Descent for the punctured universal elliptic curve, and the average number of integral points on elliptic curves

Dohyeong Kim
February 18, 2015

Abstract

We show that the average number of integral points on an elliptic curve, counted modulo the natural involution on the punctured elliptic curve, is bounded from above by 32. To prove it, we design a descent map which associates a pair of binary forms to an integral point on an elliptic curve. Other ingredients of the proof include the upper bounds for the number of solutions of a Thue equation by Evertse, and Akhtari-Okazaki, and the estimation of number of binary quartic forms by Bhargava-Shankar. Our method applies to $S$-integral points to some extent, although our present knowledge is insufficient to deduce an upper bound for the average number of them. For $S$-integral points on the universal elliptic curve, the descent map gives rise to bijection between $S$-integral points and the orbit space of a representation of an algebraic group. The number of the orbits is effectively finite, and one can use it to determine $S$-integral points on all elliptic curves which have good reduction outside of $S$. This will be worked out numerically for $S = \{2\}$.

Contents

1 Introduction 2
2 Two binary forms associated to a point on an elliptic curve 4
3 Equivalence between pairs of binary forms. 9
4 Descent for the $S$-integral points on the punctured universal elliptic curve 12
5 The example $S = \{2\}$ 13
6 Average number of integral points on elliptic curves 16
7 Remarks 22
1 Introduction

One of the approaches to investigate solutions of a diophantine equation is to relate it to a certain kind of classification problem, which is usually called the method of descent.

For example, the descent for an elliptic curve, with respect to any fixed isogeny onto it, enables one to map the Mordell-Weil group of the elliptic curve to a subgroup of a Galois cohomology group, called the Selmer group. The involved Selmer group is by definition a classifying space, which can be computed readily in practice if one restricts to isogenies of small degree, thereby obtaining an upper bound for the rank of the elliptic curve. Apart from the practical importance, Selmer group is also at the heart of several theoretical approaches to the arithmetic of elliptic curves.

Another example is Fermat’s Last Theorem. Frey’s fundamental idea to associate an elliptic curve to a solution of Fermat’s equation, regardless of its existence, paved the way to attack the problem from a completely different angle, which is not so apparent in the original form of equation. The elliptic curve associated to a solution of Fermat’s equation satisfies a number of constraints, and one shows that the constraints are so restrictive that there cannot exist such a curve.

There are more recent examples which arise from the work of Bhargava on the rational points on hyperelliptic curves, as well as the work of Bhargava and Shankar on the average rank of elliptic curves. By replacing certain diophantine problem with a classification problem, they were able to estimate the number of rational points on hyperelliptic curves, or the number of elements in Selmer groups of elliptic curves, as the curves vary in families.

Yet another such example regarding cubic Thue-Mahler equation is given in [6], which reduces the equation to a classification of elliptic curves of a given conductor.

The aim of the present paper is to devise a descent map which allows one to determine the complete set of $S$-integral points on a punctured elliptic curve by classifying orbits of a representation of an algebraic group, where $S$ is a finite set of prime numbers. We have two applications.

Firstly, we will show that the average number of integral points on elliptic curves defined over the rational numbers, counted modulo the natural involution on the punctured elliptic curve, is bounded from above by 32. Here the natural involution refers to negation with respect to the group law of the elliptic curve. Such a statement does not make sense until one specifies how the curves are ordered, as well as the $\mathbb{Z}$-model of the curve where integrality is defined. As long as the average is concerned, we shall write an elliptic curve in the form

$$E: y^2 = x^3 + ax + b, \quad a, b \in \mathbb{Z}$$

where $a, b$ are integers, such that there is no prime $p$ such that $p^4 | a$ and $p^6 | b$. Such an equation is unique for an isomorphism class of elliptic curves over $\mathbb{Q}$, and we order the curves by the size of $a^3$ and $b^2$, up to some unimportant scaling.
factors we specify later. Also, an integral point on \( E \) should be understood as an integral point with respect to the above particular model.

**Theorem 1.1** (See Theorem 6.1). The average number of integral points on an elliptic curve is bounded from above by 32, when the elliptic curves are ordered with respect to the size of their coefficients in the simplified Weierstrass equation. The number of integral points are counted modulo the natural involution. That is to say, a point \((x_0, y_0) \in E(\mathbb{Z})\) is regarded equivalent to \((x_0, -y_0)\).

Our proof will exploit two facts besides our descent map. The first ingredient is the work of Bhargava-Shankar on the number of binary quartics with bounded invariants, and the second one is the works of Evertse and Akhtari-Okazaki on the number of solutions of a quartic Thue equation. Note that our argument will not involve the ranks elliptic curves, nor arithmetic invariants of auxiliary number fields. To the best knowledge of the author, the previously known bounds for the number of integral points on a particular elliptic curve depend exponentially either on the rank of the curve, or the rank of certain ideal class group of a number field such as the two-division field of the curve. Combining this type of upper bounds for the number of integral points on a individual curves, with an analysis on the distribution of ranks, one might try to obtain an upper bound for the average number of points on elliptic curves. Indeed, Alpoge claimed [2] that the strategy works, even for the moment for the number of points. Interestingly, the method of Alpoge produces an upper bound 65.8457 for the number of integral points, counted without taking account of the natural involution. This upper bound is only slightly worse than our upper bound, when one recalls that the fixed points of the involution have negligible density. We give a more detailed comparison between two approaches in the last section.

The other application of our descent map regards the \( S \)-integral points on elliptic curves which have good reduction outside of \( S \), without averaging. When we apply our descent map to \( S \)-integral points on the punctured elliptic curves which have good reduction outside of \( S \), we are lead to an explicit bijection between the set of \( S \)-integral points on the punctured universal elliptic curve and the orbits of a certain action of an algebraic group. By the punctured universal elliptic curve, denoted by \( \mathcal{Y} \), we refer the moduli stack which classifies pairs \((Y, t)\) where \( Y \) is a punctured elliptic curve and \( t \) is a points of \( Y \).

**Theorem 1.2** (See Theorem 4.1). Let \( \mathcal{Y} \) be the punctured universal elliptic curve. There is an algebraic group \( G/\mathbb{Z} \) and a representation \( V/\mathbb{Z} \) of \( G \), such that the descent map induces a bijection

\[
\mathcal{Y}(\mathbb{Z}_S)/\{\pm 1\} \cong V(\mathbb{Z}_S)/G(\mathbb{Z}_S)
\]  

for any finite set \( S \) of prime numbers such that \( 2, 3 \in S \).

The utility of such a bijection would depend crucially on one’s knowledge about the orbits. We know that the number of orbits is finite, effectively, but it is not quite a trivial task to determine all the orbits numerically. Indeed, we are barely able to give one non-trivial example when \( S = \{2\} \), using the
work of Smart, in which he reduces the problem of solving several auxiliary equations. Nevertheless, we hope this concrete example illustrates the potential of our method to produce a systematic approach to effective Mordell conjecture for punctured elliptic curves, as necessary classification techniques will have been established.

We outline the organisation of the paper. In Section 2 we define the descent map, which associate two integral binary forms to a point on the universal elliptic curve. In Section 3 we review the basics on equivalence between pairs of binary forms. In Section 4 we use the descent map to identify $S$-integral points on universal elliptic curves with certain equivalence classes of pairs of binary forms. In Section 5 we work out the numerical example when $S = \{2\}$. In Section 6, we use the descent map together with the works of Akhtari-Okazaki and Bhargava-Shankar to obtain an upper bound for the average number of integral points on elliptic curves. In Section 7 we give some concluding remarks.

2 Two binary forms associated to a point on an elliptic curve

The aim of the present section is to define two integral binary forms associated to a point on an elliptic curve, and study its basic properties.

We begin with notations. Let $E$:

$$E: y^2z + a_1xyz + a_3yz^2 = x^3 + a_2x^2z + a_4xz^2 + a_6z^3$$

be an elliptic curve written in a generalised Weierstrass equation whose coefficients are rational integers. If $t$ is a $\mathbb{Z}$-point of $E$, then we shall write

$$t = (x_t, y_t, z_t)$$

where $x_t, y_t$, and $z_t$ are relatively prime integers.

Let $Y$ be the elliptic curve punctured at the origin. In other words, $Y$ is the open subscheme of $E$ defined by the complement of the vanishing locus of $z$. If $S$ is any finite set of primes, we denote by $\mathbb{Z}_S$ the ring of $S$-integers. Then, $\mathbb{Z}_S$-points of $Y$ can be described as

$$Y(\mathbb{Z}_S) = \{ t = (x_t, y_t, z_t) : t \in E(\mathbb{Z}), z_t \in \mathbb{Z}_S^x \}.$$  (5)

For each point $t \in Y(\mathbb{Z}_S)$, we will construct two binary forms of degree one and four respectively. We denote them by $L_t$ and $Q_t$, where the letters are chosen to suggest that they are linear and quartic forms. The variables of $L_t$ and $Q_t$ will be denoted by $u$ and $v$, so we shall often write $L_t(u, v)$ and $Q_t(u, v)$ in order to emphasise the variables. We explain the construction of $L_t(u, v)$ and $Q_t(u, v)$ below.

The construction of $L_t$ is straightforward. Independently of $t$, we let

$$L_t(u, v) = v$$  (6)
which is regarded as a linear form in variables $u$ and $v$. For the geometric reason lying under this hardly motivating definition, see Remark \ref{21}

The construction of $Q_t$ is slightly more involved, though it is a classical one which is often used in two-descent for elliptic curves. Let $\mathbb{P}^2_{xyz}$ be the projective plane with homogeneous projective coordinates $x, y, z$. Note that $E$ is given as a cubic curve in $\mathbb{P}^2_{xyz}$. For a given $t \in Y(\mathbb{Z}_s)$, the lines in $\mathbb{P}^2_{xyz}$ which pass through $t$ are parametrised by the linear forms

$$ux + vy + wz = 0$$

such that

$$ux_t + vy_t + wz_t = 0$$

is satisfied. Under the assumption that $z_t \neq 0$, such lines are parametrised by $u$ and $v$, because we can uniquely recover $w$

$$w = \frac{ux_t + vy_t}{-z_t}$$

from $u$ and $v$.

The quartic form $Q_t(u, v)$, which will be determined explicitly shortly, is characterised by the property that its four zeros represent the four lines which are the ramification points of the projection map from $E$ to the space of lines through $t$.

\begin{proposition}
The quartic form $Q_t(u, v)$ is given by

$$A^2 - 4v^2B$$

where $A$ and $B$ are given as

$$A = -z_tu^2 + z_t a_1uv + (a_2z_t + x_t) v^3$$

$$B = x_t z_t u^2 + (2y_t z_t + z_t^2 a_3) uv + (a_4z_t^2 - a_1z_t y_t + a_2z_t x_t + x_t^2) v^2.$$  

\end{proposition}

\textit{Proof.} It is a straightforward calculation. We seek for the condition that the line

$$ux + vy + wz = 0$$

is tangent to $E$. Thus, we substitute

$$y = \frac{ux + wz}{-v}$$

to

$$y^2z + a_1xyz + a_3yz^2 - (x^3 + a_2x^2z + a_4xz^2 + a_6z^3)$$

5
and obtain a cubic form $C(x, z)$ in $x$ and $z$. Using the condition that $t$ satisfies both (13) and (15), one observes that $C(x, z)$ should have a factorisation

$$C(x, z) = \frac{1}{ztv^2}(xz_t - zt) \cdot q(x, z)$$

where $q(x, z)$ is a quadratic form in $x$ and $z$ whose coefficients are quadratic in $u$ and $v$. By expanding the right hand side of (16) and equating the coefficients of it with those of $C(x, z)$, one obtains

$$q(x, z) = v^2x^2 + Axy + By^2$$

where $A$ and $B$ are polynomials given in the statement of the proposition. Since the condition that the line is ramification point of the projection map is equivalent to the condition that the discriminant of $q(x, z)$ is zero. Thus we have obtained the formula of $Q_t(u, v)$.

**Remark 2.1.** The linear form $L_t(u, v) = v$ acquires the following geometric interpretation once we view $u$ and $v$ as parameter for the lines passing through $t$. The zero of $L_t(u, v)$ is $(u, v) = (1, 0)$, which corresponds to the line

$$x - \frac{x_t}{z_t}z = 0$$

which is the line passes though $t$ and the origin of $E$.

**Remark 2.2.** In the context of two-descent for the elliptic curve $E$, $Q_t(u, v)$ represents a torsor for $E[2]$, the group of two division points of $E$.

Let us work out some numerical examples in order to ensure that the formula of $Q_t(u, v)$ is correct and to illustrate the nature of $Q_t(u, v)$. Let us consider

$$E: y^2z + yz^2 = x^3 - xz^3$$

which is the curve of conductor 37. It has no non-trivial rational point of order two. Its rank is one, and the Mordell-Weil group is generated by the point

$$P_0 = (0, 0, 1).$$

Let us take $t = n \cdot P$.

For $n = 1$, one gets

$$Q_t(u, v) = u^4 - 4uv^3 + 4v^4$$

which is irreducible.

For $n = 2$, we have $t = (1, 0, 1)$. One readily computes that

$$Q_t(u, v) = u^4 - 6u^2v^2 - 4uv^3 + v^4$$

which factors as

$$(u + v)(u^3 - u^2v - 5uv^2 + v^3)$$
verifying that the corresponding torsor is trivial. 
For \( n = 3 \), we have \( t = (−1, −1, 1) \). Similarly, we have
\[
Q_t(u, v) = u^4 + 6u^2v^2 + 4uv^3 + v^4
\]  
which is irreducible.
As a second example, consider
\[
E: y^2z = x^3 − 1681xz^2.
\]  
Since 1681 = 41^2, it has three rational points of order two, and the Mordell-Weil group has rank two, generated by
\[
P_1 = (−9, 120, 1)
\]
\[
P_2 = (841, 24360, 1).
\]
For \( t = P_1 \), we have
\[
Q_t(u, v) = u^4 + 54u^2v^2 − 960uv^3 + 6481v^4
\]  
which is irreducible.
For \( t = P_2 \), we have
\[
Q_t(u, v) = u^4 − 5046u^2v^2 − 194880uv^3 − 2115119v^4
\]  
which factors as
\[
(u^2 − 58uv − 2521v^2)(u^2 + 58uv + 839v^2)
\]  
but does not possess a linear factor.
For \( t = 2 \cdot P_1 \), one has
\[
t = (93139320, 443882159, 1728000)
\]  
and
\[
Q_t(u, v) = 43200(40u − 827v)(120u + 143v)(120u + 719v)(120u + 1619v).
\]  
It verifies that \( Q_t(u, v) \) defines the trivial torsor.
Now we turn to the key proposition regarding both \( L_t(u, v) \) and \( Q_t(u, v) \).

**Proposition 2.2.** Let \( \Delta_E \) be the discriminant of \( E \), and let \( S \) be any finite set of primes numbers. Let \( \Delta_t \) be the discriminant of binary quintic form \( L_t(u, v) \cdot Q_t(u, v) \). Then \( \Delta_t \) is a unit in \( \mathbb{Z}_S[[2\Delta_E]^{-1}] \).

**Proof.** Let \( p \) be an odd prime such that \( p \) does not divide \( \Delta_E \) and \( p \) does not belong to \( S \). In order to prove the proposition, it suffices to show that \( \Delta_t \) is prime to \( p \). We proceed in two steps.
Firstly, we will show that the discriminant of \( Q_t(u, v) \) is prime to \( p \). Let \( t \in Y(\mathbb{Z}_S) \), and let \( t_p \) be the reduction of \( t \) modulo \( p \). Let \( E_p \) be the reduction of \( E \) modulo \( p \). Consider the twisted multiplication-by-two map

\[
\theta: E_p \to E_p
\]

\[
s \mapsto -2s
\]

which is a separable morphism since \( p \) is odd. Also, the degree of \( \theta \) is four. It follows that

\[
\theta^{-1}(t_p)
\]

has four geometric points. Connecting the four geometric points with \( t_p \), we obtain four lines passing through \( t \), and these four lines are precisely represented by the zeroes of \( Q_t(u, v) \) modulo \( p \). The non-vanishing of the discriminant of \( Q_t(u, v) \) modulo \( p \) is equivalent to the condition that four lines are distinct. Suppose that two of the four lines coincide, say \( L_0 \). Then \( L_0 \) contains \( s_1, s_2 \in \theta^{-1}(t) \) which are distinct. Furthermore, \( L_0 \) is tangent to \( E_p \) at \( s_1 \) and \( s_2 \) by construction. This contradicts that \( L_0 \) and \( E_p \) intersects with multiplicity three, and we completed the proof of the first step, showing that the discriminant of \( Q_t(u, v) \) is prime to \( p \).

Now we proceed to the second step, aiming to show that \( \Delta_t \) is indeed prime to \( p \). There are two ways to show this; one is geometric, while the other is algebraic and straightforward. We first give a geometric proof. Note that the condition that \( t \) belongs to \( Y(\mathbb{Z}_S) \) implies that \( t_p \) is not the origin of the reduced elliptic curve \( E_p \). This implies that none of the four geometric points belonging to \( \theta^{-1}(t_p) \) is the origin of \( E_p \). Indeed, suppose on the contrary that \( s \) is a geometric point of \( \theta^{-1}(t_p) \) and \( s \) is the origin of \( E_p \). Then \( \theta(s) = t_p \) implies, by definition of \( \theta \), that

\[
-2s = t_p,
\]

which implies \( t_p = 0 \). It contradicts that \( t_p \) is not the origin of \( E_p \). This observation in turn implies that none of the four lines defined by the zeroes of \( Q_t(u, v) \) modulo \( p \) passes through the origin. Suppose, on the contrary, that there is a line \( L_0 \) which passes through one of the four points of \( \theta^{-1}(t_p) \), say \( s_0 \), and further passes through the origin. Note that \( s_0 \) cannot be the origin. If \( s_0 \) were the origin of \( E_p \), then \( L_0 \) is the inflection line to \( E_p \), meeting with \( E_p \) at \( s_0 \) with multiplicity three. It would contradict that \( L_0 \) passes through \( t_p \), which we had shown to be different from the origin. Thus it follows that \( L_0 \) meets \( E_p \) with multiplicity at least four, to which the origin contributes at least one, and \( t_p \) and \( s_0 \) together contributes three. It is a contradiction, showing that none of the four lines passes through the origin of \( E_p \). It follows that the line passes through \( t_p \) and the origin is not one of the four lines obtained by connecting points of \( \theta^{-1}(t_p) \) and \( t_p \). Since the linear form \( L_v(u, v) \) corresponds to the line passing through \( t_p \) and the origin, we conclude that \( \Delta_t \), the discriminant of \( L_t(u, v) \cdot Q_t(u, v) \) is prime to \( p \).
Now we give an algebraic proof for the second step. It is based on the representation of the discriminant as a product of root differences. Then, we have
\[ \Delta_t = \delta_t \cdot Q_t(1, 0)^2 \] (37)
which follows from the representation of the discriminant as square of the product of all possible differences between roots. In the first step, we showed that \( \delta_t \) is prime to \( p \), so it remains to show that \( Q_t(1, 0) \) is prime to \( p \). This follows immediately from our explicit formula for \( Q_t(u, v) \) given in Proposition 2.1.

Indeed, we have
\[ Q_t(1, 0) = z_t^2 \] (38)
which is prime to \( p \) if \( t \in Y(\mathbb{Z}_S) \).

3 Equivalence between pairs of binary forms.

There are several notions for equivalence between pairs of binary forms. The aim of the current section is to define the notion of equivalence which is relevant to our purpose.

Let \( S \) be a finite set of primes. We will always assume that
\[ 2 \in S \] (39)
in the present section, although it is only necessary when we consider the notion of minimality. Let us consider a pair \((L, Q)\) of binary forms
\[
L = b_0u + b_1v \quad (40)
\]
\[
Q = c_0u^4 + c_1u^3v + c_2u^2v^2 + c_3uv^3 + c_4v^4 \quad (41)
\]
where \( b_i \)'s and \( c_i \)'s are \( S \)-integers. We always assume that the coefficients of \( L \) and \( Q \) do not have non-trivial common divisors in \( \mathbb{Z}_S \). More precisely, we assume that the ideal of \( \mathbb{Z}_S \) generated by \( b_0 \) and \( b_1 \) is the unit ideal, and similarly the ideal of \( \mathbb{Z}_S \) generated by \( c_0, c_1, \ldots, c_4 \) is also the unit ideal.

The discriminant of \( Q \), denoted by \( \Delta_Q \), is given by
\[
\Delta_Q = c_1^2c_2^2c_3^2 - 4c_0c_3^2c_4^2 - 4c_1^3c_4^3 + 18c_0c_1c_2c_3c_4 - 27c_0^2c_3^4 - 4c_1^2c_2c_4
+ 16c_0c_2^4c_4 + 18c_1^2c_2c_3c_4 - 80c_0c_1c_2^2c_3c_4 - 6c_0c_1^2c_3^2c_4 + 144c_0^2c_2c_3^2c_4
- 27c_1^4c_4^2 + 144c_0c_1^2c_2c_4^2 - 128c_0^2c_2^2c_4^2 - 192c_1^2c_3c_4^2 + 256c_0^3c_4^4 \quad (42)
\]
and the discriminant of \( L \cdot Q \), denoted by \( \Delta \), is given by
\[
\Delta = \Delta_Q \cdot Q(-b_1, b_0)^2. \quad (43)
\]
For a fixed \( S \), we will be concerned with pairs of forms for which \( \Delta \) is \( S \)-unit. We introduce the following definition for convenience.
Definition 3.1. Let \((L, Q)\) be a pair of binary forms with \(S\)-integral coefficients as above. We say that this pair of \(S\)-admissible if \(\Delta\) is an \(S\)-unit.

Let \((L, Q)\) and \((L', Q')\) be two \(S\)-admissible pairs. There is, of course, the obvious notion of equality between them, defined by the coefficient-wise equality. A weaker notion of equality, which is more natural if we view them as elements of projective space, is the following.

Definition 3.2. Let \((L, Q)\) and \((L', Q')\) be two \(S\)-admissible pairs. We say that two pairs are projectively equivalent if there are \(\lambda_1, \lambda_2 \in \mathbb{Z}_S^*\) such that

\[
\begin{align*}
L &= \lambda_1 L' \\
Q &= \lambda_2 Q'
\end{align*}
\]

holds.

Note that this definition does make sense among \(S\)-admissible pairs, because if \(\Delta\) is the discriminant of \((L, Q)\), then the discriminant of \((\lambda_1 L, \lambda_2 Q)\) is \(\lambda_1^8 \lambda_2^2 \Delta\).

Now we introduce the desired notion of equivalence.

Definition 3.3. Let \((L, Q)\) and \((L', Q')\) be two \(S\)-admissible pairs. We say that they are \(\text{GL}_2\)-equivalent, if there is \(g \in \text{GL}_2(\mathbb{Z}_S)\) such that \((L^g, Q^g)\) is projectively equivalent to \((L', Q')\). Here \(g\) acts on \(L\) and \(Q\) by the linear change of variables.

We would like to take a closer look at the notion of \(\text{GL}_2\)-equivalence. Note that we made assumption \(2 \in S\), which has not been used so far in this section. If \(2 \in S\), then for each \(S\)-admissible pair \((L, Q)\), it is possible to find a pair \((L', Q')\), which is \(\text{GL}_2\)-equivalent form, such that

\[
\begin{align*}
L' &= v \\
Q' &= u^4 + B_2 u^2 v^2 + B_3 u v^3 + B_4 v^4
\end{align*}
\]

where \(B_2, B_3, B_4\) are integers. Furthermore, it is possible, as we will prove shortly, to choose a minimal one in the following sense.

Definition 3.4. A pair of binary forms

\[
(v, u^4 + B_2 u^2 v^2 + B_3 u v^3 + B_4 v^4)
\]

with integral coefficient is called minimal, if for every prime \(p\) such that \(p^i | B_i\) for \(i = 2, 3, 4\) simultaneously. If the form has \(S\)-integral coefficient, then it is called minimal at \(p\) for a prime \(p \notin S\), when \(p^i | B_i\) for \(i = 2, 3, 4\) does not hold simultaneously.
Proposition 3.1. Recall that 2 is always in $S$. Given any pair $(L, Q)$ of binary forms as above, it is possible to find a minimal pair
\[(v, u^4 + B_2 u^2 v^2 + B_3 uv^3 + B_4 v^4)\] which is $\text{GL}_2$-equivalent to $(L, Q)$. Such a minimal pair is unique up to replacing $B_3$ with $B_3$. In other words, such a minimal pair is unique if $B_3 = 0$, and there are precisely two such pairs if $B_3 \neq 0$.

Proof. The proof is by elementary algebra. Let $(L, Q)$ be an $S$-admissible pair given by
\[L = b_0 u + b_1 v\]
\[Q = c_0 u^4 + c_1 u^3 v + c_2 u^2 v^2 + c_3 uv^3 + c_4 v^4.\]
Since $b_0$ and $b_1$ generate the unit ideal in $\mathbb{Z}_S$, by a linear change of variables, we may assume $L = v$. Then, $c_0$ must be $S$-unit. Otherwise,
\[Q(b_1, -b_0) = c_0\]
divides $\Delta$, contradicting the $S$-admissibility of the pair. Thus, via a projective equivalence, we may assume that $c_0 = 1$. Now we have a pair
\[L = v\]
\[Q = u^4 + c_1 u^3 v + c_2 u^2 v^2 + c_3 uv^3 + c_4 v^4.\]
where the coefficients are in $\mathbb{Z}_S$. Since we assumed $2 \in S$, we are allowed make the substitution
\[u \mapsto u - \frac{c_1}{4} v\]
if necessary, so we may assume that $c_1 = 0$ as well. Since the denominators of $c_2, c_3, c_4$ are $S$-units, we may multiply an $S$-unit to $v$, and apply projective equivalence, in order to get a minimal form.

The only linear change of variables which preserve the condition that $Q$ is monic in $u$, $c_1 = 0$, and there is no prime $p$ such that $p^i | c_i$ simultaneously, is
\[(u, v) \mapsto (\lambda_1 u, \lambda_2 v)\]
where $\lambda_1$ is a fourth root of unity, and $\lambda_2$ is a unit. Thus, all possible minimal pairs which is equivalent to a given minimal form
\[L = v\]
\[Q = u^4 + u^3 v + c_2 u^2 v^2 + c_3 uv^3 + c_4 v^4.\]
can be obtained by replacing $c_3$ with $-c_3$. $\square$

Remark 3.1. It is worth noting that if we work over a general number field, then the number of possible minimal forms may grow. However, the involution $c_3 \mapsto -c_3$ on the set of minimal forms maintains an exceptional importance, since it will correspond to the negation on the elliptic curve.
4 Descent for the $S$-integral points on the punctured universal elliptic curve

We apply the results from the previous sections in order to classify $S$-integral points on the universal elliptic curve. We denote by $\mathcal{Y}$ the universal elliptic curve, whose $\mathbb{Z}_S$ points are given by

$$\mathcal{Y}(\mathbb{Z}_S) = \{(Y, P) : P \in Y(\mathbb{Z}_S), Y \text{ is punctured smooth elliptic curve over } \mathbb{Z}_S\}$$

(59)

where a smooth elliptic curve over $\mathbb{Z}_S$ means an elliptic curve over $\mathbb{Q}$ which has good reduction outside of $S$. Note that the $\mathbb{Z}_S$ points on the curve is well defined provided that the curve has good reduction outside of $S$.

There is obvious action of the group $\{\pm 1\}$ of order two on $\mathcal{Y}(\mathbb{Z}_S)$, given by

$$\pm 1 : (Y, P) \mapsto (Y, \pm P)$$

(60)

where the negation denotes the negation under the group law of the elliptic curve. As promised in the introduction, we will prove the following theorem in the present section.

**Theorem 4.1.** Assume $2, 3 \in S$. There is a bijection

$$\kappa : \mathcal{Y}(\mathbb{Z}_S)/\{\pm 1\} \rightarrow \{S\text{-admissible pairs}\}/\sim$$

(61)

where $\sim$ is the $GL_2$-equivalence relation.

**Proof.** We will prove the assertion by constructing the inverse. Let

$$(v, u^4 + B_2u^2v^2 + B_3uv^3 + B_4v^4)$$

(62)

be an $S$-admissible pair, which is minimal away from $S$. In particular, $B_2, B_3, B_4$ are $S$-integers, and the discriminant

$$-4B_2^3B_3^2 + 16B_2^3B_4 - 27B_4^4 + 144B_2B_3^2B_4 - 128B_2^2B_4^2 + 256B_4^3$$

(63)

is an $S$-unit. By defining

$$x_t = -\frac{1}{6}B_2$$

(64)

$$y_t = -\frac{1}{8}B_3$$

(65)

$$a_4 = -\frac{1}{4}(B_4 + 3x_t^2)$$

(66)

$$a_6 = y_t^2 - x_t^3 - a_4x_t$$

(67)

we obtain a curve

$$E : y^2 = x^3 + a_4x + a_6$$

(68)
which is defined over \( \mathbb{Z}_S \), and has a point \( t = (x_t, y_t) \). Note that we have to divide by 6 in order to get \( x_t \), hence rely on the assumption that \( 2, 3 \in S \). We need to show that \( E \) has good reduction outside of \( S \). By direct computation, the discriminant of \( E \) is given by

\[
2^{-8} \cdot (-4B_2^2B_3^3 + 16B_2^4B_4 - 27B_3^4 + 144B_2B_3^2B_4 - 128B_2^2B_4^2 + 256B_4^3) \quad (69)
\]

which is an \( S \)-unit by comparison to the formula (63) for the discriminant of \( Q(u, v) \).

We need to verify that the association \( (L, Q) \mapsto (E, P) \) is well-defined. As we observed earlier, there is an involution on the set of minimal pairs sending \( B_3 \) to \(-B_3\). It is clear from the formula (65) that it corresponds to the involution \( (E, P) \mapsto (E, -P) \). Thus we have constructed a section of \( \kappa \), showing its surjectivity.

To see the injectivity of \( \kappa \), recall that if \( 2, 3 \in S \), hence an elliptic curve \( E \) which has good reduction outside of \( S \) has a model of the form (68) which has good reduction outside of \( S \), and there is no prime \( p \) for which \( p^4|a_4 \) and \( p^6|a_6 \). Starting with a model of \( E \) which is minimal outside of \( S \), we will show that the pair \( (L, Q) = \kappa(E, P) \) is minimal away from \( S \). By the explicit formula of \( (L, Q) \) given in Proposition 2.1, we have

\[
Q(u, v) = u^4 - 6x_tu^2v^2 - 8y_tw^3 - (3x_t^2 + 4a_4)v^4 \quad (70)
\]

and we claim that it is minimal away from \( S \). Suppose on the contrary that there is a prime \( p \not\in S \) for which \( Q(u, v) \) is not minimal. Since \( 2, 3 \in S \),

\[
p^2|x_t \quad (71)
\]

\[
p^4|3x_t^2 + 4a_4 \quad (72)
\]

from which we conclude that \( p^4|a_4 \). Furthermore, non-minimality at \( p \) implies \( p^4|y_t \). However, by rewriting the equation of the elliptic curve in the form

\[
a_6 = y_t^2 - x_t^3 - a_4x_t \quad (73)
\]

one sees \( p^6 \) divides \( a_6 \). This contradicts the minimality of \( E \) at \( p \).

Thus we have shown that \( \kappa \) is bijection.

5 The example \( S = \{2\} \)

The aim of present section is to give a numerical example, in which one determines \( \mathcal{Y}(\mathbb{Z}_S)/\pm 1 \) from the knowledge of a set of representatives for all \( S \)-admissible pairs. Although we assumed \( 3 \in S \) in Theorem 4.1, as long as numerical examples are concerned, the assumption \( 3 \in S \) is not strictly necessary. Indeed, the map \( \kappa \) exists anyway, and for each \( S \)-admissible pair, one obtains a point of \( \mathcal{Y} \) defined over \( \mathbb{Z}_S[3^{-1}] \). One can proceed to verify whether this point is in fact defined over \( \mathbb{Z}_S \) or not, and by collecting those with an affirmative
answer, one obtains \( V(\mathbb{Z}_S)/\{\pm 1\} \). The assumption that \( 3 \in S \) assures that the last step is unnecessary.

Despite of the fact that the finiteness theorem for the number of equivalence classes of \( S \)-admissible pairs is effective, determination of it in practice can be rather challenging. In this section, we use the work of N.P. Smart who computed the all reducible binary quintic whose discriminant and \( S \)-unit with \( S = \{2\} \). All \( S \)-admissible pairs can be obtained from the work of Smart, by choosing all possible linear factors of each binary quintic.

Table 1 is produced from Table 5 of [7], which contains all reducible binary quintic forms whose discriminant is a power of 2 up to sign. In [7], the table is titled to contain all reducible binary quintic forms with 2-power discriminant, which might cause unnecessary confusion that the table is restricted to forms with positive discriminant. Thus we chose the expression that the discriminant is a power of 2 up to sign, which is equivalent to saying that the discriminant is \( S \)-unit with \( S = \{2\} \).

| \( i \) | \( f_i(u, v) \) | \( i \) | \( f_i(u, v) \) |
|-------|----------------|-------|----------------|
| 1     | \( u^3v + u^3v^3 + u^4v^3 + uv^4 \) | 2     | \( 2u^3v + 2u^3v^3 - u^2v^3 - uv^4 \) |
| 3     | \( 8u^6 - 6u^5v^2 + uv^4 \) | 4     | \( 2u^3v - 3u^2v^3 + uv^4 \) |
| 5     | \( u^5 + 4uv^4 \) | 6     | \( u^5 + 3u^2v^3 + 2uv^4 \) |
| 7     | \( u^5v + 3u^4v^3 + 2u^3v^4 \) | 8     | \( u^5 + 2u^4v^3 + 4u^3v^3 + 4uv^4 \) |
| 9     | \( u^6 + 3u^5v^2 + 2u^4v^3 + 2u^3v^4 + u^2v^5 - v^6 \) | 10    | \( u^6 - 4uv^4 \) |
| 11    | \( u^6 + 4u^5v + 4u^4v^3 + 8u^3v^5 + 4uv^6 \) | 12    | \( u^6 - 4u^5v + 8u^3v^3 - 4uv^4 \) |
| 13    | \( u^6 - 8u^5v^3 + 12u^4v^4 + 16u^3v^5 - 28v^6 \) | 14    | \( u^6 + u^5v + u^4v^3 + v^5 \) |
| 15    | \( u^6 + u^5v + u^4v^3 + v^5 \) | 16    | \( u^6 + 12u^5v + 4uv^6 \) |
| 17    | \( u^6 - 2uv^5 \) | 18    | \( u^6 + u^5v - 2uv^6 - 2v^7 \) |
| 19    | \( u^6 + 2uv^6 \) | 20    | \( u^6 + 2uv^6 \) |
| 21    | \( 4u^6v + 4u^5v^3 - 16uv^4 + 9v^6 \) | 22    | \( 3u^6v + 8u^5v^3 + 4uv^4 \) |
| 23    | \( 4u^6v - 4u^5v^2 - 16uv^4 + 9v^6 \) | 24    | \( u^6 - 4u^5v^3 + 2uv^4 \) |
| 25    | \( u^6 + 2u^5v - 4u^4v^3 - 8u^3v^5 + 2uv^4 + 4v^3 \) | 26    | \( u^6 + 4u^5v^3 + 2uv^4 \) |
| 27    | \( u^6 + 4u^5v - 4u^4v^3 - 4u^3v^5 + 2uv^4 + 2v^3 \) | 28    | \( u^6 + 9u^5v + 14u^4v^3 - 34u^3v^5 - 19uv^4 + 5v^6 \) |
| 29    | \( u^6 + 4u^5v - 6u^4v^3 - 4u^3v^5 + uv^4 \) | 30    | \( 4u^6v + 16u^5v^3 - 24u^4v^5 + 17v^6 \) |
| 31    | \( 4u^6v + 12u^5v - 28u^4v^3 - 12u^3v^5 + 41uv^4 - 17v^6 \) | 32    | \( u^6 - 8u^5v + 4u^4v^3 + 16u^3v^5 + 4uv^4 \) |
| 33    | \( u^6 - 7u^5v - 4u^4v^3 + 20u^3v^5 + 20uv^4 + 4v^5 \) | 34    | \( u^6 + 4u^5v^3 + 2uv^4 \) |
| 35    | \( u^6 + u^5v + 2u^4v^3 + 2v^3 \) | 36    | \( u^6 + 2u^5v^3 - v^6 \) |
| 37    | \( u^6 + u^5v - 2u^4v^3 - 2uv^4 - uv^6 - v^7 \) | 38    | \( u^6 - 2uv^5 - uv^6 \) |
| 39    | \( u^6 + 4u^5v^3 + 4uv^4 \) | 40    | \( u^6 + 4u^5v^3 + 4uv^4 \) |
| 41    | \( u^6 + u^5v + 4u^4v^3 + 4uv^4 - 4uv^4 - 4v^5 \) | 42    | \( u^6 + 4u^5v^3 - 6u^4v^5 + 12uv^4 - 7v^6 \) |
| 43    | \( u^6 + 3u^5v^2 - 10u^4v^3 + 18u^3v^5 - 19uv^4 + 7v^6 \) | 44    | \( u^6 - 2uv^5 + 2uv^6 \) |
| 45    | \( u^6 - 2u^5v^3 + 2v^3 \) | 46    | \( u^6 + u^5v - 2uv^5 - 2u^3v^5 + 3uv^4 + 2v^6 \) |
| 47    | \( u^6 + 4u^5v^3 + 8uv^4 \) | 48    | \( u^6 + 4uv^4 + 8v^5 \) |
| 49    | \( 5u^6 + 13u^5v + 2u^4v^3 - 14uv^4 - 3uv^5 + 5v^6 \) | 50    | \( u^6 + 6u^5v^3 + 8uv^4 + 5v^6 \) |
| 51    | \( u^6 + 4u^5v^3 - 8u^4v^5 + 4uv^4 \) | -     | \( - \) |
We wish to find all \( \{2\}\)-admissible pairs \((L, Q)\) from Table 1. For each \((L, Q)\), \(L \cdot Q\) must be equivalent to \(f_i\) for some \(i\), hence we can find all of them by considering \(f_i\) and factoring them into one linear and one quartic forms. In fact, \(f_i\) for \(1 \leq i \leq 4\) has three linear factors, and the rest have a unique linear factor.

Let us work out the case of \(i = 1\). In this case, \(f_1(u, v)\) factors as

\[
u u(u + v)(u^2 + v^2)
\]

hence there are three pairs

\[
(v, u(u + v)(u^2 + v^2)), \ (u, v(u + v)(u^2 + v^2)), \ (u + v, uv(u^2 + v^2))
\]

associated to \(f_1(u, v)\). Applying \((u, v) \mapsto (v, u)\) one sees that the first two pairs are equivalent. Transforming them into minimal forms, we obtain two pairs

\[
(L_1, Q_1) = (v, u^4 + 10u^2v^2 + 40uv^3 - 51v^4)
\]

\[
(L_2, Q_2) = (v, u^4 - v^4)
\]

in their minimal form. From \((L_1, Q_1)\), we obtain curve

\[
E_1: y^2 = x^3 + \frac{32}{3}x + \frac{1280}{27}
\]

with point

\[
t = \left(\frac{-5}{3}, -5\right)
\]

on it. Above model is not minimal at 3. The minimal equation for \(E_1\) is

\[
E_{128a1}: y^2 = x^3 + x^2 + x + 1
\]

whose label in Cremona’s Elliptic Curve Database is ”128a1”, and the coordinates of \(t\) are

\[
t = \left(\frac{-3}{4}, \frac{5}{8}\right)
\]

with respect to the minimal equation.

Similarly, from \((L_2, Q_2)\) we obtain the curve

\[
y^2 = x^3 + \frac{1}{4}x
\]

and the point \(t = (0, 0)\). Above equation is has minimal equation

\[
E_{32a1}: y^2 = x^3 + 4x
\]

whose label is ”32a1”, and \(t\) has the same coordinate \(t = (0, 0)\) with respect to the minimal equation.
In fact, \( E_{128a1} \) has more 2-integral points, one finds the list

\[
(-1, 0, 1), (-3/4, 5/8, 1), (0, 1, 1), (1, 2, 1), (7, 20, 1)
\]  

by applying the command "S_integral_points" in SAGE. Note that the list shows \( S \)-integral points modulo the action of \( \{\pm 1\} \) on the curve. We already produced the second point using \( f_1 \), and one should be able to determine the rest using the remaining \( f_i \)'s. Indeed, the four remaining points are obtained from \( i = 11, 37, 40, 41 \).

After carrying out similar calculations for all \( f_i \), we obtain Table 2. We note the reader that \( f_{30} \) and \( f_{31} \) give rise to two equivalent pairs, and \( f_{42} \) and \( f_{43} \) give rise to two equivalent pairs as well.

### Table 2:

| Label   | \( i \)  | Label     | \( i \)  | Label     | \( i \)  |
|---------|---------|-----------|---------|-----------|---------|
| "128a1" | 1,11,37,40,41 | "128a2" | 2,4,23,32,33,45,46 | "128b1" | 36       |
| "128b2" | 48      | "128c1"  | 39      | "128c2"  | 47       |
| "128d1" | 38      | "128d2"  | 44      | "256a1"  | 2,22,24,25,51 |
| "256a2" | 3,8,27,35 | "256b1"  | 2,21,28,29,50 | "256b2"  | 9,17,18   |
| "256c1" | 19      | "256c2"  | 20      | "256d1"  | 34       |
| "256d2" | 26      | "32a1"   | 1,42,43 | "32a2"   | 5,12,14   |
| "32a3"  | 7       | "32a4"   | 4,49    | "64a1"   | 13,15,16   |
| "64a2"  | 6       | "64a3"   | 3,30,31 | "64a4"   | 10       |

### 6 Average number of integral points on elliptic curves

In this section, we shift our attention to the behaviour of the number of integral points on elliptic curves on average. The aim of present section is to obtain an upper bound for the average with respect to certain ordering, exploiting our construction and the work of Bhargava-Shanker.

We are concerned about the isomorphism classes of elliptic curves over \( \mathbb{Q} \) and the number of integral points in them on average, but the notion of integral points on a curve which is merely defined over \( \mathbb{Q} \) is not well-defined until we specify an integral model. We are going to deal with this by considering the curves of the form

\[
E_{a,b}: y^2 = x^3 + ax + b
\]

where \( a, b \) are integers, and there is no prime \( p \) such that \( p^4 | a \) and \( p^6 | b \). Each isomorphism class of an elliptic curve defined over \( \mathbb{Q} \) can be uniquely written in
the above form. Such a model shall be called the simplified minimal Weierstrass model for $E$. Furthermore, an integral point on an elliptic curve in this section should be understood as an integral point in this model.

**Definition 6.1.** For an elliptic curve $E$ over $\mathbb{Q}$, $E_{a,b}$, we define the height of $E$, denoted by $H(E)$ as

$$H(E) = H(E_{a,b}) = \max \{2^{12}3^4|a|^3, 2^{14}3^{12}b^2\},$$

where $E_{a,b}$ is the simplified minimal Weierstrass model for $E$. The constants in front of $|a|^3$ and $b^2$ do not have significant meaning. We have chosen the constants so that the final estimate has a nice form.

We will order the isomorphism classes of elliptic curves over $\mathbb{Q}$ with respect to the height defined above. For any positive number $T$, define

$$N(T) = \sum_{E, H(E) < T} \sum_{t \in E_{a,b}(\mathbb{Z})/\{\pm 1\}} 1$$

which is the total number of integral points on elliptic curves of height up to $T$, counted modulo the action of $\{\pm 1\}$.

The number of elliptic curves up to height $T$ has order $T^{5/6}$, so if we bound $N(T)$ by $O(T^{5/6})$ then it amounts to proving that the number of integral points on $E_{a,b}$ is bounded on average.

**Theorem 6.1.** We have

$$N(T) < (31.5 \cdots) T^{5/6}$$

for all sufficiently large $T > 0$. In particular, the average number of integral points on elliptic curves is bounded by 32.

In the rest of the section, we give the proof for Theorem 6.1. The starting point is a map

$$\phi: (E_{a,b}, t) \mapsto ((1, 0), Q_{a,b,t}(u, v)) \in \mathbb{Z}^2 \times \text{Sym}^4(\mathbb{Z}^2)^*$$

where

$$Q_{a,b,t}(u, v) = u^4 - 6x_tu^2v^2 - 8ytuv^3 - (3x_t^2 + 4a)v^4$$

and $u, v$ are viewed as the basis of $(\mathbb{Z}^2)^*$ dual to the standard basis for $\mathbb{Z}^2$. In particular, we view $(1, 0)$ as the solution of the equation

$$Q_{a,b,t}(u, v) = 1$$

which is often called the Thue-equation associated to $Q_{a,b,t}(u, v)$. It is merely a reformulation of $\kappa$ we introduced earlier, but in this way the argument becomes more natural.

Naturally $GL_2(\mathbb{Z})$ acts on $\mathbb{Z}^2 \times \text{Sym}^4(\mathbb{Z}^2)^*$, and, moreover, the action preserves the Thue-equation. That is to say, the subset

$$\{(n, m), Q(u, v) \in \mathbb{Z}^2 \times \text{Sym}^4(\mathbb{Z}^2)^* : Q(n, m) = 1\}$$

is preserved by the action of $GL_2(\mathbb{Z})$. 17
Proposition 6.1. The map

\[ \phi: \{(E, t) : t \in E(\mathbb{Z}) \}/\{\pm 1\} \rightarrow \{(n, m), Q(u, v)) : Q(n, m) = 1\}/\sim \]

is injective, where \( \sim \) denotes the equivalence relation induced by the action of \( GL_2(\mathbb{Z}) \).

Proof. Suppose that two pairs \((E_a, b, t)\) and \((E_{a'}, b', t')\) have the same image under \( \phi \). Then we have \( \gamma \in GL_2(\mathbb{Z}) \) which fixes \((1, 0)\), and transforms

\[ Q_{a, b, t} = u^4 - 6x_tu^2v^2 - 8y_tuv^3 - (3x_t^2 + 4a)v^4 \]

into

\[ Q_{a', b', t'} = u^4 - 6x_{t'}u^2v^2 - 8y_{t'}uv^3 - (3x_{t'}^2 + 4a')v^4 \]

and it is easy to see that the identity and \((u, v) \mapsto (\pm u, \pm v)\) is only possibility for \( \gamma \). Indeed, the stabiliser subgroup of \((1, 0)\) is generated by the group of unipotent matrices, together with the transformation \((u, v) \mapsto (\pm u, \pm v)\). By comparing the coefficient of \( u^3v \), one sees that \( \gamma \) must be of the form \((u, v) \mapsto (\pm u, \pm v)\). Thus we conclude that \( a = a', b = b', \) and \( t = \pm t' \).

\[ \square \]

Remark 6.1. Note that the two pairs \(((n, m), Q(u, v)) \sim ((-n, -m), Q(u, v))\)

are equivalent via the matrix with \(-1\)'s on the diagonal.

Let

\[ Q = c_0u^4 + c_1u^3v + c_2u^2v^2 + c_3uv^3 + c_4v^4 \]

be a binary quartic form with integer coefficients. With respect to the \( GL_2(\mathbb{Z}) \) action, there are two invariants

\[ J_2 = \frac{1}{12}c_2^2 - \frac{1}{4}c_1c_3 + c_0c_4 \]

\[ J_3 = \frac{1}{216}c_3^2 - \frac{1}{48}c_1c_2c_3 + \frac{1}{16}c_0c_3^2 + \frac{1}{16}c_1^2c_4 - \frac{1}{6}c_0c_2c_4 \]

of degree two and three respectively. We define height of \( Q \) as

\[ H(Q) = \max \left\{ 2^63^4 \cdot |J_2|^3, \ 2^{10}3^{12} \cdot J_3^2 \right\} \]

where the coefficients in front of \(|J_2|^3\) and \(J_3^2\) do not have significant meanings, but chosen so that our definition of height agrees with that of [5].

Proposition 6.2. Let \( t \in E(\mathbb{Z}) \), and \( \phi((E, t)) = ((n, m), Q) \). We have

\[ H(E) = H(Q). \]
Proof. It follows from the straightforward calculation.

Having constructed $\phi$ which preserves the heights, estimation of $N(T)$ is reduced to the estimation of the pairs $(n, m), Q(u, v)$ which lies in the image of $\phi$, modulo $\text{GL}_2(\mathbb{Z})$-equivalence. We consider three types

1. $Q(u, v)$ is irreducible over $\mathbb{Q}$.
2. $Q(u, v)$ has a linear factor over $\mathbb{Q}$.
3. $Q(u, v)$ has two irreducible quadratic factors over $\mathbb{Q}$.

which are mutually disjoint. Let

$$X^i(T)$$

be the $\text{GL}_2$-orbits of binary forms of type $i$ with height less than $T$.

We consider three subtypes of $X^i_1(T)$, for $j = 0, 1, 2$, so that an element in $X^j_1(T)$ has $4 - 2j$ linear factors over $\mathbb{R}$.

Theorem 6.2. We have

$$\sum_{Q \in X^i_1(T)} 1 = \frac{2\pi^2}{405} T^{3/2} + O(T^{3/4 + \epsilon}),$$

where the sum is taken over all irreducible integral binary quartic forms up to $\text{GL}_2(\mathbb{Z})$ equivalence.

Proof. The estimation of the sum over $X^1_1(T)$ is a consequence of Theorem 1.6 of [5]. The estimation of the sum over $X^3(T)$ is given in the proof Lemma 2.3 of [5].

Proposition 6.3. For $X^2(T)$, we give the following estimation of the sum over the image of $\phi$

$$\sum_{Q \in X^2(T), Q \in \text{Im}(\phi)} 1 = O(T^{3/4}).$$

Proof. If $Q$ is in $X^2(T)$, then $Q$ factors as

$$Q = (u - rv)C(u, v)$$

where $C(u, v)$ is a binary form.

19
where \( r \) is an integer \( C(u, v) \) is binary cubic form with integral coefficients such that \( C(1, 0) = 1 \). By translation \( u \mapsto u + rv, \) \( Q \) is equivalent to the form
\[
u(v^3 + c_1v^2u + c_2vy^2 + c_3u^3)
\]
with integers \( c_1, c_2 \) and \( c_3 \). By translating \( v \mapsto v + r'u \) for some integer \( r' \) if necessary, we may assume that \( |c_1| \leq 1 \). The invariants of (109) are given as
\[
J_2 = \frac{1}{12}c_2^2 - \frac{1}{4}c_1c_3
\]
\[
J_3 = \frac{1}{216}c_3^3 - \frac{1}{48}c_1c_2c_3 + \frac{1}{16}c_2^3
\]
and \( |J_2| = O(T^{1/3}) \) and \( |J_3| = O(T^{1/2}) \). Hence the discriminant of (109) is \( O(T) \). On the other hand, the discriminant is divisible by \( c_3^2 \), hence \( |c_3| = O(T^{1/2}) \). Now \( J_2 = O(T^{1/3}) \) together with \( |c_3| = O(T^{1/2}) \) implies \( |c_2| = O(T^{1/4}) \). We conclude that there are \( O(T^{3/4}) \) possibilities for the pair \( (c_1, c_2, c_3) \).

The proof of Theorem 6.1 has two more key ingredients, besides the map \( \kappa \) defined earlier. They include the works of Evertse and Akhtari-Okazaki on the number of solutions of a given Thue-Mahler equations, and the work of Bhargava-Shankar on the number of binary quartic forms with bounded invariants, which we recall now.

Thue-Mahler equations regards a homogeneous binary form \( h(u, v) \in \mathbb{Z}[u, v] \) and a finite set \( S \) of prime numbers, to which one associates the equation
\[
h(u, v) = \pm \prod_{p \in S} p^{e_i}
\]
where \( e_i \) are non-negative integers, and \( u, v \) are relatively prime integers. We will rely on a corollary which is easily implied by the following theorem of Evertse.

**Theorem 6.3.** Let \( r \) be the degree of \( h(u, v) \), and assume that \( h(u, v) \) has at least three linearly independent linear factors over a sufficiently large number field. Let \( S \) be a finite set of prime numbers of cardinality \( s \). Then associated equation (112) has at most
\[
2 \times \tau^{3(2s+3)}
\]
solutions.

We are concerned about the case when \( h(u, v) \) is a quartic with non-zero discriminant, and \( S \) is empty. The following corollary is a direct consequence of Evertse’s theorem.

**Corollary 6.1.** Let \( Q(u, v) \) be a binary quartic form with non-zero discriminant. The equation
\[
Q(u, v) = \pm 1
\]
has at most
\[ 2 \times 7^{3.3} < 3.63 \times 10^{162} \quad (115) \]
solutions.

Despite of the large size of the upper bound, we note that it is independent of \( Q(u, v) \). On the other hand, we have a significantly better bound due to Akhtari and Okazaki, under the additional assumption that \( Q(u, v) \) is irreducible.

**Theorem 6.4.** Let \( Q(u, v) \) be an irreducible quartic equation. The associated Thue equation
\[ Q(u, v) = \pm 1 \quad (116) \]
has at most 61 solutions, provided that the discriminant of \( Q(u, v) \) is greater than an absolute constant, which is effectively computable. Here we regard a solution \((n, m)\) as the same as \((-n, -m)\). If we further assume that \( Q(u, v) \) has four linear factors defined over \( \mathbb{R} \), then it has at most 37 solutions.

Now the proof of Theorem 6.1 is straightforward. Indeed, from the injectivity of \( \phi \), one has
\[ N(T) \leq \sum_{Q \in X^1(T)} \sum_{Q(n, m) = 1} 1 + \sum_{Q \in X^2(T), Q \in \text{Im(}\phi\text{)}} \sum_{Q(n, m) = 1} 1 + \sum_{Q \in X^3(T)} \sum_{Q(n, m) = 1} 1 \quad (117) \]
where the sum over \( Q(n, m) = 1 \) means that the sum is taken over the set of pairs \((n, m)\) such that \( Q(n, m) = 1\), modulo the identification of \((n, m)\) and \((-n, -m)\). Note that (96) shows that two solutions \((n, m)\) and \((-n, -m)\) should be counted once. By Theorem 6.2 and Theorem 6.4, one has
\[ \sum_{Q \in X^1(T)} \sum_{Q(n, m) = 1} 1 \quad (118) \]
\[ = \sum_{Q \in X^1(T)} \sum_{Q(n, m) = 1} 1 + \sum_{Q \in X^2(T), Q \in \text{Im(}\phi\text{)}} \sum_{Q(n, m) = 1} 1 + \sum_{Q \in X^3(T)} \sum_{Q(n, m) = 1} 1 \quad (119) \]
\[ = 37 \cdot \frac{2 \cdot 2 \pi^2}{405} T^{5/6} + 61 \cdot \frac{16 \pi^2}{405} T^{5/6} + 61 \cdot \frac{4 \pi^2}{405} T^{5/6} + O(T^{3/4+\epsilon}) \quad (120) \]
\[ < (31.5 \cdots) T^{5/6} + O(T^{3/4+\epsilon}) \quad (121) \]
\[ < 32 \cdot T^{5/6} + O(T^{3/4+\epsilon}) \quad (122) \]
while Theorem 6.2, Proposition 6.3, and Corollary 6.1 imply that
\[ \sum_{Q \in X^2(T), Q \in \text{Im(}\phi\text{)}} \sum_{Q(n, m) = 1} 1 = O(T^{3/4}) \quad (123) \]
\[ \sum_{Q \in X^3(T)} \sum_{Q(n, m) = 1} 1 = O(T^{2/3+\epsilon}) \quad (124) \]
which has smaller order than $T^{5/6}$. We conclude that

$$N(T) < 32 \cdot T^{5/6}$$

for all sufficiently large $T > 0$.

**Remark 6.2.** We have chosen this ordering of elliptic curves and the notion of integrality mainly because in this way the counting is straightforward. Note that an integral point with respect to above the simplified minimal Weierstrass model may not be integral with respect to other models such as the global minimal model in a generalised Weierstrass equation.

# 7 Remarks

We give a few concluding remarks.

**Remark 7.1.** One naturally wonders what can be done on the number of $S$-integral points on elliptic curves. The only reason why we had to restrict to the case of integral points in the present section is that the result of Bhargava and Shankar is restricted to the binary forms with integer coefficients with respect to $GL_2(\mathbb{Z})$ transformation, rather than forms with coefficients in $\mathbb{Z}_S$ and subject to $GL_2(\mathbb{Z}_S)$-transformations. In contrast, the map $\phi$ exists without any restriction of $S$, and Theorem 6.3 is extended to the forms with $S$-integral coefficients in [4] with an upper bound which is independent of the form.

**Remark 7.2.** The method employed in [2], as mentioned in the introduction, succeeds to turn a knowledge about the distribution of ranks of elliptic curves into a knowledge about the distribution of number of integral points on elliptic curves, by using inequalities relating the ranks and the number of integral points. In particular, moments of number of integral points on elliptic curves are obtained in [2]. If we try to attack the $k$-th moments using our approach, one is lead to look for orbits of representations which parametrise the points of the complement of the discriminant locus in the $k$-th symmetric power of the punctured universal elliptic curve. We do not know such a representation yet. The method explained in the present paper was initially motivated by the author’s attempt to prove an effective version of Mordell conjecture for punctured elliptic curves, which asks one to produce an algorithm to determine the complete set of $S$-integral points of an elliptic curve.

**Remark 7.3.** In [2], their method is applied to families of elliptic curves with a fixed $j$-invariant, and the corresponding upper bound is established. We have not considered such families of elliptic curves.

**Remark 7.4.** The true average number of integral points might be much smaller. The upper bound for the number of specific Thue equation which we used, namely Theorem 6.3 from an optimistic point of view, may be improved by considering the average. Indeed, $8/9$ is proposed as a minimalist’s conjecture in [2].
Remark 7.5. From the point of view towards effective Mordell conjecture for $\mathcal{Y}$, the punctured universal elliptic curve, our method, via Theorem 4.1 reduces the computation of $\mathcal{Y}(\mathbb{Z}_S)$ to the computation of an orbit of a representation. The relevant orbits can be classified numerically by solving a finite number of quartic Thue-Mahler equations, thus can be determined effectively. It is what we did in Section 5 by taking $S = \{2\}$. Despite of this effectiveness, the present method would be aesthetically more appealing if one can develop a systematic approach to classify the orbits. In [6], a similar strategy was applied to cubic Thue-Mahler equations, in which we solved the relevant classification problem using the theory of modular forms and modularity of elliptic curves. It seems an interesting challenge to come up with an analogous classification scheme which would allow us to determine $\mathbb{Z}_S$-points of the punctured universal elliptic curve.

References

[1] S. Akhtari and R. Okazaki, Quartic Thue equations, J. Number Theory 130 (2010), no. 1, 40-60.

[2] L. Alpoge, The average number of integral points on elliptic curves is bounded., arXiv:1412.1047 [math.NT], 42 pages.

[3] J.-H. Evertse, On equations in $S$-units and the Thue-Mahler equation, Invent. Math. 75 (1984), no. 3, 561-584.

[4] J.-H. Evertse, The number of solutions of the Thue-Mahler equation, J. Reine Angew. Math. 482 (1997), 121-149.

[5] M. Bhargava and A. Shankar, Binary quartic forms having bounded invariants and the boundedness of the average rank of elliptic curves, Ann. of Math. (2) 181 (2015), no. 1, 191-242.

[6] D. Kim, A modular approach to Thue-Mahler equations, arXiv:1501.06274 [math.NT], 42 pages.

[7] N.P. Smart, S-unit equations, binary forms, and curves of genus 2, Proc. London Math. Soc. (3) 75 (1997), no. 2, 271-307