Integral points on moduli schemes of elliptic curves

Rafael von Känel

Abstract
We combine the method of Faltings (Arakelov, Paršin, Szpiro) with the Shimura–Taniyama conjecture to prove effective finiteness results for integral points on moduli schemes of elliptic curves. For several fundamental Diophantine problems, such as for example $S$-unit and Mordell equations, this gives an effective method which does not rely on Diophantine approximation or transcendence techniques.

1. Introduction
Building on the work of Wiles [69] and Taylor–Wiles [68], the Shimura–Taniyama conjecture was finally established by Breuil–Conrad–Diamond–Taylor [14]. In this paper, we combine the method of Faltings (Arakelov, Paršin, Szpiro) with the Shimura–Taniyama conjecture to prove effective finiteness results for integral points on moduli schemes of elliptic curves. For several fundamental Diophantine problems, such as for example $S$-unit and Mordell equations, this gives an effective method which does not rely on Diophantine approximation or transcendence techniques. In what follows in the introduction, we describe in more detail the content of this paper.

1.1. Integral points on moduli schemes of elliptic curves
To provide some motivation for the study of integral points on moduli schemes of elliptic curves, we discuss in the following section fundamental Diophantine equations which are related to such moduli schemes. For any $\beta \in \mathbb{Q}$, we denote by $h(\beta)$ the usual (absolute) logarithmic Weil height defined for example in [12, p. 16].

1.1.1. $S$-unit and Mordell equations
Let $S$ be a finite set of rational prime numbers. We define $N_S = 1$ if $S$ is empty and $N_S = \prod p$ with the product taken over all $p \in S$ otherwise. Let $\mathcal{O}^\times$ denote the units of $\mathcal{O} = \mathbb{Z}[1/N_S]$. First, we consider the classical $S$-unit equation

$$x + y = 1, \quad (x, y) \in \mathcal{O}^\times \times \mathcal{O}^\times. \quad (1.1)$$

The study of $S$-unit equations has a long tradition and it is known that many important Diophantine problems are encapsulated in the solutions of (1.1). For example, any upper bound for $h(x)$ which is linear in terms of $\log N_S$ is equivalent to a version of the $(abc)$-conjecture. Mahler [46], Faltings [26] and Kim [41] proved finiteness of (1.1) by completely different methods. Moreover, Baker’s method [4] or a method of Bombieri [9] both allow in principle to find all solutions of any $S$-unit equation. We will briefly discuss the methods of Baker, Bombieri, Faltings, Kim and Mahler in Paragraph 7.2.1. In addition, we now point out that...
Frey remarked in [29, p. 544] that the Shimura–Taniyama conjecture implies finiteness of (1.1). It turns out that one can make Frey’s remark in [29] effective and one obtains for example the following explicit result (see Corollary 7.2): Any solution \((x, y)\) of the \(S\)-unit equation (1.1) satisfies
\[
h(x), h(y) \leq \frac{3}{2} n_S (\log n_S)^2 + 65, \quad n_S = 2^7 N_S.
\]

(After we submitted this paper, Hector Pasten informed us about his joint work with Murty [52] in which they independently obtain a (slightly) better version [52, Theorem 1.1] of the displayed height bound by using a similar method; see below Corollary 7.2 for more details. We would like to thank Hector Pasten for informing us.) Frey uses inter alia his construction of Frey curves. This construction is without doubt brilliant, but rather ad hoc and thus works only in quite specific situations. The starting point for our generalizations are the following two observations: The solutions of (1.1) correspond to integral points on the moduli scheme \(\mathbb{P}^1_{\mathbb{Z}[1/2]} - \{0, 1, \infty\}\), and the construction of Frey curves may be viewed as an explicit Parshin construction induced by forgetting the level structure on the elliptic curves parameterized by the points of \(\mathbb{P}^1_{\mathbb{Z}[1/2]} - \{0, 1, \infty\}\).

We now discuss a second fundamental Diophantine equation which is related to integral points on moduli schemes. For any non-zero \(a \in \mathcal{O}\), one obtains a Mordell equation
\[
y^2 = x^3 + a, \quad (x, y) \in \mathcal{O} \times \mathcal{O}.
\]

We shall see in Subsection 7.3 that this Diophantine equation is a priori more difficult than (1.1). In fact, the resolution of (1.2) in \(\mathbb{Z} \times \mathbb{Z}\) is equivalent to the classical problem of finding all perfect squares and perfect cubes with given difference, which goes back at least to Bachet 1621. Mordell [49, 50], Faltings [26] and Kim [42] showed finiteness of (1.2) by using completely different proofs, and the first effective result for Mordell’s equation was provided by Baker [6]; see Paragraph 7.3.1 where we briefly discuss methods which show finiteness of (1.2). On working out explicitly the method of this paper for the moduli schemes corresponding to Mordell equations, we obtain a new effective finiteness proof for (1.2). More precisely, if \(a_S = 2^{83} 3^5 N_S^2 \prod p^{\text{min}(2, \text{ord}_p(a))}\) with the product taken over all rational primes \(p \notin S\) with \(\text{ord}_p(a) \geq 1\), then Corollary 7.4 proves that any solution \((x, y)\) of (1.2) satisfies
\[
h(x), h(y) \leq h(a) + 4a_S (\log a_S)^2.
\]

This inequality allows in principle to find all solutions of any Mordell equation (1.2) and it provides in particular an entirely new proof of Baker’s classical result [6]. Moreover, the displayed estimate improves the actual best upper bounds for (1.2) in the literature and it refines and generalizes Stark’s theorem [65]; see Subsection 7.3 for more details.

We observe that \(S \mapsto \text{Spec}(\mathbb{Z}) - S\) defines a canonical bijection between the set of finite sets of rational primes and the set of non-empty open subschemes of \(\text{Spec}(\mathbb{Z})\). In what follows in this paper (except Subsections 7.2–7.4), we will adapt our notation to the algebraic geometry setting and the symbol \(S\) will denote a base scheme.

1.1.2. Integral points on moduli schemes of elliptic curves More generally, we now consider integral points on arbitrary moduli schemes of elliptic curves. We denote by \(T\) and \(S\) non-empty open subschemes of \(\text{Spec}(\mathbb{Z})\), with \(T \subseteq S\). Let \(Y = \mathcal{M}_T\) be a moduli scheme of elliptic curves, which is defined over \(S\), and let \(|\mathcal{P}|_T\) be the maximal (possibly infinite) number of distinct level \(\mathcal{P}\)-structures on an arbitrary elliptic curve over \(T\); see Section 3 for the definitions. We denote by \(Y(T)\) the set of \(T\)-points of the \(S\)-scheme \(Y\). Let \(h_\phi\) be the pullback of the relative Faltings height by the canonical forget \(\mathcal{P}\)-map \(\phi\), defined in (3.3). Write \(\nu_T = 12^3 \prod p^2\) with the product taken over all rational primes \(p\) not in \(T\). We obtain in Theorem 7.1 the following result.
Theorem A. The following statements hold.

(i) The cardinality of $Y(T)$ is at most \( \frac{2}{3} |\mathcal{P}|_T \nu_T \prod (1 + 1/p) \) with the product taken over all rational primes $p$ which divide $\nu_T$.

(ii) If $P \in Y(T)$, then $h_\phi(P) \leq \frac{1}{2} \nu_T (\log \nu_T)^2 + 9$.

If the moduli problem $\mathcal{P}$ is given with $|\mathcal{P}|_T < \infty$, then the explicit upper bound for the height $h_\phi$ in (ii) has the following application: in principle, one can determine the abstract set $Y(T)$ up to a canonical bijection; see the discussion surrounding (3.3). Part (i) gives a quantitative finiteness result for $Y(T)$ provided that $|\mathcal{P}|_T < \infty$. In fact, most moduli schemes of interest in arithmetic, in particular, all explicit moduli schemes considered in this paper, trivially satisfy $|\mathcal{P}|_T < \infty$. However, any scheme over an arbitrary $\mathbb{Z}[\frac{1}{2}]$-scheme is a moduli scheme of elliptic curves (see Section 3) and thus there exist many open subschemes $S \subset \text{Spec}(\mathbb{Z})$ and moduli schemes $Y$ over $S$ such that $Y(S)$ is infinite.

In addition, we show that the Shimura–Taniyama conjecture $:=(ST)$ allows one to deal with other classical Diophantine problems. For example, we consider cubic Thue equations, we derive an exponential version of Szpiro’s discriminant conjecture for any elliptic curve over $\mathbb{Q}$, and we deduce an effective Shafarevich conjecture for elliptic curves over $\mathbb{Q}$.

We remark that the theory of logarithmic forms gives more general versions of the results discussed so far, see [38] and our forthcoming paper ‘Height and conductor of elliptic curves’. However, the approach via $(ST)$ has other advantages. For instance, in the two examples which we worked out explicitly, we obtained upper bounds with numerical constants that are smaller than those coming from the theory of logarithmic forms. Furthermore, in the forthcoming joint work ‘Solving $S$-unit and Mordell equations via Shimura–Taniyama conjecture’ with Matschke, we estimate more precisely the quantities appearing in our proofs to further improve our final numerical constants. This will allow us to practically resolve $S$-unit and Mordell equations with ‘small’ parameters. In fact, the practical resolution of these Diophantine equations is still a challenging problem; see, for example, Gebel–Pethö–Zimmer [30] for partial results on Mordell’s equation. We also point out that $(ST)$ has in addition the potential to find the solutions of Diophantine equations without using height bounds. For instance, we shall see in the proof of Theorem A that integral points on moduli schemes of elliptic curves correspond to elliptic curves over $\mathbb{Q}$ of bounded conductor, which in turn correspond by $(ST)$ to certain newforms of bounded level and such newforms can be computed by Cremona [20]. We refer the reader to the forthcoming paper ‘Solving $S$-unit and Mordell equations via Shimura–Taniyama conjecture’ for details.

1.1.3. Principal ideas of Theorem A We continue the notation of the previous section. Let $T \subseteq S \subseteq \text{Spec}(\mathbb{Z})$ be as above and suppose that $Y = M_P$ is a moduli scheme over $S$ with $|\mathcal{P}|_T < \infty$. To describe our finiteness proofs for $Y(T)$, we denote by $M(T)$ the set of isomorphism classes of elliptic curves over $T$. Forgetting the level structure $\mathcal{P}$ induces a canonical map (see Lemma 3.1)

$$Y(T) \longrightarrow M(T) \quad (1.3)$$

which has fibers of cardinality at most $|\mathcal{P}|_T < \infty$. Hence to show finiteness of $Y(T)$, it suffices to control $M(T)$. This can be done in two steps: (a) Finiteness of $M(T)$ up to isogenies and (b) finiteness of each isogeny class of $M(T)$. Mazur–Kenku [40] implies (b), and (a) follows from $(ST)$ [14] which provides an abelian variety $J_0(\nu_T)$ over $\mathbb{Q}$ of controlled dimension such that the generic fiber $E_{\mathbb{Q}}$ of any $[E] \in M(T)$ is a quotient

$$J_0(\nu_T) \longrightarrow E_{\mathbb{Q}}. \quad (1.4)$$
This leads to Theorem A(i). To prove the explicit height bounds in Theorem A(ii), it suffices by (1.3) to control the relative Faltings height $h(E)$ of $E_\mathbb{Q}$ (see Section 2). We first work out explicitly an estimate of Frey [28] which relies on several non-trivial results, including [40]: If $E_\mathbb{Q}$ is modular, then Frey estimates $h(E)$ in terms of the modular degree $m_f$ of the newform $f$ associated with $E_\mathbb{Q}$. The theory of modular forms allows one to bound $m_f$ in terms of the level $N_E$ of $f$, and $(ST)$ says that $E_\mathbb{Q}$ is modular. Hence one obtains an estimate for $h(E)$ in terms of $N_E$, which then leads to Theorem A(ii).

To obtain upper bounds for heights on $Y(T)$ which are different to $h_\phi$, it remains to work out height comparisons. In the two examples discussed above, one can do this explicitly by using explicit formulas for certain (Arakelov) invariants of elliptic curves.

We emphasize that the crucial ingredients for Theorem A are (1.3) and the ‘geometric’ version (1.4) of $(ST)$ which relies inter alia on the Tate conjecture proved by Faltings [26]. The other tools, such as Frey’s estimate, the theory of modular forms and the isogeny results of Mazur–Kenku [40], can be replaced by Arakelov theory and isogeny estimates; see [37]. In fact, the proof of Theorem A may be viewed as an application of the method of Faltings (Arakelov, Paršin, Szpiro) to moduli schemes of elliptic curves.

1.2. Plan of the paper

In Section 2, we recall the definition of the Faltings heights and the conductor of elliptic curves over number fields. In Section 3, we give Paršin constructions for moduli schemes of elliptic curves and in Section 4 we state a lemma which controls the variation of Faltings heights in an isogeny class of elliptic curves over $\mathbb{Q}$. In Section 5, we use the theory of modular forms to bound the modular degree. Then we prove in Section 6 an explicit height conductor inequality for elliptic curves over $\mathbb{Q}$ and we derive some applications. In Section 7, we give our effective finiteness results for integral points on moduli schemes of elliptic curves. Here we begin with the general theorem and then we consider in detail the special cases of $S$-unit and Mordell equations. In addition, we compare our results with the literature and we discuss further Diophantine applications.

We mention that the setting of certain preliminary sections will be more general than is necessary for the proofs of the main results of this paper, since we wish also to look ahead to future work (see [37], our forthcoming paper ‘Height and conductor of elliptic curves’, and our forthcoming joint work ‘Integral points on certain Shimura varieties’ with Kret).

1.3. Conventions and notations

We identify a non-zero prime ideal of the ring of integers $O_K$ of a number field $K$ with the corresponding finite place $v$ of $K$ and vice versa. We write $N_v$ for the number of elements in the residue field of $v$, we denote by $v(\alpha)$ the order of $v$ in a fractional ideal $\alpha$ of $K$ and we write $v | \alpha$ (respectively, $v \nmid \alpha$) if $v(\alpha) \neq 0$ (respectively, $v(\alpha) = 0$). If $E$ is an elliptic curve over $K$ with semi-stable reduction at all finite places of $K$, then we say that $E$ is semi-stable.

By log we mean the principal value of the natural logarithm and we define the maximum of the empty set and the product taken over the empty set as 1. For any set $M$, we denote by $|M|$ the (possibly infinite) number of distinct elements of $M$. Let $f_1$ and $f_2$ be real-valued functions on $M$. We write $f_1 \ll f_2$ if there exists a constant $c$ such that $f_1 \leq cf_2$. For any map $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, we write $f_1 \ll_{\epsilon} f_2^{f(\epsilon)}$ if for all $\epsilon > 0$ there exists a constant $c(\epsilon)$, depending only on $\epsilon$, such that $f_1 \leq c(\epsilon)f_2^{f(\epsilon)}$.

Let $S$ be an arbitrary scheme. We often identify an affine scheme $S = \text{Spec}(R)$ with the ring $R$. If $T$ and $Y$ are $S$-schemes, then we denote by $Y(T) = \text{Hom}_S(T,Y)$ the set of $S$-scheme morphisms from $T$ to $Y$ and we write $Y_T = Y \times_S T$ for the base change of $Y$ from $S$ to $T$. We say that a scheme $Y$ is defined over $S$ if $Y$ is a $S$-scheme.
2. Height and conductor of elliptic curves

Let $K$ be a number field and let $E$ be an elliptic curve over $K$. In the first part of this section, we recall the definition of the relative and the stable Faltings height of $E$. In the second part, we define the conductor $N_E$ of $E$.

2.1. Faltings heights

We begin to define the relative and stable Faltings height of $E$ following [26, p. 354]. Let $B$ be the spectrum of the ring of integers of $K$. We denote by $E$ the Néron model of $E$ over $B$, with zero section $e : B \to E$. Let $\Omega^1$ be the sheaf of relative differential 1-forms of $E/B$. We now metrize the line bundle $\omega = e^*\Omega^1$ on $B$. For any embedding $\sigma : K \hookrightarrow \mathbb{C}$, we denote by $E^\sigma$ the base change of $E$ to $\mathbb{C}$ with respect to $\sigma$. We choose a non-zero global section $\alpha$ of $\omega$. Let $\|\alpha\|_\sigma$ be the positive real number that satisfies $\|\alpha\|_\sigma^2 = \frac{i}{2} \int_{E^\sigma(\mathbb{C})} \alpha_{\sigma} \wedge \overline{\alpha_{\sigma}}$, where $\alpha_{\sigma}$ denotes the holomorphic differential form on $E^\sigma$ which is induced by $\alpha$. Then the relative Faltings height $h(E)$ of $E$ is the real number defined by

$$[K : \mathbb{Q}] h(E) = \log |\omega/\alpha\omega| - \sum \log \|\alpha\|_\sigma$$

with the sum taken over all embeddings $\sigma : K \hookrightarrow \mathbb{C}$. The product formula assures that this definition does not depend on the choice of $\alpha$. To see the behavior of $h$ under base change, we take a finite field extension $L$ of $K$. The Néron mapping property implies

$$h(E_L) \leq h(E). \quad (2.1)$$

This inequality can be strict and thus the height $h$ is in general not stable under base change. To obtain a stable height we may and do take a finite extension $L'$ of $K$ such that $E_{L'}$ is semi-stable. The stable Faltings height $h_F(E)$ of $E$ is defined as

$$h_F(E) = h(E_{L'}).$$

This definition does not depend on the choice of $L'$, since the formation of the identity components of the corresponding semi-stable Néron models commutes with the induced base change. In particular, inequality (2.1) becomes an equality when $h$ is replaced by $h_F$.

We shall state several of our results in terms of $h_F$ or $h$ and therefore we now briefly discuss important differences between these heights. From (2.1), we deduce

$$h_F(E) \leq h(E).$$

Further, as already observed, the height $h_F$ has the advantage over $h$ that it is stable under base change. On the other hand, $h_F$ has in general weaker finiteness properties. For instance, there are only finitely many $K$-isomorphism classes of elliptic curves over $K$ of bounded $h$, while $h_F$ is bounded on the infinite set given by the $K$-isomorphism classes of elliptic curves of any fixed $j$-invariant in $K$.

2.2. Conductor

We now define the conductor $N_E$ of an arbitrary elliptic curve $E$ over any number field $K$. Let $v$ be a finite place of $K$. We denote by $f_v$ the usual conductor exponent of $E$ at $v$, see, for example, [58, Subsection 2.1] for a definition. The conductor $N_E$ of $E$ is defined by

$$N_E = \prod N_v^{f_v}$$

with the product taken over all finite places $v$ of $K$. It holds that $f_v = 0$ if and only if $E$ has good reduction at $v$. We shall need an explicit upper bound for $f_v$. If $v \nmid 6$, then the wild part
\[\delta_v \text{ of } f_v \text{ satisfies } \delta_v = 0 \text{ and therefore we deduce that } f_v \leq \dim V_v(E) = 2 \text{ for } V_v(E) \text{ the rational } \ell\text{-adic Tate module of } E. \] Brumer and Kramer [15] obtained a general upper bound for \( f_v \) by refining earlier work of Serre [59, Subsection 4.9] and of Lockhart–Rosen–Silverman [45]. It follows from [15, Theorem 6.2] that
\[ f_v \leq 2 + 6v(2) \text{ if } v \mid 2 \quad \text{and} \quad f_v \leq 2 + 3v(3) \text{ if } v \mid 3. \] Furthermore, the examples in [15] show that (2.2) is best possible in a strong sense.

3. Paršin constructions: forgetting the level structure

Paršin [53] discovered a link between the Mordell and the Shafarevich conjecture which is now commonly known as Paršin construction or Paršin trick. This link gives a finite map from the set of rational points of a certain moduli space, where the moduli space is a curve of genus at least two which is defined over a number field.

In the first part of this section, we use the moduli problem formalism to obtain tautological Paršin constructions for moduli schemes of elliptic curves. In the second part, we explicitly work out this idea for \( \mathbb{P}^1 - \{0, 1, \infty\} \) and once punctured Mordell elliptic curves. This results in completely explicit Paršin constructions for these hyperbolic curves.

3.1. Moduli schemes

We begin to introduce some notation and terminology. Let \( S \) be an arbitrary scheme. An elliptic curve over \( S \) is an abelian scheme over \( S \) of relative dimension 1. A morphism of elliptic curves over \( S \) is a morphism of abelian schemes over \( S \). We denote by
\[ M(S) \]
the set of isomorphism classes of elliptic curves over \( S \). On following Katz–Mazur [39, p. 107], we write \( \text{(Ell)} \) for the category of elliptic curves over variable base-schemes: The objects are elliptic curves over schemes and the morphisms are given by cartesian squares of elliptic curves. Let \( \text{(Sets)} \) be the category of sets and let \( \mathcal{P} \) be a contravariant functor from \( \text{(Ell)} \) to \( \text{(Sets)} \).

An element \( \alpha \in \mathcal{P}(E) \) is called a level \( \mathcal{P} \)-structure on an elliptic curve \( E \) over \( S \). We say that \( \mathcal{P} \) is a moduli problem on \( \text{(Ell)} \) and we define
\[ |\mathcal{P}|_S = \sup |\mathcal{P}(E)| \] (3.1)
with the supremum taken over all elliptic curves \( E \) over \( S \). In other words, \( |\mathcal{P}|_S \) is the maximal (possibly infinite) number of distinct level \( \mathcal{P} \)-structures on an arbitrary elliptic curve over \( S \).

A scheme \( M_{\mathcal{P}} \) is called a moduli scheme (of elliptic curves) if there exists a moduli problem \( \mathcal{P} \) on \( \text{(Ell)} \) which is representable by an elliptic curve over \( M_{\mathcal{P}} \). The following lemma may be viewed as a tautological Paršin construction for moduli schemes.

**Lemma 3.1.** Suppose that \( Y = M_{\mathcal{P}} \) is a moduli scheme, defined over a scheme \( S \). If \( T \) is a \( S \)-scheme, then there is a map \( Y(T) \to M(T) \) with fibers of cardinality at most \( |\mathcal{P}|_T \).

**Proof.** We note that the statement is intuitively clear, since \( Y(T) \) is essentially the set of elliptic curves over \( T \) with ‘level \( \mathcal{P} \)-structure’ and the map is essentially ‘forgetting the level \( \mathcal{P} \)-structure’. We now verify that this intuition is correct.

By assumption, there exists a contravariant functor \( \mathcal{P} \) from \( \text{(Ell)} \) to \( \text{(Sets)} \) which is representable by an elliptic curve over \( Y \). Suppose \( E \) and \( E' \) are elliptic curves over a scheme \( Z \), with \( \alpha \in \mathcal{P}(E) \) and \( \alpha' \in \mathcal{P}(E') \). Then the pairs \( (E, \alpha) \) and \( (E', \alpha') \) are called isomorphic if there exists an isomorphism \( \varphi : E \to E' \) of objects in \( \text{(Ell)} \) with \( \mathcal{P}(\varphi)(\alpha') = \alpha \). Let \( F(Z) \) be the set of isomorphism classes of such pairs \( (E, \alpha) \). Then \( Z \mapsto F(Z) \) defines a contravariant functor
from the category of schemes to \((\text{Sets})\), which is representable by \(Y\) since \(\mathcal{P}\) is representable by an elliptic curve over \(Y\). Thus we obtain an inclusion \(Y(T) \hookrightarrow F(T)\), which composed with 

\[ F(T) \longrightarrow M(T) : \[(E, \alpha)] \mapsto [E] \]

gives a map \(Y(T) \rightarrow M(T)\). Suppose that \(\{(E_i, \alpha_i), 1 \leq i \leq n\}\) is the fiber of this map over a point in \(M(T)\). Then all \(E_i\) are isomorphic objects of \((\text{Ell})\). Therefore, after applying suitable isomorphisms of objects in \((\text{Ell})\), we may do assume that all \(E_i\) coincide. This shows that \(n \leq |\mathcal{P}|_T\) and then we conclude Lemma 3.1.

We call the map constructed in Lemma 3.1 the forget \(\mathcal{P}\)-map. To discuss some fairly general examples of moduli schemes we consider an arbitrary scheme \(Y\). If there exists an elliptic curve \(E\) over \(Y\), then \(Y = \mathcal{M}_E\) is a moduli scheme with \(\mathcal{P} = \text{Hom}(\text{Ell})(-, E)\). This shows, in particular, that any \(\mathbb{Z}[1/2]\)-scheme \(Y\) is a moduli scheme, since there exists an elliptic curve \(A\) over \(\mathbb{Z}[1/2]\) and the base change \(A_Y\) is an elliptic curve over \(Y\). Next, we discuss a classical example of a moduli problem. Let \(N \geq 1\) be an integer and consider the ‘naive’ level \(N\) moduli problem \(\mathcal{P}_N\) from \((\text{Ell})\) to \((\text{Sets})\), defined by 

\[ E/S \mapsto \{\text{S-group-scheme isomorphisms } (\mathbb{Z}/N\mathbb{Z})^2 = E[N]\}. \]

Here we view \((\mathbb{Z}/N\mathbb{Z})^2\) as a constant \(S\)-group-scheme and \(E[N]\) is the kernel of the \(S\)-homomorphism ‘multiplication by \(N\)’ on the elliptic curve \(E\) over \(S\). If \(\mathcal{P}_N(E/S)\) is non-empty and if \(S\) is connected, then we explicitly compute 

\[ \mathcal{P}_N(E/S) \cong \{\mathbb{Z}/N\mathbb{Z}\text{-bases of } (\mathbb{Z}/N\mathbb{Z})^2\}. \tag{3.2} \]

If \(N \geq 3\), then [39, Corollary 4.7.2] gives that \(\mathcal{P}_N\) is a representable moduli problem on \((\text{Ell})\), with moduli scheme \(Y(N) = \mathcal{M}_{\mathcal{P}_N}\) a smooth affine curve over \(\text{Spec}(\mathbb{Z}[1/N])\).

In the remaining of this section, we give two propositions. Their proofs consist essentially of working out explicitly Lemma 3.1 for particular moduli schemes, see the remarks given below the proofs of Propositions 3.2 and 3.4, respectively.

3.2. Explicit constructions

We introduce and recall some notation. Let \(K\) be a number field and write \(B\) for the spectrum of the ring of integers \(\mathcal{O}_K\) of \(K\). In the remaining of this section, we denote by 

\[ S \longrightarrow B \]

either a non-empty open subscheme of \(B\) or the spectrum of the function field \(K\) of \(B\) and we write \(\mathcal{O} = \mathcal{O}_S(S)\). Let \(E\) be an elliptic curve over \(S\). We denote by \(h(E)\) and by \(h_F(E)\) the relative and the stable Faltings height of the generic fiber \(E_K\) of \(E\), respectively, see Section 2 for the definitions. Let \(N_E\) be the conductor of \(E_K\) defined in Subsection 2.2 and let \(\Delta_E\) be the norm from \(K\) to \(\mathbb{Q}\) of the usual minimal discriminant ideal of \(E_K\) over \(K\). We observe that \(h(E), h_F(E), N_E\) and \(\Delta_E\) define real-valued functions on \(M(S)\).

Let \(Y = \mathcal{M}_E\) be a moduli scheme defined over \(S\), let \(T\) be a non-empty open subscheme of \(S\) and let \(\phi : Y(T) \rightarrow M(T)\) be the forget \(\mathcal{P}\)-map from Lemma 3.1. On pulling back the relative Faltings height \(h\) by \(\phi\), we obtain a height \(h_{\phi}\) on \(Y(T)\) defined by 

\[ h_{\phi}(P) = h(\phi(P)), \quad P \in Y(T). \tag{3.3} \]

The height \(h_{\phi}\) has the following properties: If \(|\mathcal{P}|_T < \infty\), then Lemma 3.1 together with Lemma 3.5 shows that there exist only finitely many \(P \in Y(T)\) with \(h_{\phi}(P)\) bounded. Furthermore, if \(\mathcal{P}\) is given with \(|\mathcal{P}|_T < \infty\), then the proof of Lemma 3.1 together with Lemma 3.5 implies that one can, in principle, determine, up to a canonical bijection, the set of \(T\)-points \(P\) of \(Y\) with \(h_{\phi}(P)\) effectively bounded.
Let $D_K$ be the absolute value of the discriminant of $K$ over $\mathbb{Q}$, let $d = [K : \mathbb{Q}]$ be the degree of $K$ over $\mathbb{Q}$ and let $h_K$ be the cardinality of the class group of $B$. We define

$$N_T = \prod N_v$$

with the product taken over all $v \in B - T$; note that $N_T = \infty$ if $S = T = \text{Spec}(K)$. Further, we say that a non-zero $\beta \in K$ is invertible on $T$ if $\beta$ and $\beta^{-1}$ are both in $\mathcal{O}_T(T)$. For any vector $\beta$ with coefficients in $K$, we denote by $h(\beta)$ the usual absolute logarithmic Weil height of $\beta$ which is defined in [12, 1.5.6].

3.2.1. $S$-unit equations We continue the notation introduced above and we now give an explicit Paršin construction for ‘$S$-unit equations’. The solutions of such equations correspond to $S$-points of

$$\mathbb{P}_S^1 - \{0, 1, \infty\} = \text{Spec}(\mathcal{O}[z, 1/(z(1 - z))]).$$

To simplify notation, we write $X = \mathbb{P}_S^1 - \{0, 1, \infty\}$. For any $P \in X(S)$, we define $h(P) = h(z(P))$. (If $Z$ is an affine $S$-scheme, $P \in Z(S)$ and $f \in \mathcal{O}_Z(Z)$, then $f(P) \in \mathcal{O}$ denotes the image of $f$ under the ring morphism $\mathcal{O}_Z(Z) \to \mathcal{O}$ which corresponds to $P : S \to Z$.) We say that a map of sets is finite if all its fibers are finite.

**Proposition 3.2.** Suppose that $T$ is an open subscheme of $S$, with $2$ invertible on $T$. Then there exists a finite map $\phi : X(S) \to M(T)$ with the following properties.

(i) Suppose $P \in X(S)$ and $[E] = \phi(P)$. Then it holds $N_E \leq 2^d 3^{5d} N_T^2$ and $h(P) \leq 6h_F(E) + 3 \log \max(1, h_F(E)) + 42$.

(ii) There is an elliptic curve $E'$ over $K$ that satisfies $h_F(E') = h_F(E)$ and $N_{E'} \leq 2^d 3^{5d} D_K^{h_{K-1}} N_T$.

(iii) If $B$ has trivial class group, then $E'$ extends to an elliptic curve over $T$ and $N_{E'} | 2^d N_T$. If $K = \mathbb{Q}$, then $h(P) \leq 6h(E') + 11$.

In this article, we shall use Proposition 3.2 only for one-dimensional $S$ and $T$. However, the height inequalities obtained in this proposition may be also of interest for $S = T = \text{Spec}(K)$. We mention that the number $6$ in these height inequalities is optimal.

To prove Proposition 3.2, we shall use inter alia the following lemma.

**Lemma 3.3.** If $E$ is an elliptic curve over $S$, then $\log \Delta_E \leq 12d(h(E) + 4/3)$.

**Proof.** For any embedding $\sigma : K \hookrightarrow \mathbb{C}$, we take $\tau_\sigma \in \mathbb{C}$ such that the base change of $E_K$ to $\mathbb{C}$ with respect to $\sigma$ takes the form $\mathbb{C}/(\mathbb{Z} + \tau_\sigma \mathbb{Z})$ and such that $\text{im}(\tau_\sigma) \geq \sqrt{3}/2$. We write $q = \exp(2\pi i \tau_\sigma)$ and $\Delta(\tau_\sigma) = q \prod_{n=1}^{\infty} (1 - q^n)^{24}$. From [63, Proposition 1.1], we obtain

$$\log \Delta_E = 12dh(E) + \sum \log |(2\pi)^{12} \Delta(\tau_\sigma) \text{im}(\tau_\sigma)^6|$$

with the sum taken over all embeddings $\sigma : K \hookrightarrow \mathbb{C}$. Here $|\cdot|$ denotes the complex absolute value. Further, on using the elementary inequalities $\log |\Delta(\tau_\sigma)/q| \leq 24|q|/(1 - |q|)$ and $|q| \leq \exp(-\pi \sqrt{3})$, we deduce the estimate

$$\log |(2\pi)^{12} \Delta(\tau_\sigma) \text{im}(\tau_\sigma)^6| \leq 16. \quad (3.4)$$

This together with the displayed formula for $\log \Delta_E$ implies the statement.

We remark that the proof shows in addition that Faltings’ delta invariant $\delta(E_C)$ [27, p. 402] of a compact connected Riemann surface $E_C$ of genus one satisfies

$$\delta(E_C) \geq -9.$$
Indeed, this follows directly from (3.4) and Faltings’ explicit formula [27, Lemma (c), p. 417] for $\delta(E_C)$. We mention that it is an important open problem to obtain explicit lower bounds, in terms of the genus, for the Faltings delta invariant of compact connected Riemann surfaces of arbitrary positive genus.

**Proof of Proposition 3.2.** We observe that if $X(S)$ is empty, then all statements are trivial. Hence we may and do assume that $X(S)$ is not empty. We denote by $Y$ the spectrum of $\mathbb{Z}[\lambda, 1/(2\lambda(1-\lambda))]$ for $\lambda$ an ‘indeterminate’. Then we observe that

$$y^2 = x(x-1)(x-\lambda)$$

defines an (universal) elliptic curve $\mathcal{E}$ over $Y$. We take $P \in X(S)$. On using that $Y_T \cong X_T$, we obtain a morphism $T \to Y_T$ induced by $P$. Let $E$ be the fiber product of $\mathcal{E}_{Y_T} \to Y_T$ with this morphism $T \to Y_T$. Then $E$ is an elliptic curve over $T$ and therefore we see that

$$P \mapsto [E]$$

defines a map $\phi : X(S) \to M(T)$. If $P' \in X(S)$ satisfies $\phi(P) = \phi(P')$, then it follows that $z(P') = (z_1 - z_2)/(z_3 - z_2)$ with pairwise distinct $z_1, z_2, z_3 \in \{0, 1, z(P)\}$. Thus $\phi$ is finite.

We now prove (i). In what follows, we write $\lambda$ for $z(P)$ to simplify notation. The $j$-invariant $j$ of the generic fiber $E_K$ of $E$ satisfies

$$j = 2^8 \frac{(\lambda^2 - \lambda + 1)^3}{(\lambda^2 - \lambda)^2}.$$ 

This implies that $2v(\lambda) = v(j) - 8v(2)$ for any finite place $v$ of $K$ with $v(\lambda) \leq -1$ and that $|\sigma(\lambda)|^2 \leq |\sigma(j)|$ for any embedding $\sigma : K \hookrightarrow \mathbb{C}$ with $|\sigma(\lambda)| \geq 2$, where $|\cdot|$ denotes the complex absolute value. We deduce

$$h(P) \leq \frac{h(j)}{2} + 5 \log 2.$$ 

Furthermore, Pellarin’s [54, p. 240] explicit calculation of the constant in Silverman’s [63, Proposition 2.1] leads to

$$h(j) \leq 12h_F(E) + 6 \log \max(1, h_F(E)) + 75.84.$$ 

This implies an upper bound for $h(P)$ as stated in (i). Next, we prove the claimed estimate for the conductor $N_E$ of $E$. This estimate holds trivially if $T = \text{Spec}(K)$, and we now assume that $T \neq \text{Spec}(K)$. In what follows, we denote by $v$ a closed point of $B$. Let $f_v$ be the conductor exponent of $E_K$ at $v$. If $v \in T$, then $E_K$ has good reduction at $v$, since $E \to T$ is smooth and projective, and we obtain $f_v = 0$. Thus the estimates in (2.2) for $f_v$ if $v \nmid 6$ combined with $f_v \leq 2$ if $v \mid 6$ lead to an upper bound for $N_E$ as stated in (i).

To show (ii), we observe that the statement is trivial if $T = \text{Spec}(K)$. Hence we may and do assume that $T \neq \text{Spec}(K)$. As in the proof of [36, Lemma 4.1], we see that Minkowski’s theorem gives an open subscheme $U$ of $B$ with the following properties. There are at most $h_K - 1$ points in $B - U$, any $v \in B - U$ satisfies $N_v^2 \leq D_K$ and the class group of $U$ is trivial. Then we may and do take coprime elements $l, m \in \mathcal{O}_U(U)$ such that

$$\lambda = l/m.$$ 

Let $E'$ be an elliptic curve over $K$ defined by the Weierstrass equation $y^2 = x(x-l)(x-m)$. We observe that $E'$ is geometrically isomorphic to $E_K$. This implies that the $j$-invariant of $E'$ coincides with $j$ and $h_F(E') = h_F(E)$. We now prove the claimed estimate for the conductor $N_{E'}$ of $E'$. Let $\Delta$ and $c_4$ be the usual quantities associated to the above Weierstrass equation of $E'$, see [64, p. 42]. They take the form

$$\Delta = 2^4(lm(l-m))^2 \quad \text{and} \quad c_4 = 2^4((l-m)^2 + lm).$$
Let \( f'_v \) be the conductor exponent of \( E' \) at \( v \). First, we assume that \( v \in U \) with \( v \not| 2 \). If \( v(\Delta) \geq 1 \), then it follows that \( v(c_4) = 0 \), since \( l, m \in \mathcal{O}_U(U) \) are coprime and \( v \not| 2 \). This implies that the above Weierstrass equation is minimal at \( v \) and then [64, p. 196] proves that \( E' \) is semi-stable at \( v \). We conclude \( f'_v \leq 1 \). Next, we assume \( v \in U \cap T \). In the proof of (i), we showed \( f_v = 0 \). This implies that \( f'_v = 0 \), since \( E_K \) is geometrically isomorphic to \( E' \) and \( E' \) is semi-stable at \( v \). On combining the above observations, we deduce

\[ N_{E'} \leq 2^{-d} N_T \prod_v N'_v \]

with the product taken over all \( v \in B \) such that \( v \in B - U \) or \( v \not| 2 \). Therefore, on using the properties of \( U \), we see that the estimates in (2.2) for \( f'_v \) if \( v \mid 6 \) combined with \( f'_v \leq 2 \) if \( v \not| 6 \) imply an upper bound for \( N_{E'} \) as claimed in (ii).

It remains to prove (iii). We note that the first assertion of (iii) is trivial if \( T = \text{Spec}(K) \). If \( B \) has trivial class group, then we can take \( U = B \) in the proof of (ii): It follows that \( f'_v = 0 \) for any closed point \( v \in T \) and that

\[ N_{E'} \leq 2^{-d} N_T \prod_v N_v^{2+6v(2)} \]

with the product taken over all \( v \in B \) with \( v \mid 2 \). This shows that \( E' \) is the generic fiber of an elliptic curve over \( T \) and that \( N_{E'} \leq 2^d N_T \). If \( K = \mathbb{Q} \), then we obtain that \( h(P) \leq 1/2 \log |\Delta| - 2 \log 2 \), and [64, p. 257] shows that \( |\Delta| \leq 2^{12} \Delta_{E'} \). Therefore, Lemma 3.3 proves (iii). This completes the proof of Proposition 3.2. \( \square \)

We remark that the elliptic curve \( E \) over \( Y \), which appears in the above proof, represents the moduli problem \( \mathcal{P} = [\text{Legendre}] \) on \( (\text{Ell}) \) defined in [39, p. 111]. The moduli scheme \( Y \) is defined over \( \text{Spec}(\mathbb{Z}[1/2]) \). If \( 2 \) is invertible on \( S \) and if \( T = S \), then it follows that \( X(S) = Y(S) \) and that the map \( \phi : X(S) \to M(T) \) in Proposition 3.2 coincides with the map \( Y(T) \to M(T) \) in Lemma 3.1. However, to obtain our explicit inequalities in Proposition 3.2 it is necessary to take into account the particular shape of \( \mathcal{P} = [\text{Legendre}] \).

### 3.2.2. Mordell equations

We continue the notation introduced above and we now give an explicit Paršin construction for Mordell equations. For any non-zero \( a \in \mathcal{O} \), we obtain that

\[ Z = \text{Spec}(\mathcal{O}[x, y]/(y^2 - x^3 - a)) \]

defines an affine Mordell curve over \( S \). To state our next result, we have to introduce some additional notation. If \( P \in Z(S) \), then we write \( h(P) = h(x(P)) \). Let \( R_K \) be the regulator of \( K \) and let \( r_K \) be the rank of the free part of the group of units \( \mathcal{O}_K^\times \) of \( \mathcal{O}_K \).

We define

\[ \kappa = \log(D_K)/2d + 79R_K(r_K!)r_K^{3/2} \log d, \]

and we observe that \( \kappa = 0 \) when \( K = \mathbb{Q} \). The origin of the constant \( \kappa \) shall be explained below Lemma 3.5. To measure the number \( a \in \mathcal{O} \), we use inter alia the quantity

\[ r_2(a) = \prod_{v \in \mathcal{O}_K} N_{v}^{\min(2, v(a))} \]

with the product taken over all closed points \( v \in S \) with \( v(a) \geq 1 \). We observe that \( \log r_2(a) \leq dh(a) \) and if \( a \in \mathcal{O}_K \), then \( r_2(a) \leq N_{K/Q}(a) \) for \( N_{K/Q} \) the norm from \( K \) to \( \mathbb{Q} \).

**Proposition 3.4.** Suppose that \( T \) is an open subscheme of \( S \), with \( 6a \) invertible on \( T \). Then there is a map \( \phi : Z(S) \to M(T) \) with the following properties.

(i) The map \( \phi \) is finite. Furthermore, if \( \pm 1 \) are the only 12th roots of unity in \( K \), then \( \phi \) is injective.
(ii) Suppose \( P \in Z(S) \) and \(|E| = \phi(P)\). Then it holds \( N_E \leq 2^{5d}3^{3d}N_S^2 \) and \( h(P) \leq \frac{1}{2}h(a) + 8h(E) + 2\log\max(1,h_F(E)) + 8\kappa + 36.\)

(iii) If, in addition, \( T = \text{Spec}(O[1/(6a)]) \), then \( N_E \leq 2^{5d}3^{3d}D_KN_S^2r_2(a)\).

To prove Proposition 3.4, we shall use a lemma which relates heights of elliptic curves. We recall that \( E_K \) denotes the generic fiber of an elliptic curve \( E \) over \( S \). Let \( W \) be a Weierstrass model of \( E_K \) over \( B \) with discriminant \( \Delta_W \), see, for example, [44, Paragraph 9.4.4] for a definition of \( W \) and \( \Delta_W \). To measure \( W \), we take the height

\[
h(W) = \frac{1}{12} \inf_{c \in O_K^\times} h(\epsilon^{12c_4^3}, \epsilon^{12c_6^2}),
\]

where \( c_4 \) and \( c_6 \) are the usual quantities of a defining Weierstrass equation of \( W \), see [64, p. 42]. It turns out that the definition of \( h(W) \) does not depend on the choice of the defining Weierstrass equation of \( W \). We obtain the following lemma.

**Lemma 3.5.** Suppose that \( E \) is an elliptic curve over \( S \). Then there exists a Weierstrass model \( W \) of \( E_K \) over \( B \) that satisfies

\[
h(W) \leq h(E) + \frac{1}{2}\log\max(1,h_F(E)) + \kappa + 7.
\]

If \( K = \mathbb{Q} \), then this lemma would follow on calculating the constants in Silverman’s [63, Proposition 2.1, Corollary 2.3]. However, the proof of [63, Corollary 2.3] does not generalize directly to arbitrary \( K \), since it uses that the ring of integers of \( \mathbb{Q} \) has class number one and unit group \( \{\pm 1\} \). To deal with arbitrary \( K \), we apply a classical theorem of Minkowski and a result which is based on estimates for certain fundamental units of \( O_K \). This leads to a dependence of the constant \( \kappa \) on \( D_K \), \( d \) and on \( R_K \), \( r_K \), \( d \).

**Proof of Lemma 3.5.** On combining [64, p. 264] with a classical result of Minkowski, we obtain a Weierstrass model \( W \) of \( E_K \) over \( B \) of discriminant \( \Delta_W \) such that

\[
\Delta_WO_K = a^{12}\mathfrak{D}
\]

for \( \mathfrak{D} \) the minimal discriminant ideal of \( E_K \) and \( a \subseteq O_K \) an ideal with \( N_{K/\mathbb{Q}}(a)^2 \leq D_K \). For any non-zero \( \beta \in O_K \), an application of [32, Lemma 3] with \( n = 12 \) gives \( \epsilon \in O_K^\times \) such that \( dh(\epsilon^{12\beta}) \leq \log N_{K/\mathbb{Q}}(\beta) + 12d\kappa - 6\log(D_K) \). (This result relies on estimates for certain fundamental units of \( O_K \).) Hence, on using (3.7), we obtain a defining Weierstrass equation of \( W \), with quantities \( c_4, c_6 \) and discriminant \( \Delta \), such that

\[
dh(\Delta) \leq \log \Delta_E + 12d\kappa.
\]

We write \( \Delta_E = \Delta_1\Delta_2 \) with \( \Delta_1 = \exp(12dh(E) - h_F(E)) \) the ‘unstable discriminant’ and \( \Delta_2 = \Delta_E\Delta_1^{-1} \) the \( d \)th power of the ‘stable discriminant’. Let \( j \) be the \( j \)-invariant of \( E_K \). Since \( c_4, c_6 \in O_K \) satisfy \( c_6^2 = c_4^3 - 1728\Delta \) and \( j = c_4^3/\Delta \), we see

\[
dh(c_4^3, c_6^2) = \log \prod (|\Delta|_\sigma \max(|j|, |j - 1728|), |\Delta|_\sigma^{-1})),
\]

and Kodaira–Néron [64, p. 200] gives \( dh(j) = \log\Delta_2 + \sum\max(1, |j|) \). Here the product and the sum are both taken over all embeddings \( \sigma : K \hookrightarrow \mathbb{C} \) and \( |\beta|_\sigma \) denotes the complex absolute value of \( \sigma(\beta) \) for \( \beta \in K \). Then, on splitting the product according to \( |\Delta|_\sigma^{-1} > |j|_\sigma + 1728 \) and \( |\Delta|_\sigma^{-1} \leq |j|_\sigma + 1728 \), we deduce from (3.8) the estimate

\[
h(c_4^3, c_6^2) \leq \log \Delta_1/d + dh(j) + 12\kappa + \log(2 \cdot 1728).
\]

Hence, on combining \( 12h(W) \leq h(c_4^3, c_6^2) \), (3.5) and \( \log \Delta_1 + 12dh_F(E) = 12dh(E) \), we see that \( W \) has the desired property. This completes the proof of Lemma 3.5. □
The proof shows in addition that one can take in Lemma 3.5 any Weierstrass model \( W \) of \( E_K \) over \( B \) with \( N_{K/\mathbb{Q}}(\Delta_W) \lesssim D_K^2 \Delta_E \). A defining Weierstrass equation of such a \( W \) is called a quasi-minimal Weierstrass equation of \( E_K \), see [64, p. 264].

**Proof of Proposition 3.4.** If \( Z(S) \) is empty, then all statements are trivial. Hence we may and do assume that \( Z(S) \) is not empty. We write \( b = -a/1728 \). Let \( Y \) be the spectrum of \( \mathbb{Z}[1/6, c_4, b, 1/b] / (1728b - c_4^2 + c_6^2) \) for \( c_4 \) and \( c_6 \) ‘indeterminates’. We observe that

\[
t^2 = s^3 - 27c_4s - 54c_6
\]

defines an (universal) elliptic curve \( E \) over \( Y \). We take \( P \in Z(S) \). On using that \( Y_T \cong Z_T \), we obtain a morphism \( T \to Y_T \) induced by \( P \). We denote by \( E \) the fiber product of \( E_{Y_T} \to Y_T \) with this morphism \( T \to Y_T \). It follows that \( E \) is an elliptic curve over \( T \) and then \( P \mapsto [E] \) defines a map \( \phi : Z(S) \to M(T) \).

To prove (i), we observe that \( E \) is a Weierstrass model of its generic fiber \( E_K \). Hence we see that if \( P' \in Z(S) \) satisfies \( \phi(P') = \phi(P) \), then there is \( u \in K \) with \( u^4x(P') = x(P) \) and \( u^6y(P') = y(P) \), and thus \( u^{12}a = a \) since \( P, P' \in Z(S) \). Therefore, we deduce (i).

We now show (ii). Let \( W \) be the Weierstrass model of \( E_K \) over \( B \) from Lemma 3.5. We denote by \( \Delta, c_4, c_6 \) the quantities of a defining Weierstrass equation of \( W \), which we constructed in the proof of Lemma 3.5. We point out that one should not confuse these \( c_4, c_6 \in \mathcal{O}_K \) with the ‘indeterminates’ which appear in the first part of the proof. On using that \( E \) is a Weierstrass model of \( E_K \) over \( T \), we obtain \( u \in K \) that satisfies

\[
b = u^{12} \Delta, \quad x(P) = u^4c_4.
\]

Thus Lemma 3.5, (3.8) and Lemma 3.3 lead to an upper bound for \( h(P) \) as stated in (ii). To estimate the conductor \( N_E \) of \( E \) we take a closed point \( v \) of \( B \). Let \( f_v \) be the conductor exponent of \( E_K \) at \( v \). If \( v \in T \), then \( f_v = 0 \) since \( E \) is a smooth projective model of \( E_K \) over \( T \), and if \( v \not\in T \), then \( f_v \leq 2 \). Thus (2.2) implies an estimate for \( N_E \) as claimed in (ii).

To prove (iii), we may and do assume that \( T = \text{Spec}(\mathcal{O}[1/(6a)]) \). Let \( U \) (respectively, \( U' \)) be the set of points \( v \in S - T \) with \( v \not\in 6 \) such that \( E_K \) has (respectively, has not) semi-stable reduction at \( v \). We define \( \Omega = \prod_{v \in U} N_v \) and \( \Omega' = \prod_{v \in U'} N_v^2 \) and then we deduce

\[
N_E \lesssim 2^a 3^d N_S^2 \cdot \Omega \cdot \Omega'.
\]

To control the unstable part \( \Omega' \), we may and do assume that \( U' \) is not empty. We take \( v \in U' \). The classification of Kodaira–Néron [64, p. 448] gives \( \nu(\Delta) \geq 2 \). Since \( P \in Z(S) \) we see that \( E \) extends to a Weierstrass model of \( E_K \) over \( S \), with discriminant \( 6^{12}b \). Hence, if \( W \) is minimal at \( v \), then \( 2 \lesssim \nu(\Delta) \lesssim \nu(b) = \nu(a) \). Further, (3.7) implies\( \prod N_v^2 \lesssim D_K \) with the product taken over all closed points \( v \in B \) with \( W \) not minimal at \( v \). We conclude

\[
\Omega' \lesssim D_K \prod N_v^2
\]

with the product taken over all \( v \in U' \) such that \( \nu(a) \geq 2 \). To estimate the stable part \( \Omega \), we use our assumption that \( T = \text{Spec}(\mathcal{O}[1/(6a)]) \). This assumption implies that any \( v \in S - T \) with \( v \not\in 6 \) satisfies \( \nu(a) \geq 1 \). Therefore, we obtain

\[
\Omega \lesssim \prod N_v
\]

with the product taken over all \( v \in U \) such that \( \nu(a) \geq 1 \). On combining the displayed inequalities, we deduce (iii). This completes the proof of Proposition 3.4.

We conclude this section with the following remarks. The elliptic curve \( E \) over \( Y \), which appears in the proof of Proposition 3.4, represents a moduli problem \([\Delta = b]\) on \((Ell)\). Here the moduli problem \([\Delta = b]\) is defined similarly as \([\Delta = 1]\) in [39, p. 70], but with 1 replaced by the number \( b \) which appears in the proof of Proposition 3.4.
The above propositions show that to solve $S$-unit and Mordell equations, it suffices to estimate effectively $h(E)$ in terms of $N_E$ for any elliptic curve $E$ over $K$. In this paper, we shall prove such estimates for $K = \mathbb{Q}$ and we refer to our forthcoming paper ‘Height and conductor of elliptic curves’ for arbitrary number fields $K$.

In the special case of $S$-unit and Mordell equations, it is possible to give ad hoc Parˇsin constructions which do not use the moduli problem formalism. For example, ‘Frey–Hellegouarch curves’ provide in principle such a construction for $S$-unit equations. However, using the moduli problem formalism gives more conceptual constructions, which generalize several known examples such as ‘Frey–Hellegouarch curves’.

4. Variation of Faltings heights under isogenies

In this section, we give a result which controls the variation of Faltings heights under isogenies of elliptic curves over $\mathbb{Q}$. For any elliptic curve $E$ over $\mathbb{Q}$, we denote by $h(E)$ the relative Faltings height of $E$ defined in Section 2. We obtain the following lemma.

**Lemma 4.1.** If $E$ and $E'$ are $\mathbb{Q}$-isogenous elliptic curves over $\mathbb{Q}$, then

$$|h(E) - h(E')| \leq \frac{1}{2} \log 163.$$ (5.2)

To prove this lemma, we combine a result of Faltings–Raynaud in [56] with the classification of cyclic $\mathbb{Q}$-isogenies of elliptic curves over $\mathbb{Q}$ of Mazur–Kenku [40, 48].

**Proof of Lemma 4.1.** If $\psi : E \to E'$ is a $\mathbb{Q}$-isogeny, then [56, Corollaire 2.1.4] gives that $|h(E) - h(E')| \leq \frac{1}{2} \log \deg(\psi)$. Let $\varphi : E \to E'$ be a $\mathbb{Q}$-isogeny of minimal degree among all $\mathbb{Q}$-isogenies $E \to E'$. This isogeny $\varphi$ is cyclic, since otherwise it factors through multiplication by an integer which contradicts the minimality of $\deg(\varphi)$. Therefore, [40, Theorem 1] gives that $\deg(\varphi) \leq 163$ and then we deduce Lemma 4.1. $\square$

5. Modular forms and modular curves

In the first part of this section, we collect results from the theory of modular forms. In the second part, we work out an explicit upper bound for the modular degree.

5.1. Cusp forms

We begin to collect standard results for cusp forms which are given, for example, in the books of Shimura [61] or Diamond and Shurman [21]. We take an integer $N \geq 1$ and we consider the classical congruence subgroup $\Gamma_0(N) \subset \text{SL}_2(\mathbb{Z})$. Let $S_2(\Gamma_0(N))$ be the complex vector space of cusp forms of weight 2 with respect to $\Gamma_0(N)$. Let $X_0(N)$ and $X(1)$ smooth, projective and geometrically connected models over $\mathbb{Q}$ of the modular curves associated to $\Gamma_0(N)$ and $\text{SL}_2(\mathbb{Z})$, respectively. Throughout Section 5, we denote by $d$ the degree of the natural projection $X_0(N) \to X(1)$. The dimension of $S_2(\Gamma_0(N))$ coincides with the genus $g$ of $X_0(N)$. Furthermore, it holds

$$g \leq d/12 \quad \text{and} \quad d = N \prod (1 + 1/p) \text{ if } N \geq 2$$ (5.1)

with the product taken over all rational primes $p$ which divide $N$. Let $f \in S_2(\Gamma_0(N))$ be a non-zero cusp form. If $\text{div}(f)$ denotes the usual rational divisor on $X_0(N) \subset \mathbb{Q}$ of $f$, then

$$\deg(\text{div}(f)) = d/6.$$ (5.2)
For any rational integer $n \geq 1$, we denote by $a_n(f)$ the $n$th Fourier coefficient of the cusp form $f$. Further, we say that $f$ is normalized if $a_1(f) = 1$.

We next review properties of the basis of $S_2(\Gamma_0(N))$ constructed by Atkin and Lehner [3, Theorem 5]. Let $S_2(\Gamma_0(N))^{\text{new}}$ be the new subspace of $S_2(\Gamma_0(N))$ and let $S_2(\Gamma_0(N))^{\text{old}}$ be the old subspace of $S_2(\Gamma_0(N))$. There is a decomposition

$$S_2(\Gamma_0(N)) = S_2(\Gamma_0(N))^{\text{new}} \oplus S_2(\Gamma_0(N))^{\text{old}},$$

which is orthogonal with respect to the Petersson inner product $(\cdot, \cdot)$ on $S_2(\Gamma_0(N))$. We say that $f$ is a newform of level $N$ if $f \in S_2(\Gamma_0(N))^{\text{new}}$ is normalized and if $f$ is an eigenform for all Hecke operators on $S_2(\Gamma_0(N))$. The set $B^{\text{new}}$ of newforms of level $N$ is an orthogonal basis of $S_2(\Gamma_0(N))^{\text{new}}$ with respect to $(\cdot, \cdot)$. Moreover, there exists a basis $B^{\text{old}}$ of $S_2(\Gamma_0(N))^{\text{old}}$ with the property that any $f \in B^{\text{old}}$ takes the form

$$f(\tau) = f_M(m\tau), \quad \tau \in \mathbb{C}, \quad \text{im}(\tau) > 0 \quad (5.3)$$

with $M \in \mathbb{Z}_{\geq 1}$ a proper divisor of $N$, with $m \in \mathbb{Z}_{\geq 1}$ a divisor of $N/M$ and with $f_M$ a newform of level $M$. Conversely, any $f \in S_2(\Gamma_0(N))$ which is of the form (5.3) is in $B^{\text{old}}$. We say that $B = B^{\text{new}} \cup B^{\text{old}}$ is the Atkin–Lehner basis for $S_2(\Gamma_0(N))$.

5.2. Modular degree

Let $f \in S_2(\Gamma_0(N))$ be a newform of level $N \geq 1$, with all Fourier coefficients rational integers. In this section, we estimate the modular degree of $f$ in terms of $N$.

We begin with the definition of the modular degree $m_f$ of $f$. Let $J_0(N) = \text{Pic}^0(X_0(N))$ be the Jacobian variety of $X_0(N)$. We denote by $\mathbb{T}_2$ the subring of the endomorphism ring of $J_0(N)$, which is generated over $\mathbb{Z}$ by the usual Hecke operators $T_n$ for all $n \geq 1$. Let $I_f$ be the kernel of the ring homomorphism $\mathbb{T}_2 \to \mathbb{Z}[\{a_n(f)\}] = \mathbb{Z}$ which is induced by $T_n \mapsto a_n(f)$. The image $I_fJ_0(N)$ of $J_0(N)$ under $I_f$ is connected and the quotient

$$E_f = J_0(N)/I_fJ_0(N) \quad (5.4)$$

is an abelian variety over $\mathbb{Q}$ of dimension $[\mathbb{Q}(\{a_n(f)\}) : \mathbb{Q}] = 1$. The cusp $\infty$ of $X_0(N)$ is a $\mathbb{Q}$-rational point of $X_0(N)$. We denote by $\iota : X_0(N) \hookrightarrow J_0(N)$ the usual embedding over $\mathbb{Q}$, which maps the cusp $\infty$ to the zero element of $J_0(N)$. On composing the embedding $\iota$ with the natural projection $J_0(N) \to J_0(N)/I_fJ_0(N)$, we obtain a finite morphism

$$\varphi_f : X_0(N) \to E_f.$$

The modular degree $m_f$ of $f$ is defined as the degree of the finite morphism $\varphi_f$.

To estimate $m_f$ we shall use properties of the congruence number $r_f$ of $f$. We recall that $r_f$ is the largest integer such that there exists a cusp form $f_c \in S_2(\Gamma_0(N))$, with rational integer Fourier coefficients, which satisfies

$$(f, f_c) = 0 \quad \text{and} \quad a_n(f) \equiv a_n(f_c) \mod (r_f), \quad n \geq 1. \quad (5.5)$$

It is known that the modular degree $m_f$ and the congruence number $r_f$ are related. For example, the arguments in Zagier’s article [70, Section 5] give

$$m_f \mid r_f. \quad (5.6)$$

We note that Zagier’s arguments are based on ideas of Ribet, see [57] and the references therein. In fact, Zagier [70, p. 381] attributes the divisibility result (5.6) to Ribet. Further, we mention that Cojocaru and Kani [18, Theorem 1.1] gave a detailed exposition of a proof of (5.6) and Agashe–Ribet–Stein generalized (5.6) in [2, Theorem 3.6].

For any real number $r$, we define $[r] = \max(m \in \mathbb{Z}, m \leq r)$, and for any integer $n$, we denote by $\tau(n)$ the number of positive integers which divide $n$. The author is grateful to Richard Taylor for proposing a strategy to prove an upper bound for $m_f$. 


Lemma 5.1. Let $N \geq 1$ be an integer. Suppose $f \in S_2(\Gamma_0(N))$ is a newform of level $N$, with all Fourier coefficients rational integers. Then the following statements hold.

(i) The modular degree $m_f$ of $f$ satisfies $\log m_f \leq \frac{1}{2} N(\log N)^2$.

(ii) More precisely, let $g$ be the genus of $X_0(N)$ and let $d$ be the degree of the natural projection $X_0(N) \to X(1)$. Then there exists a subset $J \subset \{1, \ldots, \lfloor d/6 \rfloor \}$ of cardinality $g$, which is independent of $f$, such that $m_f \leq g^{3/2} \prod_{j \in J} \tau(j) j^{1/2}$.

Proof. We first show (ii). It follows from (5.6) that $m_f \leq r_f$. To estimate $r_f$ we reduce the problem to solve (by Cramer’s rule) explicitly a system of linear Diophantine equations.

Let $I = \{1, \ldots, g\}$, let $J$ be a finite non-empty set of positive integers and put $\delta = |J|$. We write $l = \lfloor d/6 \rfloor$ and we denote by $B = \{f_i, i \in I\}$ the Atkin-Lehner basis for $S_2(\Gamma_0(N))$, see Subsection 5.1. To show that the linear morphism $F(J) = (a_j(f_i)) : \mathbb{C}^g \to \mathbb{C}^g$ is surjective for $J = \{1, \ldots, l\}$, we assume the contrary and deduce a contradiction. If $F(J)$ is not surjective for $J = \{1, \ldots, l\}$, then we obtain a non-zero $f_0 \in S_2(\Gamma_0(N))$ with Fourier expansion $\sum_{n \geq 1} a_n(f_0) g^n$. This $f_0$ vanishes at infinity with order exceeding $d/6$, contradicting (5.2). We conclude that $F(J)$ is surjective for $J = \{1, \ldots, l\}$. Therefore, we may and do take $J \subset \{1, \ldots, l\}$ such that $F = F(J)$ is an isomorphism.

We claim that $r_f \leq |\det(F)|$. To verify this claim, we take $(k_i) \in \mathbb{C}^g$ such that $f_0 \in S_2(\Gamma_0(N))$ from (5.5) takes the form $f_0 = \sum_{i} k_i f_i$. Properties of $F$ show that we may and do take $f_i = f$ and that $(f, f_i) = 0$ for any $i \geq 2$. This implies that $k_i = 0$, since $(f, f_i) = 0$ by (5.5). Therefore, on comparing Fourier coefficients, we see that $a_j(f_i) = \sum_{i \geq 2} k_i a_j(f_i)$ for all $j \in J$. Then (5.5) gives $y = (y_j) \in \mathbb{Z}^g$ such that any $x = (x_j) \in \mathbb{C}^g$ satisfies

$$\sum_{i \geq 2} k_i (f_i, x) = (f, x) = (f, x) + (y, x) r_f,$$

where $(x, h) = \sum_{i \in J} a_i(h) x_j$ for $h \in S_2(\Gamma_0(N))$ and $(y, x) = \sum_{j \in J} y_j x_j$. We write $b = (-1, 0, \ldots, 0) \in \mathbb{C}^g$. It follows that if $x = (x_j) \in \mathbb{C}^g$ satisfies $F(x) = b$, then $(f, x) = -1$, and $(f, x) = 0$ by the first equality of (5.7). Hence, the second equality of (5.7) shows that any solution $x = (x_j) \in \mathbb{C}^g$ of $F(x) = b$ satisfies

$$1 = (y, x) r_f.$$  (5.8)

The determinant $\det(F)$ of the isomorphism $F$ is non-zero. Thus Cramer’s rule gives $\xi \in \mathbb{Z}[\{a_j(f_i)\}]^g$ such that the unique solution $x = F^{-1}(b)$ of $F(x) = b$ takes the form

$$x = \xi \det(F)^{-1}.$$  (5.9)

To prove that $\det(F)^2 \in \mathbb{Z}$ we use (5.3). It gives that $a_j(f_i)$ is a coefficient of a newform. Thus it is an eigenvalue of a certain Hecke operator. This implies that all $a_j(f_i)$ are algebraic integers. Hence $\det(F)$ and all entries of $\xi$ are algebraic integers. Further, Galois conjugates of newforms are newforms of the same level. Therefore, on using properties of the basis $B$ discussed in Subsection 5.1, we see that any element $\sigma$ of the absolute Galois group of $\mathbb{Q}$ ‘permutes’ the rows of the matrix $F$. Hence, we obtain that any such $\sigma$ satisfies $\sigma(\det(F)) = \pm \det(F)$ and we deduce that $\det(F)^2 \in \mathbb{Z}$ as desired. Then the formulas (5.8) and (5.9) imply $r_f^2 | \det(F)^2$ which proves our claim $r_f \leq |\det(F)|$.

To estimate $|\det(F)|$ we use the Ramanujan–Petersson bounds for Fourier coefficients, which hold in particular for any newform, and thus for all $f_i \in B$ by (5.3). These bounds give $|a_i(f_i)| \leq \tau(j) j^{1/2}$ for all $i \in I$ and $j \in \mathbb{Z}_{\geq 1}$. Thus Hadamard’s determinant inequality leads to $|\det(F)| \leq g^{3/2} \prod_{j \in J} \tau(j) j^{1/2}$ and then the above inequalities imply (ii).

It remains to prove (i). Any elliptic curve over $\mathbb{Q}$ has conductor at least 11. Therefore, we may and do assume that $N \geq 11$. Next, we observe that any integer $n \geq 1$ satisfies the elementary inequalities: $(1/n) \sum_{k=1}^n \tau(k) \leq 1 + \log n$ and $\prod_{p \leq n} (1 + 1/p) \leq 1 + (\log n)/(2\log 2)$ with the product taken over all rational primes $p$ which divide $n$. Further, (5.1) shows that
2g \leq \lfloor d/6 \rfloor = l$ and hence the proof of (ii) implies that $m_f \leq |\det(F)| \leq (g!!)^{1/2} \prod \tau(j)$ with the product taken over the elements $j$ of a set $J \subset \{1, \ldots, l\}$ of cardinality $g$. Then the above inequalities and (5.1) lead to (i). This completes the proof of Lemma 5.1.

Frey [29, p. 544] remarked without proof that it is easy to show the asymptotic bound
\[ \log m_f \ll N \log N. \]
It seems that this estimate is still very far from being optimal. In fact, Frey [28] and Mai–Murty [47] showed that a certain polynomial upper bound for $m_f$ in terms of $N$ is equivalent to a certain version of the $abc$-conjecture.

The above proof shows in addition that the inequalities of Lemma 5.1 hold with $m_f$ replaced by the congruence number $r_f$ of $f$. We note that Murty [51, Corollary 6] used a similar method to prove a slightly weaker upper bound for $r_f$ in terms of $N$. Further, we mention that Agashe–Ribet–Stein proved in [2, Theorem 2.1] that any rational prime number $p$, with $\text{ord}_p(N) \leq 1$, satisfies $\text{ord}_p(m_f) = \text{ord}_p(r_f)$. Moreover, they conjectured in [2, Conjecture 2.2] that $\text{ord}_p(r_f/m_f) \leq \frac{1}{2} \text{ord}_p(N)$ for all rational prime numbers $p$.

### 6. Height and conductor of elliptic curves over $\mathbb{Q}$

In the first part of this section, we give explicit exponential versions of Frey’s height conjecture and of Szpiro’s discriminant conjecture for elliptic curves over $\mathbb{Q}$. We also derive an effective version of Shafarevich’s conjecture for elliptic curves over $\mathbb{Q}$. In the second part, we prove Propositions 6.1 and 6.4 on combining the Shimura–Taniyama conjecture with lemmas obtained in previous sections.

#### 6.1. Height, discriminant and conductor inequalities

Let $E$ be an elliptic curve over $\mathbb{Q}$. We denote by $N_E$ the conductor of $E$, and we denote by $h(E)$ the relative Faltings height of $E$. See Section 2 for the definitions of $N_E$ and $h(E)$. We now can state the following proposition which gives an exponential version of Frey’s height conjecture [28, p. 39] for all elliptic curves over $\mathbb{Q}$.

**Proposition 6.1.** If $E$ is an elliptic curve over $\mathbb{Q}$, then
\[ h(E) \leq \frac{1}{4} N_E (\log N_E)^2 + 9. \]

Let $K$ be a number field. On using a completely different method, which is based on the theory of logarithmic forms, we established (see our forthcoming paper ‘Height and conductor of elliptic curves’) a version of Proposition 6.1 for arbitrary elliptic curves over $K$. However, in the case of elliptic curves $E$ over $\mathbb{Q}$, this version provides only the weaker inequality
\[ h(E) \leq (25N_E)^{162}. \]

As in Subsection 3.2, we denote by $\Delta_E$ the norm of the usual minimal discriminant ideal of $E$. Our next result provides an explicit exponential version of Szpiro’s discriminant conjecture [67, p. 10] for elliptic curves over $\mathbb{Q}$.

**Corollary 6.2.** Any elliptic curve $E$ over $\mathbb{Q}$ satisfies
\[ \log \Delta_E \leq 3 N_E (\log N_E)^2 + 124. \]

**Proof.** This follows from Proposition 6.1, since $\log \Delta_E \leq 12h(E) + 16$ by Lemma 3.3.

On combining Arakelov theory for arithmetic surfaces with the theory of logarithmic forms, we obtained in [38] versions of Corollary 6.2 for all hyperelliptic (and certain more general)
curves over $K$. In the case of elliptic curves $E$ over $\mathbb{Q}$, we see that Corollary 6.2 improves the inequality $\log \Delta_E \leq (25 N_E)^{162}$ provided by [38, Theorem 3.3].

To state our next corollary, we denote by $h(W)$ the height of a Weierstrass model $W$ of $E$ over $\text{Spec}(\mathbb{Z})$, defined in (3.6). Let $S$ be a non-empty open subscheme of $\text{Spec}(\mathbb{Z})$ and

$$\nu_S = 12^4 N_S^3, \quad N_S = \prod p$$

with the product taken over all rational primes $p$ not in $S$. We say that an arbitrary elliptic curve $E$ over $\mathbb{Q}$ has good reduction over $S$ if $E$ has good reduction at all rational primes in $S$ (This definition is equivalent to the classical notion of good reduction outside a finite set $S$ of rational prime numbers. Indeed $S = \text{Spec}(\mathbb{Z})S$ has the structure of a non-empty open subscheme of $\text{Spec}(\mathbb{Z})$, and $E$ has good reduction outside $S$ if and only if $E$ has good reduction over $S$.) It turns out that the number $\nu_S$ has the property that any elliptic curve $E$ over $\mathbb{Q}$, with good reduction over $S$, has conductor $N_E$ dividing $\nu_S$. The Diophantine inequality in Proposition 6.1 leads to the following fully effective version of the Shafarevich conjecture [60] for elliptic curves over $\mathbb{Q}$.

**Corollary 6.3.** If $[E]$ is a $\mathbb{Q}$-isomorphism class of elliptic curves over $\mathbb{Q}$ with good reduction over $S$, then there exists a Weierstrass model $W$ of $E$ over $\text{Spec}(\mathbb{Z})$ that satisfies

$$h(W) \leq \frac{1}{2} \nu_S (\log \nu_S)^2.$$ 

In particular, there exist only finitely many $\mathbb{Q}$-isomorphism classes of elliptic curves over $\mathbb{Q}$ with good reduction over $S$ and these classes can be determined effectively.

**Proof.** We take a $\mathbb{Q}$-isomorphism class $[E]$ of elliptic curves over $\mathbb{Q}$, with good reduction over $S$. Lemma 3.5 gives a Weierstrass model $W$ of $E$ over $\text{Spec}(\mathbb{Z})$ that satisfies

$$h(W) \leq h(E) + \frac{1}{2} \log \max(1, h_F(E)) + 7,$$

where $h_F(E)$ is the stable Faltings height of $E$. Further, it holds that $h_F(E) \leq h(E)$ and (2.2) leads to $N_E \mid \nu_S$. Thus Proposition 6.1 implies Corollary 6.3. 

The first effective version of the Shafarevich conjecture for elliptic curves over $\mathbb{Q}$ is due to Coates [17, p. 426]. He applied the theory of logarithmic forms. This theory is also used in [36, Theorem] which provides a version of Corollary 6.3 for arbitrary hyperelliptic curves over $K$. In the case of elliptic curves over $\mathbb{Q}$, Corollary 6.3 improves the actual best bound $h(W) \leq (2N_S)^{1296}$ which was obtained in [36, Theorem].

We mention that in our forthcoming paper ‘Height and conductor of elliptic curves’, Section 2) and we give in addition effective asymptotic versions of the above results: $h(E) \ll_\epsilon N_E^{21+\epsilon}$, $\log \Delta_E \ll_\epsilon N_E^{21+\epsilon}$ and $h(W) \ll_\epsilon N_S^{21+\epsilon}$. Further, we there discuss that the exponent $21 + \epsilon$ is optimal for the known methods which are based on the theory of logarithmic forms. Thus these methods cannot produce inequalities as strong as those in Proposition 6.1, Corollaries 6.2 and 6.3.

We denote by $N(S)$ the number of $\mathbb{Q}$-isomorphism classes of elliptic curves over $\mathbb{Q}$, with good reduction over $S$. The explicit height estimate in Corollary 6.3 implies an explicit upper bound for $N(S)$. However, this bound would be exponential in terms of $\nu_S$. The following Proposition 6.4 gives an explicit upper bound for $N(S)$ which is polynomial in terms of $\nu_S$. The proof uses inter alia the Shimura–Taniyama conjecture and a result of Mazur–Kenku [40] on $\mathbb{Q}$-isogeny classes of elliptic curves.
Proposition 6.4. It holds that $N(S) \leq 3^2\nu_S \prod_{p|\nu_S} (1 + 1/p)$ with the product taken over all rational primes $p$ which divide $\nu_S$.

We now discuss bounds for $N(S)$ in the literature. The estimate $N(S) \ll \epsilon N_S^{1/2+\epsilon}$ was obtained by Brumer–Silverman [16, Theorem 1] and Poulakis established in [55, Theorem 2] an explicit upper bound for $N(S)$. One observes that Proposition 6.4 is better than Poulakis’ result when $N_S \leq 265$, and is worse when $N_S$ is sufficiently large. However, for sufficiently large $N_S$ the actual best estimate is due to Ellenberg, Helfgott and Venkatesh [23, 34]. Namely on refining the proof of [34, Theorem 4.5] with the upper bound in [23, Proposition 3.4], one obtains

$$N(S) \ll N_S^{0.1689}. \quad (6.1)$$

Furthermore, Brumer–Silverman [16] observed that one can considerably improve (6.1) on assuming (*): If $E$ is an elliptic curve over $\mathbb{Q}$, with vanishing $j$-invariant, then the $L$-function $L(E, s)$ of $E$ satisfies the ‘Generalized Riemann Hypothesis’ and the rank of $E(\mathbb{Q})$ is at most the order of vanishing of $L(E, s)$ at $s = 1$. More precisely, [16, Theorem 4] gives that (*) implies $N(S) \ll \epsilon N_S^2$; note that the ‘Generalized Riemann Hypothesis’ together with the ‘Birch and Swinnerton-Dyer conjecture’ implies (*).

We point out that the methods of Brumer–Silverman, Helfgott–Venkatesh and Poulakis are entirely different from the method which is used in the proof of Proposition 6.4. For example, to obtain Diophantine finiteness, they use the following tools: Brumer–Silverman [16] apply an estimate of Evertse–Silverman [25] based on Diophantine approximation, Helfgott–Venkatesh [34] use a bound of Hajdu–Herendi [33] relying on the theory of logarithmic forms, and Poulakis [55] applies an estimate of Evertse [24] based again on Diophantine approximation.

### 6.2. Proof of Propositions 6.1 and 6.4

Our main tool in the proof of Proposition 6.1 is the Shimura–Taniyama conjecture. Building on the work of Wiles [69] and Taylor–Wiles [68], Breuil–Conrad–Diamond–Taylor [14] proved this conjecture for all elliptic curves over $\mathbb{Q}$. The modularity result in [14] implies the following version of the Shimura–Taniyama conjecture. For any integer $N \geq 1$, let $X_0(N)$ be the modular curve defined in Subsection 5.1. Suppose $E$ is an elliptic curve over $\mathbb{Q}$ with conductor $N = N_E$. Then there exists a finite morphism

$$X_0(N) \longrightarrow E \quad (6.2)$$

can be a surjective morphism of abelian varieties

$$\psi : J_0(N) \longrightarrow E. \quad (6.3)$$

Proof of Proposition 6.1. Let $E$ be an elliptic curve over $\mathbb{Q}$ with conductor $N = N_E$.

(1) The version of the Shimura–Taniyama conjecture in (6.2) gives a finite morphism $\varphi : X_0(N) \to E$ of smooth projective curves over $\mathbb{Q}$. We recall that $J_0(N) = \text{Pic}^0(X_0(N))$ denotes the Jacobian of $X_0(N)$. By Picard functoriality, the morphism $\varphi$ induces a surjective $\mathbb{Q}$-morphism of abelian varieties $\psi : J_0(N) \to E$.\]
Let $A$ be the identity component of the kernel of $\psi$. It is an abelian $\mathbb{Q}$-subvariety of $J_0(N)$ of codimension 1. We denote by $(\cdot)^\vee = \text{Pic}^0(\cdot)$ the dual. The kernel of the inclusion morphism $i : A \hookrightarrow J_0(N)$ is geometrically connected. Thus $\text{ker}(i^\vee)$ is connected and has dimension 1 by the dimension formula. It follows that $E' = (\text{ker}(i^\vee))^\vee$ is an elliptic curve over $\mathbb{Q}$. The functor $\text{Pic}^0(\cdot)$ is exact for abelian varieties over fields. Hence, on dualizing twice, we obtain a surjective $\mathbb{Q}$-morphism $\psi : J_0(N) \to E'$ with kernel $(A^\vee)^\vee \cong A$. We deduce that $E'$ is $\mathbb{Q}$-isogenous to $E$ (For example, Poincaré’s reducibility theorem gives an abelian $\mathbb{Q}$-subvariety $E''$ of $J_0(N)$ such that addition induces a $\mathbb{Q}$-isogeny $A \times_{\mathbb{Q}} E'' \to J_0(N)$. Then on using that $A$ is a $\mathbb{Q}$-subgroup scheme of the kernels of the surjective $\mathbb{Q}$-morphisms $\psi : J_0(N) \to E$ and $\psi' : J_0(N) \to E'$, one obtains $\mathbb{Q}$-isogenies $E'' \to E$ and $E'' \to E'$ which implies that $E$ and $E'$ are $\mathbb{Q}$-isogenous.). Therefore, $E'$ has conductor $N$ and Lemma 4.1 gives

$$|h(E) - h(E')| \leq \frac{1}{2} \log 163. \quad (6.4)$$

The kernel of $\psi' : J_0(N) \to E'$ is $A$, which is connected. An elliptic curve over $\mathbb{Q}$ with this property is called an optimal quotient of $J_0(N)$ or a (strong) Weil curve. As in Subsection 5.2, we denote by $i : X_0(N) \hookrightarrow J_0(N)$ the usual embedding which maps $\infty$ to the zero element of $J_0(N)$. To simplify the exposition we write $E$ and $\varphi$ for $E'$ and $\psi' \circ i$ respectively.

(2) It is known by Frey [28, pp. 45–47] that the degree $\text{deg}(\varphi)$ of $\varphi$ is related to $h(E)$. A precise relation can be established as follows. We denote by $E$ the Néron model of $E$ over $B = \text{Spec}(\mathbb{Z})$. Since $\mathbb{Z}$ is a principal ideal domain, the line bundle $\omega = \omega_{E/B}$ on $B$ from Section 2 takes the form $\omega \cong \alpha \mathbb{Z}$ with a global differential one form $\alpha$ of $E$. Then, on recalling the definition of the relative Faltings height $h(E)$ in Section 2, we compute

$$h(E) = -\frac{1}{2} \log \left( \frac{i}{2} \int_{E(\mathbb{C})} \alpha \wedge \overline{\alpha} \right).$$

As in Subsection 5.1, we denote by $S_2(\Gamma_0(N))$ the cusp forms of weight 2 for $\Gamma_0(N)$ and by $(\cdot, \cdot)$ the Petersson inner product on $S_2(\Gamma_0(N))$. The pullback $\varphi^*\alpha$ of $\alpha$ under $\varphi$ defines a differential on $X_0(N)$. It takes the form $\varphi^*\alpha = c \cdot 2\pi if \, dz$ with $c \in \mathbb{Q}^\times$ and $f \in S_2(\Gamma_0(N))$ a newform of level $N$ with Fourier coefficients $a_n(f) \in \mathbb{Z}$ for all $n \in \mathbb{Z}_{\geq 1}$. After adjusting the sign of $\alpha$, we may and do assume that $c$ is positive. The number $c$ is the Manin constant of the optimal quotient $E$. By definition, it holds

$$(f, f) = \frac{i}{2} \int_{X_0(N)(\mathbb{C})} f \, dz \wedge \overline{f} \, \overline{dz}.$$

The elliptic curve $E_f$ over $\mathbb{Q}$, which is associated to $f$ in (5.4), is $\mathbb{Q}$-isogenous to $E$. Indeed, this follows for example from [26, Korollar 2] since by construction the $L$-functions of $E$ and $f$, of $f$ and $E_f$, and thus of $E$ and $E_f$, have the same Euler product factors for all but finitely many primes. Furthermore, $E$ is an optimal quotient of $J_0(N)$ by 1. and $E_f$ is an optimal quotient of $J_0(N)$, since the kernel $I_fJ_0(N)$ (see Subsection 5.2) of the natural projection $J_0(N) \to E_f$ is connected. Therefore, it follows that the modular degree $m_f$ of $f$, defined in Subsection 5.2, satisfies $m_f = \text{deg}(\varphi)$. Then, on using that $\varphi^*\alpha = c \cdot 2\pi if \, dz$ and on integrating over $X_0(N)(\mathbb{C})$, we see that the change of variable formula and the above displayed formulas for $h(E)$ and $(f, f)$ lead to

$$h(E) = \frac{1}{2} \log m_f - \frac{1}{2} \log (f, f) - \log(2\pi c). \quad (6.5)$$

We now estimate the quantities which appear on the right-hand side of this formula.

(3) It follows from [1, Lemme 3.7], or from [63, p. 262], that $(f, f) \geq e^{-4r}/(4\pi)$. Further, Edixhoven showed in [22, Proposition 2] that the Manin constant $c$ of the optimal quotient $E$ of
$J_0(N)$ satisfies $c \in \mathbb{Z}$ and thus we obtain that $\log(2\pi c) \geq \log(2\pi)$. Then the above lower bound for $(f, g)$, the formula (6.5) and the estimate for $m_j$ in Lemma 5.1(i) prove Proposition 6.1 for the optimal quotient $E$ of $J_0(N)$. Finally, on using the reduction in (1) and (6.4), we deduce Proposition 6.1 for all elliptic curves over $\mathbb{Q}$.

The main ingredients for the following proof of Proposition 6.4 are the Shimura–Taniyama conjecture and a result of Mazur–Kenku [40] on $\mathbb{Q}$-isogeny classes of elliptic curves over $\mathbb{Q}$.

**Proof of Proposition 6.4.** Let $E$ be an elliptic curve over $\mathbb{Q}$, with good reduction over $S$. We write $N_E$ for the conductor of $E$, and we denote by $J_0(N) = \text{Pic}^0(X_0(N))$ the Jacobian of the modular curve $X_0(N)$ for $N \geq 1$ (see Section 5). There exists a finite morphism $X_0(\nu_S) \to X_0(N_E)$ of curves over $\mathbb{Q}$, since $N_E$ divides $\nu_S$ by (2.2). Picard functoriality gives a surjective morphism $J_0(\nu_S) \to J_0(N_E)$ of abelian varieties over $\mathbb{Q}$, and as in (6.3) we see that the Shimura–Taniyama conjecture provides that $E$ is a $\mathbb{Q}$-quotient of $J_0(N_E)$. Thus there exists a surjective morphism

$$J_0(\nu_S) \to E$$

of abelian varieties over $\mathbb{Q}$. Then Poincaré’s reducibility theorem shows that $E$ is $\mathbb{Q}$-isogenous to a $\mathbb{Q}$-simple ‘factor’ of $J_0(\nu_S)$. Furthermore, the dimension of $J_0(\nu_S)$ coincides with the genus $g$ of the modular curve $X_0(\nu_S)$, and the abelian variety $J_0(\nu_S)$ has at most $g$ $\mathbb{Q}$-simple ‘factors’ up to $\mathbb{Q}$-isogenies. Therefore, we see that there exists a set of elliptic curves over $\mathbb{Q}$ with the following properties: This set has cardinality at most $g$ and for any elliptic curve $E$ over $\mathbb{Q}$, with good reduction over $S$, there exists an elliptic curve in this set which is $\mathbb{Q}$-isogenous to $E$. Further, Mazur–Kenku [40, Theorem 2] give that each $\mathbb{Q}$-isogeny class of elliptic curves over $\mathbb{Q}$ contains at most eight distinct $\mathbb{Q}$-isomorphism classes of elliptic curves over $\mathbb{Q}$. On combining the results collected above, we deduce that $N(S) \leq 8g$ and then the upper bound for $g$ in (5.1) implies Proposition 6.4.

In the following section, we shall combine Propositions 6.1 or 6.4 with the Paršin constructions from Section 3 to obtain explicit Diophantine finiteness results.

### 7. Integral points on moduli schemes

In the first part of this section, we give in Theorem 7.1 an effective finiteness result for integral points on moduli schemes of elliptic curves. In the second and third part, we refine the method of Theorem 7.1 for the moduli schemes corresponding to $\mathbb{P}^1 - \{0, 1, \infty\}$ and to once punctured Mordell elliptic curves. This leads to effective versions of Siegel’s theorem for $\mathbb{P}^1 - \{0, 1, \infty\}$ and once punctured Mordell elliptic curves, which provide explicit height upper bounds for the solutions of $S$-unit and Mordell equations. We also give additional Diophantine applications. In particular, we consider cubic Thue equations.

#### 7.1. Moduli schemes

To state our result for integral points on moduli schemes of elliptic curves, we use the notation and terminology which was introduced in Section 3.

Let $T$ and $S$ be non-empty open subschemes of Spec($\mathbb{Z}$), with $T \subseteq S$. We write $\nu_T = 12^3 \prod p^2$ with the product taken over all rational primes $p$ not in $T$. For any moduli problem $\mathcal{P}$ on $(Ell)$, we denote by $|\mathcal{P}|_T$ the maximal (possibly infinite) number of distinct level $\mathcal{P}$-structures on an arbitrary elliptic curve over $T$; see (3.1). We suppose that $Y = M_\mathcal{P}$ is a moduli scheme of elliptic curves, which is defined over $S$. Let $h_\phi$ be the pullback of the relative Faltings height by the canonical forget $\mathcal{P}$-map $\phi$, defined in (3.3).
Theorem 7.1. The following statements hold.

(i) The cardinality of $Y(T)$ is at most $\frac{2}{3}|P|_T \nu_T \prod (1 + 1/p)$ with the product taken over all rational primes $p$ which divide $\nu_T$.

(ii) If $P \in Y(T)$, then $h_\phi(P) \leq \frac{1}{2} \nu_T (\log \nu_T)^2 + 9$.

We refer to Paragraph 1.1.2 for a discussion of this theorem. In addition, we now mention that for many classical moduli problems $P$ on $(Ell)$ it is possible to express $|P|_T$ in terms of more conventional data, where $T$ is an arbitrary scheme which is connected. For example, if $P_N$ is the ‘naïve’ level $N$ moduli problem on $(Ell)$ considered in Section 3, then (3.2) shows that $|P_N|_T$ is an explicit function in terms of the level $N \geq 1$.

It is quite difficult, when not impossible, to compare Theorem 7.1 with quantitative or effective finiteness results in the literature, since these results hold in different settings. One can mention for example the quantitative result of Corvaja–Zannier [19] for hyperbolic curves which relies on Schmidt’s subspace theorem, or the effective result of Bilu [8] for certain modular curves which is based on the theory of logarithmic forms.

Proof of Theorem 7.1. To prove (i) we recall that $M(T)$ denotes the set of isomorphism classes of elliptic curves over $T$. Let $M(T)_Q$ be the set of $Q$-isomorphism classes of elliptic curves over $Q$, with good reduction over $T$. We now show that there exists a bijection

$$M(T) \cong M(T)_Q.$$ 

Any elliptic curve over $Q$ has good reduction over $T$ if and only if it is the generic fiber of an elliptic curve over $T$. Further, any elliptic curve over $T$ is the Néron model of its generic fiber (see for example [13, p. 15]), and $M(T)$ is in bijection with the set of isomorphism classes of $T$-schemes generated by elliptic curves over $T$. Hence the Néron mapping property proves that $M(T) \cong M(T)_Q$ and thus Proposition 6.4 implies

$$|M(T)| \leq \frac{2}{3} \nu_T \prod (1 + 1/p)$$

with the product taken over all rational primes $p$ which divide $\nu_T$. It follows from Lemma 3.1 that $|Y(T)| \leq |P|_T |M(T)|$ and then we deduce Theorem 7.1(i).

To show (ii), we take $P \in Y(T)$ and we write $[E] = \phi(P)$ for $\phi : Y(T) \to M(T)$ the canonical forget $P$-map from Lemma 3.1. The conductor $N_E$ of the generic fiber $E_Q$ of $E$ takes the form $N_E = \prod p^{\nu_p}$ with $\nu_p$ the conductor exponent of $E_Q$ at a rational prime $p$, see Subsection 2.2. It holds that $\nu_p \leq 2$ for $p \geq 5$ and (2.2) gives that $f_2 \leq 8$ and $f_3 \leq 5$. Furthermore, if $p \in T$, then we obtain that $\nu_p = 0$ since $E_Q$ extends to an abelian scheme over $T$. On combining the above results, we deduce that $N_E | \nu_T$. An application of Proposition 6.1 with $E_Q$ gives that $h_\phi(P) \leq \frac{1}{2} N_E (\log N_E)^2 + 9$ which together with $N_E \leq \nu_T$ implies assertion (ii). This completes the proof of Theorem 7.1.

On replacing in the proof of Theorem 7.1(i) the explicit estimate from Proposition 6.4 by the asymptotic bound (6.1) of Ellenberg–Hellgott–Venkatesh, we obtain the following version of Theorem 7.1(i): If $Y = M_P$ is a moduli scheme, defined over $S$, then

$$|Y(T)| \ll |P|_T N_T^{0.1689}$$

for $N_T$ the product of all rational primes $p$ not in $T$. Furthermore, the discussion surrounding (6.1) shows that the ‘Birch and Swinnerton-Dyer conjecture’ together with the ‘Generalized Riemann Hypothesis’ implies that for all $\epsilon > 0$ there exists a constant $c(\epsilon)$, depending only on $\epsilon$, such that $|Y(T)| \leq c(\epsilon) |P|_T N_T^\epsilon$. 


We note that the complement of $S$ in $\text{Spec}(\mathbb{Z})$ is a finite set of rational prime numbers. For the remaining of Section 7, we will adapt our notation to the classical number theoretic setting, and in (Subsections 7.2–7.4) the symbol $S$ denotes a finite set of rational prime numbers.

7.2. $\mathbb{P}^1 \setminus \{0, 1, \infty\}$: $S$-unit equations

In the first part of this section, we briefly review alternative methods which give finiteness for integral points of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$, or equivalently for the number of solutions of $S$-unit equations. In the second and third part, we establish in Corollary 7.2 an explicit upper bound in terms of $S$ for the height of the solutions of $S$-unit equations, and we compare this bound with the actual best results in the literature. In the last part, we discuss upper bounds for the number of solutions of $S$-unit equations.

Let $S$ be a finite set of rational primes, let $N_S = \prod p$ with the product taken over all $p \in S$, and let $O^\times$ be the units of $O = \mathbb{Z}[1/N_S]$. We recall the $S$-unit equation (1.1)

$$x + y = 1, \quad (x, y) \in O^\times \times O^\times.$$ 

Before we apply the method of this paper to $S$-unit equations (1.1), we briefly review in the following subsection alternative methods which give finiteness of (1.1).

7.2.1. Alternative methods

The first finiteness proof for $S$-unit equations (1.1) goes back to Mahler [46]. He used the method of Diophantine approximations (Thue–Siegel). Another proof of Mahler’s theorem was obtained by Faltings, whose general finiteness theorems in [26] cover in particular (1.1). Faltings studied semi-simple $\ell$-adic Galois representations associated to abelian varieties. Recently, Kim [41] gave a new finiteness proof of (1.1). He used Galois representations associated to the unipotent étale and de Rham fundamental group of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. The methods of Faltings, Kim and Thue–Siegel (Mahler) are a priori not effective. The first effective finiteness proof of (1.1) was given by Baker’s method, using the theory of logarithmic forms; see, for example, Györy [31] or Baker–Wüstholz [7]. (For instance, Coates explicit result [17], which was published in 1970, implies an effective height upper bound for the solutions of (1.1).) Another effective finiteness proof of (1.1) is due to Bombieri–Cohen [10]. They generalized Bombieri’s method in [9], which uses effective Diophantine approximations on the multiplicative group $G_m$ (Thue–Siegel principle). The methods of Baker and Bombieri both give explicit upper bounds for the heights of the solutions of (1.1) in terms of $S$, and they both allow to deal with $S$-unit equations in any number field. So far, the theory of logarithmic forms, which was extensively polished and sharpened over the last 47 years, produces slightly better bounds than Bombieri’s method. On the other hand, Bombieri’s method is relatively new and is essentially self-contained; see Bombieri–Cohen [11].

7.2.2. Effective resolution

To state and discuss our effective result for $S$-unit equations, we put $n_S = 2^7 N_S$. Let $h(\beta)$ be the usual absolute logarithmic Weil height of any $\beta \in \mathbb{Q}$. This height is defined for example in [12, p. 16]. We obtain the following corollary.

**Corollary 7.2.** Any solution $(x, y)$ of the $S$-unit equation (1.1) satisfies

$$h(x), h(y) \leq \frac{3}{2} n_S (\log n_S)^2 + 65.$$

**Proof.** We use the notation and terminology of Section 3. The discussion in (7.1) shows that we may and do assume that 2 is invertible on $T = \text{Spec}(O)$. Write

$$X = \mathbb{P}^1_T \setminus \{0, 1, \infty\} = \text{Spec}(O[z, 1/(z(z-1))]).$$
for \( z \) an ‘indeterminate’. We suppose that \((x, y)\) satisfies (1.1). Then we see that there exists \( P \in X(T) \) with \( z(P) = x \). Thus an application of Proposition 3.2 with \( P \) and \( T \) gives an elliptic curve \( E' \) over \( T \) that satisfies (write \( E = E' \))

\[
h(x) \leq 6h(E) + 11 \quad \text{and} \quad N_E \leq n_S.
\]

Here \( N_E \) is the conductor of \( E \) and \( h(E) \) is the relative Faltings height of \( E \), see Subsection 3.2 for the definitions. Proposition 6.1 provides that \( h(E) \leq \frac{1}{4}N_E(\log N_E)^2 + 9 \). Then the displayed inequalities imply the claimed upper bound for \( h(x) \), and then for \( h(y) \) by symmetry. This completes the proof of Corollary 7.2.

As already mentioned in the introduction, this corollary is an effective version of Frey’s remark in [29, p. 544]. (We presented Corollary 7.2 and its proof in various seminars and conferences in Princeton (September 2011 and January 2012), New York (February 2012), Michigan (March 2012), Hong Kong (June 2012), Paris (Oct. 2012) and Zurich (May 2013). After we uploaded the present paper to the arXiv in October 2013, Hector Pasten informed us about his joint work with Ram Murty ([52], submitted November 2012) which was published online in July 2013 and which was presented including the proof in a seminar in Kingston (March 2012) and in a workshop in Toronto (Nov. 2012); we thank Hector Pasten for informing us about [52]. The main results of [52] independently establish versions of Corollary 7.2 (and of Lemma 5.1, Proposition 6.1 and Corollary 6.2 which are used in the proof of Corollary 7.2) with effective bounds of the form \( \ll N \log N \), while our corresponding bounds are of the (slightly) weaker form \( \ll N(\log N)^2 \). The method used in [52] is similar to our proof of Corollary 7.2. To conclude the discussion, we point out that the results were obtained completely independently: We obtained the results of this paper without knowing anything of the related work of Hector Pasten and Ram Murty, and they obtained the results of [52] without knowing anything of our related work).

Corollary 7.2 allows in principle to find all solutions of any \( S \)-unit equation (1.1). To discuss a practical aspect of Corollary 7.2, we observe that any \( u \in \mathcal{O}^\times \) satisfies \( u = \prod p^{u_p} \) with the product taken over all \( p \in S \) and \( u_p = \text{ord}_p(u) \). Therefore, any \( S \)-unit equation may be viewed as an exponential Diophantine equation of the form

\[
\prod_{p \in S} p^{x_p} + \prod_{p \in S} p^{y_p} = 1, \quad ((x_p), (y_p)) \in \mathbb{Z}^s \times \mathbb{Z}^s
\]

for \( s = |S| \). If \((x_p), (y_p)\) satisfies this exponential Diophantine equation, then Corollary 7.2 implies that \( \max_{p \in S} |x_p| \) and \( \max_{p \in S} |y_p| \) are at most \( (3/2 \log 2)n_S(\log n_S)^2 + 94 \). On using additional tricks, we improve in von Känel and Matschke (the forthcoming paper ‘Solving \( S \)-unit and Mordell equations via Shimura–Taniyama conjecture’) the absolute constants \( 3/2 \log 2 \) and 94 and we will transform the proof of Theorem 7.2 into a practical algorithm to solve \( S \)-unit equations.

7.2.3. Comparison to known results Next, we compare Corollary 7.2 with the actual best effective results in the literature for \( S \)-unit equations (1.1). We note that (1.1) has no solutions when \( |S| = 0 \), and \((1, 1), (2, -1)\) and \((-1, 2)\) are the only solutions of (1.1) when \( |S| = 1 \). Further, we see that if (1.1) has a solution, then \( 2 \in S \). Thus, for the purpose of the comparison, we may and do assume

\[
s = |S| \geq 2 \quad \text{and} \quad 2 \in S.
\]

(7.1)

Let \((x, y)\) be a solution of the \( S \)-unit equation (1.1). The actual best explicit height upper bound for \((x, y)\) in the literature is due to Györy–Yu [32]. They used the state of the art in the
theory of logarithmic forms. In the case of (1.1), where the number field is $\mathbb{Q}$, their estimate in [32, Theorem 2] becomes

$$h(x), h(y) \leq 2^{10s+22} s^4 q \prod \log p$$

with the product taken over all rational primes $p \in S - \{q\}$ for $q = \max S$. The right-hand side of the displayed inequality is always bigger than $2^{47}$. Hence, we see that Corollary 7.2 improves [32] for sets $S$ with small $N_S$, in particular, for all sets $S$ with $N_S \leq 2^{30}$. This improvement is significant for the practical solution of $S$-unit equations, see the discussion at the end of Paragraph 7.2.2. However, the result of Stewart–Yu [66, Theorem 1], based on the actual state of the art in the theory of logarithmic forms, gives

$$h(x), h(y) \ll N_S^{1/3}(\log N_S)^3.$$ 

We observe that this inequality of Stewart–Yu is better than Corollary 7.2 for all sets $S$ with sufficiently large $N_S$. This concludes our comparison.

7.2.4. Number of solutions

To discuss explicit upper bounds for the number of solutions of $S$-unit equations (1.1), we recall that $n_S = 2^7 N_S$. In the special case of the moduli scheme $\mathbb{P}^1_{\mathbb{Z}[1/2]} - \{0, 1, \infty\}$, one can refine the proof of Theorem 7.1 and one obtains the following result.

**Corollary 7.3.** The $S$-unit equation (1.1) has at most $4n_S \prod_{p \in S} (1 + 1/p)$ solutions.

**Proof.** We use the terminology and notation introduced in Section 3. The discussion in (7.1) shows that we may and do assume that 2 is invertible on $T = \text{Spec}(\mathcal{O})$. Then there exists a bijection between the set of solutions of the $S$-unit equation (1.1) and $Y(T)$, where

$$Y = \mathbb{P}^1_{\mathbb{Z}[1/2]} - \{0, 1, \infty\}.$$ 

We now estimate the cardinality of $Y(T)$. The remark at the end of Paragraph 3.2.1 shows that $Y = M_P$ is a moduli scheme of elliptic curves, defined over $\text{Spec}(\mathbb{Z}[1/2])$, where $P = [\text{Legendre}]$ is the Legendre moduli problem on $(E|P)$. Thus Lemma 3.1 gives a map

$$\phi : Y(T) \longrightarrow M(T),$$

with all fibers having cardinality at most $|P|_T$. Here $|P|_T$ is defined in (3.1) and $M(T)$ is the set of isomorphism classes of elliptic curves over $T$. The arguments of Proposition 3.2(iii) and of Theorem 7.1 imply that the cardinality of $\phi(Y(T))$ is at most the number of $\mathbb{Q}$-isomorphism classes of elliptic curves over $\mathbb{Q}$, with conductor dividing $n_S = 2^7 N_S$. Therefore, on replacing $\nu_S$ by $n_S$ in the proof of Proposition 6.4, we deduce

$$|\phi(Y(T))| \leq \frac{2}{3} n_S \prod_{p \in S} (1 + 1/p).$$

Here we used that $2 \in S$. It follows that $Y(T)$ has at most $4n_S \prod_{p \in S} (1 + 1/p)$ elements, since the fibers of $\phi$ have cardinality at most $|P|_T \leq 6$. Then we conclude Corollary 7.3.  

We now compare Corollary 7.3 with results in the literature. Evertse [24, Theorem 1] used the method of Diophantine approximations to prove that any $S$-unit equation (1.1) has at most $3 \cdot \tau^{3+2|S|}$ solutions. We mention that Evertse’s result holds for more general unit equations in any number field, and it provides, as far as we know, the actual best upper bound in the literature for the number of solutions of (1.1). Further, we see that Evertse’s result is considerably better than Corollary 7.3 for almost all sets $S$, since $3 \cdot \tau^{3+2|S|} \ll \epsilon n_S^2$. Note there are sets $S$ for which Corollary 7.3 improves [24, Theorem 1]. For example, if $S \subseteq \{2, 3, 5, \ldots, 83, 89\}$ and if
S satisfies the reasonable assumption (7.1), then we observe that Corollary 7.3 is better than [24, Theorem 1].

7.3. Once punctured Mordell elliptic curves: Mordell equations

In the first part of this section, we briefly review alternative methods which give finiteness for integral points on once punctured Mordell elliptic curves, or equivalently for the number of $S$-integer solutions of Mordell equations. In the second and third part, we state and prove Corollary 7.4 on Mordell equations and we compare it with the actual best effective results in the literature. In the fourth and fifth part, we refine a result of Stark and we discuss explicit upper bounds for the number of solutions of Mordell equations.

We continue to denote by $S$ an arbitrary finite set of rational prime numbers and we write $O = \mathbb{Z}[1/N_S]$ for $N_S$ the product of all $p \in S$. For any non-zero $a \in O$, we recall that Mordell’s equation (1.2) is of the form

$$y^2 = x^3 + a, \quad (x, y) \in O \times O.$$ 

This Diophantine equation is a priori more difficult than $S$-unit equations (1.1). Indeed, elementary transformations reduce (1.1) to (1.2), while the known (unconditional) reductions of (1.2) to controlled $S$-unit equations require to solve (1.1) over field extensions.

7.3.1. Alternative methods As already mentioned in the introduction, the resolution of Mordell’s equation in $\mathbb{Z} \times \mathbb{Z}$ is equivalent to the classical problem of finding all perfect squares and perfect cubes with given difference. We refer to Baker’s introduction of [6] for a discussion (of partial resolutions) of this classical problem, which goes back at least to Bachet 1621. Mordell [49, 50] showed that (1.2) has only finitely many solutions in $\mathbb{Z} \times \mathbb{Z}$. He reduced the problem to Thue equations and then he applied Thue’s finiteness theorem which is based on Diophantine approximations. More generally, the completely different methods of Siegel, Faltings [26] and Kim [41, 42] give finiteness of (1.2). Siegel’s method uses Diophantine approximations, and the methods of Faltings and Kim are briefly described in Paragraph 7.2.1. We mention that these methods, which in fact allow to deal with considerably more general Diophantine problems, are all a priori not effective. See also Bombieri [9] and Kim’s discussions in [43]. The first effective finiteness result for solutions in $\mathbb{Z} \times \mathbb{Z}$ of Mordell’s equation (1.2) was provided by Baker [6]. Baker’s result is based on the theory of logarithmic forms.

7.3.2. Effective resolution We now state and prove our effective result for Mordell equations. We continue to denote by $h(\beta)$ the absolute logarithmic Weil height of any $\beta \in \mathbb{Q}$. To measure the set $S$ and the non-zero number $a \in O$, we use inter alia the quantity

$$a_S = 2^83^5N_S^2r_2(a), \quad r_2(a) = \prod p^{\min(2, \text{ord}_p(a))}$$

with the product taken over all rational primes $p \notin S$ with $\text{ord}_p(a) \geq 1$. The following corollary allows in principle to find all solutions of any Mordell equation (1.2).

**Corollary 7.4.** If $(x, y)$ satisfies Mordell’s equation (1.2), then

$$h(x), h(y) \leq h(a) + 4a_S(\log a_S)^2.$$ 

**Proof.** The proof is completely analogous to the proof of Corollary 7.2. We use the notation and terminology of Section 3. Write $T = \text{Spec}(O[1/(6a)])$ and define

$$Z = \text{Spec}(O[x_0,y_0]/(y_0^2 - x_0^3 - a))).$$
for $x_0$ and $y_0$ ‘indeterminates’. We suppose that $(x, y)$ is a solution of (1.2). Then there exists a $T$-integral point $P \in \mathbb{Z}(T)$ with $x_0(P) = x$ and thus an application of Proposition 3.4 with $K = \mathbb{Q}$, $P$ and $T$ gives an elliptic curve $E$ over $T$ that satisfies

$$h(x) \leq \frac{1}{3} h(a) + 8h(E) + 2 \log \max(1, h_F(E)) + 36 \quad \text{and} \quad NE \leq a_S.$$ 

Here $NE$ denotes the conductor of $E$, and $h(E)$ and $h_F(E)$ denote the relative and the stable Faltings height of $E$, respectively; see Subsection 3.2 for the definitions. Proposition 6.1 provides that $h(E) \leq \frac{1}{4} NE(log NE)^2 + 9$ and it holds that $h_F(E) \leq h(E)$. Therefore, the displayed inequalities lead to the claimed estimate for $h(x)$, and then for $h(y)$ since $y^2 = x^3 + a$. This completes the proof of Corollary 7.4. \hfill \Box

We already pointed out in the introduction that Corollary 7.4 provides, in particular, an entirely new proof of Baker’s classical result [6, Theorem 1].

7.3.3. Comparison to known results. In what follows, we compare our Corollary 7.4 with the actual best effective upper bounds in the literature for the solutions of Mordell’s equation (1.2). For this purpose, we note that if $a \in \mathbb{Z} - \{0\}$ and if rad$(a) = \prod_{p|a} p$ denotes the radical of $a$, then

$$r_2(a) \leq |a| \quad \text{and} \quad r_2(a) \mid \text{rad}(a)^2.$$ (7.2)

Over the last 45 years, many authors improved the explicit bound provided by Baker [6], using refinements of the theory of logarithmic forms; see Baker–Wüstholz [7] for an overview. The actual best explicit upper bound is due to Hajdu–Herendi [33], and due to Juricevic [35] in the important special case $O = \mathbb{Z}$.

We first discuss the classical case $O = \mathbb{Z}$. If $S = \{10^{181}, 4, 10^{23}, 5, 10^{19}, 6\}$ and if $a \in \mathbb{Z} - \{0\}$, then Juricevic [35] gives that any solution $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ of (1.2) satisfies

$$h(x), h(y) \leq \min_{(m,n) \in S} m|a|(|\log|a||)^n.$$ 

On using (7.2), we see that Corollary 7.4 improves this inequality and therefore our corollary establishes the actual best result for (1.2) in the classical case $O = \mathbb{Z}$.

It remains to discuss the case of arbitrary $O$. To state the rather complicated bound in [33], we have to introduce some notation. As in [33], we define

$$c_1 = \frac{32}{3} \Delta^{1/2}(8 + \frac{1}{2} \log \Delta)^4, \quad c_2 = 10^4 \cdot 256 \cdot \Delta^{2/3}, \quad \Delta = 27|a|^2.$$ 

Write $c_S = 7 \cdot 10^{38s + 86}(s + 1)^{20s + 35}q^{24} \max(1, \log q)^{4s+2}$ for $s = |S|$ and $q = \max S$. If $s \geq 1$ and if $a \in \mathbb{Z} - \{0\}$, then the result of Hajdu–Herendi [33, Theorem 2], which in fact holds more generally for any elliptic equation, gives that any solution $(x, y)$ of (1.2) satisfies

$$h(x), h(y) \leq c_S c_1 (\log c_1)^2 (c_1 + 20sc_1 + \log(\epsilon c_2)).$$ 

It follows from (7.2) that the dependence on $a \in \mathbb{Z}$ of Corollary 7.4 is of the form $|a|(|\log|a||)^2$, while [33, Theorem 2] is of the weaker form $|a|^2(|\log|a||)^10$. Further, on using again (7.2), we see that Corollary 7.4 improves [33, Theorem 2] for ‘small’ sets $S$, in particular, for all sets $S$ with $NS \leq 2^{1200}$ or with $s \leq 12$. This improvement might be useful for the practical resolution of Mordell equations (1.2), see, for example, von Känel and Matschke (the forthcoming paper ‘Solving $S$-unit and Mordell equations via Shimura–Taniyama conjecture’). However, if $NS \gg |a|$, then one cannot say which bound is better. The point is that there are sets $S$ with $NS \gg |a|$ for which our result is better than [33, Theorem 2], and vice versa. Finally, we mention that (so far) all effective results for (1.2) in the literature are based on the theory of logarithmic forms, and this theory allows us to deal with more general Diophantine equations over arbitrary number fields; see [7]. This concludes our comparison.
7.3.4. A refinement of Stark’s theorem

We now discuss a refinement of the following theorem of Stark [65, Theorem 1]: If \( a \in \mathbb{Z} - \{0\} \), then any \((x, y) \in \mathbb{Z} \times \mathbb{Z} \) with \( y^2 = x^3 + a \) satisfies
\[
h(x), h(y) \ll_{\varepsilon} |a|^{1+\varepsilon},
\]
where the implied constant is effective. This classical estimate of Stark is based on the theory of logarithmic forms. The following result is a direct consequence of Corollary 7.4.

**Corollary 7.5.** If \( \varepsilon > 0 \) is a real number, then there exists an effective constant \( c \), depending only on \( \varepsilon \), such that any solution \((x, y)\) of Mordell’s equation (1.2) satisfies
\[
h(x), h(y) \leq h(a) + c \cdot a S^{1+\varepsilon}.
\]

On using (7.2) and the fact that \( h(a) = \log |a| \) for \( a \in \mathbb{Z} - \{0\} \), we see that Corollary 7.5 generalizes and refines Stark’s theorem [65, Theorem 1] discussed above.

We remark that Stewart–Yu [66] obtained an exponential version of the abc-conjecture (abc), and (abc) is equivalent to a certain upper bound for the height of the solutions of Mordell’s equation (1.2); see, for example, [12, p. 428]. However, by elementary reasons, all known links between (abc) and height upper bounds for the solutions of (1.2) do not work any more with exponential versions. Hence, at the time of writing, it is not possible to improve Corollary 7.5 on using exponential versions of (abc).

7.3.5. Number of solutions

Next, we discuss explicit upper bounds for the number of solutions of Mordell’s equation (1.2). In the special case of the moduli scheme corresponding to (1.2), one can refine the proof of Theorem 7.1 and one obtains the following result.

**Corollary 7.6.** The number of solutions of (1.2) is at most \( \frac{2}{3} a_S S \prod_{p|a_S} (1 + 1/p) \) with the product taken over all rational primes \( p \) which divide \( a_S \).

**Proof.** In this proof, we use the terminology and notation which we introduced in Section 3. We define \( T = \text{Spec}(\mathcal{O}[1/(6a)]) \) and \( b = -a/1728 \). It follows that the number of solutions of Mordell’s equation (1.2) is at most the cardinality of \( Y(T) \), where
\[
Y = \text{Spec}(\mathbb{Z}[1/6, c_4, c_6, b, 1/b]/(1728b - c_4^3 + c_6^3))
\]
for \( c_4 \) and \( c_6 \) ‘indeterminates’. We now estimate the cardinality of \( Y(T) \). The remark at the end of Paragraph 3.2.2 gives that \( Y = M_P \) is a moduli scheme of elliptic curves, where \( P = [\Delta = b] \) is the moduli problem on (Ell). Thus, Lemma 3.1 gives a map
\[
\phi : Y(T) \rightarrow M(T)
\]
for \( M(T) \) the set of isomorphism classes of elliptic curves over \( T \). We notice that the map \( \phi \) coincides with the map \( \phi \) constructed in Proposition 3.4. Further, we see that the arguments of Proposition 3.4(iii) and of Theorem 7.1 imply that \( |\phi(Y(T))| \) is at most the number of \( \mathbb{Q} \)-isomorphism classes of elliptic curves over \( \mathbb{Q} \), with conductor dividing \( a_S \). Therefore, on replacing in the proof of Proposition 6.4 the number \( \nu_S \) by \( a_S \), we deduce
\[
|\phi(Y(T))| \leq \frac{2}{3} a_S \prod_{p|a_S} (1 + 1/p)
\]
with the product taken over all rational primes \( p \) dividing \( a_S \). Proposition 3.4(i) shows that \( \phi \) is injective and then the displayed inequality implies Corollary 7.6.

We now compare Corollary 7.6 with results in the literature. In the classical case \( \mathcal{O} = \mathbb{Z} \), the actual best explicit upper bound for the number of solutions of (1.2) is due to Poulakis [55].
We see that Corollary 7.6 is better than Poulakis’ result when \( a_S \leq 2^{180} \), and is worse when \( a_S \) is sufficiently large. However, for large \( a_S \) the actual best bound follows from Ellenberg–Helfgott–Venkatesh [23, 34]. On combining their results (see, for example, our forthcoming paper ‘Height and conductor of elliptic curves’), one obtains that the number of solutions of (1.2) is

\[ \ll c_0^2(1 + \log q)^2 \text{rad}(a)^{0.1689} \]

for \( c_0 \) an absolute constant, \( s = |S| \) and \( q = \max S \). This asymptotic bound is better than the asymptotic estimate implied by Corollary 7.6. We point out that the methods of Poulakis [55] and Helfgott–Venkatesh [34] are fundamentally different from the method of Corollary 7.6; see the end of Subsection 6.1 for a brief discussion of the Diophantine results used in the proofs of [34, 55]. To conclude our comparison, we mention that Evertse–Silverman [25] applied diophantine approximations to obtain an explicit upper bound for the number of solutions of (1.2). Their bound involves inter alia a quantity which depends on a certain class number.

7.4. Additional Diophantine applications

In this section, we discuss additional Diophantine applications of the Shimura–Taniyama conjecture. In particular, we consider cubic Thue equations.

There are many Diophantine equations which can be reduced to \( S \)-unit or Mordell equations, such as for example (super-) elliptic Diophantine equations. Usually these reductions consist of elementary, but ingenious, manipulations of explicit equations and they often require to solve \( S \)-unit and Mordell equations over controlled field extensions \( K \) of \( \mathbb{Q} \). Unfortunately, we cannot use most of the standard reductions, since our results in the previous sections only hold for \( K = \mathbb{Q} \). However, we now discuss constructions which allow to reduce certain classical Diophantine problems without requiring field extensions.

7.4.1. Thue equations

Let \( m \) be an integer and let \( f \in \mathbb{Z}[x, y] \) be an irreducible binary form of degree \( n \geq 3 \), with discriminant \( \Delta \). We consider the classical Thue equation

\[ f(u, v) = m, \quad (u, v) \in \mathbb{Z} \times \mathbb{Z}. \]  

(7.3)

The famous result of Thue, based on Diophantine approximations, gives that (7.3) has only finitely many solutions. Moreover, Baker [5] used his theory of logarithmic forms to prove an effective finiteness result for Thue equations; see [7] for generalizations.

We now suppose that \( n = 3 \). To prove (effective) finiteness results for (7.3), we may and do assume by standard reductions that (7.3) has at least one solution and that \( m \Delta \neq 0 \). Thus we obtain a smooth, projective and geometrically connected genus one curve

\[ X = \text{Proj}(\mathbb{Q}[x, y, z]/(f - mz^3)). \]

On using classical invariant theory of cubic binary forms, one obtains a finite \( \mathbb{Q} \)-morphism \( \varphi : X \to \text{Pic}^0(X) \) of degree 3 and one computes

\[ \text{Pic}^0(X) = \text{Proj}(\mathbb{Q}[x, y, z]/(y^3 x - x^3 - a z^3)), \quad a = 432 m^2 \Delta \neq 0. \]

See, for example, Silverman [62, p. 401] for details. Moreover, if \( (u, v) \) satisfies (7.3) and if \( P \) denotes the corresponding \( \mathbb{Q} \)-point of \( X \), then the definition of \( \varphi \) shows that \( x(\varphi(P)) \) and \( y(\varphi(P)) \) are both in \( \mathbb{Z} \) and \( z(\varphi(P)) = 1 \). In other words, the finite \( \mathbb{Q} \)-morphism

\[ \varphi : X \to \text{Pic}^0(X) \]

of degree 3 reduces any cubic Thue equation (7.3) to a Mordell equation (1.2) of the form \( (v')^2 = (u')^3 + a, \ (u', v') \in \mathbb{Z} \times \mathbb{Z} \). Therefore, we see that Corollary 7.6 gives a quantitative finiteness result for any cubic Thue equation (7.3). In fact, the above arguments prove more generally that any cubic Thue equation (7.3) has only finitely many solutions in \( \mathcal{O} \times \mathcal{O} \) for \( \mathcal{O} = \)}
Acknowledgements. I thank Richard Taylor for answering several questions, in particular for proposing a first strategy to prove Lemma 5.1. Many thanks go to Bao le Hung, Benjamin Matschke, Richard Taylor and Jack Thorne for motivating discussions. The results were obtained when I was a member (2011/2012) at the IAS Princeton, supported by the NSF under agreement No. DMS-0635607, and when I was an EPDI fellow (2012/2013) at the IHÉS. I am grateful to the IAS and the IHÉS for providing excellent working conditions. Also, I would like to apologize for the long delay between the first presentation (2011) of the initial results on $S$-unit and Mordell equations and the completion (2013) of the manuscript. The delay resulted from the attempt to understand the initial examples in a way which is more conceptual and which is suitable for generalizations.

References
1. A. ABBES and E. ULLMO, ‘Comparaison des métriques d’Arakelov et de Poincaré sur $X_0(N)$’, Duke Math. J. 80 (1995) 295–307.
2. A. AGASHE, K. A. RIBET and W. A. STEIN, The modular degree, congruence primes, and multiplicity one, Number theory, analysis and geometry (Springer, New York, 2012) 19–49.
3. A. O. L. ATKIN and J. LEHNER, ‘Hecke operators on $Γ_0(m)$’, Math. Ann. 185 (1970) 134–160.
4. A. BAKER, ‘Linear forms in the logarithms of algebraic numbers. I, II, III, IV’, Mathematika 13 (1966) 204–216. Mathematika 14 (1967) 102–107; Mathematika 14 (1967) 220–228; 15 (1968) 204–221.
5. A. BAKER, ‘Contributions to the theory of Diophantine equations. I. On the representation of integers by binary forms’, Philos. Trans. Roy. Soc. London Ser. A 263 (1967/1968) 173–191.
6. A. BAKER, ‘Contributions to the theory of Diophantine equations. II. The Diophantine equation $y^2 = x^3 + k$', Philos. Trans. Roy. Soc. London Ser. A 263 (1967/1968) 193–208.
7. A. BAKER and G. WÜSTHOLZ, Logarithmic forms and diophantine geometry, New Mathematical Monographs 9 (Cambridge University Press, Cambridge, 2007).
8. Y. P. BILU, ‘Baker’s method and modular curves’, A panorama of number theory or the view from Baker’s garden (Zürich 1999) (Cambridge University Press, Cambridge, 2002) 73–88.
9. E. BOMBIERI, ‘Effective Diophantine approximation on $G_m^r$, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 20 (1993) 61–89.
10. E. BOMBIERI and P. B. COHEN, ‘Effective Diophantine approximation on $G_M$. II’, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 24 (1997) 205–225.
11. E. BOMBIERI and P. B. COHEN, ‘An elementary approach to effective Diophantine approximation on $G_m$’, Number theory and algebraic geometry. London Mathematical Society Lecture Note Series 303 (Cambridge University Press, Cambridge, 2003) 41–62.
12. E. BOMBIERI and W. GUBLER, Heights in Diophantine geometry, New Mathematical Monographs 4 (Cambridge University Press, Cambridge, 2006).
13. S. BOSCH, W. LÜTKEBOHMERT and M. RAYNAUD, Néron models, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)] 21 (Springer, Berlin, 1990).
14. C. BREUIL, B. CONRAD, F. DIAMOND and R. TAYLOR, ‘On the modularity of elliptic curves over $Q$: wild 3-adic exercises’, J. Amer. Math. Soc. 14 (2001) 843–939.
15. A. BRUWER and K. KRAMER, ‘The conductor of an abelian variety’, Compositio Math. 92 (1994) 227–248.
16. A. BRUWER and J. H. SILVERMAN, ‘The number of elliptic curves over $Q$ with conductor $N$’, Manuscripta Math. 91 (1996) 95–102.
17. J. COATES, ‘An effective $p$-adic analogue of a theorem of Thue. III. The diophantine equation $y^2 = x^3 + k$’, Acta Arith. 16 (1969/1970) 425–435.
18. A. C. COJOCARU and E. KANI, ‘The modular degree and the congruence number of a weight 2 cusp form’, Acta Arith. 114 (2004) 159–167.
19. P. Corvaja and U. Zannier, ‘On the number of integral points on algebraic curves’, *J. reine angew. Math.* 565 (2003) 27–42.
20. J. E. Cremona, *Algorithms for modular elliptic curves*, 2nd edn (Cambridge University Press, Cambridge, 1997).
21. F. Diamond and J. Shurman, *A first course in modular forms*, Graduate Texts in Mathematics 228 (Springer, New York, 2005).
22. B. Edixhoven, ‘On the Manin constants of modular elliptic curves’, *Arithmetic algebraic geometry (Texel, 1989)*, Progress in Mathematics 89 (Birkhäuser, Boston, MA, 1991) 25–39.
23. J. S. Ellenberg and A. Venkatesh, ‘Reflection principles and bounds for class group torsion’, *Int. Math. Res. Not.*, 2007, Art. ID rnm002, 18.
24. J.-H. Evertse, ‘On equations in $S$-units and the Thue–Mahler equation’, *Invent. Math.* 75 (1984) 561–584.
25. J.-H. Evertse and J. H. Silverman, ‘Uniform bounds for the number of solutions to $Y^n = f(X)$’, *Math. Proc. Cambridge Philos. Soc.* 100 (1986) 237–248.
26. G. Faltings, ‘Endlichkeitssätze für abelsche Varietäten über Zahlkörpern’, *Invent. Math.* 73 (1983) 349–366.
27. G. Faltings, ‘Calculus on arithmetic surfaces’, *Ann. of Math.* (2) 119 (1984) 387–424.
28. G. Frey, ‘Links between solutions of $A - B = C$ and elliptic curves’, *Number theory (Ulm, 1987)*, Lecture Notes in Mathematics 1380 (Springer, New York, 1989) 31–62.
29. G. Frey, ‘On ternary equations of Fermat type and relations with elliptic curves’, *Modular forms and Fermat’s last theorem (Boston MA 1995)* (Springer, New York, 1997) 527–548.
30. J. Gerel, A. Pethő and H. G. Zimmer, ‘On Mordell’s equation’, *Compositio Math.* 110 (1998) 335–367.
31. K. Györy, ‘On the number of solutions of linear equations in units of an algebraic number field’, *Comment. Math. Helv.* 54 (1979) 583–600.
32. K. Györy and K. Yu, ‘Bounds for the solutions of $S$-unit equations and decomposable form equations’, *Acta Arith.* 123 (2006) 9–41.
33. L. Hajdu and T. Herendi, ‘Explicit bounds for the solutions of elliptic equations with rational coefficients’, *J. Symbolic Comput.* 25 (1998) 361–366.
34. H. A. Helfgott and A. Venkatesh, ‘Integral points on elliptic curves and 3-torsion in class groups’, *J. Amer. Math. Soc.* 19 (2006) 527–550.
35. R. Juricevic, ‘Explicit estimates of solutions of some Diophantine equations’, *Funct. Approx. Comment. Math.* 38 (part 2) (2008) 171–194.
36. R. von Känel, ‘An effective proof of the hyperelliptic Shafarevich conjecture’, Preprint, 2013, arXiv:1310.6727, 25 pages.
37. R. von Känel, ‘The effective Shafarevich conjecture for abelian varieties of GL$2$-type’, Preprint, 2013, Sections 1.2, 4.5, 8 and 9 of arXiv:1310.7263, 41 pages.
38. R. von Känel, ‘On Szpiro’s Discriminant Conjecture’, *Int. Math. Res. Not.*, 2013, 35 pages, Published online: doi:10.1093/imrn/rnt079.
39. N. M. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, Annals of Mathematics Studies 108 (Princeton University Press, Princeton, NJ, 1985).
40. M. A. Kenku, ‘On the number of $\mathbb{Q}$-isomorphism classes of elliptic curves in each $\mathbb{Q}$-isogeny class’, *J. Number Theory* 15 (1982) 199–202.
41. M. Kim, ‘The motivic fundamental group of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ and the theorem of Siegel’, *Invent. Math.* 161 (2005) 629–656.
42. M. Kim, ‘$p$-adic $L$-functions and Selmer varieties associated to elliptic curves with complex multiplication’, *Ann. of Math.* (2) 172 (2010) 751–759.
43. M. Kim, ‘Remark on fundamental groups and effective Diophantine methods for hyperbolic curves’, Number theory, analysis and geometry (Springer, New York, 2012) 355–368.
44. Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Graduate Texts in Mathematics 6 (Oxford University Press, Oxford, 2002). Oxford Science Publications.
45. P. Lockhart, M. Rosen and J. H. Silverman, ‘An upper bound for the conductor of an abelian variety’, *J. Algebraic Geom.* 2 (1993) 569–601.
46. K. Mahler, ‘Zur Approximation algebraischer Zahlen. I’, *Math. Ann.* 107 (1933) 691–730.
47. L. Mai and M. R. Murty, ‘The Phragmén–Lindelöf theorem and modular elliptic curves’, *The Rademacher legacy to mathematics (University Park, PA, 1992)*, Contemporary Mathematics 166 (American Mathematical Society, Providence, RI, 1994) 335–340.
48. B. Mazur, ‘Rational isogenies of prime degree’, *Invent. Math.* 44 (1978) 129–162.
49. L. J. Mordell, ‘Note on the integer solutions of the equation $Ey^2 = Ax^4 + Bx^2 + Cx + D$’, *Messenger Math.* 51 (1922) 169–171.
50. L. J. Mordell, ‘On the integer solutions of the equation $3y^2 = ax^4 + bx^2 + cx + d$’, *Proc. Lond. Math. Soc.* 21 (1923) 415–419.
51. M. R. Murty, ‘Bounds for congruence primes’, *Automorphic forms automorphic representations and arithmetic (Fort Worth, TX, 1996)*, Proceedings of Symposia in Pure Mathematics 66 (American Mathematical Society, Providence, RI, 1999) 177–192.
52. M. R. Murty and H. Pasten, ‘Modular forms and effective Diophantine approximation’, *J. Number Theory* 133 (2013) 3739–3754.
53. A. N. Paršin, ‘Algebraic curves over function fields. I’, *Izv. Akad. Nauk SSSR Ser. Mat.* 32 (1968) 1191–1219.
54. F. Pellarin, ‘Sur une majoration explicite pour un degré d’isogénie liant deux courbes elliptiques’, Acta Arith. 100 (2001) 203–243.
55. D. Poulakis, ‘Corrigendum to the paper: ‘The number of solutions of the Mordell equation’ [Acta. Arith. 88 (1999) 173–179]’, Acta Arith. 92 (2000) 387–388.
56. M. Raynaud, ‘Hauteurs et isogénies’, Astérisque (1985) 199–234. Seminar on arithmetic bundles: the Mordell conjecture (Paris, 1983/1984).
57. K. A. Ribet, ‘Mod p Hecke operators and congruences between modular forms’, Invent. Math. 71 (1983) 193–205.
58. J.-P. Serre, ‘Facteurs locaux des fonctions zêta des variétés algébriques’, Séminaire DPP 19 (1969–1970).
59. J.-P. Serre, ‘Sur les représentations modulaires de degré 2 de Gal(̄Q/Q)’, Duke Math. J. 54 (1987) 179–230.
60. I. R. Shafarevich, ‘Algebraic number fields’, Proc. Internat. Congr. Mathematicians (Stockholm, Inst. Mittag-Leffler, Djursholm, 1962) 163–176.
61. G. Shimura, Introduction to the arithmetic theory of automorphic functions, Publications of the Mathematical Society of Japan 11 (Iwanami Shoten, Publishers, Tokyo, 1971), Kanô Memorial Lectures 1.
62. J. H. Silverman, ‘Integer points and the rank of Thue elliptic curves’, Invent. Math. 66 (1982) 395–404.
63. J. H. Silverman, ‘Heights and elliptic curves’, Arithmetic geometry (Storrs Conn., 1984) (Springer, New York, 1986) 253–265.
64. J. H. Silverman, The arithmetic of elliptic curves, 2nd edn, Graduate Texts in Mathematics 106 (Springer, Dordrecht, 2009).
65. H. M. Stark, ‘Effective estimates of solutions of some Diophantine equations’, Acta Arith. 24 (1973) 251–259. Collection of articles dedicated to Carl Ludwig Siegel on the occasion of his seventy-fifth birthday, III.
66. C. L. Stewart and K. Yu, ‘On the abc conjecture. II’, Duke Math. J. 108 (2001) 169–181.
67. L. Szpiro (ed.), Séminaire sur Les Pinceaux de Courbes Elliptiques (Société Mathématique de France, Paris, 1990); À la recherche de ‘Mordell effectif’. Papers from the seminar held in Paris, 1988, Astérisque No. 183 (1990).
68. R. Taylor and A. Wiles, ‘Ring-theoretic properties of certain Hecke algebras’, Ann. of Math. (2) 141 (1995) 553–572.
69. A. Wiles, ‘Modular elliptic curves and Fermat’s last theorem’, Ann. of Math. (2) 141 (1995) 443–551.
70. D. Zagier, ‘Modular parametrizations of elliptic curves’, Canad. Math. Bull. 28 (1985) 372–384.

Rafael von Känel
MPIM Bonn
Vivatsgasse 7
53111 Bonn
Germany
rvk@mpim-bonn.mpg.de