolic $\gamma \in \Gamma$ satisfies $\tilde{\gamma} = I$. It follows that there is $\gamma \in \Gamma$ hyperbolic with

$$2\rho_\Gamma = d_h(i, \tilde{\gamma} \cdot i) = \frac{1}{2}d_h(i, \tilde{\beta} \cdot i) = \log |\lambda_\beta| \quad (i = \sqrt{-1} \in h)$$

with $\beta$ as in the previous lemma. The bound now follows from Corollary 2 in [87]. (We remark that even weaker bounds would suffice for our purposes, but we choose to use Voutier’s result for the sake of concreteness.) □

8.3. Proof of the norm comparison. Recall that on $h$ (with complex variable $z = x + iy$) we consider the volume form

$$d\mu_h(z) = \frac{dx \wedge dy}{y^2}.$$ 

On the open unit disc $\mathbb{D}$ (with complex variable $w = u + vi$), let us consider

$$d\mu_\mathbb{D}(w) = \frac{4du \wedge dv}{(1 - (u^2 + v^2))^2}.$$ 

These volume forms come from the usual hyperbolic metrics $d_h$ and $d_\mathbb{D}$ on $h$ and $\mathbb{D}$ respectively, with constant curvature $-1$. For $z \in h$, $w \in \mathbb{D}$ and $r > 0$, we let $B_h(z, r)$ and $B_\mathbb{D}(w, r)$ be the corresponding balls of hyperbolic radius $r$ centered at $z$ and $w$ respectively.

For any $\tau \in h$, the biholomorphic map

$$c_\tau : \mathbb{D} \to h, \quad c_\tau(w) = \Re(\tau) - \Im(\tau)i \cdot \frac{w + i}{w - i}$$

is an isometry for $d_\mathbb{D}$ and $d_h$. Furthermore, it satisfies $c_\tau(0) = \tau$ and $c_\tau^*d\mu_h = d\mu_\mathbb{D}$.
Lemma 8.4. Let \( t > 0 \) and let \( B(0, t) \) be the Euclidean ball in \( \mathbb{C} \) centered at 0 with radius \( t \). Let \( h \) be holomorphic on a neighborhood of \( B(0, t) \). Then
\[
\pi t^2 |h(0)|^2 \leq \int_{B(0, t)} |h(z)|^2 dx \wedge dy.
\]

Proof. This is immediate from expanding \( h \) as a power series and integrating \( h \cdot \bar{h} \). \qed

Lemma 8.5. Let \( f : \mathfrak{h} \to \mathbb{C} \) be holomorphic, let \( \tau \in \mathfrak{h} \) and let \( r > 0 \). Then
\[
|f(\tau)|^2 \Im(\tau)^2 \leq \frac{e^{2r}}{4\pi(tanh(r/2))^2} \int_{B(\tau, r)} |f(z)|^2 y^2 d\mu_h(z).
\]

Proof. Let \( f_\tau = c_\tau^* f \). It is a holomorphic function on \( \mathbb{D} \) and we have
\[
\int_{B_h(\tau, r)} |f(z)|^2 y^2 d\mu_h(z) = \int_{B_0(0, r)} |f_\tau(w)|^2 \Im(c_\tau(w))^2 d\mu_\mathbb{D}(w)
\]
\[
= \int_{B_0(0, r)} |f_\tau(w)|^2 \Im(c_\tau(w))^2 \frac{4du \wedge dv}{(1 - (u^2 + v^2))^2}
\]
\[
\geq 4 \int_{B_0(0, r)} |f_\tau(w)|^2 \Im(c_\tau(w))^2 du \wedge dv.
\]
Note that
\[
\inf \{ \Im(c_\tau(w)) : w \in B_\mathbb{D}(0, r) \} = \inf \{ \Im(z) : z \in B_h(\tau, r) \} = e^{-r} \Im(\tau)
\]
and that if \( B(0, t) \) denotes the Euclidean ball of radius \( t \) centered at 0, then we have \( B(0, t) = B_{\mathbb{D}}(0, r) \) for \( t = \tanh(r/2) \). So, from the previous lemma we get
\[
\int_{B_h(\tau, r)} |f(z)|^2 y^2 d\mu_h(z) \geq 4e^{-2r} \Im(\tau)^2 \int_{B(0, \tanh(r/2))} |f_\tau(w)|^2 du \wedge dv
\]
\[
\geq 4\pi e^{-2r}(\tanh(r/2))^2 \Im(\tau)^2 |f(\tau)|^2.
\]
\qed

Proof of Theorem \[8.7\]. Choose \( \tau_0 \in \mathfrak{h} \) such that \( |h(\tau_0)|\Im(\tau_0) = \|h\|_{U,g,\infty} \) and apply Lemma \[8.5\] with \( \tau = \tau_0 \) and \( 2r = 1/(\log(6n))^3 \). This choice of \( r \) together with Lemma \[8.3\] ensure
\[
\int_{B_h(\tau, r)} |h(z)|^2 y^2 d\mu_h \leq \|h\|^2_{U,g,2}.
\]
\qed

9. Finite part: Differentials on integral models

9.1. Relative differentials on Shimura curves. We now return to the case \( F = \mathbb{Q} \).

Given a compact open subgroup \( U \subseteq \mathbb{B}^\times \) contained in \( O_{\mathbb{B}}^\times \), we write \( m_U \) for its level and \( X_0^U(M, U) \) for the Shimura curve associated to the compact open subgroup \( U \cap U_0^D(M) \). Here, \( N = DM \) is and admissible factorization, \( D \) is the discriminant of \( B \), and we always assume \( m_U \) coprime to \( D \).
Let $\mathcal{X}^D_0(M, U)$ be the standard integral model for $X^D_0(M, U)$ over $\mathbb{Z}[m_U^{-1}]$, constructed as coarse moduli scheme for the moduli problem of abelian surfaces with quaternionic multiplication (i.e. fake elliptic curves) and $U^0_0(M) \cap U$-structure. (This direct moduli approach to integral models is available as we are working over $\mathbb{Q}$.)

Let $p \nmid m_U$ be a prime. For $p\not| D$ the reduction of $\mathcal{X}^D_0(M, U)$ is of Cerednik-Drinfeld type, for $p\not| M$ it is of Deligne-Rapoport type, and for $p \nmid N m_U$ the integral model has good reduction.

Let $\mathcal{X}^D_0(M, U)^0$ be the smooth locus of $\mathcal{X}^D_0(M, U) \to \text{Spec} \mathbb{Z}[m_U^{-1}]$. It is obtained from $\mathcal{X}^D_0(M, U)$ by removing the supersingular points in the special fibres of characteristic dividing $D$, and removing supersingular points and non-reduced components of the special fibres of characteristic dividing $M$ (non-reduced components in characteristic $p$ only occur when $p^2 \not| M$).

There is a natural forgetful $\mathbb{Z}[m_U^{-1}]$-map

$$\pi^D_{M,U} : \mathcal{X}^D_0(M, U) \to \mathcal{X}^D_0(1, U).$$

Given $N = DM$, we say that $U$ is good enough for $(D, M)$ if the following conditions are satisfied:

1. $m_U$ is coprime to $N$;
2. $\text{rn}(U) = \hat{\mathbb{Z}}$;
3. $U \subseteq U^D_1(m)$ for some $m \geq 4$ with $(D, m) = 1$.

See [35] for details on these conditions. Thus, when $U$ is good enough for $(D, M)$, we have that both $\mathcal{X}^D_0(1, U)/\mathbb{Z}[m_U^{-1}]$ and $\mathcal{X}^D_0(M, U)/\mathbb{Z}[m_U^{-1}]$ are fine moduli spaces for the associated moduli problems (by (i) and (iii)), and they have geometrically connected generic fibre (by (i) and (ii)).

**Theorem 9.1.** Suppose that $U$ is good enough for $(D, M)$. The integer $M$ annihilates the sheaf of relative differentials $\Omega^1_{\mathcal{X}^D_0(M, U)}$ on $\mathcal{X}^D_0(M, U)^0$.

Assuming this result for the moment, we obtain:

**Corollary 9.2.** Suppose that $U$ is good enough for $(D, M)$. On $\mathcal{X}^D_0(M, U)^0$, the canonical morphism of sheaves

$$\left(\pi^D_{M,U}\right)^* \Omega^1_{\mathcal{X}^D_0(1, U)/\mathbb{Z}[m_U^{-1}]} \to \Omega^1_{\mathcal{X}^D_0(M, U)/\mathbb{Z}[m_U^{-1}]}$$

is injective and its cokernel is annihilated by $M$.

**Proof.** The map is injective because it is non-zero, and on the smooth locus the two sheaves are invertible.

The assertion about the cokernel follows from the fundamental exact sequence of relative differentials and the previous theorem. \hfill $\square$

### 9.2. The fibres at primes dividing $M$

Suppose that $U \subseteq O^\times_B$ is good enough for $(D, M)$.

Let $p$ be a prime dividing $M$, and let $n, m$ be positive integers defined by $p \nmid m$ and $M = p^n m$. The geometric fibre of $\mathcal{X}^D_0(M, U)$ at $p$ is described as follows (cf. Chapter 13 in [35], suitably adapted to moduli of fake elliptic curves; see [35] for this adaptation):

Put $k = \mathbb{F}_p$. The fibre $\mathcal{X}^D_0(M, U) \otimes k$ is formed by $n + 1$ copies of $\mathcal{X}^D_0(m, U) \otimes k$ with suitable multiplicities, crossing at supersingular points. More precisely, for any pair of integers $a, b \geq 0$ with $a + b = n$, there is a geometrically integral closed sub-scheme $F_{a,b}$
of $\mathcal{X}_0^D(M,U) \otimes k$, isomorphic to the curve $\mathcal{X}_0^D(m,U) \otimes k$ and occurring with multiplicity $\varphi(p^{\min\{a,b\}})$ in $\mathcal{X}_0^D(M,U) \otimes k$. Then the irreducible components of $\mathcal{X}_0^D(M,U) \otimes k$ are precisely the curves $F_{a,b}$, with multiplicity $\varphi(p^{\min\{a,b\}})$, crossing at the super-singular points of $\mathcal{X}_0^D(M,U) \otimes k$. Note that $F_{n,0}$ and $F_{0,n}$ occur with multiplicity 1.

The map $\mathcal{X}_0^D(M,U) \to \mathcal{X}_0^D(m,U)$ induces maps $f_{a,b} : F_{a,b} \to \mathcal{X}_0^D(m,U) \otimes k$. The morphism $f_{n,0}$ is an isomorphism, and $f_{0,n}$ has degree $p^n$.

For the sake of exposition, let us recall that in the simplest case $U = \mathbb{F}_q$, $D = 1$, $m = 1$, the above mentioned facts correspond to Kronecker’s congruence for the modular polynomials $\Phi_{p^n}$, namely

$$\Phi_{p^n}(X,Y) \equiv \prod_{a,b \geq 0, a+b=n, c=\min\{a,b\}} (X^{p^a-c} - Y^{p^b-c})^{\varphi(p^n)} \mod p.$$  

9.3. Annihilation of relative differentials.

**Lemma 9.3.** If $U$ is good enough for $(D,M)$, then the sheaf $\Omega^1_{\pi^D_{M,U}}$ is supported on the special fibres at primes dividing $M$, i.e.,

$$\Omega^1_{\pi^D_{M,U}} |_{\mathcal{X}^D_0(M,U)[M^{-1}]} = 0.$$  

**Proof.** Note that $\mathcal{X}_0^D(M,U) \otimes \mathbb{Z}[M^{-1}] = \mathcal{X}_0^D(1,U')$ with $U' = U_0(M) \cap U$. Since $U$ is good enough for $(D,M)$, we get that $U'$ is good enough for $(D,1)$. Since $U$ and $U'$ are contained in $U_0^D(\ell)$ with $\ell \geq 5$, the morphism $\mathcal{X}_0^D(1,U') \to \mathcal{X}_0^D(1,U)[M^{-1}]$ induced by the inclusion $U' \subseteq U$ is étale, hence the result. \qed

**Proof of Theorem [77].** Write $M = p^n m$ with $p$ a prime, $p \nmid m$ and $n \geq 1$. Consider the forgetful map $\pi : \mathcal{X}_0^D(M,U) \to \mathcal{X}_0^D(m,U)$ and the factorization $\pi^D_{M,U} = \pi^D_{m,U} \pi$. We have the exact sequence

$$\pi^* \Omega^1_{\pi^D_{M,U}} \to \Omega^1_{\pi^D_{m,U}} \to \Omega^1_{\pi} \to 0.$$  

Note that upon inverting $p$, the map $\pi$ becomes the forgetful $\mathbb{Z}[(pmU)^{-1}]$-morphism

$$\mathcal{X}_0^D(m,U') \to \mathcal{X}_0^D(m,U) \otimes \mathbb{Z}[p^{-1}]$$  

with $U' = U \cap U_0^D(p^n)$, which is étale because $U$ is good enough. So, $\Omega^1_{\pi}[p^{-1}] = 0$. Furthermore, it also follows that away from characteristic $p$ we have $\pi^{-1} \mathcal{X}_0^D(m,U)[1/p] = \mathcal{X}_0^D(M,U)[1/p]$. Thus, if one knows that $m$ annihilates $\Omega^1_{\pi^D_{m,U}} |_{\mathcal{X}^D_0(m,U)[p]}$, then the previous exact sequence would show that $m$ also annihilates $\Omega^1_{\pi^D_{M,U}} |_{\mathcal{X}^D_0(M,U)[p]}$ away from characteristic $p$. To complete the argument, it would suffice that $p^n$ annihilates $\Omega^1_{\pi^D_{m,U}} |_{\mathcal{X}^D_0(p^n U)[p]}$ after inverting $m$, that is, $\Omega^1_{\pi^D_{p^n U}} |_{\mathcal{X}^D_0(p^nU)[p]}$ where $U'' = U \cap U_0^D(m)$.

From the previous analysis, we see —by induction on the number of distinct prime factors of $M$— that it suffices to prove the result in the particular case $M = p^n$ for $p$ a prime and $n \geq 1$. So let’s assume that this is the case.

By Lemma [93] we only need to show that $p^n$ annihilates $\Omega^1_{\pi^D_{p^n U}} |_{\mathcal{X}^D_0(p^n U)[p]}$ étale-locally on the fibre at $p$ of $\mathcal{X}_0^D(p^n U)$, that is, for points on the ordinary locus of the components $F_{n,0}$ and $F_{0,n}$ of the special fibre at $p$.  

36
Write $k = \mathbb{F}_p$, let $K$ be an algebraic closure of $k$ and let $W$ be the ring of Witt vectors of $K$. Let $y_0 \in \mathcal{X}_0^D(p^n, U)^0$ be a closed point with residue characteristic $p$ and let $x_0 \in \mathcal{X}_0^D(1, U)^0$ be its image under $\pi_{p^n, U}$. Let $y$ be a $K$-valued point in $\mathcal{X}_0^D(p^n, U)^0 \otimes W$ lying above $y_0$ and let $x$ be its image in $\mathcal{X}_0^D(1, U)^0 \otimes W$. Let $\mathcal{A} = W[[T]]$ so that we have an isomorphisms of complete local rings involving the completed strict henselization of $\mathcal{O}_{\mathcal{X}_0^D(1, U)^0, x_0}$ (cf. p.133-134 in [18]):

\[(\mathcal{O}_{\mathcal{X}_0^D(1, U)^0, x_0}^{s,h})^\wedge \simeq \hat{\mathcal{O}}_{\mathcal{X}_0^D(1, U)^0 \otimes W, x} \simeq \mathcal{A}.\]

Since $x$ is in the ordinary locus, Serre-Tate theory (cf. section 8.9 in [18], see also [11]) for an adaptation to deformations of fake elliptic curves gives $\mathcal{A}$ the structure of a $\mathbb{Z}[q, q^{-1}]$-algebra by letting $p$ be the Serre-Tate parameter of the pull-back of the universal family over $\mathcal{X}_0^D(1, U)$ (recall that $U$ is good enough). Denote the image of $q$ in $\mathcal{A}$ again by $q$, so that $q \in \mathcal{A}^\wedge$.

By the isomorphism (9.1), we have that $(\mathcal{O}_{\mathcal{X}_0^D(p^n, U)^0, y_0}^{s,h})^\wedge$ is a finite $\mathcal{A}$-algebra under the pull-back map. By Theorem 13.6.6 in [18], this $\mathcal{A}$-algebra structure can be described as follows:

\[(\mathcal{O}_{\mathcal{X}_0^D(p^n, U)^0, y_0}^{s,h})^\wedge \simeq \mathcal{B} := \begin{cases} \mathcal{A} & \text{if } y_0 \in F_{n,0} \cap \mathcal{X}_0^D(p^n, U)^0 \\ \mathcal{A}[Z]/(Zp^n - q) & \text{if } y_0 \in F_{0,n} \cap \mathcal{X}_0^D(p^n, U)^0. \end{cases}\]

Finally, let us check that $p^n$ annihilates the étale stalk $\Omega_{\mathcal{X}_0^D(p^n, U)^0, y_0}^{1, \text{et}}$. It suffices to check this after completion. We have

\[(\Omega_{\mathcal{X}_0^D(p^n, U)^0, y_0}^{1, \text{et}})^\wedge \simeq \Omega_{\mathcal{X}_0^D(p^n, U)^0, y_0}^{1, \text{et}} \otimes (\mathcal{O}_{\mathcal{X}_0^D(p^n, U)^0, y_0}^{s,h})^\wedge \simeq \Omega_{\mathcal{X}_0^D(p^n, U)^0, y_0}^{1, \text{et}} \otimes (\mathcal{O}_{\mathcal{X}_0^D(p^n, U)^0, y_0}^{s,h})^\wedge \simeq \Omega_{\mathcal{A}^\wedge/\mathcal{A}}^{1}.

If the ordinary point $y_0$ belongs to $F_{n,0}$, then the previous module is $\Omega_{\mathcal{A}^\wedge/\mathcal{A}}^{1} = (0)$. On the other hand, if $y_0 \in F_{0,n}$ then using the fact that $\frac{d}{dz}(Zp^n - q) = p^n Zp^{n-1}$ we find

\[\Omega_{\mathcal{A}^\wedge/\mathcal{A}}^{1} \simeq \frac{\mathcal{A}[Z]}{(Zp^n - q, p^n Zp^{n-1})}.
\]

Since $p^nq = Z \cdot p^n Zp^{n-1} - p^n \cdot (Zp^n - q)$ and $q \in \mathcal{A}^\wedge$, we see that the previous module is a quotient of $\mathcal{A}[Z]/p^n \mathcal{A}[Z]$, hence, it is annihilated by $p^n$. \hfill \Box

10. Bounds for the Manin constant

10.1. The Manin constant. Given an elliptic curve $A$ over $\mathbb{Q}$ with conductor $N$, which is an optimal quotient $q : J_0(N) \to A$ with associated normalized newform $f \in S_2(N)$, we write $c_f$ for its Manin constant (cf. Section 3). Thus, letting $\omega_A$ be a global Néron differential for $A$, the pull-back of $\omega_A$ under $\mathfrak{h} : X_0(N) \to A$ is $2\pi i c_f f(z)dz$. Here, the map $X_0(N) \to A$ is $\phi = q_j N$. Multiplying $\omega_A$ by $-1$ if necessary, we assume that $c_f$ is positive.

Edixhoven [26] proved that $c_f$ is a non-zero integer. After the work of Mazur [60] and Abbes, Ullmo, and Raynaud [11] we know that if $v_p(N) \leq 1$ then $v_p(c_f) = 0$, except, perhaps, for $p = 2$ in which case the assumption $v_2(N) \leq 1$ only gives $v_2(c_f) \leq 1$. (See [2] and the references therein for more results on the Manin constant.) This last caveat at $p = 2$ has been removed by recent work of Cesnavicius [15], so that now one knows that for every prime $p$ the following implication holds:

$v_p(N) \leq 1 \Rightarrow p \nmid c_f$. 37
Since the conductor of the elliptic curve $A$ is $N$, we know that in the relevant cases
\begin{equation}
 v_p(N) \leq \begin{cases} 
 8 & \text{if } p = 2 \\
 5 & \text{if } p = 3 \\
 2 & \text{if } p \geq 5.
\end{cases}
\end{equation}

It is desirable to have control on $v_p(c_f)$ at all primes, not just when $v_p(N) \leq 1$. However, not much is known about $v_p(c_f)$ in the general case. See [26] for some partial results when $p \geq 7$.

We will prove:

**Theorem 10.1.** Let $S$ be a finite set of primes and let $p$ be a prime number. There is a constant $\mu_{S,p}$ depending only on $S$ and $p$, such that for every optimal elliptic curve $A$ over $\mathbb{Q}$ with semi-stable reduction outside $S$ and with associated newform $f \in S_2(N)$, we have
\[ v_p(c_f) \leq \mu_{S,p}. \]

The following is an immediate consequence of the previous theorem and the known results about the Manin constant at primes with $v_p(N) \leq 1$.

**Corollary 10.2.** Let $S$ be a finite set of primes. There is a constant $M_S$ depending only on $S$ such that for every optimal elliptic curve $A$ defined over $\mathbb{Q}$ with semi-stable reduction away from $S$ and with associated newform $f \in S_2(N)$, we have $c_f \leq M_S$.

The proof of these results is motivated by the existing literature, especially [2, 14, 26]. Our main new contribution is the idea of working with towers of suitable modular curves with additional level structure to get good integral models, and our method to deal with the case when the additional level structure makes the relevant eigenform an old form.

10.2. **Setup for the proof of Theorem 10.1.** Let us fix a set of primes $S$, a prime number $p$, a positive integer $n \geq 2$, and an auxiliary prime number $\ell \geq 5$ different from $p$. We will prove:

**Theorem 10.3.** There is a bound $B = B(S,p,n,\ell)$ such that for any given optimal elliptic curve $A$ with conductor $N = p^n m$ and with associated newform $f \in S_2(N)$, satisfying that $p \nmid m$ and that $m$ squarefree away from $S$, one has $v_p(c_f) \leq B$.

Note that the existence of such an $A$ and our assumption $n \geq 2$ force $p \in S$.

For notational convenience, we will fix an optimal elliptic curve $A$ as in Theorem 10.3 for the rest of this section. So we need to state explicitly the parameters on which each bound depends, and we will do so by adding appropriate subscripts to the asymptotic notation $\ll$, $O(-), \asymp$.

Theorem 10.1 will follow by fixing the choice $\ell = 5$ unless $p = 5$ in which case we take $\ell = 7$, and from the fact that $n \leq 8$ by (10.1). The cases of semi-stable reduction at $p$ (that is, $n = 0$ or 1) follow from the existing literature.

Since $\ell$ has to be chosen uniformly bounded, we are forced to consider the cases $\ell \mid N$ and $\ell \nmid N$ (equivalently, $\ell \mid m$ and $\ell \nmid m$) separately, the second being the more laborious.

Let $R = \mathbb{Z}_{(p)}$. Given an $R$-module $M$, we define
\[ v(M) = \begin{cases} 
 \infty & \text{if no power of } p \text{ annihilates } M_{\text{tor}}; \\
 \min\{k \geq 0 : p^k \cdot M_{\text{tor}} = (0)\} & \text{otherwise.}
\end{cases} \]
Since $R$ is DVR, whenever $M$ is finitely generated we have that $v(M)$ is finite and there is an element $x \in M$ such that $v(M) = v(x)$.

When $N$ is a free $R$-module, we say that $x \in N$ is primitive if $x \neq 0$ and $v(N/\langle x \rangle) = 0$.

More generally, for a $\mathbb{Z}$-module $G$, we write $v(G) := v(G[p^\infty])$. We observe that when $G$ is a finitely generated $\mathbb{Z}$-module, $v(G)$ is the $p$-adic valuation of the exponent of the finite group $G_{\text{tor}}$.

10.3. A projective system of curves. In this section, $m$ will always denote a positive integer coprime to $p$ (possibly divisible by $\ell$), and we will consider the curves $X_m := X_{U(0,p^n)}$ in $U_m(\ell)$. The notation $X_m$ (and related notation to be introduced below) will be used with this meaning only in the present Section [10].

The curves $X_m$ are defined over $\mathbb{Q}$ and are geometrically irreducible. The cusp $i\infty$ defines a $\mathbb{Q}$-rational point in $X_m$ (possibly after conjugation of the open compact group $U_1(\ell)$, depending on conventions; cf. Variant 8.2.2 in [24]). The integral model over $\mathbb{Z}[1/\ell]$ provided by the theory of Deligne-Rapoport [23], Katz-Mazur [48] and Casson [13] can be base-changed to $R$, obtaining an integral model that we denote by $\mathcal{X}_m/R$. Then $\mathcal{X}_m$ is regular (since $\ell \geq 5$) and $\mathcal{X}_m \to \Spec(R)$ is flat and proper. Hence, $\mathcal{X}_m \to \Spec(R)$ is Gorenstein, and Grothendieck’s duality theory (cf. [23]) applies. The relative dualizing sheaf is denoted by $\omega_m$ and it is invertible. Furthermore, for $m|m'$ both coprime to $p$, the forgetful map $\mathcal{X}_m' \to \mathcal{X}_m$ is etale and the pull back of $\omega_m$ is $\omega_{m'}$.

Let $J_m$ be the Jacobian of $X_m$ over $\mathbb{Q}$, and let $\mathcal{J}_m$ be the Néron model over $R$.

Since $\mathcal{X}_m/R$ has sections (e.g. the one induced by the cusp $i\infty$) and has some fibre components with multiplicity 1 (from the standard description of the special fibre at $p$) we see from Theorem 1, Sec. 9.7 [8] that $\Pic^0_{\mathcal{X}_m/R}$ is a scheme, the canonical map $\Pic^0_{\mathcal{X}_m/R} \to \mathcal{J}_m$ to the identity component of $\mathcal{J}_m$ is an isomorphism, and there are canonical identifications $H^1(\mathcal{X}_m, \mathcal{O}_{\mathcal{X}_m}) = \Lie(\Pic^0_{\mathcal{X}_m/R}) \simeq \Lie(\mathcal{J}_m)$.

Dualizing we obtain the $R$-isomorphisms

$$H^0(\mathcal{J}_m, \Omega^1) \simeq \Lie(\mathcal{J}_m)^\vee \simeq H^1(\mathcal{X}_m, \mathcal{O}_{\mathcal{X}_m})^\vee \simeq H^0(\mathcal{X}_m, \omega_m).$$

The cusp $i\infty$ defines a $\mathbb{Q}$-rational point of $X_0(p^n)$, hence, an $R$-section $[i\infty]$ on $\mathcal{X}_m$. Let $\mathcal{X}_m$ be the open set of $\mathcal{X}_m$ obtained by deleting from $\mathcal{X}_m$ the fibre components that do not meet $[i\infty]$. In this way, $[i\infty]$ induces an $R$-morphism $j_m : \mathcal{X}_m \to \Pic^0(\mathcal{X}_m/R) = \mathcal{J}_m \subseteq \mathcal{J}_m$. On $\mathcal{X}_m$ we have a canonical isomorphism between $\Omega^1_{\mathcal{X}_m/R}$ and $\omega_m$, hence we obtain by pull-back

$$j_m^* : H^0(\mathcal{J}_m, \Omega^1) \to H^0(\mathcal{X}_m, \omega_m).$$

One can check (say, by base change to $\mathbb{C}$) that this map factors through (10.2). In particular

$$v(\text{coker}(j_m^*)) = v \left( \frac{H^0(\mathcal{X}_m, \omega_m)}{H^0(\mathcal{X}_m, \omega_m)} \right).$$

On the étale projective system $\{\mathcal{X}_m\}_{p|m}$ the invertible sheaves $\omega_m$ are compatible by pull-back as explained above. One can check that the theory of Conrad (cf. [20], specially Theorem B.3.2.1) for comparing integral structures applies in this slightly modified setting, which gives:

39
Theorem 10.4. As $m$ varies over integers coprime to $p$, we have

\begin{equation}
(10.3) \quad v(\coker(j^*)) = v\left( \frac{H^0(\mathcal{X}_{0,m}^\infty, \Omega_m^1)}{R \cdot \phi^*\omega} \right) \ll_{p,n,\ell} 1.
\end{equation}

In our application, note that we will be taking $n \leq 8$ and $\ell = 5$ or 7, so the implicit constant will be bounded just in terms of $p$.

10.4. Some reductions. Let us write $X_{0,m} = X_0(p^n m)$, $J_{0,m} = J_0(p^n m)$, and consider the standard integral model $\mathcal{X}_{0,m} = \mathcal{X}_0(p^n m) \otimes R$ as well as the Néron model over $R$ of $J_{0,m}$, which we denote by $\mathfrak{J}_{0,m}$.

Let $A$ be an elliptic curve of conductor $p^n m$ and assume that we have an optimal quotient $q : J_{0,m} \to A$. Let $\mathfrak{A}$ be the Neron model of $A$ over $R$, and let $\omega \in H^0(\mathfrak{A}, \Omega^1)$ be a Neron differential.

Consider the embedding $j : X_{0,m} \to J_{0,m}$ induced by the cusp $i\infty$, and the modular parameterization $\phi = q j$. These extend to maps

$$
\mathcal{X}_{0,m}^\infty \to \mathfrak{J}_{0,m} \to \mathfrak{A}
$$

(that we still call $j$, $q$, $\phi$) by the Néron mapping property, where $\mathcal{X}_{0,m}^\infty$ is obtained from $\mathcal{X}_{0,m}$ by deleting the fibre components that do not meet the section $[i\infty]$. The special fibre of $\mathcal{X}_{0,m}^\infty$ is irreducible and the $p$-adic valuation of the Manin constant $c_f$ is the vanishing order $\phi^*\omega$ along it, as a section of the line bundle $\Omega^1_{\mathcal{X}_{0,m}/R}$. Then one has

$$
v_p(c_f) = v\left( \frac{H^0(\mathcal{X}_{0,m}^\infty, \Omega^1)}{R \cdot \phi^*\omega} \right).
$$

That the expression on the right agrees with the vanishing order of $\phi^*\omega$ on the special fibre of $\mathcal{X}_{0,m}^\infty$ as a section of $\Omega^1_{\mathcal{X}_{0,m}/R}$, is seen by considerations on $q$-expansions along the section $[i\infty]$.

Unfortunately, the geometry of $\mathcal{X}_{0,m}$ is not convenient (in particular, duality theory is an issue). So we relate the previous expression to $\mathcal{X}_m$ instead, using the forgetful degeneracy map $\alpha : \mathcal{X}_m \to \mathcal{X}_{0,m}$. We have

\begin{equation}
(10.4) \quad v_p(c_f) = v\left( \frac{H^0(\mathcal{X}_{0,m}^\infty, \Omega^1)}{R \cdot \phi^*\omega} \right) \leq v\left( \frac{H^0(\mathcal{X}_m^\infty, \Omega^1)}{R \cdot (\phi \alpha)^*\omega} \right)
\end{equation}

because $\alpha$ maps the cusp at infinity to the cusp at infinity, so, it restricts to $\mathcal{X}_m^\infty \to \mathcal{X}_{0,m}^\infty$.

However, there is the inconvenience (for later in our argument) that when $\ell \nmid m$, the modular form attached to $A$ is no longer new for the group $U_0(p^n m) \cap U_1(\ell)$, and in that case we are led to also consider the standard second degeneracy map $\beta : \mathcal{X}_m \to \mathcal{X}_{0,m}$ induced on complex points by the map $z \mapsto \ell z$ on $\mathfrak{h}$.

The map $\beta$ sends the cusp $i\infty$ to itself, so it restricts to $\mathcal{X}_m^\infty \to \mathcal{X}_{0,m}^\infty$, giving

\begin{equation}
(10.5) \quad v_p(c_f) = v\left( \frac{H^0(\mathcal{X}_{0,m}^\infty, \Omega^1)}{R \cdot \phi^*\omega} \right) \leq v\left( \frac{H^0(\mathcal{X}_m^\infty, \Omega^1)}{R \cdot (\phi \beta)^*\omega} \right).
\end{equation}

When $\ell \nmid m$, it is important to note (say, by looking at $q$-expansions) that $(\phi \alpha)^*\omega$ and $(\phi \beta)^*\omega$ are $R$-linearly independent.
10.5. **The case** \( \ell \mid m \). Suppose that \( \ell | m \). Then the newform \( f \in S_2(p^m) \) attached to \( A \) for the group \( U_0(p^m) \) continues to be new for the group \( U_0(p^m) \cap U_1(\ell) \). Then we have an optimal quotient \( \theta : J_m \to C \) with \( C \) an elliptic curve over \( \mathbb{Q} \) isogenous to \( A \). By optimality of \( \theta \), there is an isogeny \( \pi : C \to A \) over \( \mathbb{Q} \) such that the following diagram of morphisms over \( \mathbb{Q} \) commutes

\[
\begin{array}{ccc}
J_m & \xrightarrow{\theta} & C \\
\downarrow{\alpha_*} & & \downarrow{\pi} \\
J_{0,m} & \xrightarrow{q} & A.
\end{array}
\]

(10.6)

Here, \( \alpha_* \) is induced by the degeneracy map \( \alpha : X_m \to X_{0,m} \) under Albanese functoriality.

Let \( C \) be the Néron model of \( C \) over \( R \). We have the commutative diagram of \( R \)-morphisms

\[
\begin{array}{ccc}
\mathcal{X}_m^\infty & \xrightarrow{j_m} & \mathcal{J}_m \\
\downarrow{\alpha} & & \downarrow{\theta} \\
\mathcal{X}_{0,m}^\infty & \xrightarrow{j} & \mathcal{J}_{0,m} \\
\end{array}
\]

(10.6)

Dualizing (10.6), we find that \( \alpha^\vee = \alpha^* \) (induced by Picard functoriality) extends \( \pi^\vee \) after composition with the inclusions \( q^\vee \) and \( \theta^\vee \). Thus, \( \operatorname{deg} \pi \) divides \( \operatorname{deg} \alpha \), and the latter is

\[
[U_0(p^m) : U_0(p^m) \cap U_1(\ell)] \leq [U(1) : U_1(\ell)].
\]

Hence

\[
v_p(\operatorname{deg} \pi) \ll_\ell 1.
\]

Let \( a \geq 0 \) be an integer such that \( \tilde{\omega} := p^{-a} \pi^* \omega \) is a Néron differential on \( C \), and note that \( a \leq v_p(\operatorname{deg} \pi) \ll_\ell 1 \). Then we have

\[
v_p(c_f) \leq v \left( \frac{H^0(\mathcal{X}_m^\infty, \Omega^1)}{R \cdot (\phi \alpha)^* \omega} \right) = v \left( \frac{H^0(\mathcal{X}_m^\infty, \Omega^1)}{R \cdot \phi^*_m \tilde{\omega}} \right) + a
\]

where \( \phi_m = \theta j_m \). Note that

\[
v \left( \frac{H^0(\mathcal{X}_m^\infty, \Omega^1)}{R \cdot \phi^*_m \tilde{\omega}} \right) \leq v(\operatorname{coker}(j_m)) + v \left( \frac{H^0(\mathcal{J}_m, \Omega^1)}{R \cdot \theta^* \tilde{\omega}} \right)
\]

because \( j_m^* : H^0(\mathcal{J}_m, \Omega^1) \to H^0(\mathcal{X}_m^\infty, \Omega^1) \) is injective. So by (10.3) we deduce

\[
v_p(c_f) \leq v \left( \frac{H^0(\mathcal{J}_m, \Omega^1)}{R \cdot \theta^* \tilde{\omega}} \right) + O_{p,n,\ell}(1).
\]

(10.7)

Let \( \chi : T_{U_0(p^m) \cap U_1(\ell)} \to \mathbb{Z} \) be the system of Hecke eigenvalues attached to the optimal quotient \( \theta \). Let \( P_{\chi} : H^0(\mathcal{J}_m, \Omega^1)_q \to H^0(\mathcal{J}_m, \Omega^1)_q \) be the orthogonal projection onto the \( \chi \)-component with respect to the Petersson inner product. Observe that \( \theta^* \tilde{\omega} \in H^0(\mathcal{J}_m, \Omega^1)_\chi \), and from the equation \( \theta \theta^\vee = [\operatorname{deg} \phi_m] \in \operatorname{End}(C) \) we deduce that the following diagram commutes:

\[
\begin{array}{c}
R \cdot \tilde{\omega} \\
\downarrow{\theta^*} \\
H^0(\mathcal{J}_m, \Omega^1)
\end{array} \xrightarrow{\phi_m} \begin{array}{c}
R \cdot \tilde{\omega} \\
\downarrow{\theta^*} \\
H^0(\mathcal{J}_m, \Omega^1)
\end{array} \xrightarrow{(\deg \phi_m)} \begin{array}{c}
R \cdot \tilde{\omega} \\
\downarrow{\theta^*} \\
H^0(\mathcal{J}_m, \Omega^1)
\end{array}
\]

where \( \phi_m = \theta j_m \).
The rightmost arrow is induced by \((\theta^v)\bullet\) and the fact that \(\ker((\theta^v)\bullet) = (H^0(\mathcal{F}_m, \Omega^1))^\perp\).

The bottom-right term in the diagram can be replaced by \(P_{\chi}(H^0(\mathcal{F}_m, \Omega^1))\), which is an \(R\)-module of rank 1, and the bottom arrow can be replaced by \(P_{\chi}\). Chasing the image of \(\tilde{\omega}\) we deduce

\[
(10.8) \quad v_p(\deg \phi_m) \geq v \left( \frac{H^0(\mathcal{F}_m, \Omega^1)}{R \cdot \theta^v \tilde{\omega}} \right) + v \left( \frac{P_{\chi}(H^0(\mathcal{F}_m, \Omega^1))}{H^0(\mathcal{F}_m, \Omega^1)^\perp} \right).
\]

In [3] Theorem 3.6 (a), it is shown that the modular exponent (which equals the modular degree in the elliptic curve case) divides the congruence exponent (defined in terms of Fourier expansions at \(i\infty\)) for groups of the form \(U_0(N)\) and \(U_1(N)\). The same proof works for intermediate subgroups such as \(U_0(p^n m) \cap U_1(\ell)\). So, the integer \(\deg \phi_m\) divides the congruence exponent, which by definition is the exponent of

\[
\frac{S_m(\mathbb{Z})}{(S_m(\mathbb{Z}))^\perp + S_m(\mathbb{Z})^\perp}
\]

where \(S_m(\mathbb{Z})\) is the subgroup of \(S_2(U_0(p^n m) \cap U_1(\ell))\) consisting of modular forms with Fourier coefficients at \(i\infty\) in \(\mathbb{Z}\). By the \(q\)-expansion principle, we have a canonical isomorphism \(S_m(\mathbb{Z}) \otimes R = H^0(\mathcal{F}_m^\infty, \Omega^1) = H^0(\mathcal{F}_m^\infty, \omega_m)\), which together with (10.2) gives

\[
v_p(\deg \phi_m) \leq v \left( \frac{H^0(\mathcal{F}_m^\infty, \omega_m)}{H^0(\mathcal{F}_m^\infty, \omega_m)^\perp + H^0(\mathcal{F}_m^\infty, \omega_m)^\perp} \right)
\leq v \left( \frac{H^0(\mathcal{F}_m^\infty, \omega_m)}{H^0(\mathcal{F}_m^\infty, \omega_m)^\perp + H^0(\mathcal{F}_m^\infty, \omega_m)^\perp} \right) + v \left( \frac{H^0(\mathcal{F}_m^\infty, \omega_m)}{H^0(\mathcal{F}_m^\infty, \omega_m)^\perp + H^0(\mathcal{F}_m^\infty, \omega_m)^\perp} \right)
\leq v \left( \frac{P_{\chi}(H^0(\mathcal{F}_m, \Omega^1))}{H^0(\mathcal{F}_m, \Omega^1)^\perp} \right) + v(\ker(j_m)) .
\]

It follows from (10.3) and (10.8) that

\[
v(\frac{H^0(\mathcal{F}_m, \Omega^1)}{R \cdot \theta^v \tilde{\omega}}) \ll_{p,n,\ell} 1
\]

which by (10.7) proves

\[
v_p(c_f) \ll_{p,n,\ell} 1 .
\]

This concludes the proof of Theorem 10.3 in the case \(\ell | m\). Note that the bound is independent of \(S\), and we actually obtain

\[
v_p(c_f) \leq 2v(\ker(j_m)) + v_p(\deg \pi) .
\]

10.6. The case \(\ell \nmid m\). Now we assume that \(\ell \nmid m\) (and, as always, \(\ell \neq p\)). Then the newform \(f \in S_2(p^n m)\) attached to \(A\) for the group \(U_0(p^n m)\) is no longer new for the group \(U_0(p^n m) \cap U_1(\ell)\). In fact, let \(\chi : T_{U_0(p^n m) \cap U_1(\ell)} \to \mathbb{Z}\) be the system of Hecke eigenvalues attached to this old form, then \(\dim_{\mathbb{Q}} H^0(J_m, \Omega^1)^\perp = 2\), which is explained by the two degeneracy maps \(\alpha, \beta : X_m \to X_{0,m}\).

Nevertheless, attached to \(\chi\) we have an optimal quotient \(\theta : J_m \to \Sigma\) over \(\mathbb{Q}\), where \(\Sigma\) is an abelian surface isogenous to \(A \times A\). By optimality of \(\theta\) and considering the relevant
cotangent spaces, we see that there is an isogeny $\pi : \Sigma \to A \times A$ over $\mathbb{Q}$ making the following diagram commutative:

$$
\begin{array}{ccc}
J_m & \overset{\theta}{\longrightarrow} & \Sigma \\
\downarrow \alpha \times \beta & & \downarrow \pi \\
J_{0,m}^{2} & \overset{q \times q}{\longrightarrow} & A \times A.
\end{array}
$$

(10.9)

Lemma 10.5. There is an integer $u \ll_{S,p,\ell} 1$ such that $p^u$ annihilates the $p$-primary part of the kernel of $\pi$.

Proof. Let $\sigma : X_m \to X_0(p^n m \ell)$ be the forgetful map, and write $\alpha_0, \beta_0 : X_0(p^n m \ell) \to X_{0,m}$ for the two degeneracy maps. We have $\alpha = \alpha_0 \sigma$ and $\beta = \beta_0 \sigma$, hence, the following diagram commutes:

$$
\begin{array}{ccc}
J_0(p^n m \ell) & \overset{\sigma^*}{\longrightarrow} & J_m & \overset{\theta^*}{\longleftarrow} & \Sigma^\vee \\
\alpha_0 + \beta_0 & \uparrow & \alpha_0 + \beta_0 & \uparrow & \pi^\vee \\
J_{0,m}^2 & \overset{q^\vee \times q^\vee}{\leftarrow} & J_{0,m}^2 & \overset{\alpha^* + \beta^*}{\leftarrow} & A \times A.
\end{array}
$$

(10.10)

The kernel of $\pi^\vee$ is Cartier-dual to that of $\pi$, so, it suffices to prove the claim for $\ker(\pi^\vee)$ instead.

The maps $q^\vee \times q^\vee$ and $\theta^\vee$ are injective as they are duals of optimal quotients, so, it suffices to bound a power of $p$ that annihilates the $p$-primary part of $\ker(\alpha^* + \beta^*) \cap q^\vee(A) \times q^\vee(A)$.

Note that $\sigma_\alpha \sigma^* = [\deg \sigma] \in \text{End}(J_0(p^n m \ell))$, so $v(\ker(\sigma^*)) \ll_{\ell} 1$. Hence

$$
v(\ker((\alpha^* + \beta^*) \circ (q^\vee \times q^\vee))) = v(\ker((\alpha_0^* + \beta_0^*) \circ (q^\vee \times q^\vee))) + O(1)
$$

and moreover $\ker((\alpha_0^* + \beta_0^*) \circ (q^\vee \times q^\vee))$ is isomorphic to

$$
Z = \ker(\alpha_0^* + \beta_0^*) \cap (q^\vee(A) \times q^\vee(A)).
$$

By Ihara’s lemma in Ribet’s formulation [68, 70], we have that $\ker(\alpha_0^* + \beta_0^*)$ is Eisenstein (in Mazur’s terminology), so that for every prime $r \nmid pm \ell$ one has that $T_r$ acts as $r + 1$ on it. On the other hand, when $r \nmid pm \ell$ we have that $T_r$ acts as $\chi(T_r) = a_r(A)$ on $A \times A \subseteq J_{0,m}^2$. Thus, if $Z$ has a (geometric) point of exact order $p^e$, we see that for all primes $r \nmid pm \ell$ we have $a_r(A) \equiv r + 1 \mod p^e$. By Lemma [65] we obtain $e \ll_{S,p,\ell} 1$, which concludes the proof. □

The modular exponent $n_S$ of the optimal quotient $\theta : J_m \to \Sigma$ is defined as in [3], namely, as the exponent of the group $\ker(\theta^\vee)$. The theory of [3] applies to the abelian surface $\Sigma$ as in the first example of Section 3 in loc. cit., with only some minor modifications due to the fact that we are working with the group $U_0(p^n m) \cap U_1(\ell)$ rather than a group of the form $U_0(N)$ or $U_1(N)$.

Consider the map

$$
\tau = \pi \theta : J_m \to A \times A.
$$

Then $\tau^\vee = \pi \theta^\vee \pi^\vee$, so that $v(\ker(\tau^\vee)) = v_p(n_S) + O_{S,p,\ell}(1)$ by the previous lemma.

On the other hand, from the description $\tau = (q \times q) \circ (\alpha \times \beta) = (qa_\alpha) \times (q\beta_\alpha)$ we see that

$$
\tau^\vee = [\deg(\phi) \deg(\alpha)]_A \times [\deg(\phi) \deg(\beta)]_A = [\deg(\phi) \deg(\alpha)]_{A \times A}.
$$

(10.11)
It follows that

\[(10.12) \quad v_p(\deg(\phi) \deg(\alpha)) = v_p(\tilde{n}_\Sigma) + O_{S,p,\ell}(1).\]

Take any \(\tilde{\omega} \in H^0(\mathcal{O}^2, \Omega^1)\) primitive. We make two observations about \(\tilde{\omega}\).

First, we have

\[(10.13) \quad v\left(\frac{H^0(\mathcal{O}^2, \Omega^1)}{R \cdot (\tau_{m})^* \tilde{\omega}}\right) \leq v(\coker(\phi)) + v\left(\frac{H^0(\mathcal{O}^2, \Omega^1)}{R \cdot \tau \tilde{\omega}}\right).\]

Secondly, we have

\[(10.14) \quad v_p(\deg(\phi) \deg(\alpha)) \geq v\left(\frac{H^0(\mathcal{O}^2, \Omega^1)}{R \cdot \tau \tilde{\omega}}\right) + v\left(\frac{P_\chi(H^0(\mathcal{O}^2, \Omega^1))}{H^0(\mathcal{O}^2, \Omega^1) \chi \cap \mathbb{Q} \cdot \tau \tilde{\omega}}\right).\]

This last bound is proved by recalling \([10.11]\), chasing \(\tilde{\omega}\) in the diagram

\[
\begin{array}{ccc}
H^0(\mathcal{O}^2, \Omega^1) & \stackrel{(\tau \tilde{\omega})^*}{\longrightarrow} & H^0(\mathcal{O}^2, \Omega^1) \\
\downarrow & & \uparrow \\
H^0(\mathcal{O}^2, \Omega^1) & \longrightarrow & \frac{H^0(\mathcal{O}^2, \Omega^1)}{(\tau \tilde{\omega})^*} \\
\end{array}
\]

noticing that the rightmost arrow is injective, replacing the bottom right term in the diagram by \(P_\chi(H^0(\mathcal{O}^2, \Omega^1))\), and noticing that \(H^0(\mathcal{O}^2, \Omega^1) \cap \mathbb{Q} \cdot \tau \tilde{\omega} = H^0(\mathcal{O}^2, \Omega^1) \chi \cap \mathbb{Q} \cdot \tau \tilde{\omega}\). Here, \(P_\chi\) is the orthogonal projection onto the \(\chi\)-isotypical component.

Furthermore, since \(\tau^*\) maps \(H^0(\mathcal{O}^2, \Omega^1)\) into \(H^0(\mathcal{O}^2, \Omega^1) \chi\) with torsion cokernel, we see that there is some \(\tilde{\omega}_0 \in H^0(\mathcal{O}^2, \Omega^1)\) primitive such that

\[
v\left(\frac{P_\chi(H^0(\mathcal{O}^2, \Omega^1))}{H^0(\mathcal{O}^2, \Omega^1) \chi \cap \mathbb{Q} \cdot \tau \tilde{\omega}_0}\right) = v\left(\frac{P_\chi(H^0(\mathcal{O}^2, \Omega^1))}{H^0(\mathcal{O}^2, \Omega^1) \chi}\right)
\]

We fix a choice of such an \(\tilde{\omega}_0\), and for it we obtain from \([10.12], (10.13), (10.14)\) and \([10.3]\) that

\[(10.15) \quad v\left(\frac{P_\chi(H^0(\mathcal{O}^2, \Omega^1))}{H^0(\mathcal{O}^2, \Omega^1) \chi}\right) + v\left(\frac{H^0(\mathcal{O}^2, \Omega^1)}{R \cdot (\tau_{m})^* \tilde{\omega}_0}\right) \leq v_p(\tilde{n}_\Sigma) + O_{S,p,\ell}(1).
\]

By \([3]\) as in the case \(\ell|m\) above, the modular exponent \(\tilde{n}_\Sigma\) divides the congruence exponent associated to \(\Sigma\), and one deduces

\[
v_p(\tilde{n}_\Sigma) \leq v\left(\frac{P_\chi(H^0(\mathcal{O}^2, \Omega^1))}{H^0(\mathcal{O}^2, \Omega^1) \chi}\right) + v(\coker(\phi))
\]

which together with \([10.15]\) gives

\[(10.16) \quad v\left(\frac{H^0(\mathcal{O}^2, \Omega^1)}{R \cdot (\tau_{m})^* \tilde{\omega}_0}\right) \ll_{S,p,\ell} 1.
\]

Recall that we had a Neron differential \(\omega \in H^0(\mathcal{O}^2, \Omega^1)\). Let \(\pi_1, \pi_2\) be the two projections \(\mathcal{O}^2 \to \mathcal{O}\) and let \(\omega_i = \pi_i^* \omega\). Then \(\omega_1, \omega_2\) generate \(H^0(\mathcal{O}^2, \Omega^1)\) as an \(R\)-module and we can
write $\tilde{\omega}_0 = r_1\omega_1 + r_2\omega_2$ with $r_1, r_2 \in R$, at least one of them a unit. We observe that
\[(\tau j_m)^*\tilde{\omega}_0 = r_1(\tau j_m)^*\omega_1 + r_2(\tau j_m)^*\omega_2 \]
\[= r_1(\tau j_m)^*\pi_1^*(\omega) + r_2(\tau j_m)^*\pi_2^*(\omega) \]
\[= r_1(\pi_1\tau j_m)^*(\omega) + r_2(\pi_2\tau j_m)^*(\omega) \]
\[= r_1(q\alpha_s j_m)^*(\omega) + r_2(q\beta_s j_m)^*(\omega) \]
\[= r_1(qj\alpha)^*(\omega) + r_2(qj\beta)^*(\omega) \]
\[= r_1(\phi\alpha)^*(\omega) + r_2(\phi\beta)^*(\omega). \]

We claim that
\[(10.17) \quad \min \left\{ v \left( \frac{H^0(\mathcal{D}_m^\infty, \Omega^1)}{R \cdot (\phi\alpha)^*(\omega)} \right), v \left( \frac{H^0(\mathcal{D}_m^\infty, \Omega^1)}{R \cdot (\phi\beta)^*(\omega)} \right) \right\} \leq v \left( \frac{H^0(\mathcal{D}_m^\infty, \Omega^1)}{R \cdot (\tau j_m)^*\tilde{\omega}_0} \right). \]

In fact, this follows by interpreting the three terms appearing in the expression as the vanishing orders of the corresponding differential forms along the (irreducible) special fibre of $\mathcal{D}_m^\infty$.

Finally, from (10.4) and (10.5) we obtain
\[v_p(c_f) \leq \min \left\{ v \left( \frac{H^0(\mathcal{D}_m^\infty, \Omega^1)}{R \cdot (\phi\alpha)^*(\omega)} \right), v \left( \frac{H^0(\mathcal{D}_m^\infty, \Omega^1)}{R \cdot (\phi\beta)^*(\omega)} \right) \right\}. \]

By (10.16) and (10.17) we get
\[v_p(c_f) \ll_{S,p,n,\ell} 1 \]
which concludes the proof of Theorem 10.3 in the case $\ell \mid m$. This completes the proof of Theorem 10.3; hence, of Theorem 10.1.

11. Counting imaginary quadratic extensions of $\mathbb{Q}$

11.1. The counting result. We will be interested in having a suitable Heegner point in $X_0^D(M)$. This will be achieved by showing the existence of sufficiently many imaginary quadratic extensions $K/\mathbb{Q}$ satisfying certain technical conditions.

For a Dirichlet character $\chi$, we let $L(s, \chi)$ be its $L$-function. If $D$ is a fundamental discriminant, we write $K_D$ for the quadratic number field of discriminant $D$, and $\chi_D$ for the non-trivial quadratic Dirichlet character associated to $K_D$. The counting result is the following.

**Theorem 11.1.** There is a uniform constant $\kappa$ such that the following holds:

Let $\theta > 0$. For $x > 1$ and coprime positive integers $D$ and $M$, let $S_\theta(D, M, x)$ be the set of positive integers $d$ satisfying the following conditions:

(i) $x < d \leq 2x$
(ii) $d \equiv -1 \mod 4$ and $d$ is squarefree (hence, $-d$ is a fundamental discriminant);
(iii) $p$ splits in $K_{-d}$ for each prime $p|M$
(iv) $p$ is inert in $K_{-d}$ for each prime $p|D$
(v) $\#Cl(K_{-d}) > d^{0.5-\theta}$
(vi) $|\frac{1}{2}L(1, \chi_{-d})| < \kappa \log \log d$.

Then, writing $N = DM$, we have that for $x \gg_\theta 1$
\[
\#S_\theta(D, M, x) = \frac{x}{2^\omega(N) + 1} \zeta(2) \prod_{p|2N} (1 + \frac{1}{p})^{\frac{1}{45}} + O(x^{1/2}N^{1/4}(\log N)^{1/2})
\]
where the implicit constant is absolute.

**Corollary 11.2.** Let \( \theta > 0 \) and \( \delta > 0 \). With the notation of Theorem \[11.1\], for \( N \gg_{\theta, \delta} 1 \) and \( x > N^{0.5+\delta} \), we have
\[
\#S_\theta(D, M, x) > x^{1-\delta}.
\]

11.2. **Preliminaries on \( L \)-functions.** We need two analytic results about \( L(s, \chi_D) \) where \( D \) is a fundamental discriminant.

The next result is Corollary 2.5 in [54].

**Proposition 11.3.** Let \( \epsilon > 0 \). There is a number \( c_\epsilon > 0 \) depending only on \( \epsilon \) such that the bound
\[
\left| \frac{L'}{L}(1, \chi_D) \right| < c_\epsilon \log \log |D|
\]
holds for all but \( O_\epsilon(x^\epsilon) \) fundamental discriminants \( D \) with \( |D| < x \).

We also need Siegel’s classical bound for the class number. The following explicit version is due to Tatuzawa [82].

**Proposition 11.4.** Let \( 0 < \epsilon < 1/12 \). For \(-d\) a negative fundamental discriminant with \( d > e^{1/\epsilon} \), the bound
\[
\#Cl(\mathbb{Q}(\sqrt{-d})) > 0.2 \cdot d^{0.5-\epsilon}
\]
holds with at most one exception.

11.3. **Lemmas on squarefrees.**

**Lemma 11.5.** Let \( m > 1 \) be a positive integer and let \( \chi \) be a non-principal Dirichlet character to the modulus \( m \) (not necessarily primitive). For \( x \geq 1 \) we have
\[
\left| \sum_{d \leq x} \mu^2(d) \chi(d) \right| \leq 5x^{1/2} m^{1/4} (\log m)^{1/2}.
\]

*Proof.* When \( x \leq m^{1/2} \log m \) we see that the required sum is bounded by
\[
x \leq x^{1/2} m^{1/4} (\log m)^{1/2}
\]
so we may assume that \( x > m^{1/2} \log m \).

Let \( y \leq x^{1/2} \) be a positive integer. We have
\[
\sum_{d \leq x} \mu^2(d) \chi(d) = \sum_{d \leq x} \chi(d) \sum_{k^2 | d} \mu(k) = \sum_{k \leq x^{1/2}} \mu(k) \chi(k^2) \sum_{r \leq x/k^2} \chi(r) = S_1 + S_2
\]
where \( S_1 \) corresponds to terms with \( k \leq y \), and \( S_2 \) corresponds to those with \( y < k \leq x^{1/2} \).

By Polya-Vinogradov
\[
|S_1| \leq 2y m^{1/2} \log m,
\]
while for \( S_2 \) we have the trivial bound
\[
|S_2| \leq x \sum_{k > y} \frac{1}{k^2} < \frac{x}{y}.
\]
Take \( y = \lceil x^{1/2}m^{-1/4}(\log m)^{-1/2} \rceil \) and note that
\[
\frac{x^{1/2}}{m^{1/4}(\log m)^{1/2}} \leq y \leq \frac{2x^{1/2}}{m^{1/4}(\log m)^{1/2}}
\]
because \( x > m^{1/2} \log m \). The result follows. \( \square \)

Lemma 11.6. Let \( a \) be an odd residue class modulo 4, let \( m \) be a positive integer and let \( x > 1 \). Then
\[
\sum_{\substack{d \leq x \\ (d,m) = 1 \\ d \equiv a \pmod{4}}} \mu^2(d) = \frac{x}{2\zeta(2)} \prod_{p|2m} \left( 1 + \frac{1}{p} \right) + O(x^{1/2}m^{1/4}(\log m)^{1/2}).
\]

Proof. Let \( \psi_0 \) and \( \psi \) be the principal and the non-principal characters modulo 4 respectively. By the previous lemma we have
\[
2 \sum_{\substack{d \leq x \\ (d,m) = 1 \\ d \equiv a \pmod{4}}} \mu^2(d) = \sum_{d \leq x \atop (d,m) = 1} \mu^2(d) \psi_0(d) + \psi(a) \sum_{d \leq x \atop (d,m) = 1} \mu^2(d) \psi(d)
\]
\[
= \sum_{d \leq x \atop (d,2m) = 1} \mu^2(d) + O(x^{1/2}m^{1/4}(\log m)^{1/2}).
\]
The number of positive integers coprime to \( A \) up to a bound \( y \) is
\[
\sum_{b \mid A} \mu(b) = \sum_{b \mid A} \mu(b) \left( \frac{y}{b} + O(1) \right) = \frac{\phi(A)y}{A} + O(2^{\omega(A)}).
\]
Taking \( A = 2m \) and writing \( \mu^2(d) = \sum_{b \mid d} \mu(b) \) we find
\[
\sum_{d \leq x \atop (d,2m) = 1} \mu^2(d) = \sum_{b \leq x^{1/2} \atop (b,2m) = 1} \mu(b) \sum_{c \leq x/b^2 \atop (c,2m) = 1} 1
\]
\[
= \frac{\phi(2m)x}{2m} \sum_{b \leq x^{1/2} \atop (b,2m) = 1} \frac{\mu(b)}{b^2} + O(2^{\omega(m)}x^{1/2})
\]
\[
= \frac{\phi(2m)x}{2m} \sum_{b=1 \atop (b,2m) = 1}^{\infty} \frac{\mu(b)}{b^2} + O(x^{1/2} + 2^{\omega(m)}x^{1/2}).
\]
The infinite series is no other than \( 1/\zeta(2) \) with the Euler factors for primes dividing \( 2m \) removed, hence the result because \( 2^{\omega(m)} \ll \epsilon m^\epsilon \). \( \square \)

11.4. Proof of the counting result. We will restrict ourselves to \( d > 0 \) squarefree satisfying \( d \equiv 3 \pmod{4} \), in which case \( -d \) is a negative fundamental discriminant. Thus, \( \chi_{-d} \) is the non-principal Dirichlet character associated with the imaginary quadratic field \( K_{-d} = \mathbb{Q}(\sqrt{-d}) \), and the discriminant of \( K_{-d} \) is precisely \( -d \).
Under our assumption on \( d \), the character \( \chi_{-d} \) is the Kronecker symbol \((\frac{-d}{\cdot})\). It has conductor \( d \) and is determined by the following conditions on primes: for \( p \) odd, \( \chi_{-d}(p) \) is the Legendre symbol \((\frac{-d}{p})\), and for \( p = 2 \) we have

\[
\chi_{-d}(2) = \begin{cases} 
1 & \text{if } d \equiv 7 \pmod{8} \\
-1 & \text{if } d \equiv 3 \pmod{8}.
\end{cases}
\]

Equivalently, the values of \( \chi_{-d} \) at primes are determined by

\[
\chi_{-d}(p) = \begin{cases} 
0 & \text{if } p \text{ ramifies in } K_{-d} \\
1 & \text{if } p \text{ splits in } K_{-d} \\
-1 & \text{if } p \text{ is inert in } K_{-d}.
\end{cases}
\]

**Proof of Theorem 11.1.** Let \( S'(D, M, x) \) be the set of positive integers \( d \) satisfying conditions (i), (ii), (iii), and (iv) of the statement of Theorem 11.1. For each pair \( (D, M) \) and each prime \( p | N = DM \), define the numbers

\[
e_p = \epsilon_p(D, M) = \begin{cases} 
1 & \text{if } p | M \\
-1 & \text{if } p | D
\end{cases}
\]

and define \( \epsilon_b \) for all divisors \( b \) of \( N \) as \( \epsilon_b = 0 \) if \( b \) is not squarefree, and multiplicatively otherwise (using the numbers \( \epsilon_p \)). Then we have

\[
\#S'(D, M, x) = 2^{-\omega(N)} \sum_{x < d \leq 2x, (d, N) = 1, d \equiv 3 \pmod{4}} \mu(d)^2 \prod_{p | N} (1 + \epsilon_p \chi_{-d}(p))
\]

\[
= 2^{-\omega(N)} \sum_{x < d \leq 2x, (d, N) = 1, d \equiv 3 \pmod{4}} \mu(d)^2 \sum_{b | N} \epsilon_b \chi_{-d}(b).
\]

Writing \( 1_N \) for the principal character to the modulus \( N \), we find

\[
\#S'(D, M, x) = 2^{-\omega(N)} \sum_{b | N} \epsilon_b \sum_{x < d \leq 2x, d \equiv 3 \pmod{4}} \mu^2(d) \left(\frac{-d}{b}\right) \cdot 1_N(d).
\]

By Lemma 11.6 the contribution to the previous expression coming from \( b = 1 \) is

\[
2^{-\omega(N)} \sum_{x < d \leq 2x, (d, N) = 1, d \equiv 3 \pmod{4}} \mu^2(d) = x \frac{2^{-\omega(N)+1} \zeta(2) \prod_{p | 2N} \left(1 + \frac{1}{p}\right)}{2^{\omega(N)+1} \zeta(2) \prod_{p | 2N} \left(1 + \frac{1}{p}\right)} + O(x^{1/2} N^{1/4} (\log N)^{1/2}).
\]

When \( b \neq 1 \) is a squarefree divisor of \( N \), the function \( d \mapsto \left(\frac{-d}{b}\right) \cdot 1_N(d) \) is a Dirichlet character whose modulus divides \( 4N \), and it is non-principal because \( b \) is squarefree. Thus, by Lemma 11.5 we get

\[
2^{-\omega(N)} \sum_{b \neq 1, b | N} \epsilon_b \sum_{x < d \leq 2x, d \equiv 3 \pmod{4}} \mu^2(d) \left(\frac{-d}{b}\right) \cdot 1_N(d) \ll x^{1/2} N^{1/4} (\log N)^{1/2}.
\]
Therefore we find
\[
\#S'(D, M, x) = \frac{x}{2^{\omega(N)+1} \zeta(2) \prod_{p|2N} \left(1 + \frac{1}{p}\right)} + O \left(x^{1/2} N^{1/4} (\log N)^{1/2}\right).
\]

Finally, by Proposition 11.3 with \(\epsilon = 1/2\) and Proposition 11.4 with \(\epsilon = \theta/2\) we see that for \(x \gg \theta\) we have
\[
\#S'(D, M, x) - \#S(\theta, D, M, x) \ll x^{1/2}
\]
which concludes the proof, with the constant \(\kappa = c_{1/2}\) from Proposition 11.3.

\[\square\]

12. Arakelov degrees

The purpose of this brief section is to introduce some notation related to Arakelov degrees of metrized line bundles. This will be used at various places later in the paper.

12.1. Metrized line bundles on arithmetic curves. Let \(L\) be a number field and write \(S_L = \text{Spec} \mathcal{O}_L\). If \(\mathcal{M}\) is an invertible sheaf (also called line bundle) on \(S_L\), then \(\mathcal{M} = H^0(S_L, \mathcal{M})\) is a projective \(\mathcal{O}_L\)-module of rank 1, and since \(S_L\) is affine we have \(\mathcal{M} = \mathcal{M}^\vee\). In this way, invertible sheaves on \(S_L\) correspond to projective \(\mathcal{O}_L\)-modules of rank 1.

For every embedding \(\sigma : L \to \mathbb{C}\) we let \(|-|_\sigma\) be the absolute value on \(L\) induced by \(\sigma\). A metrized line bundle on \(S_L\) is an invertible sheaf \(\mathcal{M}\) (or equivalently, its associated projective \(\mathcal{O}_L\)-module \(\mathcal{M}\)) together with the following data: for each embedding \(\sigma : L \to \mathbb{C}\), a norm \(\|\cdot\|_\sigma\) on \(M_\sigma := \mathcal{M} \otimes_\sigma \mathbb{C}\) compatible with the absolute value \(|-|_\sigma\). Thus, a metrized invertible sheaf on \(S_L\) is a pair \(\hat{\mathcal{M}} = (\mathcal{M}, \{|-\|_\sigma\})\) with the notation as before.

12.2. Arakelov degree. The Arakelov degree \(\hat{\deg}_L \hat{\mathcal{M}}\) of a metrized line bundle \(\hat{\mathcal{M}}\) on \(S_L\) is defined as follows: take any non-zero \(\eta \in \mathcal{M} = H^0(S_L, \mathcal{M})\), then
\[
\hat{\deg}_L \hat{\mathcal{M}} := \log \# (\mathcal{M}/\langle \eta \rangle) - \sum_{\sigma : L \to \mathbb{C}} \log \|\eta\|_\sigma,
\]
which is independent of the choice of \(\eta \neq 0\). The Arakelov degree \(\hat{\deg}_L\) is additive on tensor products of metrized line bundles, so, it has an obvious extension to metrized \(\mathbb{Q}\)-line bundles.

We conclude by making two observations that will be useful in later sections of the paper. First, for any choice of \(0 \neq \eta \in \mathcal{M}\) we have
\[
\hat{\deg}_L \hat{\mathcal{M}} \geq - \sum_{\sigma : L \to \mathbb{C}} \log \|\eta\|_\sigma.
\]
Secondly, note that if \(\mathcal{N}\) is an invertible sub-sheaf of \(\mathcal{M}\), and the metrized line bundle \(\hat{\mathcal{N}}\) is defined by restricting the metrics at infinity of the metrized line bundle \(\hat{\mathcal{M}}\), then we have
\[
\hat{\deg}_L \hat{\mathcal{N}} \leq \hat{\deg}_L \hat{\mathcal{M}}.
\]
13. Arakelov height of Heegner points

13.1. Heegner hypothesis for \((D, M)\). Let \(D\) and \(M\) be coprime positive integers with \(D\) squarefree and \(\omega(D)\) even. We say that a quadratic number field \(K/\mathbb{Q}\) satisfies the Heegner hypothesis for the pair \((D, M)\) if every prime \(p|D\) is inert in \(K\) and every prime \(p|M\) splits in \(K\). In particular, the primes dividing \(DM\) are unramified in \(K\).

If \(K\) satisfies the Heegner hypothesis for \((D, M)\), there is an embedding \(\psi : K \to B\) (with \(B\) the quaternion algebra of discriminant \(D\)) which is optimal for the Eichler order \(R_0^D(M) \subseteq O_B\) of reduced discriminant \(M\), in the sense that \(\psi^{-1}(R_0^D(M)) = O_K\).

Fixing a choice of \(\psi\) this leads to a point \(\tau_K \in \mathfrak{h}\) and the corresponding point \(P_K \in X(K^{ab})\) (cf. Paragraph 4.4) with \(X = \lim_{\to U} X_U\) (cf. Paragraph 4.2). Its image \(P_{K,D,M} := P_{K,U_0^D(M)}\) in \(X_0^D(M)\) has residue field \(H_K\), the Hilbert class field of \(K\).

We call these points \((D, M)\)-Heegner points. They are a particular case of the \(D\)-Heegner points discussed in Paragraph 4.4.

13.2. Reduction of Heegner points. We write \(\mathcal{X}_0^D(M) = \mathcal{X}_0^D(M, U_0^D(1))\) for the normal integral model of \(X_0^D(M)\), flat, projective over \(\mathbb{Z}\), introduced in Paragraph 4.4, and we let \(\mathcal{X}_0^D(M)^0 = \mathcal{X}_0^D(M, U_0^D(1))^0\) be the smooth locus of its structure map as in loc cit.

**Lemma 13.1.** Let \(K\) be a quadratic imaginary extension of \(\mathbb{Q}\) satisfying the \((D, M)\)-Heegner hypothesis. Let \(U\) be a compact open subgroup of \(O_\mathbb{Z}^\times\) with \(m_U\) coprime to \(DM\). Let \(C\) be the closure of \(P_{K,U_0^D(M),ab}\) in the surface \(\mathcal{X}_0^D(M, U)\). Then \(C\) is contained in \(\mathcal{X}_0^D(M, U)^0\).

**Proof.** We observe that \(\mathcal{X}_0^D(M, U)^0\) is the preimage of \(\mathcal{X}_0^D(M)^0\) via the forgetful morphism \(\mathcal{X}_0^D(M, U) \to \mathcal{X}_0^D(M)\). Thus, we may assume that \(U = U^D(1)\). Furthermore, we only need to study the intersection of \(C\) with fibres at \(p\) with \(p|DM\).

If \(p|D\) then \(P_{K,U_0^D(M)}\) does not reduce to a supersingular point because \(p\) does not ramify in \(K\), and this suffices.

If \(p|M\) then \(P_{K,U_0^D(M)}\) does not reduce to a supersingular point because \(p\) splits in \(K\) and by same argument as in p. 256 in [37]. It remains to show that \(C\) does not meet a non-reduced component of the fibre at \(p\) (this case only occurs when \(M\) is not squarefree).

Since \(K\) satisfies the Heegner hypothesis for \((D, M)\), the residue field of \(P_{K,U_0^D(M)}\) is \(H_K\). As \(p\) does not ramify in \(K\), it does not ramify in \(H_K\) and we can base change to \(H_K\) obtaining an étale cover near the fibre at \(p\). Now the multi-section \(C\) is the image of a section \(C_{H_K}\) of the structure map to \(\text{Spec} O_{H_K}\). Blowing-up the supersingular points of characteristic \(p\) we may work on a regular surface, and now Corollary 1.32 in p. 388 of [50] shows that \(C_{H_K}\) does not meet a non-reduced fibre.

We remark that for \(D = 1\), this is proved in Proposition (3.1) in [37]. One could also adapt the argument there to analyze the case \(p|M\) in the previous proof.

13.3. Metrized canonical sheaf. The surface \(\mathcal{X}_U = \mathcal{X}_0^D(1, U)\) is semi-stable over \(\mathbb{Z}[m_U^{-1}]\), and when \(U\) is good enough in the sense of Paragraph 4.4, then \(\mathcal{X}_U\) is regular. Let \(\omega_U\) be the relative dualizing sheaf of \(\mathcal{X}_U \to \text{Spec} \mathbb{Z}[m_U^{-1}]\). Then \(\omega_U\) is an invertible sheaf.

For each inclusion \(U \subseteq V\) of open compact subgroups with \(U\) and \(V\) good enough, the map \(\pi_V^U : \mathcal{X}_U \to \mathcal{X}_V\) is étale and

\[
(\pi_V^U)^\ast \omega_V = \omega_U
\]
by Ch. 6 of [56], more precisely, Lemma 4.26 and Theorem 2.32 in loc. cit.

For each $U$ good enough, the metrized canonical sheaf $\hat{\omega}_U$ on the surface $\mathcal{X}_U$ (over $\mathbb{Z}[m_U^{-1}]$) is defined as the line sheaf $\omega_U$ endowed with the following metric $\| - \|_U$ at infinity (defined away from cuspidal points, if any):

The base change of $\mathcal{X}_U$ to $\mathbb{C}$ is $X_U^\text{en}$ and the relative dualizing sheaf becomes $\Omega_{X_U^\text{en}}^1$. The decomposition (14.1) has only one term (because $U$ is good enough), so we can take $g_a = 1$ and drop it from the related notation. The uniformization $\xi_U : \mathfrak{h} \to \tilde{\Gamma}_U \backslash \mathfrak{h}$ is unramified, and the metric on $\Omega_{X_U^\text{en}}^1$ is defined by

$$\|\alpha_P\|_U = 2|f(\tau)|\Im(\tau)$$

for $P \in \tilde{\Gamma}_U \backslash \mathfrak{h} \subseteq X_U^\text{en}$, $\alpha$ a regular section of $\Omega_{X_U^\text{en}}^1$ near $P$, $\tau \in \mathfrak{h}$ a pre-image of $P$ under $\xi_U$, and $f$ holomorphic near $\tau$ such that $\xi_U \cdot g_a \alpha = f(z)dz$ on an appropriate neighborhood of $\tau \in \mathfrak{h}$.

When $U$ is not good enough, we do not define a metrized canonical sheaf. We observe that the metric that we put on $\omega_U$ is not the Arakelov canonical metric, but instead, the hyperbolic metric. When $D = 1$ this metric has singularities at the cusps; otherwise, it is smooth.

At least two other alternative approaches to define metrized canonical bundles on integral models of Shimura curves are available in the literature. In [52] one works on the stack-theoretical integral model of $X_0^D(1)$, while in [93] one proceeds as above with two technical differences: the integral models for $U$ sufficiently small used there come from the theory of integral models of curves of genus $g \geq 2$, and then quotient maps are used to define a version of $\hat{\omega}_U$ as metrized $\mathbb{Q}$-line bundles even when $U$ is not sufficiently small. The geometric properties are deduced by relating them to integral models of auxiliary Shimura curves. The construction of integral models in [93] has the technical advantage of being available even beyond the case of Shimura curves over $\mathbb{Q}$, where a direct modular interpretation is no longer possible.

We will be interested on the Arakelov height of Heegner points with respect to the metrized line bundles $\hat{\omega}_U$ (suitably defined in the next paragraph). The three methods for defining a metrized canonical sheaf are in fact equivalent for this purpose, as we will explain.

13.4. Arakelov height. Let $P$ be an algebraic point in $X = \varprojlim X_U$. For each $U$ we denote by $P_U$ the image of $P$ in $X_U$.

We define the Arakelov-theoretical height of $P$ with respect to the metrized canonical bundles, denoted by $h_{Ar}(P)$, as follows:

Take any finite collection $\mathcal{U} = \{U_j\}_{j=1}^r$ of good enough open compact subgroups such that the numbers $m_j = m_{U_j}$ satisfy $\gcd(m_1,\ldots,m_r) = 1$. Let $U = \cap_{j=1}^r U_j$ and note that $m_U$ has the same prime factors as $m_1 \cdots m_r$, and it is good enough too. Let $F_{P,U}$ be a field containing the residue field of $P_U$ and let $S_{P,U} = \text{Spec} O_{F_{P,U}}$. The point $P_U$ extends to a map $s_U : S_{P,U}[m_U^{-1}] \to \mathcal{X}_U$. Similarly, the points $P_{U_j}$ extend to maps $s_j : S_{P,U}[m_j^{-1}] \to \mathcal{X}_{U_j}$, which are compatible in the sense that

$$s_j|_{S_{P,U}[m_U^{-1}]} = \pi_{U_j}^U \circ s_U.$$

As the $U_j$ are good enough and $\gcd(m_j,m_l) = 1$, from the previous equation and (13.1) we deduce that the line sheaves $s_j^* \omega_{U_j}$ (each on $S_{P,U}[m_j^{-1}]$, respectively) glue together, defining a line sheaf on $S_{P,U}$ which we denote by $\omega_{P,U}$. The metrics at infinity induce a metric on
ω_{P,U} for each \( \sigma : F_{P,U} \to \mathbb{C} \). Thus, we obtain a metrized projective \( O_{F_{P,U}} \)-module \( \mathfrak{M} \) of rank 1, namely, \( \mathfrak{M} = H^0(S_{P,U}, \omega_{P,U}) \) with the induced metrics at infinity, which we denote by \( \| - \|_{\mathfrak{M}, \sigma} \) for each embedding \( \sigma : F_{P,U} \to \mathbb{C} \). This allows us to define \( h_{Ar}(P) \) as the (normalized) Arakelov degree of \( \mathfrak{M} \):

\[
(13.3) \quad h_{Ar}(P) = \deg_{F_{P,U}} \mathfrak{M} / [F_{P,U} : \mathbb{Q}]
\]

or more explicitly

\[
(13.4) \quad h_{Ar}(P) = \frac{1}{[F_{P,U} : \mathbb{Q}]} \left( \log \# (\mathfrak{M} / \langle \eta \rangle) - \sum_{\sigma : F_{P,U} \to \mathbb{C}} \log \| \eta \|_{\mathfrak{M}, \sigma} \right)
\]

for any non-zero \( \eta \in \mathfrak{M} \).

Note that we can always choose the required \( U_j \), for instance, for any \( r \geq 2 \), we can take \( U_j = U_j^D(\ell_j) \) with \( \ell_j \geq 5 \) distinct primes not dividing \( D \). Furthermore, the number \( h_{Ar}(P) \) only depends on \( P \); it is independent of the choice of \( \{ U_j \}_j \), the choice of field \( F_{P,U} \), and the choice of non-zero \( \eta \in \mathfrak{M} \).

For Heegner points, one has the following explicit formula:

**Theorem 13.2.** Let \( K \) be a quadratic imaginary field satisfying the Heegner hypothesis for \( D \). Let \( P_K \) be the associated Heegner point in \( \mathcal{X} = \varprojlim X_U \). Let \( d_K \) be the absolute value of the discriminant of \( K \), and let \( \chi_K \) be the non-trivial primitive Dirichlet character attached to \( K \). Then

\[
(13.5) \quad h_{Ar}(P_K) = - \frac{L'(0, \chi_K)}{L(0, \chi_K)} + \frac{1}{2} \log (d^{-1}_K D).
\]

Here, \( L(s, \chi) = \sum_{n \geq 1} \chi(n)n^{-s} \) (for \( \Re(s) > 1 \)). In the next two paragraphs, we explain how this height formula follows from the existing literature in the case \( D > 1 \). The case \( D = 1 \) will not be used in our work, but we remark that the result is still correct in that case, and it can be deduced directly from the Chowla-Selberg formula.

13.5. **The Chowla-Selberg formula, after Gross, Colmez, Kudla-Rapoport-Yang.** By work of Gross \[35\] and Colmez \[18\], the classical Chowla-Selberg formula can be understood as a formula for the semi-stable Falting’s height of a CM elliptic curve. Kudla, Rapoport and Yang \[52\] \[53\] used this fact to give a height formula for Heegner points on Shimura curves with respect to a suitably defined metrized canonical bundle. Keeping track of normalizations, Theorem 13.2 with \( D > 1 \) can be deduced from the results of Kudla-Rapoport-Yang taking into account that the quantity \( c \) in Section 10 of \[52\] in our case is \( c = 1 \), because we only consider Heegner points subject to the Heegner hypothesis for \( D \).

For the sake of exposition, let us briefly recall the method of proof in \[52\]. They work on \( \mathcal{X}_0^D(1) \), the moduli stack over \( \mathbb{Z} \) associated to \( \mathcal{X}_0^D(1) \). First they show that the metrized canonical sheaf on \( \mathcal{X}_0^D(1) \) (the relative dualizing sheaf with metrics coming from the complex uniformization) can be identified, up to an explicit factor in the metrics at infinity, with the metrized Hodge bundle coming from the universal family of fake elliptic curves on \( \mathcal{X}_0^D(1) \) (cf. Section 3 \[52\]). The Arakelov height of an algebraic point \( P \) on \( \mathcal{X}_0^D(1) \) relative to the metrized Hodge bundle coincides with the Faltings height of the fake elliptic curve \( A_P \) associated to \( P \), thus, the same holds (up to an explicit factor) for the height relative to the
metrized canonical sheaf. On the other hand, when $P$ is a CM point, $A_P$ is isogenous to $E_P^2$ for certain CM elliptic curve $E_P$, thus the Faltings height of $A_P$ equals $2h(E_P)$ up to a factor coming from the isogeny, which is made explicit in Theorem 10.7 [52]. Finally, $h(E_P)$ is expressed in terms of the logarithmic derivative of $L(s, \chi)$ by the classical Chowla-Selberg formula, as mentioned before.

The translation to our setting is possible because, upon adding level structure of a good enough $U$, the corresponding stack is in fact our scheme $\mathcal{X}_U$ over $\mathbb{Z}[m_U^{-1}]$, and the pull-back of the metrized canonical sheaf on $\mathcal{X}_0^D(1)$ coincides with our $\hat{\omega}_U$ up to suitable normalization factors on the metrics.

13.6. **The Yuan-Zhang height formula.** Another (more direct) way to deduce Theorem 13.2 from the existing literature is using a recent general result of Yuan and Zhang, namely, Theorem 1.5 in [93], along with their theory of integral models. The Yuan-Zhang theorem works for quaternionic Shimura curves over totally real fields in general, not just over $\mathbb{Q}$.

Using [93], the translation to our setting with $D > 1$ is almost immediate:

The finite parts of our metrized sheaves $\hat{\omega}_U$ agree with the arithmetic Hodge bundle as defined in [93] when $U$ is good enough, although in loc. cit. the definition is given in more generality (cf. Theorem 4.7 (2) loc. cit. and equation (13.1) above). The metrics at infinity in loc. cit. also agree (cf. Equation (13.2) above and Theorem 4.7 (3) loc. cit.).

The integral model for $X_U$ over $\mathbb{Z}[m_U^{-1}]$ used in [93], for $U$ small enough, is the minimal regular model. It is unique as $X_U$ has genus $g_U \geq 2$. It can be obtained from our $\mathcal{X}_U$ by repeated blow-up and contraction away from the smooth locus $\mathcal{X}_U^0$. By Lemma 13.1 above, if $K$ satisfies the Heegner hypothesis for $D$ (in particular, the discriminant of $K$ is coprime to $D$ as required in [93]) the closure of the associated Heegner point in $\mathcal{X}_0^D(1)$ is contained in $\mathcal{X}_0^D(1)^0$, hence, the closure of $P_{K,U}$ in $\mathcal{X}_U$ is contained in $\mathcal{X}_U^0$. Therefore, this difference on integral models does not affect the height of Heegner points.

At this point, we mention that later, in Section 18, we will directly use the Yuan-Zhang theory of integral models and height formula over totally real fields.

13.7. **Functional equation.** Given an imaginary quadratic field $K$, the primitive quadratic character $\chi_K$ is odd, has conductor $d_K$, and the functional equation for $L(s, \chi_K)$ is given by $\xi(1-s, \chi_K) = \xi(s, \chi_K)$ where

$$\xi(s, \chi_K) = \left(\frac{d_K}{\pi}\right)^{(s+1)/2} \Gamma \left(\frac{s+1}{2}\right) L(s, \chi_K).$$

Thus,

$$-\frac{L'}{L}(1-s, \chi_K) = \log \left(\frac{d_K}{\pi}\right) + \frac{1}{2} \left( \frac{1}{\Gamma'} \left(\frac{s+1}{2}\right) + \frac{1}{\Gamma'} \left(\frac{2-s}{2}\right) \right) + \frac{L'}{L}(s, \chi_K)$$

and we see that an equivalent way to state formula (13.5) is

$$h_{Ar}(P_K) = \frac{L'}{L}(1, \chi_K) + \frac{1}{2} \log \left(\frac{d_K}{\pi}\right) - (\gamma + \log(2\pi))$$

where $\gamma$ is the Euler-Mascheroni constant. This seems more natural from the point of view of analytic number theory because $\frac{L'}{L}(1, \chi_K)$ is expected to be small as $K$ varies. For instance, from [43] one deduces the following:
Proposition 13.3. As $K$ varies over quadratic imaginary fields satisfying the Heegner hypothesis for $D$, if GRH holds for $L(s, \chi_K)$ then we have
\[ h_{Ar}(P_K) = \frac{1}{2} \log(d_K D) + O(\log \log d_K). \]

The implicit constant in the error term can be taken as $2 + \epsilon$ for any $\epsilon > 0$ (for $d_K \gg 1$).

See also [19] for some related applications of analytic number theory to estimate the height of CM abelian varieties. In loc. cit., however, analytic lower bounds for the height are found, while we need analytic upper bounds, which seems to be a more difficult problem from the point of view of $L$-functions.

We will not use Proposition 13.3. For our purposes, we will have some freedom to choose the quadratic imaginary field $K$, and an unconditional estimate for $h_{Ar}(P_K)$ of essentially the same strength can be deduced from Theorem 11.1.

14. Integrality and lower bounds for the $L_2$-norm

14.1. Notation and result. We are now ready to present a key application of our work in the previous sections. As always, $D$ and $M$ stand for coprime positive integers, with $D$ squarefree with an even number of prime factors. Also, we write $N = DM$. In addition, we will assume $D > 1$; in fact, for the results in this section the case $D = 1$ is already known by using methods based on $q$-expansions, which are not available in the present case $D > 1$.

We will be working with open compact subgroups of the form $U = U_0^D(m) \cap U_1^D(m')$ with $m, m'$ coprime, both coprime to $D$, in which case $C(U) = (1)$ so that $X_U^{an}$ is connected. Thus, we can choose $g_\alpha = 1$ in the decomposition $4.1$ and write $\tilde{\Gamma}_U = \tilde{\Gamma}_{U,g_\alpha}$. Similarly, with this choice we may drop the subscript $g$ in the notation of the $L_2$ and supremum norms of Section 8.

For $U$ as above, we write $\xi_U : \mathfrak{h} \to \tilde{\Gamma}_U \backslash \mathfrak{h} = X_U^{an}$ for the complex uniformization. We have an injective map into the space $S_U$ of holomorphic of weight 2 modular forms for $U$
\[ \Psi_U : H^0(X_U, \Omega^1) \to S_U \]
given by the condition that the image of a section $\alpha$ is the modular form $\Psi_U(\alpha) \in S_U$ satisfying that $\xi_U^* \alpha = \Psi_U(\alpha)dz$ with $z$ the complex variable in $\mathfrak{h}$ (cf. Paragraph 4.1).

Recall the normal integral model $\mathcal{Z}_0^D(M) = \mathcal{Z}_0^D(M, U_1^D(1))$, flat and projective over $\mathbb{Z}$, introduced in Paragraph 9.1. We let
\[ \mathcal{I}_0^D(M) = \Psi_{U_0^D(M)}(H^0(\mathcal{Z}_0^D(M)^0, \Omega_{\mathcal{Z}_0^D(M)/\mathbb{Z}}^1)) \subseteq S_0^D(M). \]
Thus, $\mathcal{I}_0^D(M)$ defines a notion of $\mathbb{Z}$-integrality in $S_0^D(M)$.

We note that when $D = 1$ (which we are not considering here), every element of $\mathcal{I}_0^D(M)$ has Fourier expansion (at $i\infty$) with Fourier coefficients in $2\pi i \mathbb{Z}$, but the converse usually fails, see [27] for details.

The main result in this section is:

**Theorem 14.1** (Integral forms are not too small). Given $\epsilon > 0$, if $N \gg \epsilon^1$ and if $N = DM$ is an admissible factorization with $D > 1$, then for every $f \in \mathcal{I}_0^D(M)$ integral non-zero modular form for $U_0^D(M)$ we have
\[ -\log \|f\|_{U_0^D(M), 2} \leq \left( \frac{5}{6} + \epsilon \right) \log N + \frac{1}{2} \log M. \]
We remark that in this result either $D$ or $M$ can be small (or even remain bounded) as long as the product $N = DM$ is sufficiently large in terms of $\epsilon$.

14.2. Lower bound for the height of Heegner points.

**Proposition 14.2.** Let $K$ be a quadratic imaginary field satisfying the Heegner hypothesis for $(D, M)$ and let $P_K \in X(K^{ab})$ be a Heegner point associated to $K$. Let $H$ be a finite extension of $K$ such that $P_{K, U^0_D(M)^U}$ is $H$-rational, and for each $\sigma : H \to \mathbb{C}$ choose $\tau_{K, \sigma} \in \mathfrak{h}$ such that $\xi_{U^0_D(M)}(\tau_{K, \sigma}) = P_{K, U^0_D(M)}^\sigma$.

Let $\alpha \in H^0(\mathcal{X}_0^D(M)^0, \Omega^1_{\mathcal{X}_0^D(M)/\mathbb{Z}})$ be a non-zero section, write $f = \Psi_{U^0_D(M)}(\alpha)$ and suppose that $f$ does not vanish at the points $\tau_{K, \sigma}$. We have

$$h_{A^r}(P_K) \geq -\log(2M) - \frac{1}{[H : \mathbb{Q}] \sum_{\sigma : H \to \mathbb{C}} \log(\|f(\tau_{K, \sigma})\|\mathfrak{F}(\tau_{K, \sigma})).$$

**Proof.** Let $\ell_1, \ell_2 \geq 5$ be two distinct primes not dividing $N = DM$. Let $U_j = U^1_D(\ell_j)$ for $j = 1, 2$, let $U = U_1 \cap U_2$, and let $H'$ be a finite extension of $H$ such that $P_{K, U^0_D(M)^U}$ is $H'$-rational. Note that $H_K \subseteq H'$ because $P_{K, U^0_D(M)^U}$ maps to $P_{K, U^0_D(M)^U}$, and similarly, $H'$ contains the residue field of $P_{K, U}$. In the notation of Paragraph 13.4 we let $\mathfrak{M} = \{U_1, U_2\}$, $P = P_K$ and $F_{P, U} = H'$, obtaining the metrized $O_{H'}$-module $\mathfrak{M}$ of rank 1 with

$$h_{A^r}(P_K) = \frac{1}{[H' : \mathbb{Q}] \deg_{H'} \mathfrak{M}.}$$

We would like to use the section $\alpha$ to construct a non-zero element of $\mathfrak{M}$, so that $h_{A^r}(P_K)$ can be computed as in (13.4). For this we construct a second metrized $O_{H'}$-module $\mathfrak{N}$ where $\alpha$ canonically induces an element, and such that $\mathfrak{M}$ maps to $\mathfrak{N}$ with controlled torsion cokernel (and respecting the metrics).

Let $S_{H'} = \text{Spec} O_{H'}$. Let $s_U : S_{H'}[(\ell_1, \ell_2)^{-1}] \to \mathcal{X}_U$ be the morphism associated to $P_{K, U}$, and for $j = 1, 2$ let $s_j : S_{H'}[\ell_j^{-1}] \to \mathcal{X}_{U_j}$ be the morphism associated to $P_{K, U_j}$. Similarly, we also have morphisms $s_U^j : S_{H'}[(\ell_1, \ell_2)^{-1}] \to \mathcal{X}_0^D(M, U)$ and $s_j^1 : S_{H'}[\ell_j^{-1}] \to \mathcal{X}_0^D(M, U_j)$ induced by the points $P_{K, U^0_D(M)^U}$ and $P_{K, U^0_D(M)^{U_j}}$ $(j = 1, 2)$ respectively. These are compatible in the obvious way with the forgetful maps among the six surfaces.

Let $\omega, \omega_j, \omega', \omega_j'$ be the sheaves of relative differentials for $\mathcal{X}_U/\mathbb{Z}[(\ell_1, \ell_2)^{-1}], \mathcal{X}_{U_j}/\mathbb{Z}[\ell_j^{-1}]$, $\mathcal{X}_0^D(M, U)/\mathbb{Z}[(\ell_1, \ell_2)^{-1}], \mathcal{X}_0^D(M, U_j)/\mathbb{Z}[\ell_j^{-1}]$ (for $j = 1, 2$) respectively. They are not invertible in general, although they are indeed invertible sheaves on the smooth loci of the corresponding structure maps.

The sheaves $s_j^* \omega_j'$ on $S_{H'}[\ell_j^{-1}]$ glue along $s^* \omega'$ to define a sheaf $\omega_{P_{K, U}}'$ on $S_{H'}$ in the same way that the line sheaves $s_j^* \omega_j$ determine $\omega_{P_{K, U}}$ on $S_{H'}$ (cf. Paragraph 13.4). This is because the forgetful maps $\tau_{U^0_D(M)^U}^1 : \mathcal{X}_0^D(M, U) \to \mathcal{X}_0^D(M, U_j)$ are étale (as $U_j$ is sufficiently small), which can be used instead of (13.1) in order to check compatibility. Furthermore, $\omega_{P_{K, U}}'$ and $\omega_{P_{K, U}}$ are invertible sheaves on $S_{H'}$, by Lemma 13.1.

We define \( \mathfrak{N} = H^0(\text{Spec} O_{H'}, \omega_{P_{K, U}}') \), which is a projective $O_{H'}$-module of rank 1, endowed with the metrics coming from (13.2) and the complex uniformization of $X_0^D(M, U)^{an}$. 55
The non-zero global section \( \alpha \in H^0(\mathcal{D}_0^U(M)^0, \Omega^1_{\mathcal{D}_0^U(M)/\mathbb{Z}}) \) induces (via pull-back) compatible sections on
\[
H^0(\mathcal{D}_0^U(M, U)^0, \Omega^1_{\mathcal{D}_0^U(M, U)/\mathbb{Z}[\ell, \ell_2]^{-1}}) = H^0(\mathcal{D}_0^U(M, U)^0, \omega')
\]
and
\[
H^0(\mathcal{D}_0^U(M, U_j)^0, \Omega^1_{\mathcal{D}_0^U(M, U_j)/\mathbb{Z}[\ell_j]^{-1}}) = H^0(\mathcal{D}_0^U(M, U_j)^0, \omega_j')
\]
\((j = 1, 2)\) which determine an element \( \beta \in \mathfrak{N} \).

Since \( P_{K, U_0^D(M)} \) is \( H \)-rational, we have for each \( \sigma : H' \to \mathbb{C} \) and each \( \tau'_{K, \sigma} \in \mathfrak{h} \) mapping to \( P_{K, U_0^D(M) \cap U}^\sigma \)
\[
\pi_{U_0^D(M)}^\sigma(U) (\xi_{U_0^D(M) \cap U}^{\tau'_{K, \sigma}}) = \pi_{U_0^D(M)}^\sigma (P_{K, U_0^D(M) \cap U}^\sigma) = P_{K, U_0^D(M)}^\sigma = \xi_{U_0^D(M) (\tau_{K, \sigma} \mid \mathfrak{h})}.
\]

Since \( f \) is in the image of \( \Psi_{U_0^D(M)} \), we see that condition that \( f \) does not vanish at the points \( \tau_{K, \sigma} \) for \( \sigma : H \to \mathbb{C} \) implies that \( \beta \neq 0 \).

We observe that the construction of \( \mathfrak{M} \) factors through “adding level \( U_0^D(M) \)” in the sense that
\[
s^* \omega_j = s_j^*(\pi_{U_j}^{U_0^D(M) \cap U} (\omega_j)) \]
and similarly for \( s^* \omega \). Furthermore, on the smooth loci we have exact sequences
\[
0 \to (\pi_{U_0^D(M) \cap U}^{U_0^D(M)^0})^* \omega|_{\mathcal{D}_0^D(M, U)^0} \to \omega'|_{\mathcal{D}_0^D(M, U)^0} \to \mathcal{C} \to 0
\]
and
\[
0 \to (\pi_{U_j}^{U_0^D(M) \cap U_j})^* \omega_j|_{\mathcal{D}_0^D(M, U_j)^0} \to \omega_j'|_{\mathcal{D}_0^D(M, U_j)^0} \to \mathcal{C}_j \to 0
\]
with \( \mathcal{C} \), \( \mathcal{C}_j \) annihilated by the integer \( M \), thanks to Corollary [2,22].

By (14.2) and Lemma [13.1], these exact sequences on smooth loci induce a map of \( O_{H'} \)-modules \( \iota : \mathfrak{M} \to \mathfrak{N} \) which is still injective because the maps \( \pi_{U_j}^{U_0^D(M) \cap U_j} \) and \( \pi_{U_0^D(M) \cap U}^{U_0^D(M)^0} \) are \( \text{étale} \), and also by the non-vanishing assumption on \( f \).

Thus, the cokernel of \( \iota \) is annihilated by the integer \( M \). Furthermore, the map \( \iota \) is locally described by pull-back, and the forgetful maps are compatible with the complex uniformization of the relevant curves, so, the map \( \iota \) respects the metrics at infinity. Thus, \( \iota \) induces an inclusion \( \widehat{\mathfrak{M}} \subseteq \widehat{\mathfrak{N}} \) of metrized \( O_{H'} \)-modules, whose cokernel is annihilated by \( M \).

So, \( M \beta \in \mathfrak{M} \) is a non-zero element, and from (13.4) we obtain
\[
h_{Ar}(P_K) = \frac{1}{[H' : \mathbb{Q}]} \left( \log \# (\mathfrak{M} / (M \beta)) - \sum_{\sigma : H' \to \mathbb{C}} \log \| M \beta \|_{\mathfrak{M}, \sigma} \right)
\]
\[
\geq \frac{1}{[H' : \mathbb{Q}]} \sum_{\sigma : H' \to \mathbb{C}} \log \| M \beta \|_{\mathfrak{M}, \sigma}
\]
\[
= - \log(M) - \frac{1}{[H' : \mathbb{Q}]} \sum_{\sigma : H' \to \mathbb{C}} \log \| \beta \|_{\mathfrak{M}, \sigma}.
\]

For each \( \sigma : H' \to \mathbb{C} \) choose \( \tau'_{K, \sigma} \in \mathfrak{h} \) such that \( \xi_{U_0^D(M) \cap U}^{\tau'_{K, \sigma}} = P_{K, U_0^D(M) \cap U}^\sigma \), then we have that the last expression is
\[
(14.3) \quad - \log(M) - \frac{1}{[H' : \mathbb{Q}]} \sum_{\sigma : H' \to \mathbb{C}} \log \left( 2 \cdot |f(\tau'_{K, \sigma})| \Im(\tau'_{K, \sigma}) \right).
\]
Also from (14.1) and the fact that \( f \) is in the image of \( \Psi_{U_0^D(M)} \), it follows that (14.3) is equal to

\[
- \log(2M) - \frac{1}{|H : \mathbb{Q}|} \sum_{\sigma : \mathfrak{h} \to \mathbb{C}} \log(|f(\tau_K,\sigma)|^2) - \Theta(\tau_K,\sigma))
\]

which proves the result.

\[\blacksquare\]

14.3. \textit{L}_2\text{-norm of integral modular forms.}

\textbf{Lemma 14.3.} Let \( P \) be an algebraic point of \( X_0^D(M) \) (non-cuspidal, as \( D > 1 \)), and let \( P_j \) for \( j = 1, \ldots, r \) be the Galois conjugates of \( P \), with \( P_1 = P \), say. Let \( \tau_j \in \mathfrak{h} \) be such that \( \xi_{U_0^D(M)}(\tau_j) = P_j \) for each \( 1 \leq j \leq r \). Let \( \alpha \in H^0(X_0^D(M), \Omega_{X_0^D(M)/\mathbb{Q}}^1) \) and let \( f = \Psi_{U_0^D(M)}(\alpha) \) be the associated modular form. Then \( f \) vanishes at one of the \( \tau_j \) if and only if it vanishes at all of them.

Furthermore, given \( \eta > 0 \), if \( N \gg 1 \) and \( N = DM \), then \( f \) has at most \( N^{1+\eta} \) zeros on \( \mathfrak{h} \) up to \( \tilde{\Gamma}_{U_0^D(M)} \)-equivalence.

\textbf{Proof.} The zeros of \( f \) on \( \mathfrak{h} \) map via \( \xi_{U_0^D(M)} \) to two types of points on \( X_0^D(M) \): points in the support of the divisor of \( \alpha \), and the elliptic points on \( X_0^D(M) \). The first is Galois-stable, while the second can be seen as the branch locus of

\[
\pi_{U_0^D(M)}^* \Omega_{X_0^D(M)}^1 : X_0^D(M, U_1^D(\ell)) \to X_0^D(M)
\]

for any prime \( \ell > 3 \) coprime to \( N = DM \), and this branch locus is also Galois-stable. Hence the first claim.

For the second claim, an upper bound is given by the number of zeros of \( f \) up to \( \tilde{\Gamma}_{U_0^D(M)} \)-equivalence, and this number is at most the degree of the zero divisor of \( \beta = (\pi_{U_0^D(M)}^* \Omega_{X_0^D(M)}^1)^* \alpha \) on \( X_0^D(M, U_1^D(\ell)) \) since the complex uniformization \( \xi_{U_0^D(M)} \) is unramified. The degree of the divisor of \( \beta \) is \( 2g - 2 \), where \( g \) the genus of \( X_0^D(M, U_1^D(\ell)) \).

By standard dimension formulas, the genus satisfies

\[
\log g = \log N + O(\log \log \log N) + O(\log \ell).
\]

Since it is possible to take \( \ell \ll \log N \), we get the result. \[\blacksquare\]

\textbf{Proof of Theorem 14.1.} Fix \( \epsilon > 0 \). Let \( \theta, \delta > 0 \) be defined by \( \delta = \min\{1/4, \epsilon\} \) and \( \theta = \delta/5 \).

Take \( N \gg_\theta, \delta 1 \) with the same implicit constant as in Corollary 11.2 and take \( N = DM \) an admissible factorization with \( D > 1 \). We also require \( N \gg_\delta 1 \) so that the implicit constant is admissible for Lemma 14.3 with \( \eta = \delta/5 \). After these two conditions, note that we are only requiring that \( N \gg_\epsilon 1 \).

By Corollary 11.2 with \( x = N^{\frac{3}{2} + \delta} > N^{\frac{1}{2} + \delta} \), we have

\[
\# S_0(D, M, N^{\frac{3}{2} + \delta}) > N^{(\frac{3}{4} + \delta)(1 - \delta)}.
\]

For each fundamental discriminant \(-d\) with \( d \in S_0(D, M, N^{\frac{3}{2} + \delta}) \) note that \( K_{-d} \) satisfies the Heegner hypothesis for \((D, M)\). Let \( P_{-d} = P_{K_{-d}, U_0^D(M)} \) be the associated Heegner point in \( X_0^D(M) \) and note that the number of Galois conjugates of \( P_{-d} \) is

\[
[H_{K_{-d}} : \mathbb{Q}] = 2[H_{K_{-d}} : K_{-d}] > 2d^{\frac{1}{2} - \theta} > x^{\frac{1}{2} - \theta} = N^{(\frac{3}{4} + \delta)(\frac{1}{2} - \theta)}
\]
because the residue field of $P_{-d}$ over $K$ is the Hilbert class field $H_{K-d}$, and by item (v) in the definition of $S_{\theta}(D, M, x)$ in Theorem [11.1]. Hence, the number of points in the set

$$\{P_{-d}^*: d \in S_{\theta}(D, M, N^{\frac{2}{3}+\delta}), \sigma : H_{K_d} \to \mathbb{C}\}$$

is at least

$$N^{(\frac{2}{3}+\delta)(\frac{1}{2}-\theta)+\frac{(2+\delta)(1-\delta)}{\theta}} = N^{1+(\frac{1}{2}-\delta)-\frac{(2+\delta)}{\theta}} > N^{1+\frac{1}{2}\delta-\theta} > N^{1+\frac{1}{2}\delta}$$

because $\delta < 1/3$ and $\theta < \delta/4$.

By Lemma [14.3] with $\eta = \delta/5$, we deduce from the previous estimate that there is $d_0 \in S_{\theta}(D, M, N^{\frac{2}{3}+\delta})$ such that if we let $K = K_{d_0}$, the Heegner point $P_{K,U_0^D(M)}^\sigma$ satisfies the non-vanishing condition with respect to $f$ required in Proposition [14.2]. Namely, letting $\tau_{K,\sigma} \in \mathfrak{h}$ be a point mapping to $P_{K,U_0^D(M)}^\sigma$ for each $\sigma : H_K \to \mathbb{C}$, we have that $f$ does not vanish at these points.

Thus, Proposition [14.2] gives

$$h_{Ar}(P_K) \geq -\log(2M) - \frac{1}{[H_K : \mathbb{Q}]} \sum_{\sigma : H_K \to \mathbb{C}} \log(|f(\tau_{K,\sigma})| \Im(\tau_{K,\sigma}))$$

$$\geq -\log(2M) - \log \|f\|_{U_0^D(M),\infty}. $$

By Theorem [8.1] (as $D > 1$) applied to $f$ and the compact open subgroup $U_0^D(M) \cap U_1^D(\ell)$ for some prime $5 \leq \ell \leq 2 \log N$ coprime to $N$, we deduce that

$$\log \|f\|_{U_0^D(M),\infty} = \log \|f\|_{U_0^D(M) \cap U_1^D(\ell),\infty}$$

$$\leq \log \|f\|_{U_0^D(M) \cap U_1^D(\ell),2} + O(1)$$

$$= \log \|f\|_{U_0^D(M),2} + O(\log \log N)$$

with absolute implicit constants. Hence,

$$-\log \|f\|_{U_0^D(M),2} \leq \log M + h_{Ar}(P_K) + O(\log \log N)$$

with an absolute implicit constant.

Recall that $K$ has discriminant $-d_0$ with $d_0 \in S_{\theta}(D, M, N^{\frac{2}{3}+\delta})$. By Theorem [13.2] in its formulation [13.6], we obtain that for some uniform constant $\kappa$ as in Theorem [11.1] (where $S_{\theta}(D, M, x)$ was defined)

$$-\log \|f\|_{U_0^D(M),2} \leq \log M + \frac{L'}{L}(1, \chi_K) + \frac{1}{2} \log(d_0D) + O(\log \log N)$$

$$\leq \log M + \kappa \log \log d_0 + \frac{1}{2} \log(d_0D) + O(\log \log N).$$

As $d_0 \in S_{\theta}(D, M, N^{\frac{2}{3}+\delta})$ we have $d_0 \leq 2N^{\frac{2}{3}+\delta}$, which gives

$$-\log \|f\|_{U_0^D(M),2} \leq \log M + \frac{1}{2} \log D + \left(\frac{1}{3} + \frac{\delta}{2}\right) \log N + O(\log \log N)$$

$$= \frac{1}{2} \log M + \left(\frac{5}{6} + \frac{\delta}{2}\right) \log N + O(\log \log N)$$

with an absolute implicit constant, provided that $N \gg 1$. Since $\delta \leq \epsilon$, the result follows. □
15. Linear forms in logarithms

15.1. An application of linear forms in logarithms. As a preparation for the proof of Theorem 1.1 we prove the following simple consequence of the theory of linear forms in $p$-adic logarithms.

**Proposition 15.1.** Let $\epsilon > 0$. There is a constant $C_\epsilon > 1$ such that for all triples $a, b, c$ of coprime positive integers with $a + b = c$, we have

\[
\frac{d(abc)}{\log d(abc)} < C_\epsilon^\nu \nu^{2\nu^2} \text{rad}(abc)^{1+\epsilon \nu}
\]

where $\nu = \omega(abc)$ is the number of distinct prime divisors of $abc$. In particular, if we consider $\nu$ as fixed, then for those triples $a, b, c$ we have

\[
d(abc) \ll \nu \epsilon \nu \text{rad}(abc)^{1+\epsilon}.
\]

The reader will note that we are actually aiming for a bound of the following sort for $abc$-triples:

\[
(15.1) \quad d(abc) \ll \text{rad}(abc)^\kappa
\]

for some fixed $\kappa$. Proposition 15.1 falls short of proving such an estimate because it only applies for fixed (or bounded) value of $\omega(abc)$, but nevertheless, it will be useful to get an improved value of $\kappa$. Namely, Proposition 15.1 will be used for $abc$-triples with at most $\nu$ prime factors (for some suitable choice of $\nu$), while the theory developed in previous sections of this paper will be used on the general case of $abc$ triples with more than $\nu$ prime factors. We remark that for our purposes, the previous proposition is not really necessary and one can show a version of Theorem 1.1 (i.e. (15.1) with a slightly worse exponent $\kappa$) without it.

**Proof.** Consider coprime positive integers $a, b, c$ with $a + b = c$. Let $\nu_a = \omega(a)$ be the number of distinct prime factors of $a$, and let $L_a = \prod_{p|a}(1 + \log p)$. We make analogous definitions for $b$ and $c$. If $p$ is a prime dividing $a$, we note that

\[
v_p(abc) = v_p(a) = v_p\left(\frac{b}{c} - 1\right).
\]

The latter quantity is well-suited for the theory of linear forms in $p$-adic logarithms. For instance, the corollary in p. 245 in [90] gives

\[
(15.2) \quad v_p(abc) \ll (\nu_b + \nu_c)^{3(\nu_b + \nu_c)} \cdot p \cdot L_b L_c \log(L_b L_c) \log d(bc)
\]

where we used the inequality $d(m) \geq \max_{q|m} v_q(m) + 1$ ($q$ varying over primes). Multiplying the analogous bounds for all $p|abc$ we get

\[
d(abc) = \prod_{p|abc} (v_p(abc) + 1)
\]

\[
\ll \nu^{6(\nu_b + \nu_c)} \text{rad}(abc)(L_a L_b L_c)^{2\nu} (\log d(abc))^\nu
\]

\[
\leq \nu^{2\nu^2} \text{rad}(abc) \cdot \prod_{p|abc} (1 + \log p)^{2\nu} (\log d(abc))^\nu.
\]

The result now follows from

\[
\prod_{p|m} (1 + \log p) \ll \epsilon \text{rad}(m)^\epsilon
\]
which has an implicit constant that only depends on $\epsilon$. Note that we will take $\nu$-th power of the previous estimate, which explains the term $C'_{\epsilon}^\nu$ in the final result.

We also the following estimate for $v_2(abc)$, which also relies on the theory of linear forms in logarithms.

**Lemma 15.2.** Let $\epsilon > 0$. There is a constant $C'_{\epsilon} > 1$ such that for all triples $a, b, c$ of coprime positive integers with $a + b = c$ we have

$$v_2(abc) < C'_{\epsilon} \cdot \text{rad}(abc)^\epsilon.$$

**Proof.** For $abc$ triples with $\omega(abc) = 2$ the result is a consequence of Mihailescu’s solution to Catalan’s conjecture [61] (this particular case could be addressed by linear forms in logarithms too, cf. [84]).

So we may assume $\omega(abc) \geq 3$. Then the prime $p = 2$ satisfies condition (16) in [79], namely, $2 < \exp((\log \omega(abc))^2)$. Noticing that for $abc$ triples we know [77]

$$\log \log(abc) \leq \log(\text{rad}(abc)^{15}) + O(1) \ll \log \text{rad}(abc),$$

our claim follows from (21), (22) and (23) in [79].

15.2. **Heuristics on the applicability of linear forms in logarithms.** The bound (15.2) is not the sharpest available result in the literature, and it was used because it is simple to state and enough for our purposes. Can one get (15.1) by using instead the best available bounds on linear forms in $p$-adic logarithms? We feel skeptical about this, although strictly speaking we don’t have a proof that this is not possible. Nevertheless, here is a heuristic justification:

To the best of the author’s knowledge, the sharpest improvements to (15.2) are due to Yu [91], see in particular the quantities $C_1$ and $C_2$ given in p.189 loc. cit. For coprime positive integers $a, b, c$ with $a + b = c$, let us write $\nu_a = \omega(a)$, $\Lambda_a = \prod_{p|a} \log p$ and similarly for $b$ and $c$. Combining the best aspects of these two quantities $C_1$ and $C_2$, and optimistically ignoring a few factors (which can possibly be problematic!), the main theorem in p. 190 [91] points in the direction that the methods can at best give a bound of the form

$$v_p(abc) = v_p(b/c - 1) \ll (\epsilon \nu_a \nu_b \nu_c) \cdot \frac{p}{(\log p)^{\nu_a + \nu_b + \nu_c}} \cdot \Lambda_a \cdot \Lambda_b \cdot \Lambda_c$$

for $p|a$, with $\alpha > 0$ some absolute constant.

For the sake of clarity, note that the factor $e^{\alpha n}$ (as opposed to $n^{\alpha n}$) is the best aspect suggested by $C_2$, while the denominator on the right hand side of (15.3) is the best aspect suggested by $C_1$, but as far as we know, these two improvements are not available simultaneously.

Nevertheless, let us examine the strength of the hypothetical bound (15.3). Multiplying as $p$ varies and using the analogous bounds for $b$ and $c$, one would get in this optimistic scenario that

$$\prod_{p|abc} v_p(abc) < e^{2\alpha (\nu_a \nu_b + \nu_a \nu_c + \nu_b \nu_c)} \frac{\text{rad}(abc)}{\Lambda_a^{\nu_b + \nu_c + 2} \Lambda_b^{\nu_a + \nu_c + 2} \Lambda_c^{\nu_a + \nu_b + 2}} \times (\Lambda_a \Lambda_b)^{\nu_c} (\Lambda_a \Lambda_c)^{\nu_b} (\Lambda_b \Lambda_c)^{\nu_a}$$

$$= e^{2\alpha (\nu_a \nu_b + \nu_a \nu_c + \nu_b \nu_c)} \frac{\text{rad}(abc)}{(\Lambda_a \Lambda_b \Lambda_c)^2}.$$
This bound is not better than
\[ \prod_{p|abc} v_p(abc) < e^{2\alpha(n_a n_b + n_a n_c + n_b n_c)} \text{rad}(abc)^{1/2}. \]

When \(a, b, c\) have a comparable number of prime factors (which \textit{a priori} is a possible scenario), this would only give
\[ (15.4) \quad \prod_{p|abc} v_p(abc) < e^{\beta \omega(abc)^2} \text{rad}(abc)^{1/2} \]
for some constant \(\beta\). From here it is not clear to the author how to get (15.1), because
\[ \omega(N) = \Omega \left( \frac{\log \text{rad}N}{\log \log \text{rad}N} \right) \]
and in fact, it would already be a problem if for some fixed \(\delta > 0\) one has
\[ \omega(abc) > (\log \text{rad}(abc))^{\frac{1}{2} + \delta}. \]

In order to make this heuristic more precise, let us describe a \textit{hypothetical} type of \(abc\) triples, whose existence would be consistent with (15.4) and with any bound of the type
\[ (15.5) \quad \log c \ll \text{rad}(abc)^{\kappa} \]
with \(\kappa\) fixed (as the ones obtained by Stewart, Tijdeman, and Yu \cite{77, 78, 79}), yet it would contradict any bound of the form (15.1). To simplify notation, write \(\nu = \omega(abc)\), \(M = abc\) and \(R = \text{rad}(abc)\). We require the following:
(i) \(\nu\) is large,
(ii) the prime factors of \(M\) are among the first \(2\nu\) primes,
(iii) all the primes \(p|M\) satisfy
\[ v_p(M) < \exp \left( \frac{\beta \log R}{2 \log \log R} \right) \]
with \(\beta > 0\) as in (15.4), and
(iv) a positive proportion of the primes \(p|M\), say at least half of them, satisfy
\[ v_p(M) > \exp \left( (\log \log R)^2 \right). \]

Note that by (i) and (ii)
\[ (15.6) \quad \frac{\log R}{\log \log R} < 2\nu. \]
Then by (iii) and (15.6) we have
\[ \log c < \log M \leq \left( \max_{p|M} v_p(M) \right) \log R \leq \exp \left( \frac{\beta \log R}{\log \log R} \right) \ll c R^\epsilon \]
which is consistent with (15.5). Also, observe that by (iii) and (15.6) we have
\[ \prod_{p|M} v_p(M) \leq \left( \max_{p|M} v_p(M) \right)^\nu \leq \exp \left( \frac{\beta \nu \log R}{2 \log \log R} \right) \leq e^{\beta \nu^2} \]
which is consistent with (15.4), even without the factor \( \text{rad}(abc)^{1/2} \). Finally, note that by (iv) and (15.6) we have

\[
\prod_{p|D} v_p(M) \geq \exp \left( \frac{\eta}{2} (\log_2 R)^2 \right) \geq \exp \left( \frac{1}{4} (\log R) \log_2 R \right) = \text{Rad}(\log_2 R)/4
\]

(here, \( \log_2 X = \log \log X \)) which is not consistent with (15.1), for any value of \( \kappa \).

Of course, such hypothetical \( abc \)-triples satisfying (i), (ii), (iii), and (iv) do not exist because our Theorem 11 actually proves a version of (15.1). Nevertheless, this heuristic analysis suggests (to the author) that the approach of \( p \)-adic linear forms in logarithms has a limitation in this direction. In any case, regardless of the applicability of the theory of linear forms in logarithms in the context of (15.1), it is worth noticing that the theory in this paper provides a completely new approach to establishing \( abc \)-type bounds.

16. Bounds for the product of exponents of the minimal discriminant

16.1. A general estimate for elliptic curves.

Theorem 16.1. Let \( S \) be a finite set of primes and let \( \epsilon > 0 \). For all but finitely many elliptic curves \( E/\mathbb{Q} \) semi-stable away from \( S \), the following holds:

Let \( N = N_E \) be the conductor of \( E \) and let \( \Delta_E \) be the minimal discriminant of \( E \). Consider an admissible factorization \( N = DM \) (in particular, \( D \) is supported away from \( S \)). Then

\[
\prod_{p|D} v_p(\Delta_E) < N^{\frac{8}{3} + \epsilon} M.
\]

Proof. We may assume \( D > 1 \). With the notation as in Theorem 6.1, we have

\[
(16.1) \quad \log \left( \prod_{p|D} v_p(\Delta_E) \right) = \log \delta_{1,N} - \log \delta_{D,M} + O_S \left( \frac{\log D}{\log \log D} \right).
\]

Here, we recall that \( \delta_{D,M} \) is the modular degree of the optimal quotient \( q_{D,M} : J_0^D(M) \to A_{D,M} \) with \( A_{D,M} \) an elliptic curve isogenous to \( E \) over \( \mathbb{Q} \).

Using the cusp \( i\infty \) we have an embedding \( j_N : X_0(N) \to J_0(N) \) so that the composite map \( \phi_N : X_0(N) \to A_{1,N} \) is a classical optimal modular parameterization and has degree exactly \( \delta_{1,N} \). By Frey’s formula (3.2) we get

\[
(16.2) \quad \log \delta_{1,N} = 2 \log(2\pi |c_f|) + 2 \log(\|f\|_{L^2_0(N)}) + 2h(A_{1,N})
\]

where \( c_f \) is the (positive) Manin constant of the modular parameterization \( \phi_N \) and \( f \in S_2(N) \) is the normalized cuspidal Hecke newform associated to \( E \).

Let \( \ell \) be a prime not dividing \( N \). From Proposition 5.1, we recall that there is a suitable morphism \( j_{D,M,\ell} : X_0^D(M) \to J_0^D(M) \) defined over \( \mathbb{Q} \) such that the composite map \( \phi_{D,M,\ell} = q_{D,M,\ell} j_{D,M,\ell} : X_0^D(M) \to A_{D,M} \) is non-constant of degree \( (\ell + 1 - a_\ell(E))^2 \delta_{D,M} \). Let \( \mathcal{C}_{D,M} \) be the Néron model of \( A_{D,M} \) over \( \mathbb{Z} \), then by the Néron mapping property \( \phi_{D,M,\ell} \) extends to a \( \mathbb{Z} \)-morphism of integral models on the smooth locus

\[
\phi_{D,M,\ell} : \mathcal{C}_{0}^D(M)^0 \to \mathcal{C}_{D,M}.
\]

Let \( \omega_{A_{D,M}} \) be a Néron differential of \( A_{D,M} \) (unique up to sign) and let

\[
\alpha_{D,M} = \phi_{D,M,\ell}^* \omega_{A_{D,M}} \in H^0(\mathcal{C}_{0}^D(M)^0, \Omega^1_{\mathcal{C}_{0}^D(M)/\mathbb{Z}})
\]
i.e. the image of $\phi_{D,M}^* \omega_{A_{D,M}}$ under $\phi_{D,M}^* \Omega_{D,M/Z}^1 \rightarrow \Omega_{J_0^D(M)/Z}^1$. Then

$$f_{D,M} = \Psi_{U_0^D(M)}(\alpha_{D,M}) \in \mathcal{S}_0^D(M)$$

is integral and by the same argument as the proof of (3.2) we get

$$(16.3) \log \left( (\ell + 1 - a_\ell(E))^2 \delta_{D,M} \right) = \log \deg \phi_{D,M,\ell} = 2 \log(\|f_{D,M}\|_{U_0^D(M),2}) + 2h(A_{D,M}).$$

By (16.1), (16.2), and (16.3) we deduce

$$(16.4) \log \left( \prod_{p|D} v_p(\Delta_E) \right) = 2 \log |c_f| + 2 \log(\|f\|_{U_0^D(N),2}) - 2 \log(\|f_{D,M}\|_{U_0^D(M),2})$$

$$+ 2h(A_{1,N}) - 2h(A_{D,M})$$

$$+ 2 \log(\ell + 1 - a_\ell(E)) + O_S \left( \frac{\log D}{\log \log D} \right).$$

The auxiliary prime $\ell$ can be chosen $\ell \leq 2 \log N$. Also, since $A_{1,N}$ and $A_{D,M}$ are both isogenous to $E$, they are connected by an isogeny of degree $\leq 163$, so that $2|h(A_{1,N}) - h(A_{D,M})| \leq \log 163$. It follows that

$$(16.4) \log \left( \prod_{p|D} v_p(\Delta_E) \right) = 2 \log |c_f| + 2 \log(\|f\|_{U_0^D(N),2}) - 2 \log(\|f_{D,M}\|_{U_0^D(M),2})$$

$$+ O_S \left( \log \log N + \frac{\log D}{\log \log D} \right).$$

Since $f$ is a normalized newform for $\Gamma_0(N)$, from [58, 62] we get $\|f\|^2_{U_0^D(N),2} \ll N \log N$ which gives

$$2 \log(\|f\|_{U_0^D(N),2}) \leq \log N + O(\log \log N).$$

On the other hand, since $f_{D,M} \in \mathcal{S}_0^D(M)$ is integral and non-zero, by Theorem 14.1 we obtain

$$-2 \log(\|f_{D,M}\|_{U_0^D(M),2}) \leq \left( \frac{5}{3} + \frac{\epsilon}{2} \right) \log N + \log M$$

provided that $N \gg \epsilon 1$. Therefore, for $N \gg \epsilon S 1$, we obtain

$$\log \left( \prod_{p|D} v_p(\Delta_E) \right) \leq 2 \log |c_f| + \left( \frac{8}{3} + \epsilon \right) \log N + \log M.$$

Finally, by Corollary 10.2 we see that $\log |c_f| \ll S 1$, hence the result. \qed

Let us observe that if one only works in the semi-stable case (as in Theorem 16.2 below) then we don’t need Corollary 10.2 in the previous proof. The existing literature on the Manin constant in the semi-stable case would suffice.

We also remark that a computation in Section 2.2.1 of [67] also combines the classical Ribet-Takahashi formula with Frey’s equation (3.2), in a way somewhat similar to what we did to derive equation (16.4) in the previous argument. However, the computation in [67] occurs in a different context ($p$-integrality of ratios of Petersson norms) and in a less precise form, omitting the contribution of Eisenstein primes and requiring semi-stability. Of course,
for our purposes it is crucial to have control on each prime, since we aim to prove inequalities. Frey curves always have \( p = 2 \) as an Eisenstein prime, so we cannot ignore this issue.

16.2. The semistable case.

**Theorem 16.2.** Given \( \epsilon > 0 \), there is a constant \( K_\epsilon > 0 \) such that for every semi-stable elliptic curve \( E \) over \( \mathbb{Q} \) with conductor \( N_E \) and minimal discriminant \( \Delta_E \), one has

\[
\prod_{p|N_E} v_p(\Delta_E) < K_\epsilon \cdot N_E^{4.5+\epsilon}.
\]

**Proof.** If \( N_E = p \) is prime then \( v_p(\Delta_E) \leq 5 \) (cf. [74]) which implies the result. For the remaining cases, we will use Theorem 16.1.

If \( N_E \) is the product of an even number of primes, we can take \( D = N_E \) and \( M = 1 \) in Theorem 16.1 obtaining

\[
\prod_{p|N_E} v_p(\Delta_E) \ll \epsilon N_E^{8/3+\epsilon}.
\]

Finally, if \( N_E = p_1 p_2 \cdots p_n \) is the product of \( n \) primes with \( n \geq 3 \) odd, we apply Theorem 16.1 with \( D = N_E/p_j \) and \( M = p_j \) as \( j \) varies, then we multiply the resulting bounds and take \((n-1)\)-st roots. This gives

\[
\prod_{p|N_E} v_p(\Delta_E) \ll \epsilon \left( N_E^{8/3+\epsilon} N_E \right)^{1/(n-1)} < N_E^{9/2+2\epsilon}.
\]

\( \square \)

16.3. Proof of the \( d(abc) \) Theorem.

**Theorem 16.3.** Given \( \epsilon > 0 \), there is a constant \( K_\epsilon > 0 \) such that the following holds:

Given coprime positive integers \( a, b, c \) with \( a + b = c \), we have

\[
\langle d(abc) \rangle < K_\epsilon \cdot \text{rad}(abc)^{8/3+\epsilon}.
\]

**Proof.** Let \( a, b, c \) be as in the statement, and recall that for a positive integer \( n \) the number of positive divisors of \( n \) is \( d(n) = \prod_{p|n} (v_p(n) + 1) \).

Let \( \nu \geq 2. \) If \( \omega(abc) \leq \nu \) we obtain \( d(abc) \ll_\nu \text{rad}(abc)^2 \) (say) by Proposition 15.1. So we can assume that \( n := \omega(abc) > \nu. \)

Let \( E \) be the Frey curve with affine equation \( y^2 = x(x-a)(x+b) \) associated to the triple \( a, b, c \). Let \( \Delta \) be its minimal discriminant and \( N \) its conductor. We note that for \( p > 2 \), the conditions \( p|abc \) and \( p|N \) are equivalent, and moreover for such a prime \( p \) we have

\[
v_p(N) = 1 \text{ and } v_p(\Delta) = 2v_p(abc).
\]

For \( p = 2 \), however, we always have \( 2|abc \), while 2 does not need to divide \( N \), and when it does, its exponent can be larger than 1. Nevertheless, we still have \( N \approx \text{rad}(abc) \), and by Lemma 15.2 we know

\[
(16.5) \quad v_2(abc) \ll_\epsilon \text{rad}(abc)^\epsilon.
\]

Write \( N = 2^r p_1 \cdots p_m \) where \( 0 \leq r \leq 8 \), \( p_i > 2 \) are distinct primes, and \( m = n \) if \( r = 0 \), or \( m = n - 1 \) if \( 1 \leq r \leq 8 \). In either case, \( m \geq \nu. \)
If \( m \) is even, we let \( D = p_1 \cdots p_m \) and \( M = 2^r \). With these choices, Theorem 16.1 gives that
\[
\prod_{p|abc \atop p \neq 2} (v_p(abc) + 1) \leq \prod_{j=1}^{m} (2v_{p_j}(abc)) = \prod_{j=1}^{m} v_{p_j}(\Delta_E) \ll \epsilon N^{\frac{m}{2}+\epsilon}.
\]
This, together with (16.5) gives
\[
d(abc) \ll \epsilon N^{\frac{m}{2}+\epsilon}
\]
proving the result in the case that \( m \) is even.

If \( m \) is odd, for any choice of \( 1 \leq j_0 \leq m \) we let \( D = p_1 \cdots p_m/p_{j_0} \) and \( M = 2^r p_{j_0} \) and similarly we obtain
\[
\prod_{1 \leq j \leq m \atop j \neq j_0} (2v_{p_j}(abc)) \ll \epsilon N^{\frac{m}{2}+\epsilon} p_{j_0}.
\]
Varying over \( j_0 \), multiplying these \( m \) inequalities, and taking \((m-1)\)-st roots, we obtain
\[
\prod_{p|abc \atop p \neq 2} (v_p(abc) + 1) \leq \prod_{j=1}^{m} (2v_{p_j}(abc)) \ll \epsilon N^{\frac{m}{2}+\epsilon} N^{\frac{1}{m-1}}.
\]
The result now follows from (16.5) and the fact that \( m \geq \nu \), where \( \nu \) can be taken in advance as large as needed.

17. Counting quadratic extensions of a totally real number field

In this section we give suitable analogues of the results in Section 11 in the setting of totally real number fields.

In this section \( n \) will denote the degree of the number field under consideration, and we will write \( c_1(n), c_2(n), ... \) for certain strictly positive quantities that only depend on the integer \( n \); some \( c_i(n) \) will need to be large, while others will be needed small.

17.1. Auxiliary quadratic extensions.

Lemma 17.1. Let \( F \) be a totally real number field with \([F : \mathbb{Q}] = n\) and discriminant \( d_F \).
Let \( \mathfrak{I} \) and \( \mathfrak{S} \) be non-zero coprime ideals of \( \mathcal{O}_F \). There is a totally real quadratic extension \( L/F \) satisfying that each prime \( p \) dividing \( \mathfrak{I} \) is inert in \( \mathcal{O}_L \), each prime \( p \) dividing \( \mathfrak{S} \) is split in \( L \), and with
\[
\text{Norm}(\text{Disc}(L/F)) < c_1(n) \cdot (d_F \text{Norm}(\mathfrak{I} \mathfrak{S}))^{c_2(n)}.
\]

Proof. The construction of \( L \) without the condition on the size of \( \text{Norm}(\text{Disc}(L/K)) \) is classical: one chooses a monic quadratic polynomial \( f(x) = x^2 + \alpha x + \beta \in \mathcal{O}_F[x] \) which is irreducible modulo \( p \) for each \( p|\mathfrak{I} \), is the product of two distinct linear factors modulo \( p \) for each \( p|\mathfrak{S} \), and with positive discriminant under each real embedding. (These are simply congruence conditions on the coefficients \( \alpha, \beta \).) Then any root of \( f(x) \) generates a field \( L \) over \( F \) with the desired properties.

Using geometry of numbers (or more elementary arguments) one can control the size under each real embedding of the coefficients \( \alpha, \beta \) of \( f \) in the previous construction. The discriminant of \( f(x) \) is divisible by the relative discriminant \( L/F \), and the result follows. \( \square \)
Given a number field $F$ and quadratic extensions $M_1, M_2$ of $F$, we define the quadratic extension $M_3 = M_1 \ast M_2$ of $F$ as follows: If $M_1 = M_2$ we let $M_3 = M_1$. Otherwise, the compositum $M_1M_2$ has degree 4 over $F$ and contains three quadratic extensions of $F$, namely, $M_1, M_2$ and the third one is defined to be $M_3 = M_1 \ast M_2$.

**Lemma 17.2.** Let $F$ be a totally real number field, let $\mathfrak{I}$ and $\mathfrak{S}$ be non-zero coprime ideals in $O_F$ and let $L/F$ be a totally real quadratic extension such that every prime dividing $\mathfrak{I}$ is inert in $O_L$ and every prime dividing $\mathfrak{S}$ splits in $O_L$. Let $K/Q$ be an imaginary quadratic extension such that every rational prime below $3\mathfrak{I}\mathfrak{S}$ splits in $K$. Then the field $M = L \ast (FK)$ satisfies the following:

- $M/F$ is quadratic
- $M$ is totally imaginary
- every prime of $O_F$ dividing $\mathfrak{I}$ is inert in $O_M$
- every prime of $O_F$ dividing $\mathfrak{S}$ splits in $O_M$.

**Proof.** This follows by observing that any two of $L, FK, M$ generate the compositum $LK$. \(\square\)

We will also need the next lemma, which follows from Frobenius reciprocity.

**Lemma 17.3.** Let $F_1 \supseteq F_2 \supseteq F_3$ be a tower of number fields with absolute Galois groups $G_1, G_2, G_3$ respectively. Let $\rho_1$ be a finite dimensional complex Galois representation on $G_1$. Suppose that $\text{Ind}_{G_2}^{G_1}\rho_1$ is the restriction to $G_2$ of a Galois representation on $G_3$. Then $\rho_1$ is a direct summand of the restriction to $G_1$ of a Galois representation on $G_3$.

17.2. The counting result. For simplicity (and to avoid ambiguity with the use of the word “solvable”), we call an extension $F/L$ tower of cyclic if there is a tower of field extensions $L = L_0 \subseteq L_1 \subseteq ... \subseteq L_m = F$ such that each $L_i/L_{i-1}$ is cyclic. In particular, $F/L$ is not required to be Galois.

**Theorem 17.4.** Let $F$ be a totally real number field with $[F : Q] = n$ and discriminant $d_F$. Let $\mathfrak{I}$ and $\mathfrak{S}$ be non-zero coprime ideals of $O_F$ and let $x > c_3(n) \cdot (d_F \cdot \text{Norm}(\mathfrak{I}\mathfrak{S}))^{c_4(n)}$. Suppose that either $F/Q$ is tower of cyclic, or that GRH holds for the $L$-functions of quadratic Hecke characters over $F$.

The number of (pairwise non-isomorphic) quadratic extensions $M/F$ satisfying

(i) $M$ is totally imaginary,
(ii) each prime dividing $\mathfrak{I}$ is inert in $O_M$,
(iii) each prime dividing $\mathfrak{S}$ splits in $O_M$,
(iv) $x < \text{Norm}(\text{Disc}(M/F)) < c_5(n)x$, and
(v) $\left| \frac{L'}{L}(1, \chi_M) \right| \leq c_6(n) (\log d_F + \log \log x)$ where $\chi_M$ is the quadratic Hecke character over $F$ associated to the extension $M/F$

is at least $x^{c_7(n)}$.

**Proof.** We construct the desired fields $M$ as $M = L \ast K$ with $L$ as in Lemma [17.1] (whose discriminant has controlled size) and $K/Q$ as in Lemma [17.2]. Then conditions (i), (ii) and (iii) are satisfied.

Using Corollary [11.2] we can produce enough quadratic extensions $K/Q$ so that now (iv) is also satisfied. Thus, it only remains to show that upon discarding a negligible number of quadratic extensions $K/Q$, we can also achieve (v).
First, under the assumption of GRH, it turns out that (v) always holds by results in [43, 44], which concludes the proof in that case. From now on we assume that $F/Q$ is tower of cyclic and we continue the proof unconditionally.

Given $L/F$ and $K/Q$ quadratic extensions as above, let $\xi$ and $\psi$ be the corresponding quadratic Hecke characters over $F$ and $Q$ respectively. Then for $M = L^*(FK)$, the quadratic Hecke character over $F$ is

$$\chi_M = \xi \cdot \psi_F$$

where $\psi_F = \psi \circ \text{Norm}_{F/Q}$ is the base change of $\psi$ from $Q$ to $F$.

By automorphic induction in the solvable case (cf. [4]; see [16] for an exposition), we have that $\xi$ induces an automorphic representation on $GL[F:Q]$ over $Q$, which we denote by $\pi_{\xi}$. This automorphic representation is induced from cuspidal. Observe that upon including $c_8(n)$ more prime factors in the ideals $\mathfrak{I}$ and $\mathfrak{G}$ with norms bounded by $(d_F \text{Norm}(\mathfrak{IS}))^{o(n)}$, we can ensure that $\pi_{\xi}$ is in fact cuspidal. Indeed, the splitting behavior of those primes in the extension $L/F$ determines enough values of $\xi$ as to ensure that the Galois representation $\rho_{\xi}$ attached to $\xi$ is not a direct summand of the restriction of a Galois representation coming from a field properly contained in $F$. We conclude that $\pi_{\xi}$ is cuspidal by Lemma 17.3 and the fact that automorphic induction corresponds to induction of Galois representations when there is a complex Galois representation attached to the corresponding automorphic forms.

From [46] we have the following equality of $L$-functions

$$L(s, \chi_M) = L(s, \xi \otimes \psi_F) = L(s, \pi_{\xi} \otimes \psi).$$

Finally, the bound (v) is proved for all but a negligible number of choices of $K$ exactly as in [54]. That is, proving a version of Proposition 2.3 loc. cit. with $k = 1$ for $L(s, \chi_M)$, and then proceeding as in Corollary 2.5 loc. cit., using the zero density estimate in Corollary 1.4 of [66] for $L(s, \pi_{\xi} \otimes \psi)$ as $\psi$ varies, instead of Heath-Brown’s result from [10]. The additional term $\log d_F$ in (v) comes from [88] Lemma 1.11(b). □

18. A modular approach to Szpiro’s conjecture over number fields

In this section we introduce some tools that make it possible to approach Szpiro’s conjecture over totally real fields using modular parameterizations coming from Shimura curves.

As in the previous section, $n$ will denote the degree of a number field, and we will continue to write $c_i(n)$ for strictly positive quantities that only depend on $n$. The numeration will be consistent with that of the previous section, although this is only intended to stress the fact that these “constants” (depending only on $n$) can change from line to line.

Several of the necessary ideas have been already presented in detail over $Q$ in previous sections of this paper, so, here the discussion will be more concise.

18.1. Faltings height of elliptic curves over number fields. In this paragraph we discuss the Faltings height of elliptic curves in more generality than in Section 3.

Let $L$ be a number field and let $E$ be an elliptic curve over $L$. Let $E_L$ be the Néron model of $E$ over $O_L$. For each embedding $\sigma : L \to \mathbb{C}$ we let $E_\sigma = E \otimes_\sigma \mathbb{C}$ be the complexification of $E$ under the embedding $\sigma$. We define the metric $\| - \|_\sigma$ on $H^0(E_\sigma, \Omega^1_{E_\sigma/\mathbb{C}})$ by

$$\|\alpha\|^2_\sigma := \frac{i}{2} \int_{E_\sigma} \langle \alpha \wedge \overline{\alpha} \rangle.$$

This metric is well-defined due to the following proposition.

Proposition 18.1. Let $\sigma, \tau : L \to \mathbb{C}$ be embeddings. Then

$$\|\alpha\|^2_\sigma = \|\alpha\|^2_\tau + (\text{res}_\tau \text{deg}_\sigma) \langle \alpha \rangle_{\sigma/\tau}.$$
The Faltings height of $E$ over $L$, denoted by $h(E)$, is defined as the normalized Arakelov degree of the metrized rank 1 projective module $H^0(\mathcal{E}, \Omega^1_{\mathcal{E}/O_L})$ with the previous metrics at infinity. Namely, taking any non-zero $\beta \in H^0(\mathcal{E}, \Omega^1_{\mathcal{E}/O_L})$ one defines

$$h(E) = \frac{1}{[L : \mathbb{Q}]} \left( \log \#(H^0(\mathcal{E}, \Omega^1_{\mathcal{E}/O_L})/O_L\beta) - \sum_{\sigma : L \to \mathbb{C}} \log \|\beta\|_\sigma \right).$$

Note that this is not the stable Faltings height, and, in general, it can change after enlarging $L$. However, when $A$ is semi-stable over $L$ then $h(A)$ is invariant under base change to a finite extension of $L$.

Silverman [75] proved an alternative formula for $h(E)$ that we now recall. The modular $j$-function and the Ramanujan cusp form $\Delta$ are normalized so that

$$j(z) = q^{-1} + 744 + ...$$

$$\Delta(z) = q \prod_{n \geq 1} (1 - q^n)^{24}, \quad q = e^{2\pi iz}.$$ 

For $E$ an elliptic curve over $L$ and $\sigma : L \to \mathbb{C}$ and embedding, we choose $\tau_{E,\sigma} \in \mathfrak{h}$ satisfying that $j(\tau_{E,\sigma})$ is the $j$-invariant of $E \otimes_{\sigma} \mathbb{C}$. Then Silverman’s formula is

$$(18.1) \quad h(E) = \frac{1}{12[L : \mathbb{Q}]} \left( \log \Delta_E - \sum_{\sigma : L \to \mathbb{C}} \log \left| \Delta(\tau_{E,\sigma}) \cdot \Im(\tau_{E,\sigma})^6 \right| \right) - \log(2\pi)$$

where $\Delta_E$ is the norm of the minimal discriminant ideal of $E$ over $L$. (Note that in [75] there is a minor typo in the definition of $\Delta(\tau)$ that gives $+ \log(2\pi)$ instead of $- \log(2\pi)$; this has been corrected in a number of places). One deduces:

**Lemma 18.1.** With the previous notation,

$$\frac{1}{[L : \mathbb{Q}]} \log \Delta_E < 12h(E) + 16.$$ 

We will need another description of $h(E)$. Suppose that $E$ is semi-stable over $F$ and let $\mathcal{E}$ be the corresponding semi-stable model over $O_L$, such that the smooth locus of its structure map $\mathcal{E}$ is the Néron model of $E$. Let $\omega$ be the relative dualizing sheaf of $\mathcal{E}$; in particular $\omega$ is an invertible sheaf ($E$ is semi-stable) and its restriction to $\mathcal{E}$ is $\Omega^1_{\mathcal{E}/O_L}$.

We make $\omega$ into a metrized line bundle $\hat{\omega}$ with the following metric $\| - \|_\sigma$ associated to an embedding $\sigma : L \to \mathbb{C}$:

Let $x \in E_{\sigma}$ and let $\lambda \in \Omega^1_{E_{\sigma}/\mathbb{C}}|_x$. Let $\alpha \in H^0(E_{\sigma}, \Omega^1_{E_{\sigma}/\mathbb{C}})$ be the unique invariant differential with $\alpha_x = \lambda$. Then we define

$$\|\lambda\|_{\sigma,x}^2 = \frac{i}{2} \int_{E_{\sigma}} \alpha \wedge \overline{\alpha}.$$ 

We remark that, in general, the metrized line bundle $\hat{\omega}$ is not the Arakelov canonical metrized line bundle, which one might denote by $\hat{\omega}^{Ar}$. The metrics of $\hat{\omega}$ differ from the metrics of $\hat{\omega}^{Ar}$ by a scalar multiple.

**Theorem 18.2.** Let $E$ be a semi-stable elliptic curve over $L$ and keep the previous notation. Let $P$ be an algebraic point of $E$ satisfying that its closure in $\mathcal{E}$ is contained in $\mathcal{E}$. Let $L'/L$
be a finite extension over which \( P \) is defined and let \( S = \text{Spec} O_L \) and \( S' = \text{Spec} O_{L'} \). Let 
\[ s : S' \to S \] 
be the \( S \)-map attached to \( P \). Then
\[ \deg_{L'} s^* \hat{\omega} = [L' : \mathbb{Q}] \cdot h(E). \]

**Proof.** This is proved in the same way as Proposition 7.2 in [17], which assumes that \( P \) is rational and uses \( \hat{\omega}^{Ar} \) instead of \( \hat{\omega} \). The passage from rational to algebraic points of the type considered here is straightforward, provided that in loc. cit. one starts with \((P, \hat{\omega}^{Ar})\) (instead of starting with \((P, P)\) and using adjunction). On the other hand, replacing \( \hat{\omega}^{Ar} \) by \( \hat{\omega} \) modifies the metrics in a way made explicit by Definition 4.1 and Proposition 4.6 in loc. cit. Taking this contribution at infinity into account, one obtains the Faltings height of \( E \) in the form \([18.1]\).

18.2. Shimura curves. Let \( F \) be a totally real number field of degree \( n \) over \( \mathbb{Q} \) with \( \tau_1, \ldots, \tau_n \) its real embeddings. Let \( \mathfrak{N} \) be a non-zero ideal in \( O_F \) which is the product of \( \nu \) distinct prime ideals. Suppose that \( n + \nu \) is odd and let \( B \) be a quaternion \( F \)-algebra of discriminant \( \mathfrak{N} \) and having \( \tau_1 \) as its only split place at infinity. Furthermore, when \( F = \mathbb{Q} \) we assume \( \nu > 0 \). Write \( B = B \otimes \mathbb{Z} \) and fix a choice of maximal order \( O_B \) in \( B \). Let \( B_F^+ \) be the set of units in \( B \) with totally positive reduced norm.

Associated to this data, for each open compact \( U \subseteq B^\times \) there is a compact Shimura curve \( X_U \) defined over \( F \), whose complex points (via \( \tau_1 \)) are given by
\[ X^{an}_U = B_F^+ \backslash B^\times / U. \]

The curve \( X_U \) is irreducible over \( F \), although \( X^{an}_U \) is not necessarily irreducible and its connected components are parametrized by the class group \( C_U = F_F^+ / C_U \) where \( C_U \) denotes the reduced norm and \( F_F^+ \) is the multiplicative group of totally positive elements of \( F \). We also keep the notation from Paragraph \([8.1]\).

The following expression for the hyperbolic volume of the Shimura curve with \( U = O_B^\times \), is a special case of a result of Shimizu \([72]\).

**Proposition 18.3.** The number of connected components of \( X^{an}_{O_B^\times} \) is \( h_F^+ \), the narrow class number of \( F \). Each component has the form \( \tilde{\Gamma}_{O_B^\times, g} \backslash h \) for suitable \( g \in B^\times \), and they have hyperbolic area
\[ \text{Vol} \left( \tilde{\Gamma}_{O_B^\times, g} \backslash h \right) \asymp_n d^{3/2}_{F} \prod_{\mathfrak{p} \mid \mathfrak{N}} (\text{Norm}(\mathfrak{p}) - 1). \]

The Jacobian of \( X_U \) over \( F \) is denoted by \( J_U \). We denote by \( \mathbb{T}_U \) the ring of Hecke correspondences on \( X_U \) and by \( \mathbb{T}_U \) the ring of Hecke operators acting on \( J_U \). They are generated by the Hecke correspondences \( T^c_n \) (resp. Hecke operators \( T_n \)) for all \( n \) integral ideals of \( O_F \) coprime to \( \mathfrak{N} \) and coprime to the places where \( U \) is not maximal. Then \( \mathbb{T}_U \) acts also on holomorphic differentials of \( J_U \) and \( X_U \). See \([36]\) for details. Note that the action of \( \mathbb{T}_n^c \) on divisors permutes the components of \( X^{an}_U \) according to the action of \( n \) on \( C(U) \).

The Shimura construction associated to a system of Hecke eigenvalues \( \chi : \mathbb{T}_U \to \mathbb{C} \) has been carried out in \([91]\) and the theory is similar to that of Shimura curves over \( \mathbb{Q} \). In particular, when \( \chi \) is \( \mathbb{Z} \)-valued and new, the associated optimal quotient \( q : J_U \to A \) satisfies that \( A \) is an elliptic curve. We define the modular degree of \( A \) arising in this way, as the integer
\[ \delta_\chi = qq^\vee \in \text{End}(A). \]
uniformity as the number field varies. It turns out that the bounds for the degree of an
over number fields of bounded degree has not yet been proved, while we would like some
bound with Theorem 18.4. However, a version of the "isogeny theorem" for elliptic curves
Jacobian of \( X \) affording the modular parameterization \( \phi \), the modular degree \( \delta \) case
E two conditions, is modular of Shimura level 1 in the sense that we just defined.

Let us say that \( E \) is modular of Shimura level 1 if in addition one can choose \( \mathfrak{N} \) as the conductor ideal of \( E \) and \( U = O_{\mathfrak{B}}^\times \). In this case \( E \) is semi-stable (so, its conductor
is squarefree) and the parity of the number of primes of bad reduction of \( E \) (i.e. of the prime
factors of its conductor) is opposite to the parity of \( n = [F : \mathbb{Q}] \). The usual modularity conjecture for elliptic curves over totally real number fields along with the Jacquet-Langlands correspondence, imply that in fact each elliptic curve over \( F \) whose conductor satisfies these two conditions, is modular of Shimura level 1 in the sense that we just defined.

If \( E \) is modular of Shimura level 1 we simply write \( X_1 = X_{O_{\mathfrak{B}}^\times} \) for the Shimura curve
affording the modular parameterization \( X_1 \to E \) of the definition. We observe that in this
case \( E \) is isogenous to \( \tilde{A} \) for an optimal elliptic curve quotient \( q : J_1 \to A \), where \( J_1 \) is the
Jacobian of \( X_1 \) over \( F \). The associated system of Hecke eigenvalues is denoted by \( \chi_E \) and
the modular degree \( \delta_{\chi_E} \) is simply written \( \delta_E \).

One would like to control the minimal degree of an isogeny \( A \to E \) and to combine this
bound with Theorem 18.4. However, a version of the "isogeny theorem" for elliptic curves
over number fields of bounded degree has not yet been proved, while we would like some
uniformity as the number field varies. It turns out that the bounds for the degree of a
minimal isogeny (with controlled field of definition) proved by Gaudron and Rémond are enough for our purposes, and one obtains:

**Theorem 18.5.** Let $E$ be an elliptic curve over a totally real number field $F$ and assume that $E$ is modular of Shimura level $E$. Then, with the previous notation, there is a non-constant map

$$\phi_E : X_1 \to E$$

defined over $F$, whose degree satisfies

$$h_F^+ \delta_E \leq \deg \phi_E \leq c_{12}(n) \max \{1, h(E)\}^{c_{13}(n)} h_F^+ \delta_E$$

where $h(E)$ denotes the (logarithmic) Faltings height of $E$.

(Note that the factor $(\log \text{Norm}(\mathfrak{M}))^{c_{11}(n)}$ from Theorem 18.4 has been absorbed by the factor $\max \{1, h(E)\}^{c_{13}(n)}$, in view of Lemma 18.1)

18.4. **Arakelov height of Heegner points after Yuan and Zhang.** In this paragraph we briefly recall some of the main points of the theory developed by Yuan and Zhang in [93].

As $U$ varies over open compact subgroups of $O_B^\times$ (notation as above) one obtains a projective system $\{X_U\}_U$ of curves over $F$ mapping to $X_{O_B^\times}$.

For a positive integer $m$, write $U(m) = (1 + mO_B)^\times$. The next is Proposition 4.1 in [93].

**Proposition 18.6.** If $m \geq 3$ and $U \subset U(m)$, then for every $g \in \mathbb{B}^\times$ the group $\tilde{\Gamma}_{U,g}$ acts freely on $h$, and the genus of every geometric component of $X_U$ is at least 2.

Write $X_m$ for the curve $X_{U(m)}$, in particular $X_1 = X_{O_B^\times}$ agrees with our previous notation. Note that when $m_1 | m_2$ we have a natural map $X_{m_2} \to X_{m_1}$. The next result follows from the first part of Section 4.2 in [93].

**Theorem 18.7.** For $m \geq 3$ coprime to $\text{Norm}(\mathfrak{M})$, there is a canonical regular integral model $\mathcal{X}_m$ for $X_m$, which is flat and projective over $O_F$. Furthermore, $\mathcal{X}_m[1/m]$ is semi-stable over $O_F[1/m]$. More precisely, for a prime ideal $p$ of $O_F$ not dividing $m$ one has that the special fibre at $p$ is

(i) relative Mumford, if $p | \mathfrak{D}$;

(ii) smooth, if $p \nmid \mathfrak{D}$.

For any $m_0$ coprime to $\text{Norm}(\mathfrak{M})$ an integral model $\mathcal{X}_{m_0}$ for $X_{m_0}$ is obtained as follows: Choose any $m \geq 3$ coprime to $\text{Norm}(\mathfrak{M})$ with $m_0 | m$, then $\mathcal{X}_{m_0}$ is defined as the quotient of $\mathcal{X}_m$ by the natural action of the finite group $U(m_0)/U(m)$. The scheme $\mathcal{X}_{m_0}$ is independent of the choice of $m$ (up to isomorphism), it is normal, and it is flat and projective over $O_F$. In particular, this construction applies to $X_1$.

For $m_1 | m_2$ positive integers coprime to $\text{Norm}(\mathfrak{M})$, the map $X_{m_2} \to X_{m_1}$ extends to an $O_F$-morphism $\mathcal{X}_{m_2} \to \mathcal{X}_{m_1}$.

In [93] integral models $\mathcal{X}_U$ for $X_U$ are constructed in more generality, but for our purposes the case of $X_m$ with $m$ a positive integer coprime to $\text{Norm}(\mathfrak{M})$ suffices. In what follows, such an integer $m$ will be called admissible.

For each admissible $m$ let $\mathcal{L}_m$ be the metrized Hodge bundle on $\mathcal{X}_m$, as constructed in [93] Section 4.2. This is a metrized $\mathbb{Q}$-line bundle on $\mathcal{X}_m$ in general, and when $m \geq 3$ it is a metrized line bundle on $\mathcal{X}_m$. Its finite part (i.e. forgetting the metrics) is denoted by $\mathcal{L}_m$. The next result follows from Theorem 4.7 [93].


Theorem 18.8. The metrized Hodge bundles satisfy the following properties:

(i) for \( m_1 | m_2 \) admissible integers, the pull-back of \( \tilde{L}_{m_1} \) by \( X_{m_2} \to X_{m_1} \) is \( \tilde{L}_{m_2} \);

(ii) for \( m \geq 3 \) admissible, we have that \( \mathcal{L}_m[1/m] \) is the relative dualizing sheaf of \( \mathcal{Z}_m[1/m] \to \text{Spec } O_F[1/m] \);

(iii) For each embedding \( \sigma : F \to \mathbb{R} \subseteq \mathbb{C} \), the metric on the restriction of \( \mathcal{L} \) to \( X_{m} \otimes_\sigma \mathbb{C} \) is induced via complex uniformization by the metric \( |dz| = 2\Im(z) \) on differential forms on \( \mathfrak{h} \).

A totally imaginary quadratic extension \( K/F \) is said to satisfy the Heegner hypothesis for \( N \) if every prime \( p \mid N \) is inert in \( K \). In particular, such an extension \( K/F \) has relative discriminant \( \mathfrak{d}_{K/F} \) coprime to \( N \).

If \( K \) satisfies the Heegner hypothesis for \( \mathfrak{N} \), for each open compact \( U \) we have a Heegner point \( P_{K,U} \) in \( X_U \). These can be chosen to form a compatible system of algebraic points in the projective system \( \{ X_U \}_U \). The point \( P_K = P_{K,O_\mathfrak{N}} \) is defined over the Hilbert class field of \( K \), and in general, all the \( P_{K,U} \) are defined over abelian extensions of \( K \).

Theorem 1.5 in [93] gives a formula for the Arakelov-theoretical height

\[
h_{Ar}(P_K) := \frac{1}{[H_K : F]} \deg_{O_{H_K}} s^\ast \tilde{L}_1
\]

of \( P_K \) on \( \mathcal{Z}_1 \) with respect to \( \tilde{L}_1 \), where \( s : \text{Spec } O_{H_K} \to \mathcal{Z}_1 \) is the multi-section attached to \( P_K \). The result is:

**Theorem 18.9.** Suppose that \( F \) has at least two ramified places in \( \mathbb{B} \). Let \( K/F \) be a totally imaginary quadratic extension satisfying the Heegner hypothesis for \( \mathfrak{N} \). Let \( \chi_K \) be the quadratic Hecke character of \( F \) associated to the extension \( K/F \). Then we have

\[
h_{Ar}(P_K) = -\frac{L'}{L}(0, \chi_K) + \frac{1}{2} \log \frac{\text{Norm}(\mathfrak{N})}{\text{Norm}(\mathfrak{d}_{K/F})}.
\]

This can be rewritten as

\[
h_{Ar}(P_K) = \frac{L'}{L}(1, \chi_K) + \frac{1}{2} \log \text{Norm}(\mathfrak{d}_{K/F} \mathfrak{N}) + \log d_F - n \cdot (\gamma + \log(2\pi)).
\]

The first formula is Theorem 1.5 in [93], while the second is simply a consequence of the functional equation of \( L(s, \chi_K) \).

18.5. **An approach to Szpiro’s conjecture over totally real number fields.**

**Theorem 18.10.** Let \( F \) be a totally real number field of degree \( n \) over \( \mathbb{Q} \). Let \( E \) be an elliptic curve over \( F \) and suppose that \( E \) is modular of Shimura level 1. Let \( \mathfrak{N} \) be the conductor ideal of \( E \), let \( \mathbb{B} \) be a quaternion \( F \)-algebra of discriminant \( \mathfrak{N} \) with exactly one split place at infinity, and let \( X_1 \) be the associated Shimura curve.

Let \( \psi : X_1 \to E \) be a non-constant morphism over \( F \) (which exists, as \( E \) is modular of Shimura level 1). Suppose that \( F/\mathbb{Q} \) is tower of cyclic, or that GRH holds for the \( L \)-function of quadratic Hecke characters over \( F \).

Then we have

\[
h(E) \leq \frac{1}{2} \log \text{deg } \psi + c_{14}(n) \log(d_F \text{Norm}(\mathfrak{N}))
\]

72
and 
\[
\log \Delta_E \leq 6n \log \deg \psi + c_{15}(n) \log(d_{F, \text{Norm}}(\mathfrak{N})).
\]

**Proof.** For each admissible \( m \), let \( \psi_m : X_m \to E \) be the composition of \( X_m \to X_1 \) with \( \psi \). For \( m \geq 3 \) admissible, the map \( \psi_m \) extends to an \( O_F[1/m] \)-morphism
\[
\psi_m : X_m[1/m] \to \mathcal{E}[1/m]
\]
by the Néron mapping property, where \( \mathcal{E} \) is the Néron model of \( E \) and \( X_m[1/m] \) is the smooth locus of \( X_m \to \text{Spec} O_F[1/m] \), obtained from \( X_m[1/m] \) by removing the singular points of the special fibres at primes dividing \( m \).

Choose two different admissible prime numbers \( p, q \), different from 2 and small in the sense that
\[
p, q < 10 \log \text{Norm}\mathfrak{N}.
\]
Let \( \mathcal{Y}_{pq} \) be the open set of \( \mathcal{X}_{pq} \) obtained as the union of the preimages of \( \mathcal{X}[1/p] \subseteq \mathcal{X}_{p} \) and \( \mathcal{X}_{q}[1/q] \subseteq \mathcal{X}_{q} \). Then by gluing we obtain an \( O_F \)-morphism
\[
\psi_{pq} : \mathcal{Y}_{pq} \to \mathcal{E}.
\]

Let \( K/F \) be a totally imaginary quadratic extension of \( F \) satisfying the Heegner hypothesis. Furthermore, assume that the associated Heegner point \( P_{K,U(pq)} \) is not a ramification point of \( \psi_{pq} \). Let \( H \) be a number field containing \( K \) over which \( P_{K,U(pq)} \) is defined. Write \( S_H = \text{Spec} O_H \) and let \( s : S_H \to \mathcal{X}_{pq} \) be the \( O_F \)-map obtained from the algebraic point \( P_{K,U(pq)} \).

Observe that \( s \) exists because \( \mathcal{X}_{pq} \) is projective over \( O_F \), and it has image contained in \( \mathcal{Y}_{pq} \) because all primes dividing \( \mathfrak{N} \) are inert in \( K \) (cf. p. 242 in [92]).

By a patching argument using \( \psi_p \) and \( \psi_q \), we get the canonical injective sheaf morphism (cf. the notation in Paragraphs 18.4 and 18.5)
\[
u : \psi_{pq}^*(\omega|_\mathcal{E}) \to \mathcal{L}_{pq}|_{\mathcal{Y}_{pq}}
\]
which allows us to see the first as a sub-sheaf of the second, on \( \mathcal{Y}_{pq} \). The sheaf \( \psi_{pq}^*(\omega|_\mathcal{E}) \) has the metrics induced from the metrized relative dualizing sheaf \( \mathcal{D} \) of the semi-stable elliptic curve \( E \), while the Hodge bundle has its own metric. We now compare these metrics.

Let \( \sigma : F \to \mathbb{C} \) be any embedding and take a complex point \( x \in X_{pq} \otimes_\mathbb{Q} \mathbb{C} \). Let \( \mu \in \psi_{pq}(\Omega^1_{E_\sigma/\mathbb{C}})|_x \) and let \( \lambda = \Omega^1_{E_\sigma/\mathbb{C}}|_{\psi_{pq}(x)} \) be such that \( \psi_{pq}(\lambda) = \mu \). Let \( \alpha \in H^0(E_\sigma, \Omega^1_{E_\sigma/\mathbb{C}}) \) be the invariant differential with \( \alpha|_{\psi_{pq}(x)} = \lambda \). Then by definition we have
\[
u(\mu)|_x = \frac{i}{2} \int_{E_\sigma} \alpha \wedge \overline{\alpha}.
\]

We now estimate \( \| u(\mu) \|_{\sigma, x} \) according to the metrics of the Hodge bundle. Note that
\[
u(\mu) = u(\psi_{pq}^*(\lambda)) = u(\psi_{pq}^*(\alpha|_{\psi_{pq}(x)})) = u(\psi_{pq}^*\alpha)|_x = (\psi_{pq}^*\alpha)|_x
\]
and therefore for suitable \( g \in \mathbb{B} \) according to the connected component in which \( x \) is, we have
\[
u(\mu)|_x \leq 2\| \Psi_{U(pq),g} \psi_{pq}^*\alpha \|_{U(pq),g, \sigma} \leq c_{16}(n)\| \Psi_{U(pq),g} \psi_{pq}^*\alpha \|_{U(pq),g,2, \sigma}
\]
where \( \Psi_{U, g} \) is defined as in Paragraph 14.6 and the second inequality is by Theorem 8.1.

By pulling back the \( (1,1) \)-form \( \alpha \wedge \overline{\alpha} \) from (18.2), we now observe that
\[
u(\mu)|_x \leq \deg(\psi_{pq})\| u(\mu) \|^2_{\sigma, x}
\]
(this is only a bound rather than an equality because the left hand side just considers one geometric component of $X_{U(pq)}$) which gives the norm comparison

\begin{equation}
\|u(\mu)\|_{\sigma,x} \leq c_{17}(n)(\deg \psi_{pq})^{1/2}\|\mu\|_{\sigma,x}.
\end{equation}

Pulling back $u$ by $s$ we get a sheaf morphism

$$u'(\psi_{pq}s)^*\omega = (\psi_{pq}s)^* (\omega|_s) \to s^*\mathbb{L}_{pq}|_{\psi_{pq}} = s^*\mathbb{L}_{pq}$$

which still is injective because $P_{K,U(pq)}$ is not in the ramification locus of $\psi_{pq}$, by assumption. Injectivity, together with the norm comparison (18.3) gives

$$\hat{\deg}_{SH}(\psi_{pq}s)^*\hat{\omega} \leq \hat{\deg}_{SH}s^*\mathbb{L}_{pq} + \frac{[H : \mathbb{Q}]}{2} \log \deg \psi_{pq} + c_{18}(n)[H : \mathbb{Q}].$$

Dividing by $[H : \mathbb{Q}]$, recalling Theorem 18.2, the definition of $h_{Ar}(P_K)$, and the fact that the metrized Hodge bundles are compatible with pull-back (Theorem 18.8), one obtains

$$h(E) \leq \frac{1}{n}h_{Ar}(P_K) + \frac{1}{2}\log \deg \psi_{pq} + c_{19}(n).$$

By Theorem 18.9 we get

\begin{equation}
(n \cdot h(E) \leq \frac{n}{2}\log \deg \psi_{pq} + \frac{L'}{L}(1, \chi_K) + \frac{1}{2}\log \text{Norm}(\mathfrak{o}_{K/F}\mathfrak{M}) + \log d_F + c_{20}(n).
\end{equation}

The number of complex ramification points of $\psi_m$ is at most $2g_m - 2$ where $g_m$ is the sum of the genera the geometric components of $X_m$. This is at most the total (hyperbolic) volume of $X_m$ divided by $2\pi$, which is at most

$$\frac{1}{2\pi}[U(1) : U(m)] \cdot Vol(X_m^{an}) \leq c_{21}(n)(md_F)^{c_{22}(n)}\text{Norm}(\mathfrak{M})$$

since $X_m^{an}$ has $h_F^+$ components, and using the volume formula from Proposition 18.3.

Hence, by Theorem 17.4 and recalling that $p$ and $q$ are small, there is some totally imaginary quadratic extension $K/F$ satisfying the Heegner hypothesis for $\mathfrak{M}$, such that $P_{K,U(pq)}$ is not a ramification point of $\psi_{pq}$, and satisfying the estimates

$$\text{Norm}(\mathfrak{o}_{K/F}) < c_{23}(n)(d_F\text{Norm}(\mathfrak{M}))^{c_{24}(n)}$$

and

$$\left|\frac{L'}{L}(1, \chi_K)\right| < c_{25}(n) (\log d_F + \log \log \text{Norm}(\mathfrak{M})).$$

From (18.3) and using the fact that the degree of $X_m \to X_1$ is at most $m^{c_{26}(n)}$ (and that $p$ and $q$ are small), we finally deduce

$$h(E) \leq \frac{1}{2}\log \deg \psi + c_{27}(n)\log(d_F\text{Norm}(\mathfrak{M})).$$

The estimate for the discriminant ideal of $E$ follows by Lemma 18.1 \[\square\]

Theorem 18.10 motivates the following conjecture:

**Conjecture 18.11.** Suppose that $F$ is a totally real number field of degree $n$ over $\mathbb{Q}$, and that $E/F$ is an elliptic curve which is modular of Shimura level 1. Then

$$\log \delta_E < c_{28}(n)\log(d_F\text{Norm}(\mathfrak{M})).$$
Note that the conjecture only concerns modular elliptic curves of Shimura level 1, and it involves the canonically defined quantity $\delta_E$, instead of using the degree of some arbitrary choice of modular parametrization $X_1 \to E$. Conjecture 18.11 is of interest because under suitable conditions it implies Szpiro’s conjecture, as we discuss in the next paragraph.

18.6. Bounds for $\delta_E$ and Szpiro’s conjecture. The purpose of this paragraph is to spell-out the precise relation between Szpiro’s conjecture and our proposed conjectural bound for the quantity $\delta_E$, namely, Conjecture 18.11.

First we do this in a particularly convenient case where several technical assumptions can be removed: the case when $F$ is real quadratic. In particular, here one does not need to assume modularity as a hypothesis, because it is proved in the relevant cases.

**Theorem 18.12.** Assume Conjecture 18.11 for real quadratic fields and their totally real quadratic extensions. Then there is an absolute constant $c > 0$ such that the following holds:

For every real quadratic field $F$ and every semi-stable elliptic curve $E$ over $F$ which does not have everywhere good reduction, we have

$$h(E) \leq c \log (d_F N_E)$$

and

$$\Delta_E \leq (d_F N_E)^c.$$

**Proof.** If $E$ has an odd number of places of bad reduction, then it is modular of Shimura level 1 (according to our definition) by the modularity results from [29] (see also [92, 94] to deduce geometric modularity from automorphy). In this case we have, for the corresponding Shimura curve $X_1$, a modular parameterization

$$\phi : X_1 \to E$$

afforded by Theorem 18.5, whose degree satisfies

$$\log \deg(\phi) \leq \log \delta_E + \log h_F^+ + O_n (1 + \log \max\{1, h(E)\}).$$

By Theorem 18.10 we get (unconditionally)

$$h(E) \leq \frac{1}{2} \log \delta_E + O_n (\log \max\{1, h(E)\} + \log (d_F N_E)).$$

At this point we apply Conjecture 18.11 to get the result.

When $E$ has a non-zero even number of primes of bad reduction over $F$, we base-change to a suitable totally real quadratic extension $F'/F$, such that exactly one prime of bad reduction of $E$ is inert and all the other are split. Since $E$ is semi-stable over $F$, now it has an odd number of bad places over $F'$ and by [25] it is modular of Shimura level 1. As $E/F$ is semi-stable, its Faltings height is the same after base change. Furthermore, $F'$ can be chosen with $\log d_{F'} \ll \log (d_F N_E)$ by Lemma 17.1. As the extension $F'/\mathbb{Q}$ is tower of cyclic, the same argument applies. \(\square\)

Finally, here is a version of the previous result for totally real number fields beyond the quadratic case, under some additional simplifying assumptions. This result constitutes our modular approach to Szpiro’s conjecture over totally real number fields mentioned in the introduction.
Theorem 18.13. Assume Conjecture 18.11. Assume that elliptic curves over totally real number fields are modular (in the sense of [33]). Then the following holds:

Let $F$ be a totally real number field of degree $n$ over $\mathbb{Q}$. Suppose that $F/\mathbb{Q}$ is tower of cyclic or that GRH holds for quadratic Hecke characters over number fields. Then for every semi-stable elliptic curve $E$ over $F$ which does not have everywhere good reduction, we have

$$h(E) \leq c_{29}(n) \log(d_F N_E)$$

and

$$\Delta_E \leq (d_F N_E)^{c_{30}(n)}.$$ 

The proof is similar to that of Theorem 18.12. Of course one does not need the full modularity conjecture, and even the known results on potential modularity would suffice in some cases.

18.7. Application: Unconditional exponential bounds. Finally, we remark that unconditional exponential bounds for $\delta_E$ over a totally real field $F$ (and thanks to our results, for Szpiro’s conjecture) can be proved by the same methods of Sections 5 and 7. Namely, it is straightforward to extend any of the two proofs of Theorem 5.3 to the totally real setting, and then such an extension is used to prove an analogue of Theorem 7.2. To do the latter, one should use results on effective multiplicity one for $GL_2$—such as those in [10]—in order to distinguish systems of Hecke eigenvalues using only few Hecke operators. To count the number of systems of Hecke eigenvalues used in the proof, a crude bound is the genus of $X_{O_{\mathbb{R}}}^\times$, which can be estimated by Shimizu’s volume formula.

For instance, a simple unconditional consequence is the following:

Theorem 18.14. Let $F$ be a totally real number field and let $\epsilon > 0$. Suppose that $F/\mathbb{Q}$ is tower of cyclic. For all modular elliptic curves $E$ over $F$ of Shimura level 1, we have

$$h(E) \ll_{F, \epsilon} N_E^{1+\epsilon}$$

and

$$\log \Delta_E \ll_{F, \epsilon} N_E^{1+\epsilon}.$$ 

With more work, one can be more precise about the dependence of the implicit constant, but the estimates will not be better than the exponential bounds that we have obtained in the case $F = \mathbb{Q}$.

19. Acknowledgments

Part of this project was carried out while I was a member at the Institute for Advanced Study in 2015-2016, and then continued at Harvard. This research was supported by a Benjamin Peirce Fellowship at Harvard, and at the IAS it was partially supported by a Schmidt Fellowship and by the IAS from their NSF grant agreement No. DMS-1128155. While working on this project, I greatly benefited from conversations with Enrico Bombieri, Noam Elkies, Nicholas Katz, Barry Mazur, Peter Sarnak, Richard Taylor, and Shou-Wu Zhang. I sincerely thank them for generously sharing their knowledge and expertise. In particular, the argument in Paragraph 5.6 originates in an idea of R. Taylor in the case of classical modular curves, and the connection between injectivity radius and Lehmer’s conjecture used in Section 8 was pointed out to me by P. Sarnak.
It was also extremely useful for me to attend the joint IAS-Princeton algebraic number theory seminar during my stay at the IAS, as it covered relevant technical material on Shimura curves. I thank Christopher Skinner and Richard Taylor for organizing it.

Feedback from Natalia Garcia-Fritz and Ricardo Menares regarding the presentation of this article is gratefully acknowledged.

Last but not least, I thank Enrico Bombieri and Natalia Garcia-Fritz for their encouragement, which allowed me to push this project further than I initially had in mind.

References

[1] A. Abbes, E. Ullmo, À propos de la conjecture de Manin pour les courbes elliptiques modulaires. Compositio Math. 103 (1996), no. 3, 269-286.
[2] A. Agashe, K. Ribet, W. Stein, The Manin constant. Pure Appl. Math. Q. 2 (2006), no. 2, Special Issue: In honor of John H. Coates. Part 2, 617-636.
[3] A. Agashe, K. Ribet, W. Stein, The modular degree, congruence primes, and multiplicity one. Number theory, analysis and geometry, 19-49, Springer, New York, (2012).
[4] J. Arthur, L. Clozel, Simple algebras, base change, and the advanced theory of the trace formula. Annals of Mathematics Studies, 120. Princeton University Press, Princeton, NJ, 1989. xiv+230 pp. ISBN: 0-691-08517-X; 0-691-08518-8
[5] A. Atkin, J. Lehner, Hecke operators on $\Gamma_0(N)$. Math. Ann. 185 (1970) 134-160.
[6] A. Baker, Experiments on the abc-conjecture. Publ. Math. Debrecen 65 (2004), no. 3-4, 253-260.
[7] V. Blomer, R. Holowinsky, Bounding sup-norms of cusp forms of large level. Inventiones mathematicae 179.3 (2010): 645-681.
[8] S. Bosch, W. Lütkemüller, M. Raynaud, Néron models. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], 21. Springer-Verlag, Berlin, 1990. x+325 pp. ISBN: 3-540-50587-3
[9] C. Breuil, B. Conrad, F. Diamond, R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises. Journal of the American Mathematical Society, 14 (4): 843-939
[10] F. Brumley, Effective multiplicity one on $GL_N$ and narrow zero-free regions for Rankin-Selberg $L$-functions. Amer. J. Math. 128 (2006), no. 6, 1455-1474.
[11] K. Buzzard, Integral models of certain Shimura curves. Duke Mathematical Journal 87.3 (1997): 591.
[12] H. Carayol, Sur les représentations $l$-adiques associées aux formes modulaires de Hilbert. Annales scientifiques de l'École Normale Supérieure, Série 4 : Volume 19 (1986) no. 3 , p. 409-468
[13] K. Cesnavičius, A modular description of $X_0(n)$. Preprint arXiv: 1511.07475 (2015).
[14] K. Cesnavičius, The Manin-Stevens constant in the semistable case. Preprint arXiv: 1604.02165 (2016).
[15] K. Cesnavičius, The Manin constant in the semistable case. Preprint arXiv: 1703.02951 (2017).
[16] L. Clozel, Base change for $GL(n)$. Proceedings of the International Congress of Mathematicians. Vol. 1. (1986).
[17] A. Cojocaru, E. Kani, The modular degree and the congruence number of a weight 2 cusp form. Acta Arith. 114 (2004), no. 2, 159-167.
[18] P. Colmez, Périodes des variétés abéliennes à multiplication complexe. Ann. Math., 138 (1993), 625-683.
[19] P. Colmez, Sur la hauteur de Faltings des variétés abéliennes à multiplication complexe. Compositio Math. 111 (1998), no. 3, 359-368.
[20] B. Conrad, Integrality of modular forms via intersection theory. Appendix to: $p$-adic $L$-functions and the converse filtration on Chow groups, by M. Bertolini, H. Darmon, K. Prasanna, to appear in Crelle.
[21] D. Cox, W. Parry, Genera of congruence subgroups in $\mathbb{Q}$-quaternion algebras. J. Reine Angew. Math. 351 (1984), 66-112.
[22] S. Das, J. Sengupta, $L^\infty$ norms of holomorphic modular forms in the case of compact quotient. Forum Mathematicum. Vol. 27. No. 4. (2015).
[23] P. Deligne, M. Rapoport, Les schémas de modules de courbes elliptiques. Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972), pp. 143-316. Lecture Notes in Math., Vol. 349, Springer, Berlin, 1973.
[24] F. Diamond J. Im, *Modular forms and modular curves*. Seminar on Fermat’s Last Theorem (Toronto, ON, 1993-1994), 39-133, CMS Conf. Proc., 17, Amer. Math. Soc., Providence, RI, 1995.

[25] L. Dieulefait, N. Freitas, *Base change for elliptic curves over real quadratic fields*. Comptes Rendus Mathematique 353.1 (2015): 1-4.

[26] B. Edixhoven, *On the Manin constants of modular elliptic curves*. Arithmetic algebraic geometry (Texel, 1989), 25-39, Progr. Math., 89, Birkhäuser Boston, Boston, MA, (1991).

[27] B. Edixhoven, *Comparison of integral structures on spaces of modular forms of weight two, and computation of spaces of forms mod 2 of weight one*. (English, French summary) With appendix A (in French) by Jean-François Mestre and appendix B by Gabor Wiese. J. Inst. Math. Jussieu 5 (2006), no. 1, 1-34.

[28] É. Fouvry, M. Nair, G. Tenenbaum, *L’ensemble exceptionnel dans la conjecture de Szpiro*. (French, English, French summary) Bull. Soc. Math. France 120 (1992), no. 4, 485-506.

[29] N. Freitas, B. V. Le Hung, S. Siksek, *Elliptic curves over real quadratic fields are modular*. Inventiones mathematicae 201.1 (2015): 159-206.

[30] G. Freixas i Montplet, *The Jacquet-Langlands correspondence and the arithmetic Riemann-Roch theorem for pointed curves*. Int. J. Number Theory 8 (2012), no. 1, 1-29.

[31] G. Frey, *Links between solutions of $A-B=C$ and elliptic curves*. Number theory (Ulm, 1987), 31-62, Lecture Notes in Math., 1380, Springer, New York, (1989).

[32] É. Gaudron, G. Rémond, *Polarisations et isogénies*. Duke Mathematical Journal 163.11 (2014): 2057-2108.

[33] S. Gelbart, *Elliptic curves and automorphic representations*. Advances in Mathematics 21, no. 3 (1976): 235-292.

[34] D. Goldfeld, J. Hoffstein, *On the number of Fourier coefficients that determine a modular form*. A tribute to Emil Grosswald: number theory and related analysis, 385-393, Contemp. Math., 143, Amer. Math. Soc., Providence, RI, (1993).

[35] B. Gross, *Arithmetic on elliptic curves with complex multiplication*. LNM 776, Springer-Verlag, New York, (1980).

[36] B. Gross, *Local orders, root numbers, and modular curves*. Amer. J. Math. 110 (1988), no. 6, 1153-1182.

[37] B. Gross, D. Zagier, *Heegner points and derivatives of $L$-series*. Invent. Math. 84 (1986), no. 2, 225-320.

[38] K. Győry, *On the abc conjecture in algebraic number fields*. Acta Arith. 137 (2009), no. 3, 253-276.

[39] E. Heilbronn, A. Kraus, *Courbes de Fermat: résultats et problèmes*. J. Reine Angew. Math. 548 (2002) 167-234.

[40] D. R. Heath-Brown, *A mean value estimate for real character sums*. Acta Arith. 72 (1995), no. 3, 295-275.

[41] D. Helm, *On maps between modular Jacobians and Jacobians of Shimura curves*. Israel J. Math. 160 (2007), 61-117.

[42] J. Hoffstein, P. Lockhart, *Coefficients of Maass forms and the Siegel zero*. With an appendix by D. Goldfeld, J. Hoffstein and D. Lieman. Ann. of Math. (2) 140 (1994), no. 1, 161-181.

[43] Y. Ihara, *On the Euler-Kronecker constants of global fields and primes with small norms*. Algebraic geometry and number theory, 407-451, Progr. Math., 253, Birkhäuser Boston (2006).

[44] Y. Ihara, K. Murty, M. Shimura, *On the logarithmic derivatives of Dirichlet $L$-functions at $s = 1$*. Acta Arith. 137 (2009), no. 3, 253-276.

[45] H. Iwaniec, E. Kowalski, *Analytic number theory*. American Mathematical Society Colloquium Publications, 53. American Mathematical Society, Providence, RI 2004. xii+615 pp. ISBN: 0-8218-3633-1

[46] H. Jacquet, *On the base change problem: after J. Arthur and L. Clozel*. Number theory, trace formulas and discrete groups (Oslo, 1987), 111-124, Academic Press, Boston, MA, 1989.

[47] R. de Jong, *On the Arakelov theory of elliptic curves*. Enseign. Math. (2) 51 (2005), no. 3-4, 179-201.

[48] N. Katz, B. Mazur, *Arithmetic Moduli of Elliptic Curves*. Annals of Mathematics Studies, 108. Princeton University Press, Princeton, NJ, 1985. xiv+514 pp. ISBN: 0-691-08349-5; 0-691-08352-5

[49] R. von Kanel *Integral points on moduli schemes of elliptic curves*. Trans. London Math. Soc. 1 (2014), no. 1, 85-115.

[50] R. von Kanel, B. Matschke, *Solving S-unit, Mordell, Thue, Thue-Mahler and generalized Ramanujan-Nagell equations via Shimura-Taniyama conjecture*. Preprint arXiv: 1605.06079 (2016)
[51] M. Kenku, *On the number of Q-isomorphism classes of elliptic curves in each Q-isogeny class*. J. Number Theory 15 (1982), no. 2, 199-202.

[52] S. Kudla, M. Rapoport, T. Yang, *Derivatives of Eisenstein series and Faltings heights*. Compos. Math. 140 (2004), no. 4, 887-951.

[53] S. Kudla, M. Rapoport, T. Yang, *Modular forms and special cycles on Shimura curves*. Annals of Mathematics Studies, 161. Princeton University Press, Princeton, NJ, 2006. x+373 pp. ISBN: 978-0-691-12551-0; 0-691-12551-1

[54] Y. Lamzouri, *The distribution of Euler-Kronecker constants of quadratic fields*. J. Math. Anal. Appl. 432 (2015), no. 2, 632-653.

[55] W. Li, *L-series of Rankin type and their functional equations*. Math. Ann. 244 (1979), no. 2, 135-166.

[56] Q. Liu, *Algebraic geometry and arithmetic curves*. Oxford Graduate Texts in Mathematics, 6. Oxford Science Publications. Oxford University Press, Oxford, 2002. xvi+576 pp. ISBN: 0-19-850284-2

[57] C. Maclachlan, A. Reid, *The arithmetic of hyperbolic 3-manifolds*. Graduate Texts in Mathematics, 219. Springer-Verlag, New York, 2003. xiv+463 pp. ISBN: 0-387-98386-4

[58] L. Mai, R. Murty, *The Phragmén-Lindelöf theorem and modular elliptic curves*. The Rademacher legacy to mathematics (University Park, PA, 1992), 335-340, Contemp. Math., 166, Amer. Math. Soc., Providence, RI, (1994).

[59] G. Martin, *Dimensions of the spaces of cusp forms and newforms on \( \Gamma_0(N) \) and \( \Gamma_1(N) \).* J. Number Theory 112 (2005), no. 2, 298-331.

[60] B. Mazur, *Rational isogenies of prime degree* (with an appendix by D. Goldfeld). Invent. Math. 44 (1978), no. 2, 129-162.

[61] P. Mihăilescu, *Primary cyclotomic units and a proof of Catalan’s conjecture*. J. Reine Angew. Math. 572 (2004), 167-195.

[62] R. Murty, *Bounds for congruence primes*. Automorphic forms, automorphic representations, and arithmetic (Fort Worth, TX, 1996), 177-192, Proc. Sympos. Pure Math., 66, Part 1, Amer. Math. Soc., Providence, RI, (1999).

[63] R. Murty, H. Pasten, *Modular forms and effective Diophantine approximation*. J. Number Theory 133 (2013), no. 11, 3739-3754.

[64] J.-L. Nicolas, G. Robin, *Majorations explicites pour le nombre de diviseurs de N*. Canad. Math. Bull. 26 (1983), no. 4, 485-492.

[65] J. Oesterlé, *Nouvelles approches du “théorème” de Fermat*. Séminaire Bourbaki 30 (1988): 165-186.

[66] R. Oliver, J. Thorner. *Effective log-free zero density estimates for automorphic L-functions and the Sato-Tate conjecture*. Preprint arXiv: 1505.03122v4

[67] K. Prasanna, *Completeness of a ratio of Petersson norms and level-lowering congruences*. Annals of Mathematics, Second Series, Vol. 163, No. 3 (May, 2006), pp. 901-967.

[68] K. Ribet, *Congruence relations between modular forms*. Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983), 503-514, PWN, Warsaw, 1984.

[69] K. Ribet, *On modular representations of Gal(\( \overline{\mathbb{Q}}/\mathbb{Q} \)) arising from modular forms*. Invent. Math. 100 (1990), no. 2, 431-476.

[70] K. Ribet, *On the component groups and the Shimura subgroup of \( J_0(N) \).* Séminaire de Théorie des Nombres, 1987-1988 (Talence, 1987-1988), Exp. No. 6.

[71] K. Ribet, S. Takahashi, *Parametrizations of elliptic curves by Shimura curves and by classical modular curves*. Elliptic curves and modular forms (Washington, DC, 1996). Proc. Nat. Acad. Sci. U.S.A. 94 (1997), no. 21, 11110-11114.

[72] H. Shimizu, *On zeta functions of quaternion algebras*. Ann. of Math. (2) 81 1965 166-193.

[73] G. Shimura, *Construction of class fields and zeta functions of algebraic curves*. Annals of Mathematics, Second Series, Vol. 85, No. 1 (1967), pp. 58-159.

[74] A. Silverberg, *Open questions in arithmetic algebraic geometry*. Arithmetic algebraic geometry (Park City, UT, 1999), 83-142, IAS/Park City Math. Ser., 9, Amer. Math. Soc., Providence, RI, 2001.

[75] J. Silverman, *Heights and elliptic curves*. Arithmetic geometry (Storrs, Conn., 1984), 253-265, Springer, New York, (1986).

[76] B. de Smit, *ABC Triples*. (Sept 28, 2012) www.math.leidenuniv.nl/~desmit/abc

[77] C. L. Stewart, R. Tijdeman, *On the Oesterlé-Masser conjecture*. Monatsh. Math. 102 (1986), 251-257.
[78] C. L. Stewart, K. R. Yu, *On the abc conjecture*. Math. Ann. 291 (1991), no. 2, 225-230.

[79] C. L. Stewart, K. R. Yu, *On the abc conjecture. II*. Duke Math. J. 108 (2001), no. 1, 169-181.

[80] A. Surroca, *Sur l’effectivité du théorème de Siegel et la conjecture abc*. J. Number Theory 124 (2007), no. 2, 267-290.

[81] S. Takahashi, *Degrees of parametrizations of elliptic curves by Shimura curves*. J. Number Theory 90 (2001), no. 1, 74-88.

[82] T. Tatuzawa, *On Siegel’s theorem*, Japan. J. Math. 21 (1951), 163-178.

[83] R. Taylor, A. Wiles, *Ring-theoretic properties of certain Hecke algebras*, Annals of Mathematics. Second Series, (1995) 141 (3): 553-572.

[84] R. Tijdeman, *On the equation of Catalan*. Acta Arith. 29 (2): 197-209.

[85] E. Ullmo, *Hauteur de Faltings de quotients de $J_0(N)$, discriminants d’algèbres de Hecke et congruences entre formes modulaires*. Amer. J. Math. 122 (2000), no. 1, 83-115.

[86] P. Vojta, *Diophantine approximations and value distribution theory*. Lecture Notes in Mathematics, 1239. Springer-Verlag, Berlin (1987). x+132 pp. ISBN: 3-540-17551-2

[87] P. Voutier, *An effective lower bound for the height of algebraic numbers*. Acta Arith. 74 (1996), no. 1, 81-95.

[88] A. Weiss, *The least prime ideal*. J. Reine Angew. Math. 338 (1983), 56-94.

[89] A. Wiles, *Modular elliptic curves and Fermat’s last theorem*. Annals of Mathematics. Second Series, (1995) 141 (3): 443-551.

[90] K. R. Yu, *Linear forms in $p$-adic logarithms. III*. Compositio Math. 91 (1994), no. 3, 241-276.

[91] K. R. Yu, *$p$-adic logarithmic forms and group varieties. III*. Forum Math. 19 (2007), no. 2, 187-280.

[92] X. Yuan, S. Zhang, W. Zhang, *The Gross-Zagier formula on Shimura curves*. Annals of Mathematics Studies, No. 184, Princeton University Press, 2013.

[93] X. Yuan, S.-W. Zhang, *On the averaged Colmez Conjecture*. Preprint arXiv: 1507.06903 (2015)

[94] S.-W. Zhang, *Heights of Heegner points on Shimura curves*. Ann. of Math. (2) 153 (2001), no. 1, 27-147.

[95] S.-W. Zhang, *Gross-Schoen cycles and dualising sheaves*. Invent. Math. 179 (2010), no. 1, 1-73.

Department of Mathematics
Harvard University
Science Center
1 Oxford Street
Cambridge, MA 02138, USA
E-mail address, H. Pasten: hpasten@gmail.com