ACTA ARITHMETICA
Online First version

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

by

Dohyeong Kim (Ann Arbor, MI)

1. Introduction. The goal of the present article is to show that the average number of integral points on the curves

$$Y_{a,b}\colon y^2 = x^3 + ax + b, \qquad a, b \in \mathbb{Z},\ 4a^3 + 27b^2 \neq 0, \tag{1.1}$$

is bounded from above by 2.1 × 10^8. The points are counted modulo the natural involution $(x, y) \mapsto (x, -y)$, which is of course equivalent to the negative with respect to the group law of the underlying elliptic curve. The average is taken with respect to the height

$$H(Y_{a,b}) := \max\{2^{12}3^4|a|^3, 2^{14}3^{12}b^2\} \tag{1.2}$$

where the reasons behind the numbers multiplied by $|a|^3$ and b^2 with be explained later.

For any positive real number T, let

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

be the total number of $\mathbb{Z}$-valued points on curves $Y_{a,b}$ of height bounded by T. We also let $R(T)$ be the number of curves of height at most T.

Theorem 1.1. *For all sufficiently large T, we have*

$$\frac{N(T)}{R(T)} < 2.1 \times 10^8. \tag{1.3}$$

Although our primary interest is in curves of the form (1.1), we will develop some techniques that are applicable to a slightly wider range of equations. Namely, we will consider any curve

$$Y\colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \qquad a_1, a_2, a_3, a_4, a_6 \in \mathbb{Z}, \tag{1.4}$$

2010 *Mathematics Subject Classification*: 11D, 11G05, 11D59.
Key words and phrases: integral points, elliptic curve, Thue equation.
Received 9 March 2016; revised 25 November 2017.
Published online 27 April 2018.

DOI: 10.4064/aa8458-11-2017

represented by a generalised Weierstrass equation, and study the set of S-integral points on it, where S is a finite set of prime numbers. We always assume that $Y/\mathbb{Q}$ is nonsingular. Our main strategy is to reduce the study of S-integral points on the curve of the form (1.4) to that of solutions of certain quartic Thue–Mahler equations.

In fact, the above strategy is not entirely new; the possibility of such a reduction was known at least to Mordell. In Chapter 27 of his book [7], he proves that the set of integral solutions of the equation

$$ey^2 = ax^3 + bx^2 + cx + d, \qquad a, b, c, d, e \in \mathbb{Z}$$

is finite, under the assumption that the cubic polynomial on the right hand side does not have repeated roots, by reducing it to the combination of two finiteness results on the number of binary quartic forms with given invariants and the number of solutions of a quartic Thue equation. More precisely, Mordell showed that the x and z coordinates of the affine surface

$$ey^2 = ax^3 + bx^2z + cxz^2 + dz^3 \tag{1.5}$$

can be parametrised by a pair of explicit quartic forms. Geometrically speaking, this shows that the above affine surface is unirational. Perhaps some readers might be reminded about the well-known result which says that any smooth cubic surface is geometrically rational.

In our approach, a key role is played by an explicit map, called the descent map, which is generically an isomorphism between open subsets of two GIT type spaces. One is the universal elliptic curve modulo the natural involution, and the other is the orbit space of pairs of binary forms of degree 1 and 4. The S-integral points on elliptic curves are parametrised by the complement of the zero section of the universal elliptic curve, namely the punctured universal elliptic curve, while the binary quartic forms together with a solution of the associated Thue–Mahler equation are parametrised by an open subset of the latter.

It turns out that the binary quartic form that we associate to a point on an elliptic curve via the descent map is equivalent to the quartic form which is used by Mordell in order to parametrise the z-coordinate of the affine surface (1.5). In some sense, our method is essentially that of Mordell, and our contribution is to appropriately repackage his method so that it is suitable for our purpose, and that one can connect it to a few deep results that could not have been available to him.

Having established the descent map in an appropriate form, we can obtain the average number of integral points on curves of the form $Y_{a,b}$ without too much difficulty. Indeed, the work [3] of Bhargava–Shankar provides the asymptotic growth of the average number of integral binary quartic forms with given invariants, and the works [1, 4] of Akhtari–Okazaki and Evertse provide absolute upper bounds for the number of solutions of a quartic Thue equation. Combining these, we will be able to prove the desired upper bound. In fact, the normalisation of $H(Y_{a,b})$ is chosen in a way which is compatible with the choice made by Bhargava–Shankar.

We briefly discuss our terminology. Equations (1.1) and (1.4) have underlying (projective) elliptic curves, and their $\mathbb{Z}_S$-solutions may be abusively called $\mathbb{Z}_S$-points on those curves. Here, an S-integral point on an elliptic curve should be understood as a scheme-theoretic $\mathbb{Z}_S$-point on the punctured elliptic curve. Of course, the notion of S-integral point coincides with that of rational point for a projective curve, and the study of S-integral points is meaningful only for the punctured elliptic curve. Since rational points on elliptic curves are not our current subject matter, our abuse of terminology should not cause too much confusion.

Going back to our discussion on the technical aspects of the present article, note that our argument does not involve the ranks of elliptic curves, nor the arithmetic invariants of auxiliary number fields. To the best knowledge of the author, the previously known bounds such as [6, 8] for the number of integral points on a particular elliptic curve depend exponentially either on the rank of the curve, or the rank of a certain ideal class group of a number field such as the two-division field of the curve. Combining this type of upper bounds 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 [2] considered a family, which is almost but not exactly identical to ours, of elliptic curves, and claimed that this strategy yields 65.8457 as an upper bound. His family consists of the curves $Y_{a,b}$ as above, but with an additional condition that $Y_{a,b}$ is minimal; there is no prime p such that both $p^4 \mid a$ and $p^6 \mid b$ hold.

Another way of attacking integral points on elliptic curves is to apply Baker's method. Since the upper bound resulting from Baker's method grows rather rapidly in terms of the coefficients of the equation, it does not seem easy to obtain an upper bound for $N(T)/R(T)$ out of it.

As mentioned earlier, the descent map is generically an isomorphism, and this has an implication about S-integral points on elliptic curves. In fact, the descent map turns out to be an isomorphism over $\mathbb{Z}[1/6]$. If an elliptic curve $E/\mathbb{Q}$ has good reduction outside S, then the S-integral points on E can be defined using the smooth model of E over $\mathbb{Z}_S$, the ring of S-integers. Let us temporarily denote by E an elliptic curve over $\mathbb{Q}$ which has good reduction outside S, and by t a $\mathbb{Z}_S$-point on E minus the origin. Using the descent map, we will obtain a bijection between the set of all equivalence classes of pairs (E, t) and the set of orbits of pairs of binary forms, provided that both 2 and 3 are contained in S. For arbitrary S, the descent map does not necessarily induce a bijection, but it remains to be injective, whence it can be used to compute all such pairs (E, t). We numerically demonstrate this for $S = \{2\}$.

We explain why the height function (1.2) is normalised in such a way. From the point of view of the present article, the set of integral points of the punctured universal elliptic curve is mapped onto a subset of the set of orbits of pairs (L, Q) of binary forms of degree 1 and 4. In particular, giving a height on the latter set always pulls back to the former to define a compatible height function. Once we declare the height of the pair (L, Q) to be the height of Q considered by [3], we are forced to work with the pull-back of it, which is precisely our normalisation of (1.2). The magnitude of the constants in (2) is a reason why the constant in Theorem 1.1 is admittedly very large. Most likely 2.1×10^8 is not optimal, but our method falls short of obtaining a better upper bound.

It turns out that our descent map is not surjective when we work over $\mathbb{Z}$. The reason for this is that quartic forms arising in this way have invariants divisible by 4. Since we will be counting all binary quartic forms without any divisibility conditions on their invariants, we will be overestimating $N(R)/R(T)$. Our method may be sharpened using a finer estimate on the number of quartic forms with invariants satisfying certain divisibility conditions, which we do not pursue here.

We outline the organisation of the paper. In Section 2, we review some basic properties of the notion of equivalence between pairs of binary forms. In Section 3, we define the descent map, which associates two integral binary forms to a point on the punctured universal elliptic curve. In Section 4, we use the descent map to identify S-integral points on the punctured universal elliptic curve with certain equivalence classes of pairs of binary forms. In Section 5, we work out a numerical example with $S = \{2\}$. In Section 6, we use the descent map together with the works of Akhtari–Okazaki, Evertse, and Bhargava–Shankar to establish the desired upper bound for the average number of integral points on elliptic curves.

We close the introduction with two remarks. Firstly, one naturally wonders what can be said about the average number of S-integral points on elliptic curves. An obstacle is due to the fact that the result of Bhargava and Shankar is restricted to binary forms with integer coefficients with respect to $\mathrm{GL}_2(\mathbb{Z})$-transformations, rather than forms with coefficients in $\mathbb{Z}_S$ that are subject to $\mathrm{GL}_2(\mathbb{Z}_S)$-transformations. On the other hand, the descent map exists without any restriction of S, and Theorem 6.9 is extended to the forms with S-integral coefficients in [5] with an upper bound which is independent of the form. Another obstacle is that one of the properties of the height function that holds true for $S = \emptyset$ does not generalise to the situation with $S \neq \emptyset$. More precisely, when $S = \emptyset$, the height function on the quartic form, when pulled back to the punctured universal elliptic curve via the descent map, gives rise to a height function which is vertical in the sense that the height only depends on the elliptic curve but not on the point on it. However, whenever $S \neq \emptyset$, the height function on the punctured universal elliptic curve is no longer vertical. We do not know how to overcome these obstacles.

Secondly, one also wonders what would be the true average number of integral points, if it exists, on curves of the form $Y_{a,b}$. While we are relying on the absolute upper bound for the number of solutions of a Thue equation (Theorem 6.12), the average number of solutions of a Thue equation may well be smaller. More precisely, we can order the equivalence classes of integral binary quartic forms in the way of Bhargava–Shankar, where the equivalence relation is taken with respect to the obvious $\mathrm{GL}_2(\mathbb{Z})$-transformations, and ask whether the average number of solutions of the associated Thue equations is smaller than the absolute bounds employed in this article. An affirmative answer, which is seemingly missing at present, will allow one to effortlessly improve the upper bound obtained in this article.

2. 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 for our purpose.

Let S be any finite set of primes. Let us consider a pair (L, Q) of binary forms

$$L = b_0u + b_1v, \qquad Q = c_0u^4 + c_1u^3v + c_2u^2v^2 + c_3uv^3 + c_4v^4,$$

where b_i's and c_i's are S-integers. We always assume that the coefficients of L and Q do not have nontrivial 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 $\varDelta_Q$, is given by

$$\begin{aligned}\varDelta_Q = {} & c_1^2c_2^2c_3^2 - 4c_0c_2^3c_3^2 - 4c_1^3c_3^3 + 18c_0c_1c_2c_3^3 - 27c_0^2c_3^4 - 4c_1^2c_2^3c_4 \\ & + 16c_0c_2^4c_4 + 18c_1^3c_2c_3c_4 - 80c_0c_1c_2^2c_3c_4 - 6c_0c_1^2c_3^2c_4 + 144c_0^2c_2c_3^2c_4 \\ & - 27c_1^4c_4^2 + 144c_0c_1^2c_2c_4^2 - 128c_0^2c_2^2c_4^2 - 192c_0^2c_1c_3c_4^2 + 256c_0^3c_4^3,\end{aligned}$$

and the discriminant of $L \cdot Q$, denoted by $\varDelta$, is given by

$$\varDelta = \varDelta_Q \cdot Q(-b_1, b_0)^2.$$

For a fixed S, we will be concerned with pairs of forms for which $\varDelta$ is an S-unit. We introduce the following notion of admissibility to simplify the exposition.

Definition 2.1. Let (L, Q) be a pair of binary forms with S-integral coefficients as above. We say that this pair is *S-admissible* if Δ 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 coefficientwise equality. A weaker notion of equality, which is more natural if we view them as elements of projective space, is the following.

Definition 2.2. Let (L, Q) and (L', Q') be S-admissible pairs. We say that the pairs are *projectively equivalent* if there are $\lambda_1, \lambda_2 \in \mathbb{Z}_S^\times$ such that $(L, Q) = (\lambda_1 L', \lambda_2 Q')$.

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

Now we introduce the desired notion of equivalence.

Definition 2.3. Let (L, Q) and (L', Q') be S-admissible pairs. We say that they are GL_2-*equivalent* if there is $g \in \mathrm{GL}_2(\mathbb{Z}_S)$ 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 now take a closer look at the notion of GL_2-equivalence under the assumption that $2 \in S$. If $2 \in S$, then for each S-admissible pair (L, Q) it is possible to find a pair (L', Q') which is GL_2-equivalent to (L, Q) and such that

$$L' = v, \qquad Q' = u^4 + B_2 u^2 v^2 + B_3 u v^3 + B_4 v^4,$$

where B_2, B_3, B_4 are integers, rather than S-integers. Furthermore, it is possible, as we will prove shortly, to choose a minimal pair in the following sense.

Definition 2.4. 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 there is no prime prime p such that $p^i \mid B_i$ for $i = 2, 3, 4$ simultaneously. If the form has S-integral coefficients, it is called *minimal at p* for a prime $p \notin S$ when $p^i \mid B_i$ for $i = 2, 3, 4$ do not hold simultaneously.

Proposition 2.5. *Recall that $2 \in S$. Given any pair (L, Q) of binary forms as above, it is possible to find a minimal pair*

$$(v, u^4 + B_2 u^2 v^2 + B_3 u v^3 + B_4 v^4) \tag{2.1}$$

which is 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

$$L = b_0 u + b_1 v, \qquad Q = c_0 u^4 + c_1 u^3 v + c_2 u^2 v^2 + c_3 u v^3 + c_4 v^4.$$

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 an S-unit. Otherwise, the non-S-unit

$$Q(b_1, -b_0) = c_0$$

would divide Δ, contradicting the S-admissibility of the pair. Thus, via projective equivalence, we may assume that $c_0 = 1$. Now we have a pair

$$L = v, \qquad Q = u^4 + c_1 u^3 v + c_2 u^2 v^2 + c_3 u v^3 + c_4 v^4,$$

where the coefficients are in $\mathbb{Z}_S$. Since we assume $2 \in S$, we are allowed to 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 v by an S-unit, and apply projective equivalence, in order to get a minimal form.

The only linear change of variables which preserves the minimality is given by $(u, v) \mapsto (\pm u, v)$. Thus, there is only one other pair equivalent to a given minimal form

$$L = v, \qquad Q = u^4 + c_2 u^2 v^2 + c_3 u v^3 + c_4 v^4,$$

and this is obtained by replacing c_3 by $-c_3$. The proof of the proposition is complete. ■

Remark 2.6. 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 remains of exceptional importance, since it will correspond to the negative on the elliptic curve.

3. 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 their basic properties.

We begin with notation. Let

$$E\colon y^2 z + a_1 xyz + a_3 yz^2 = x^3 + a_2 x^2 z + a_4 xz^2 + a_6 z^3 \tag{3.1}$$

be an elliptic curve represented by 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 the $\mathbb{Z}_S$-points of Y can be described as

$$Y(\mathbb{Z}_S) = \{t = (x_t : y_t : z_t)\colon t \in E(\mathbb{Z}),\ z_t \in \mathbb{Z}_S^\times\}. \tag{3.2}$$

Of course, the points of $Y(\mathbb{Z}_S)$ bijectively correspond to the solutions of the affine equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \tag{3.3}$$

so one can view (3.2) as an alternative description for the set of solutions of (3.3) in $\mathbb{Z}_S$. In our exposition, we will mainly use (3.2).

For each $t \in Y(\mathbb{Z}_S)$, 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 indicate that they are linear and quartic forms, respectively. 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

$$L_t(u, v) = v,$$

which is regarded as a linear form in the variables u and v. For the geometric reason underlying this hardly motivating definition, see Remark 3.2.

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 coordinates x, y, and 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 (projectively) parametrised by the linear forms

$$ux + vy + wz = 0 \quad \text{such that} \quad ux_t + vy_t + wz_t = 0.$$

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

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

The quartic form $Q_t(u, v)$, which will be explicitly determined 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.

Proposition 3.1. *The quartic form $Q_t(u, v)$ is given by*

$$A^2 - 4v^2B \tag{3.4}$$

where

$$A = -z_tu^2 + z_ta_1uv + (a_2z_t + x_t)v^2, \tag{3.5}$$

$$B = x_tz_tu^2 + (2y_tz_t + z_t^2a_3)uv + (a_4z_t^2 - a_1z_ty_t + a_2z_tx_t + x_t^2)v^2. \tag{3.6}$$

Proof. This follows from a straightforward calculation. We need to find the algebraic condition that is equivalent to the geometric one that the line

$$ux + vy + wz = 0 \tag{3.7}$$

is tangent to E. We insert $y = \frac{ux+wz}{-v}$ in

$$y^2z + a_1xyz + a_3yz^2 - (x^3 + a_2x^2z + a_4xz^2 + a_6z^3) \tag{3.8}$$

and obtain a cubic form $C(x, z)$ in x and z. Using the condition that t satisfies both (3.7) and (3.8), one observes that $C(x, z)$ should have a factorisation

$$C(x, z) = \frac{1}{z_t v^2}(xz_t - zx_t) \cdot q(x, z) \tag{3.9}$$

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 (3.9) and equating the coefficients of it with those of $C(x, z)$, one obtains

$$q(x, z) = v^2x^2 + Axz + Bz^2$$

where A and B are polynomials given in the statement of the proposition. The condition that the line is a ramification point of the projection map is equivalent to the discriminant of $q(x, z)$ being zero. From this, one obtains the formula for $Q_t(u, v)$. ■

Remark 3.2. The linear form $L_t(u, v) = v$ acquires the following geometric interpretation once we view u and v as parameters for the lines passing through t. The zero of $L_t(u, v)$ is $(u, v) = (1, 0)$, which corresponds to the line

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

passing through t and the origin of E.

Remark 3.3. 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.

Remark 3.4. A more precise form of Proposition 3.1 will be given in Theorem 4.1.

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

$$E\colon y^2z + yz^2 = x^3 - xz^3, \tag{3.10}$$

which is a curve of conductor 37. It has no nontrivial 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 $t = n \cdot P_0$.

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

$$(u+v)(u^3 - u^2v - 5uv^2 + v^3),$$

showing that the corresponding torsor is trivial.

For $n = 3$, we have $t = (-1, -1, 1)$. Similarly,

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

which is irreducible.

As a second example, consider

$$E\colon y^2z = x^3 - 1681xz^2. \tag{3.11}$$

Since $1681 = 41^2$ is a square, it has three rational points of order two. Also, it turns out that the Mordell–Weil group has rank two, generated by

$$P_1 = (-9, 120, 1), \qquad P_2 = (841, 24360, 1).$$

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

$$Q_t(u,v) = 43200(40u - 827v)(120u + 143v)(120u + 719v)(120u + 1619v).$$

This proves that $Q_t(u,v)$ defines the trivial torsor, as expected.

Now we turn to the key proposition regarding both $L_t(u,v)$ and $Q_t(u,v)$.

Proposition 3.5. *Let Δ_E be the discriminant of E, and let S be any finite set of primes numbers. Let Δ_t be the discriminant of binary quintic form $L_t(u,v) \cdot Q_t(u,v)$. Then Δ_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 Δ_E and p does not belong to S. In order to prove the proposition, it suffices to show that Δ_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

$$\theta\colon E_p \to E_p, \qquad s \mapsto -2s,$$

which is a separable morphism since p is odd. Also, the degree of θ is four. It follows that $\theta^{-1}(t_p)$ has four geometric points. Here a geometric point means one defined over an algebraic closure of the finite field with p elements. Connecting the four geometric points to t_p, we obtain four lines passing through t_p, and these four lines are precisely represented by the zeros of $Q_t(u,v)$ modulo p. The nonvanishing of the discriminant of $Q_t(u,v)$ modulo p is equivalent to the condition that the four lines are distinct. Suppose that two of the four lines coincide, say L_0. Then L_0 contains $s_1, s_2 \in \theta^{-1}(t_p)$ 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 do not have more than three points of intersection counted with multiplicity. Hence we have 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. It is based on the representation of the discriminant as a product of root differences. Indeed, if we let δ_t be the discriminant of Q_t, then we find that

$$\Delta_t = \delta_t \cdot Q_t(1,0)^2 \tag{3.12}$$

from the representation of the discriminant as square of the product of all possible differences between roots. In the first step, we showed that δ_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 3.1, from which we see that the number

$$Q_t(1,0) = z_t^2$$

is prime to p if $t \in Y(\mathbb{Z}_S)$. ∎

The argument using the dull algebraic identity (3.12) can be replaced with the following geometric argument. We would like to show geometrically that any of the four lines defined by $Q_t = 0$ equals the line defined by $L_t = 0$, after reduction modulo p. Let us begin with the following lemma.

Lemma 3.6. *None of the four geometric points belonging to $\theta^{-1}(t_p)$ is the origin of E_p.*

Proof. 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 θ, that

$$-2s = t_p,$$

which implies $t_p = 0$. This 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 zeros 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 both t_p and the origin. Note that s_0 cannot be the origin by the lemma. It follows that L_0 and E_p have at least five points of intersection counted with multiplicity, to which the origin contributes at least three, and s_0 contributes two. This is clearly absurd.

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 punctured universal elliptic curve, whose $\mathbb{Z}_S$-points are given by

$$\mathcal{Y}(\mathbb{Z}_S) = \{(Y,P)\colon P \in Y(\mathbb{Z}_S),\ Y \text{ is a punctured smooth elliptic curve over } \mathbb{Z}_S\} \tag{4.1}$$

where a smooth elliptic curve over $\mathbb{Z}_S$ means an elliptic curve over $\mathbb{Q}$ which has good reduction outside S. Note that the set of $\mathbb{Z}_S$-points on such an elliptic curve E over $\mathbb{Q}$ can be unambiguously defined in this way, because for every prime $p \notin S$, E has a unique minimal and smooth model over $\mathbb{Z}_p$, the ring of p-adic integers.

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

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

where the negative denotes the negative 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*

$$\kappa\colon \mathcal{Y}(\mathbb{Z}_S)/\{\pm 1\} \to \{S\text{-}admissible\ pairs\ \}/\!\sim \tag{4.2}$$

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

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

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

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

$$-4B_2^3B_3^2 + 16B_2^4B_4 - 27B_3^4 + 144B_2B_3^2B_4 - 128B_2^2B_4^2 + 256B_4^3 \tag{4.3}$$

is an S-unit. By defining

$$\begin{aligned} x_t &= -\tfrac{1}{6}B_2, & a_4 &= -\tfrac{1}{4}(B_4 + 3x_t^2), \\ y_t &= -\tfrac{1}{8}B_3, & a_6 &= y_t^2 - x_t^3 - a_4 x_t, \end{aligned} \tag{4.4}$$

we obtain a curve

$$E\colon y^2 = x^3 + a_4 x + a_6 \tag{4.5}$$

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 we have to rely on the assumption that $2, 3 \in S$. We need to show that E has good reduction outside S. By direct computation, the discriminant of E is given by

$$2^{-8} \cdot (-4B_2^3B_3^2 + 16B_2^4B_4 - 27B_3^4 + 144B_2B_3^2B_4 - 128B_2^2B_4^2 + 256B_4^3),$$

which is an S-unit by comparison with the formula (4.3) for the discriminant of $Q(u, v)$.

We need to verify that the association $(L, Q) \mapsto (E, P)$ is well-defined. As observed earlier, there is an involution on the set of minimal pairs sending B_3 to $-B_3$. It is clear from (4.4) that it corresponds to the involution $(E, P) \mapsto (E, -P)$. Thus we have constructed a section of κ, showing its surjectivity.

To see the injectivity of κ, recall that $2, 3 \in S$, hence an elliptic curve E which has good reduction outside S has a model of the form (4.5) which has good reduction outside S, and there is no prime p for which $p^4 \mid a_4$ and $p^6 \mid a_6$. Starting with a model of E which is minimal outside 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 3.1, we have

$$Q(u, v) = u^4 - 6x_t u^2v^2 - 8y_t uv^3 - (3x_t^2 + 4a_4)v^4,$$

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

$$p^2 \mid x_t, \qquad p^4 \mid 3x_t^2 + 4a_4,$$

from which we conclude that $p^4 \mid a_4$. Furthermore, non-minimality at p implies $p^3 \mid y_t$. However, by rewriting the equation of the elliptic curve in the form

$$a_6 = y_t^2 - x_t^3 - a_4 x_t$$

one sees that p^6 divides a_6. This contradicts the minimality of E at p.

Thus, we have shown that κ is a bijection. ■

Remark 4.2. One may wonder what can be said about $\kappa(Y(\mathbb{Z}_S))$ for a fixed curve Y. A difficulty in characterizing $\kappa(Y(\mathbb{Z}_S))$ among the equivalence classes of all admissible pairs arises from the existence of twists. If Y' is a quadratic twist of Y, then it is not possible to separate $\kappa(Y(\mathbb{Z}_S))$ from $\kappa(Y'(\mathbb{Z}_S))$ using invariants.

5. The example $S = \{2\}$. The aim of the 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 the S-admissible pairs. Although we assumed $2, 3 \in S$ in Theorem 4.1, as long as numerical examples are concerned, the assumption $2, 3 \in S$ is not strictly necessary. Indeed, the map κ exists anyway, and for each S-admissible pair, one obtains a point of $\mathcal{Y}$ defined over $\mathbb{Z}_S[6^{-1}]$. 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 $\mathcal{Y}(\mathbb{Z}_S)/\{\pm 1\}$.

Despite 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 Smart [9], who computed all binary quintics, reducible or not, whose discriminant is an S-unit where $S = \{2\}$. In particular, all S-admissible pairs can be obtained from [9], by choosing all possible linear factors of each binary quintic.

Table 1 is produced from Table 5 of [9], which contains all reducible binary quintic forms whose discriminant is a power of 2 up to sign, i.e., an S-unit with $S = \{2\}$. In [9], the title of Table 5 says that the quintic forms listed have discriminant a power of 2, but also the forms with discriminant minus a power of 2 have been included. 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 an S-unit with $S = \{2\}$.

We wish to find all $\{2\}$-admissible pairs (L, Q) from Table 1. For each (L, Q), the quintic form $L \cdot Q$ must be equivalent to f_i for some i, hence we can find all of them by finding all possible factorisations of f_i into a linear form and quartic forms. In fact, f_i for $1 \leq i \leq 4$ has three linear factors, and the others have a unique linear factor. It turns out that every reducible quintic form in the table admits a linear factor.

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

$$vu(u + v)(u^2 + v^2),$$

hence there are three pairs

$$(v, u(u + v)(u^2 + v^2)), \qquad (u, v(u + v)(u^2 + v^2)), \qquad (u + v, uv(u^2 + v^2))$$

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

$$(L_1, Q_1) = (v, u^4 + 10u^2v^2 + 40uv^3 - 51v^4), \qquad (L_2, Q_2) = (v, u^4 - v^4),$$

in their minimal forms. From (L_1, Q_1), we obtain the curve

$$E_1\colon y^2 = x^3 + \tfrac{32}{3}x + \tfrac{1280}{27}$$

Table 1. Reducible quintics whose discriminant is a power of 2 up to sign

i	$f_i(u,v)$	i	$f_i(u,v)$
1	$u^4v+u^3v^2+u^2v^3+uv^4$	2	$2u^4v+2u^3v^2-u^2v^3-uv^4$
3	$8u^5-6u^3v^2+uv^4$	4	$2u^5-3u^3v^2+uv^4$
5	u^5+4uv^4	6	$u^5+3u^3v^2+2uv^4$
7	$u^4v+3u^2v^3+2v^5$	8	$u^5+2u^4v+4u^3v^2+4u^2v^3+4uv^4$
9	$u^5+3u^4v+2u^3v^2+2u^2v^3+uv^4-v^5$	10	u^5-4uv^4
11	$u^5+4u^4v+4u^3v^2+8u^2v^3+4uv^4$	12	$u^5-4u^4v+8u^2v^3-4uv^4$
13	$u^4v-8u^3v^2+12u^2v^3+16uv^4-28v^5$	14	$u^5+u^4v+uv^4+v^5$
15	u^5+uv^4	16	$u^5+12u^3v^2+4uv^4$
17	u^4v-2v^5	18	$u^5+u^4v-2uv^4-2v^5$
19	u^5-2uv^4	20	u^4v+2v^5
21	u^5+2uv^4	22	$3u^5+8u^4v+4u^3v^2+4uv^4$
23	$4u^4v+4u^2v^3-16uv^4+9v^5$	24	$u^5-4u^3v^2+2uv^4$
25	$u^5+2u^4v-4u^3v^2-8u^2v^3+2uv^4+4v^5$	26	$u^4v-4u^2v^3+2v^5$
27	$u^5+u^4v-4u^3v^2-4u^2v^3+2uv^4+2v^5$	28	$u^5+9u^4v+14u^3v^2-34u^2v^3-19uv^4+5v^5$
29	$u^5+4u^4v-6u^3v^2-4u^2v^3+uv^4$	30	$4u^4v+16u^3v^2-12u^2v^3-24uv^4+17v^5$
31	$4u^5+12u^4v-28u^3v^2-12u^2v^3+41uv^4-17v^5$	32	$u^5-8u^4v+4u^3v^2+16u^2v^3+4uv^4$
33	$u^5-7u^4v-4u^3v^2+20u^2v^3+20uv^4+4v^5$	34	$u^5+4u^3v^2+2uv^4$
35	$u^4v+4u^2v^3+2v^5$	36	$u^4v-2u^2v^3-v^5$
37	$u^5+u^4v-2u^3v^2-2u^2v^3-uv^4-v^5$	38	$u^5-2u^3v^2-uv^4$
39	$u^5+4u^3v^2-4uv^4$	40	$u^4v+4u^2v^3-4v^5$
41	$u^5+u^4v+4u^3v^2+4u^2v^3-4uv^4-4v^5$	42	$u^4v+4u^3v^2-6u^2v^3+12uv^4-7v^5$
43	$u^5+3u^4v-10u^3v^2+18u^2v^3-19uv^4+7v^5$	44	$u^5-2u^3v^2+2uv^4$
45	$u^4v-2u^2v^3+2v^5$	46	$u^5+u^4v-2u^3v^2-2u^2v^3+2uv^4+2v^5$
47	$u^5+4u^3v^2+8uv^4$	48	$u^4v+4u^2v^3+8v^5$
49	$5u^5+13u^4v+2u^3v^2-14u^2v^3-3uv^4+5v^5$	50	$u^4v+6u^2v^3+8uv^4+5v^5$
51	$u^5+4u^4v+4u^3v^2-8u^2v^3+4uv^4$		

with point

$$t=(-\tfrac{5}{3},-5)$$

on it. The above model is not minimal at 3. The minimal equation for E_1 is

$$E_{128\mathrm{a}1}\colon y^2=x^3+x^2+x+1,$$

whose label in Cremona's Elliptic Curve Database is "128a1", and the coordinates of t are

$$t=(-\tfrac{3}{4},\tfrac{5}{8})$$

with respect to the minimal equation.

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

$$y^2=x^3+\tfrac{1}{4}x$$

and the point $t=(0,0)$. The above equation corresponds to the minimal equation

$$E_{32\mathrm{a}1}\colon y^2=x^3+4x,$$

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 the 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 can be obtained from $i = 11, 37, 40, 41$.

By carrying out similar calculations for all f_i, we obtain Table 2. We note that f_{30} and f_{31} give rise to two equivalent pairs, and so do f_{42} and f_{43}.

Table 2. Correspondence between f_i’s and elliptic curves

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. The average number of integral points on elliptic curves.

In this section, we turn to the main goal of the paper, namely the average number of integral points on elliptic curves. Recall that we are considering curves of the form

$$Y_{a,b}\colon y^2 = x^3 + ax + b \tag{6.1}$$

such that $a, b \in \mathbb{Z}$ and $4a^3 + 27b^2 \neq 0$. The curves $Y_{a,b}$ will be ordered by height, normalised in the following way.

Definition 6.1. Define the *height* of $Y_{a,b}$ to be

$$H(Y_{a,b}) = \max\{2^{12}3^4|a|^3, 2^{14}3^{12}b^2\}. \tag{6.2}$$

For any positive real number T, define $R(T)$ to be the number of curves $Y_{a,b}$ up to height T.

Lemma 6.2. *For sufficiently large T, we have*

$$R(T) < 2^{-11}3^{-22/3}T^{5/6} < 1.55 \times 10^{-7} \times T^{5/6}.$$

Proof. This follows from the observation that there are $O(T^{1/3})$ pairs (a,b) satisfying $4a^3+27b^2=0$ and $\max\{2^{12}3^4|a|^3, 2^{14}3^{12}b^2\} < T$. ■

For any positive number T, define

$$N(T) = \sum_{Y_{a,b}, H(Y_{a,b})<T} \sum_{t\in Y_{a,b}(\mathbb{Z})/\{\pm 1\}} 1, \tag{6.3}$$

which is the total number of integral points on the curves of the form $Y_{a,b}$ up to height T, counted modulo the action of $\{\pm 1\}$. Accordingly, the ratio $N(T)/R(T)$ will be called the average number of integral points on these curves up to height T. By Lemma 6.2, finding an upper bound for the average number of integral points reduces to finding a bound for $N(T)$.

Theorem 6.3. *We have*

$$N(T) < (31.5\ldots)T^{5/6} \tag{6.4}$$

for all sufficiently large $T>0$. In particular, the average number of integral points on curves of the form $Y_{a,b}$, namely the ratio $N(T)/R(T)$, is bounded by 2.1×10^8 *for all sufficiently large $T>0$. It is counted modulo the natural involution on the underlying elliptic curves.*

In the rest of the section, we give the proof for Theorem 6.3. The starting point is a map

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

where

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

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 the map κ we introduced earlier, but in this way the argument becomes more natural.

Naturally $\mathrm{GL}_2(\mathbb{Z})$ acts on $\mathbb{Z}^2\times\mathrm{Sym}^4(\mathbb{Z}^2)^*$, and the action preserves solutions of the Thue equations. That is, the subset

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

is preserved by the action of $\mathrm{GL}_2(\mathbb{Z})$.

Proposition 6.4. *The map*

$$\phi\colon \{(E,t)\colon t\in E(\mathbb{Z})\}/\{\pm 1\} \to \{((n,m),Q(u,v))\colon Q(n,m)=1\}/\sim \tag{6.5}$$

is injective, where $\sim$ denotes the equivalence relation induced by the action of $\mathrm{GL}_2(\mathbb{Z})$.

Proof. Suppose that $(Y_{a,b}, t)$ and $(Y_{a',b'}, t')$ have the same image under ϕ. Then we have $\gamma \in \mathrm{GL}_2(\mathbb{Z})$ which fixes $(1,0)$ and transforms

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

into

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

It is easy to see that the identity and $(u, v) \mapsto (\pm u, \pm v)$ are the only possibilities for γ. Indeed, the stabiliser of $(1,0)$ in $\mathrm{GL}_2(\mathbb{Z})$ is generated by the group of unipotent matrices, together with the transformation $(u, v) \mapsto (\pm u, \pm v)$. By comparing the coefficients of u^3v, one sees that γ must be of the form $(u, v) \mapsto (\pm u, \pm v)$. Thus we conclude that $a = a'$, $b = b'$, and $t = \pm t'$. Of course, the possible minus sign in the equality $t = \pm t'$ means the negative with respect to the group law of the underlying elliptic curve. ∎

Remark 6.5. Note that

$$((n,m), Q(u,v)) \sim ((-n,-m), Q(u,v)) \tag{6.6}$$

via the matrix with -1's on the diagonal.

We briefly recall the invariant theory of binary quartic forms. 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 action of $\mathrm{GL}_2(\mathbb{Z})$, there are two invariants

$$\begin{aligned} J_2 &= \tfrac{1}{12}c_2^2 - \tfrac{1}{4}c_1c_3 + c_0c_4, \\ J_3 &= \tfrac{1}{216}c_2^3 - \tfrac{1}{48}c_1c_2c_3 + \tfrac{1}{16}c_0c_3^2 + \tfrac{1}{16}c_1^2c_4 - \tfrac{1}{6}c_0c_2c_4, \end{aligned}$$

of degree two and three respectively. We define the height of Q by

$$H(Q) = \max\{2^6 3^4 \cdot |J_2|^3,\ 2^{10} 3^{12} \cdot J_3^2\}$$

where the coefficients of $|J_2|^3$ and J_3^2 are chosen so that our definition of height agrees with that of [3]. Note that our normalisation of J_2 and J_3 differs from the corresponding invariants I and J given in [3, §2].

Proposition 6.6. *Let $t \in Y_{a,b}(\mathbb{Z})$ and $\phi((Y_{a,b}, t)) = (L, Q)$. Then*

$$H(Y_{a,b}) = H(Q). \tag{6.7}$$

In other words, ϕ preserves the heights.

Proof. This follows by straightforward calculation. Indeed,

$$Q = u^4 - 6x_t u^2v^2 - 8y_t uv^3 - (3x_t^2 + 4a)v^4$$

and we have the relation $y_t^2 = x_t^3 + ax_t + b$, from which one deduces $J_2(Q) = 4a$ and $J_3(Q) = 4b$. Thus

$$H(Q) = \max\{2^{12}3^4|a|^3, 2^{14}3^{12}b^2\} = H(Y_{a,b}). \ ∎$$

As we have the injective map ϕ which preserves the heights, the estimation of $N(T)$ is reduced to the estimation of the number of pairs $((n,m),Q(u,v))$ which lie in the image of ϕ, modulo $\mathrm{GL}_2(\mathbb{Z})$-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 collection of $\mathrm{GL}_2(\mathbb{Z})$-orbits of binary forms of type i whose height is less than T.

We consider three subcollections of $X^1_j(T)$, $j=0,1,2$, defined by the condition that an element in $X^1_j(T)$ has exactly $4-2j$ linear factors over $\mathbb{R}$.

Theorem 6.7. *We have*

$$\sum_{Q\in X^1_0(T)} 1 = \frac{2\pi^2}{405}T^{5/6} + O(T^{3/4+\epsilon}), \tag{6.8}$$

$$\sum_{Q\in X^1_1(T)} 1 = \frac{16\pi^2}{405}T^{5/6} + O(T^{3/4+\epsilon}), \tag{6.9}$$

$$\sum_{Q\in X^1_2(T)} 1 = \frac{4\pi^2}{405}T^{5/6} + O(T^{3/4+\epsilon}), \tag{6.10}$$

$$\sum_{Q\in X^3(T)} 1 = O(T^{2/3+\epsilon}), \tag{6.11}$$

where the sum is taken over all irreducible integral binary quartic forms with respect to $\mathrm{GL}_2(\mathbb{Z})$*-equivalence.*

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

Proposition 6.8. *For* $X^2(T)$*,*

$$\sum_{Q\in X^2(T),\, Q\in \mathrm{Im}(\phi)} 1 = O(T^{3/4}). \tag{6.12}$$

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

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

where r is an integer and $C(u,v)$ is a binary cubic form with integral coefficients such that $C(1,0)=1$. By means of the translation $u\mapsto u+rv$, Q is equivalent to the form

$$u(v^3+c_1v^2u+c_2vu^2+c_3u^3) \tag{6.13}$$

with integers c_1, c_2 and c_3. Using the translation $v \mapsto v + r'u$ for some integer r' if necessary, we may assume that $|c_1| \leq 1$. The invariants of (6.13) are given as

$$J_2 = \tfrac{1}{12}c_2^2 - \tfrac{1}{4}c_1c_3, \qquad J_3 = \tfrac{1}{216}c_2^3 - \tfrac{1}{48}c_1c_2c_3 + \tfrac{1}{16}c_3^2,$$

and $|J_2| = O(T^{1/3})$ and $|J_3| = O(T^{1/2})$. Hence the discriminant of (6.13) 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 $|c_3| = O(T^{1/2})$ imply $|c_2| = O(T^{1/4})$. We conclude that there are $O(T^{3/4})$ possibilities for the triple (c_1, c_2, c_3). ∎

We also need to invoke the works of Evertse and Akhtari–Okazaki on the number of solutions of a given Thue–Mahler equation, which we recall now. A Thue–Mahler equation is about 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

$$h(u,v) = \pm \prod_{p_i \in S} p_i^{e_i} \tag{6.14}$$

where e_i are nonnegative integers, and u, v are relatively prime integers. A Thue–Mahler equation with $S = \emptyset$ is called a *Thue equation.* We will rely on a corollary which is easily implied by the following theorem of Evertse.

Theorem 6.9 ([4, Corollary 2]). *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 the associated equation* (6.14) *has at most*

$$2 \times 7^{r^3(2s+3)} \tag{6.15}$$

solutions.

We are concerned about the case when $h(u,v)$ is a quartic with nonzero discriminant, and S is empty. The following corollary is a direct consequence of Evertse's theorem.

Corollary 6.10. *Let $Q(u,v)$ be a binary quartic form with nonzero discriminant. The equation*

$$Q(u,v) = \pm 1 \tag{6.16}$$

has at most

$$2 \times 7^{4^3 \cdot 3} < 3.63 \times 10^{162} \tag{6.17}$$

solutions.

Remark 6.11. Theorem 6.9 will not be used in what follows, and we only need its consequence, Corollary 6.10, for the proof of the main theorem.

Despite 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 6.12. *Let $Q(u,v)$ be an irreducible quartic equation. The associated Thue equation*

$$Q(u,v) = \pm 1 \tag{6.18}$$

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 6.3 is straightforward. Indeed, from the injectivity of ϕ, one has

$$N(T) \leq \sum_{Q\in X^1(T)} \sum_{Q(n,m)=1} 1 + \sum_{Q\in X^2(T)} \sum_{Q(n,m)=1} 1 + \sum_{Q\in X^3(T)} \sum_{Q(n,m)=1} 1$$

where the sum over $Q(n,m) = 1$ means the following: 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 (6.6) shows that two solutions (n,m) and $(-n,-m)$ should be counted once. By Theorems 6.7 and 6.12, one has

$$\begin{aligned}
&\sum_{Q\in X^1(T)} \sum_{Q(n,m)=1} 1 \\
&\quad = \sum_{Q\in X^1_0(T)} \sum_{Q(n,m)=1} 1 + \sum_{Q\in X^1_1(T)} \sum_{Q(n,m)=1} 1 + \sum_{Q\in X^1_2(T)} \sum_{Q(n,m)=1} 1 \\
&\quad = 37 \cdot \frac{2\pi^2}{405} T^{5/6} + 61 \cdot \frac{16\pi^2}{405} T^{5/6} + 61 \cdot \frac{4\pi^2}{405} T^{5/6} + O(T^{3/4+\epsilon}) \\
&\quad < (31.5\ldots) T^{5/6} + O(T^{3/4+\epsilon}),
\end{aligned}$$

while Theorem 6.7, Proposition 6.8, and Corollary 6.10 imply that

$$\sum_{Q\in X^2(T)} \sum_{Q(n,m)=1} 1 = O(T^{3/4}), \qquad \sum_{Q\in X^3(T)} \sum_{Q(n,m)=1} 1 = O(T^{2/3+\epsilon}),$$

both of which have smaller orders than $T^{5/6}$. We conclude that

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

for all sufficiently large $T > 0$. Combining this with Lemma 6.2, we obtain the desired upper bound on the average number of integral points on $Y_{a,b}$.

Acknowledgements. This work was supported by IBS-R003-D1. The author thanks anonymous referees for helpful comments.

References

[1] S. Akhtari and R. Okazaki, *Quartic Thue equations*, J. Number Theory 130 (2010), 40–60.

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

[3] 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), 191–242.

[4] J.-H. Evertse, *On equations in S-units and the Thue–Mahler equation*, Invent. Math. 75 (1984), 561–584.

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

[6] 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.

[7] L. J. Mordell, *Diophantine Equations*, Pure Appl. Math. 30, Academic Press, London, 1969.

[8] J. Silverman, *A quantitative version of Siegel's theorem: integral points on elliptic curves and Catalan curves*, J. Reine Angew. Math. 378 (1987), 60–100.

[9] N. P. Smart, *S-unit equations, binary forms, and curves of genus 2*, Proc. London Math. Soc. (3) 75 (1997), 271–307.

Dohyeong Kim
Department of Mathematics
University of Michigan
2074 East Hall
530 Church Street
Ann Arbor, MI 48109-1043, U.S.A.
E-mail: dohyeong@umich.edu