JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 16, Number 4, Pages 981–990
S 0894-0347(03)00433-8
Article electronically published on July 8, 2003

HILBERT'S TENTH PROBLEM AND MAZUR'S CONJECTURE FOR LARGE SUBRINGS OF $\mathbb{Q}$

BJORN POONEN

1. Introduction

Hilbert's Tenth Problem, in modern terms, was to find an algorithm (Turing machine) to decide, given a polynomial equation $f(x_1, \dots, x_n) = 0$ with coefficients in $\mathbb{Z}$, whether it has a solution with $x_1, \dots, x_n \in \mathbb{Z}$. Y. Matijasevič [Mat70], building on earlier work of M. Davis, H. Putnam, and J. Robinson [DPR61], showed that no such algorithm exists. If one replaces $\mathbb{Z}$ in both places by a different commutative ring R (let us assume its elements can be and have been encoded for input into a Turing machine), one obtains a different question, called *Hilbert's Tenth Problem over R*, whose answer depends on R. These problems are discussed in detail in [DLPVG00].

In particular, the answer for $R = \mathbb{Q}$ is unknown. Hilbert's Tenth Problem over $\mathbb{Q}$ is equivalent to the general problem of deciding whether a variety over $\mathbb{Q}$ has a rational point. One approach to proving that Hilbert's Tenth Problem over $\mathbb{Q}$ has a negative answer would be to deduce this from Matijasevič's theorem for $\mathbb{Z}$, by showing that $\mathbb{Z}$ is diophantine over $\mathbb{Q}$ in the following sense:

Definition 1.1. Let R be a ring, and let $A \subseteq R^m$. Then A is *diophantine over R* if and only if there exists a polynomial f in $m + n$ variables with coefficients in R such that

$$A = \{\, a \in R^m \mid \exists x \in R^n \text{ such that } f(a, x) = 0 \,\}.$$

On the other hand, Mazur conjectures that if X is a variety over $\mathbb{Q}$, then the topological closure of $X(\mathbb{Q})$ in $X(\mathbb{R})$ has only finitely many components [Maz92], [Maz95]. This would imply that $\mathbb{Z}$ is not diophantine over $\mathbb{Q}$. More generally, Cornelissen and Zahidi [CZ00] have shown that Mazur's Conjecture implies that there is no diophantine model of $\mathbb{Z}$ over $\mathbb{Q}$.

Definition 1.2. A *diophantine model* of $\mathbb{Z}$ over $\mathbb{Q}$ is a set $A \subseteq \mathbb{Q}^n$ that is diophantine over $\mathbb{Q}$ with a bijection $\mathbb{Z} \to A$ under which the graphs of addition and multiplication on $\mathbb{Z}$ correspond to subsets of $A^3 \subseteq \mathbb{Q}^{3n}$ that are diophantine over $\mathbb{Q}$.

This is important, because the existence of such a diophantine model, together with Matijasevič's Theorem, would imply a negative answer for Hilbert's Tenth Problem over $\mathbb{Q}$.

Received by the editors December 8, 2002.

2000 *Mathematics Subject Classification.* Primary 11U05; Secondary 11G05.

Key words and phrases. Hilbert's Tenth Problem, elliptic curve, Mazur's Conjecture, diophantine definition.

This research was supported by NSF grant DMS-0301280 and a Packard Fellowship.

This paper studies Hilbert's Tenth Problem over rings between $\mathbb{Z}$ and $\mathbb{Q}$. Such rings are in bijection with subsets of the set $\mathcal{P}$ of prime numbers. Namely, given $S \subseteq \mathcal{P}$, one has the ring $\mathbb{Z}[S^{-1}]$, and conversely, given a subring R between $\mathbb{Z}$ and $\mathbb{Q}$, one has $R = \mathbb{Z}[S^{-1}]$ where $S = \mathcal{P} \cap R^{\times}$.

Using quadratic forms as in J. Robinson's work, one can show that for any prime p the ring $\mathbb{Z}_{(p)}$ of rational numbers with denominators prime to p is diophantine over $\mathbb{Q}$ [KR92, Proposition 3.1]. A short argument using this shows that for finite S, Hilbert's Tenth Problem over $\mathbb{Z}[S^{-1}]$ has a negative answer.

In this paper, we give the first examples of *infinite* subsets S of $\mathcal{P}$ for which Hilbert's Tenth Problem over $\mathbb{Z}[S^{-1}]$ has a negative answer. In fact, we show that there exist such S of natural density 1, so in one sense, we are approaching a negative answer for $\mathbb{Q}$. (See Section 6 for the definition of natural density.) Previously, Shlapentokh proved that if K is a totally real number field or a totally complex degree-2 extension of a totally real number field, then there exists a set of places S of K of Dirichlet density arbitrarily close to $1 - [K : \mathbb{Q}]^{-1}$ such that if $\mathcal{O}_{K,S}$ is the subring of elements of K that are integral at all places outside S, then Hilbert's Tenth Problem over $\mathcal{O}_{K,S}$ has a negative answer [Shl97], [Shl00], [Shl02]. But for $K = \mathbb{Q}$, this gives nothing beyond Matijasevič's Theorem.

More generally, we prove the following:

Theorem 1.3. *There exist disjoint recursive sets of primes T_1 and T_2, both of natural density* 0*, such that for any set S of primes containing T_1 and disjoint from T_2, the following hold:*

1. *There exists an affine curve E' over $\mathbb{Z}[S^{-1}]$ such that the topological closure of $E'(\mathbb{Z}[S^{-1}])$ in $E'(\mathbb{R})$ is an infinite discrete set.*
2. *The set of positive integers with addition and multiplication admits a diophantine model over $\mathbb{Z}[S^{-1}]$.*
3. *Hilbert's Tenth Problem over $\mathbb{Z}[S^{-1}]$ has a negative answer.*

Remark 1.4.

- (i) Arguably (3) is the most important of the three parts. We have listed the parts in the order they will be proved.
- (ii) A subset $T \subseteq \mathbb{Z}$ is *recursive* if and only if there exists an algorithm (Turing machine) that takes as input an integer t and outputs YES or NO according to whether $t \in T$.
- (iii) We use natural density instead of Dirichlet density in order to have a slightly stronger statement. See [Ser73, VI.4.5] for the definition of Dirichlet density and its relation to natural density.

Previously, Shlapentokh [Shl03] used norm equations to prove that there exist sets $S \subseteq \mathcal{P}$ of Dirichlet density arbitrarily close to 1 for which there exists an affine variety X over $\mathbb{Z}[S^{-1}]$ such that the closure of $X(\mathbb{Z}[S^{-1}])$ in $X(\mathbb{R})$ has infinitely many connected components. (She also proved an analogous result for localizations of the ring of integers of totally real number fields and totally complex degree-2 extensions of totally real number fields. For number fields with exactly one conjugate pair of nonreal embeddings, she obtained an analogous result, but with density only 1/2.)

Question 4.1 of [Shl03] asked whether over $\mathbb{Q}$ one could do the same for some $S \subseteq \mathcal{P}$ of Dirichlet density exactly 1. Part (1) of our Theorem 1.3 gives an affirmative answer (take $S = \mathcal{P} - T_2$). In fact, it was the attempt to answer Shlapentokh's

Question 4.1 that inspired this paper, so the author thanks her for asking the right question.

The rest of this paper is devoted to proving Theorem 1.3. The strategy is to take an elliptic curve E over $\mathbb{Q}$ such that $E(\mathbb{Q})$ is generated by one point P of infinite order and to construct T_1 (resp. T_2) so that certain prime multiples (resp. at most finitely many other integer multiples) of P have coordinates in $\mathbb{Z}[S^{-1}]$. Using Vinogradov's result on the equidistribution of the prime multiples of an irrational number modulo 1, we can prescribe the approximate locations of the prime multiples of P in $E(\mathbb{R})$. If we prescribe them so that their y-coordinates approximate the set of positive integers sufficiently well, then approximate addition and approximate squaring on the set A of these y-coordinates make A into a diophantine model of the positive integers.

Shlapentokh and the author plan eventually to write a joint paper generalizing Theorem 1.3 to other number fields and to places other than the real place.

2. Elliptic curve setup

Let E be an elliptic curve over $\mathbb{Q}$ of rank 1. To simplify the arguments, we will assume moreover that $E(\mathbb{Q}) \simeq \mathbb{Z}$, that $E(\mathbb{R})$ is connected, and that E does not have complex multiplication. For example, these conditions hold for the smooth projective model of $y^2 = x^3+x+1$. Let P be a generator of $E(\mathbb{Q})$. Fix a Weierstrass equation $y^2 = x^3 + ax + b$ for E, where $a, b \in \mathbb{Z}$.

Let $E' = \operatorname{Spec} \mathbb{Z}[S_{\text{bad}}^{-1}][x,y]/(y^2 - (x^3 + ax + b))$, where S_{bad} is a finite set of primes such that E' is smooth over $\mathbb{Z}[S_{\text{bad}}^{-1}]$. In particular, $2 \in S_{\text{bad}}$. Enlarge S_{bad} if necessary so that $P \in E'(\mathbb{Z}[S_{\text{bad}}^{-1}])$.

3. Denominators of x-coordinates

For nonzero $n \in \mathbb{Z}$, let $d_n \in \mathbb{Z}_{>0}$ be the prime-to-S_{bad} part of the denominator of $x(nP)$; that is, d_n is the product one obtains if one takes the prime factorization of the denominator of $x(nP)$ and omits the powers of primes in S_{bad}. Define $d_0 = 0$. The notation $m \mid n$ means $n \in m\mathbb{Z}$.

Lemma 3.1.

(a) *For any $r \in \mathbb{Z}$, the set $\{\, n \in \mathbb{Z} : r \mid d_n \,\}$ is a subgroup of $\mathbb{Z}$.*

(b) *There exists $c \in \mathbb{R}_{>0}$ such that $\log d_n = (c - o(1))n^2$ as $n \to \infty$ (cf.* [Sil88, Lemma 8]*).*

Proof. (a) We may reduce to the case where $r = p^e$ for some prime $p \notin S_{\text{bad}}$ and $e \in \mathbb{Z}_{>0}$. Thus it suffices to show that the set $E_e := \{\, Q \in E(\mathbb{Q}_p) : v_p(x(Q)) \leq -e \,\} \cup \{O\}$ is a subgroup of $E(\mathbb{Q}_p)$, where $v_p \colon \mathbb{Q}_p^* \to \mathbb{Z}$ is the p-adic valuation. The set E_1 is the kernel of the reduction map $E(\mathbb{Q}_p) \to E(\mathbb{F}_p)$ (extend E to a smooth proper curve over $\mathbb{Z}_p$ to make sense of this). Since $p > 2$, the formal logarithm $\lambda \colon E_1 \to p\mathbb{Z}_p$ is an isomorphism [Sil92, IV.6.4]. By [Sil92, IV.5.5, IV.6.3], $v_p(\lambda(Q)) = v_p(z(Q))$ for all $Q \in E_1$, where $z = -x/y$ is the standard parameter for the formal group. By [Sil92, pp. 113–114], $x = -2z^{-2} + \cdots \in \mathbb{Z}_p((z))$, so $v_p(x(Q)) = -2v_p(z(Q)) = -2v_p(\lambda(Q))$. Thus G_e corresponds under λ to $p^{\lceil e/2 \rceil}\mathbb{Z}_p$ and is hence a subgroup.

(b) The number $\log d_n$ is the logarithmic height $h(nP)$, except that in the sum defining the height, the terms corresponding to places in $S_{\text{bad}} \cup \{\infty\}$ have been omitted. A standard diophantine approximation result (see Section 7.4 of [Ser97]) implies that each such term contributes at most a fraction $o(1)$ of the height, as $n \to \infty$. If $\hat{h}$ is the canonical height, then $h(nP) = \hat{h}(nP) + O(1) = \hat{h}(P)n^2 + O(1)$. Take $c = \hat{h}(P)$, which is positive, since P is not torsion. □

Remark 3.2. The bottom of p. 306 in [Aya92] relates the denominators of $x(nP)$ in lowest terms to the sequence of values of division polynomials evaluated at P. The study of divisibility properties of the latter sequence is very old: results were claimed in the 19th century by Lucas (but apparently not published), and proofs were given in [War48].

For $n \in \mathbb{Z}$, let S_n be the set of prime factors of d_n. If $m, n \in \mathbb{Z}$, then (m, n) denotes their greatest common divisor.

Corollary 3.3. *If $m, n \in \mathbb{Z}$, then $S_{(m,n)} = S_m \cap S_n$. In particular, if $(m, n) = 1$, then S_m and S_n are disjoint.*

Proof. Lemma 3.1(a) implies the first statement. Since $P \in E'(\mathbb{Z}[S_{\text{bad}}^{-1}])$, we have $S_1 = \emptyset$, and the second statement follows. □

Lemma 3.4. *If ℓ and m are primes, and* $\max\{\ell, m\}$ *is sufficiently large, then $S_{\ell m} - (S_\ell \cup S_m)$ is nonempty.*

Proof. If $p \mid d_m$, or equivalently $v_p(x(mP)) < 0$, then using the formal logarithm λ as in the proof of Lemma 3.1(a), we obtain

$$v_p(d_{\ell m}) = -v_p(x(\ell mP)) = 2v_p(\lambda(\ell mP)) = 2v_p(\ell\lambda(mP)) = v_p(\ell^2 d_m).$$

If $S_{\ell m} - (S_\ell \cup S_m)$ were empty, then for each $p \mid d_{\ell m}$ we could apply either this result or the analogue with ℓ and m interchanged and hence deduce $d_{\ell m} \mid \ell^2 m^2 d_\ell d_m$. This contradicts Lemma 3.1(b) if $\max\{\ell, m\}$ is sufficiently large. □

Remark 3.5. Our Lemma 3.4 is a special case of Lemma 9 of [Sil88] (except for the minor differences that [Sil88] requires E to be in minimal Weierstrass form and considers the full denominator instead of its prime-to-S_{bad} part). The method of proof is the same. These results may be viewed as elliptic analogues of Zsigmondy's Theorem; see [Eve02].

4. Definition of T_1 and T_2

For each prime number ℓ, let a_ℓ be the smallest $a \in \mathbb{Z}_{>0}$ such that $d_{\ell^a} > 1$. By Lemma 3.1(b), a_ℓ exists, and $a_\ell = 1$ for all ℓ outside a finite set L of primes. Baker's method [Ser97, Chapter 8] lets us compute the finite set $E'(\mathbb{Z}[S_{\text{bad}}^{-1}])$, so the set L and the values a_ℓ for $\ell \in L$ are computable.

Let $p_\ell = \max S_{\ell^a}$ where $a = a_\ell$. For primes ℓ and m (possibly equal), Lemma 3.4 lets us define $p_{\ell m} = \max\left(S_{\ell m} - (S_\ell \cup S_m)\right)$ when $\max\{\ell, m\}$ is sufficiently large. Let $\ell_1 < \ell_2 < \dots$ be a sequence of primes outside L. (The ℓ_i will be constructed in Section 7 with certain properties, but for now these properties are not relevant.) The plan will be to force $\ell_i P \in E'(\mathbb{Z}[S^{-1}])$ for all i, by requiring each S_{ℓ_i} to be contained in S. On the other hand, we must require other primes to lie outside S to make sure that not too many other multiples of P end up in $E'(\mathbb{Z}[S^{-1}])$.

Let $T_1 = S_{\text{bad}} \cup \bigcup_{i \ge 1} S_{\ell_i}$. Let T_2^a be the set of p_ℓ for all primes $\ell \notin \{\ell_1, \ell_2, \dots\}$. If ℓ_1 is sufficiently large, we may define $T_2^b = \{\, p_{\ell_i \ell_j} : 1 \le j \le i \,\}$ and $T_2^c = \{\, p_{\ell \ell_i} : \ell \in L, i \ge 1 \,\}$. Finally, let $T_2 = T_2^a \cup T_2^b \cup T_2^c$.

5. Properties of T_1 and T_2

Proposition 5.1. *The sets T_1 and T_2 are disjoint.*

Proof. By definition of p_ℓ and $p_{\ell m}$, $S_{\text{bad}} \cap T_2 = \emptyset$. If $\ell \neq \ell_i$, then $(\ell^{a_\ell}, \ell_i) = 1$, so Corollary 3.3 implies $p_\ell \notin S_{\ell_i}$. Thus $T_1 \cap T_2^a = \emptyset$. If $i \notin \{j,k\}$, then Corollary 3.3 implies $p_{\ell_j \ell_k} \notin S_{\ell_i}$, while if $i \in \{j,k\}$, then $p_{\ell_j\ell_k} \notin S_{\ell_i}$ by definition of $p_{\ell_j\ell_k}$. Thus $T_1 \cap T_2^b = \emptyset$. If $i \neq j$, then Corollary 3.3 implies $p_{\ell\ell_j} \notin S_{\ell_i}$, while if $i = j$, then $p_{\ell\ell_i} \notin S_{\ell_i}$ by definition of $p_{\ell_j\ell_k}$. Thus $T_1 \cap T_2^c = \emptyset$. □

Proposition 5.2. *If S contains T_1 and is disjoint from T_2, then $E'(\mathbb{Z}[S^{-1}])$ is the union of $\{\, \pm\ell_i P : i \geq 1 \,\}$ and some subset of the finite set $\left\{\, rP : r \mid \prod_{\ell \in L} \ell^{a_\ell - 1} \,\right\}$.*

Proof. Because the equation of E relates the x- and y-coordinates, a point nP belongs to $E'(\mathbb{Z}[S^{-1}])$ if and only if $S_n \subseteq S$. In particular, $S_{\ell_i} \subseteq T_1 \subseteq S$, so $\pm\ell_i P \in E'(\mathbb{Z}[S^{-1}])$.

Any point outside

$$\{\, \pm\ell_i P : i \geq 1 \,\} \cup \left\{\, rP : r \mid \prod_{\ell \in L} \ell^{a_\ell - 1} \,\right\}$$

is nP for some n divisible by one of the following:

- ℓ^{a_ℓ} for some ℓ not in the sequence $\ell_1, \ell_2, \ldots$,
- $\ell_i \ell_j$ for some $1 \leq j \leq i$, or
- $\ell\ell_i$ for some $\ell \in L$ and $i \geq 1$.

Lemma 3.1(a) implies then that S_n contains a prime of T_2^a, T_2^b, or T_2^c, respectively, so $S_n \not\subseteq S$. □

6. Natural density

The *natural density* of a subset $T \subseteq \mathcal{P}$ is defined as

$$\lim_{X \to \infty} \frac{\#\{p \in T : p \leq X\}}{\#\{p \in \mathcal{P} : p \leq X\}},$$

if the limit exists. One defines *upper natural density* similarly, using lim sup instead of lim.

Lemma 6.1. *If $\alpha \in \mathbb{R} - \mathbb{Q}$, then $\{\, \ell\alpha \bmod 1 : \ell \text{ is prime} \,\}$ is equidistributed in $[0,1]$. That is, for any interval $I \subseteq [0,1]$, the set of primes ℓ for which $(\ell\alpha \bmod 1)$ belongs to I has natural density equal to the length of I.*

Proof. See p. 180 of [Vin54]. □

Let $y(\ell P) \in \mathbb{Q}$ denote the y-coordinate of $\ell P \in E(\mathbb{Q})$.

Corollary 6.2. *If $I \subseteq \mathbb{R}$ is an interval with nonempty interior, then the set of primes ℓ for which $y(\ell P) \in I$ has positive natural density.*

Proof. Since $E(\mathbb{R})$ is a connected compact 1-dimensional Lie group over $\mathbb{R}$, we can choose an isomorphism $E(\mathbb{R}) \to \mathbb{R}/\mathbb{Z}$ as topological groups. Since P is of infinite order, its image in $\mathbb{R}/\mathbb{Z}$ is represented by an irrational number. The subset of $E(\mathbb{R})$ having y-coordinate in I corresponds to a nontrivial interval in $\mathbb{R}/\mathbb{Z}$. Now apply Lemma 6.1. □

7. Construction of the ℓ_i

For prime ℓ, define

$$\mu_\ell = \sup_{X \in \mathbb{Z}_{\geq 2}} \frac{\#\{p \in S_\ell : p \leq X\}}{\#\{p \in \mathcal{P} : p \leq X\}}.$$

The supremum is attained for some $X \leq \max S_\ell$, so μ_ℓ is computable for each ℓ.

Lemma 7.1. *For any $\epsilon > 0$, the natural density of $\{\, \ell : \mu_\ell > \epsilon \,\}$ is 0.*

Proof. For $X \in \mathbb{R}$, let $\pi(X) := \#\{p \in \mathcal{P} : p \leq X\}$. If ℓ is a prime and $\mu_\ell > \epsilon$, then we can choose $X_\ell \in \mathbb{Z}_{\geq 2}$ such that

$$\frac{\#\{p \in S_\ell : p \leq X_\ell\}}{\pi(X_\ell)} > \epsilon.$$

For $M \in \mathbb{Z}_{\geq 2}$, let U_M be the set of primes ℓ such that $\mu_\ell > \epsilon$ and $X_\ell \in [M, 2M)$. If $\ell \in U_M$, then

$$\#\{p \in S_\ell : p \leq 2M\} \quad \geq \quad \#\{p \in S_\ell : p \leq X_\ell\} \quad > \quad \epsilon\pi(X_\ell) \quad \geq \quad \epsilon\pi(M).$$

But the S_ℓ are disjoint by Corollary 3.3, so

$$\pi(2M) \quad \geq \quad \sum_{\ell \in U_M} \#\{p \in S_\ell : p \leq 2M\} \quad \geq \quad \epsilon\pi(M)\#U_M.$$

Thus by the Prime Number Theorem, $\#U_M = O(1)$ as $M \to \infty$. If $2^{k-1} \leq N < 2^k$, then

$$\#\{\, \ell : \mu_\ell > \epsilon \text{ and } X_\ell \leq N \,\} \leq \#U_2 + \#U_4 + \#U_8 + \cdots + \#U_{2^k} = O(k) = O(\log N)$$

as $N \to \infty$. If $\mu_\ell > \epsilon$, then by definition of X_ℓ,

$$\pi(X_\ell) < \frac{\#S_\ell}{\epsilon} \leq \frac{\log_2 d_\ell}{\epsilon} = O(\ell^2)$$

as $\ell \to \infty$ by Lemma 3.1(b), so $X_\ell = O(\ell^2 \log \ell)$ by the Prime Number Theorem. Combining the previous two sentences shows that

$$\begin{aligned} \#\{\, \ell \leq Y : \mu_\ell > \epsilon \,\} &= \#\{\, \ell \leq Y : \mu_\ell > \epsilon \text{ and } X_\ell \leq O(Y^2 \log Y) \,\} \\ &= O(\log O(Y^2 \log Y)), \end{aligned}$$

which is $o(\pi(Y))$ as $Y \to \infty$. $\square$

Define the ℓ_i inductively as follows. Given $\ell_1, \ldots, \ell_{i-1}$, let ℓ_i be the smallest prime outside L such that all of the following hold:

(1) $\ell_i > \ell_j$ for all $j < i$,
(2) $\mu_{\ell_i} \leq 2^{-i}$,
(3) $p_{\ell_i \ell_j} > 2^i$ for all $j \leq i$,
(4) $p_{\ell \ell_i} > 2^i$ for all $\ell \in L$, and
(5) $|y(\ell_i P) - i| \leq 1/(10i)$.

Proposition 7.2. *The sequence $\ell_1, \ell_2, \ldots$ is well defined and computable.*

Proof. By induction, we need only show that for each i, there exists ℓ_i as above. By Corollary 6.2, the set of primes satisfying (5) has positive natural density. By Lemma 7.1, (2) fails for a set of natural density 0. Therefore it will suffice to show that (1), (3), and (4) are satisfied by all sufficiently large ℓ_i.

For fixed $j \le i$, the primes $p_{\ell_i \ell_j}$ for varying values of ℓ_i are distinct by Corollary 3.3, so eventually they are greater than 2^i. The same holds for $p_{\ell \ell_i}$ for fixed $\ell \in L$. Thus by taking ℓ_i sufficiently large, we can make all the $p_{\ell_i \ell_j}$ and $p_{\ell \ell_i}$ greater than 2^i.

Each ℓ_i can be computed by searching primes in increasing order until one is found satisfying the conditions. □

8. Recursiveness of T_1 and T_2

The set $\{\ell_1, \ell_2, \dots\}$ is recursive, since it is a strictly increasing sequence whose terms can be computed in order. This is needed for the proofs in this section.

Proposition 8.1. *The set T_1 is recursive.*

Proof. Since S_{bad} is finite, it suffices to give an algorithm for deciding whether a prime $p \notin S_{\text{bad}}$ belongs to $\bigcup_{i \ge 1} S_{\ell_i}$. We have $p \in \bigcup_{i \ge 1} S_{\ell_i}$ if and only if $p \mid d_{\ell_i}$ for some i, which holds if and only if the order n_p of the image of P in $E(\mathbb{F}_p)$ divides ℓ_i for some i. The order n_p can be computed, and $n_p \ne 1$, since $P \in E'(\mathbb{Z}[S_{\text{bad}}^{-1}])$. So we simply check whether $n_p \in \{\ell_1, \ell_2, \dots\}$. □

Lemma 8.2. *If ℓ is prime, then $\ell \mid \#E(\mathbb{F}_{p_\ell})$.*

Proof. By definition of p_ℓ, the point $\ell^{a_\ell} P$ reduces to 0 in $E(\mathbb{F}_{p_\ell})$ but $\ell^{a_\ell - 1} P$ does not. □

Proposition 8.3. *The set T_2^a is recursive.*

Proof. If $p \in T_2^a$, then $p = p_\ell$ for some $\ell \notin \{\ell_1, \ell_2, \dots\}$, and then $\ell \mid \#E(\mathbb{F}_p)$ by Lemma 8.2. Therefore to test whether a prime $p \notin S_{\text{bad}}$ belongs to T_2^a, compute $\#E(\mathbb{F}_p)$ and its prime factors: one has $p \in T_2^a$ if and only if there is a prime factor ℓ such that $\ell \notin \{\ell_1, \ell_2, \dots\}$ and $p_\ell = p$. □

Proposition 8.4. *The sets T_2^b and T_2^c are recursive.*

Proof. By condition (3) in the definition of ℓ_i, if a prime p belongs to T_2^b, it must equal $p_{\ell_i \ell_j}$ for some $1 \le j \le i$ with $2^i < p$. Thus to test whether a prime p belongs to T_2^b, simply compute $p_{\ell_i \ell_j}$ for $1 \le j \le i < \log_2 p$.

The proof that T_2^c is recursive is similar, using condition (4). □

Thus T_1 and T_2 are recursive.

9. The densities of T_1 and T_2

Proposition 9.1. *The set T_1 has natural density* 0.

Proof. For fixed $r \in \mathbb{Z}_{>0}$, the set $\bigcup_{i>r} S_{\ell_i}$ differs from T_1 in only finitely many primes, so it suffices to show that the former has upper natural density tending to 0 as $r \to \infty$. By definition of μ_{ℓ_i}, the upper natural density is bounded by $\sum_{i>r} \mu_{\ell_i} \le \sum_{i>r} 2^{-i} = 2^{-r}$, which tends to 0 as $r \to \infty$. □

Proposition 9.2. *The sets T_2^b and T_2^c have natural density* 0.

Proof. Suppose $2^m \le X < 2^{m+1}$. By condition (3) defining ℓ_i, the only primes of the form $p_{\ell_i \ell_j}$ that might be $\le X$ are those with $1 \le j \le i \le m$. There are at most $O(m^2) = O((\log X)^2)$ of these, which is negligible compared to $\pi(X)$. Thus T_2^b has natural density 0. The proof for T_2^c is similar. □

The rest of this section is devoted to the proof that T_2^a has natural density 0. Recall that T_2^a consists of primes of the form p_ℓ. If the sequence of p_ℓ grew faster than the sequence of primes ℓ, then T_2^a would have density 0. But Lemma 8.2 implies only that p_ℓ is at least about the size of ℓ. The strategy for strengthening this bound will be to show that numbers of the form $\#E(\mathbb{F}_p)$ are typically divisible by many primes. For $n \in \mathbb{Z}_{>0}$, let $\omega(n)$ be the number of distinct prime factors of n.

Lemma 9.3. *For any $t \geq 1$, the natural density of $\{\, p : \omega(\#E(\mathbb{F}_p)) < t \,\}$ is* 0.

Proof. For a prime ℓ, let $E[\ell]$ denote the group of points of order dividing ℓ on E. Then $\ell \mid \#E(\mathbb{F}_p)$ if and only if the image of the Frobenius element at p under $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \operatorname{Aut} E[\ell]$ has a nonzero fixed vector. Since E does not have complex multiplication, the image of $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ in $\prod_\ell \operatorname{Aut} E[\ell]$ is open. (This follows from [Ser72].) Thus $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \prod_{\ell \notin L'} \operatorname{Aut} E[\ell]$ is surjective for some finite $L' \subseteq \mathcal{P}$. A calculation shows that the fraction of elements of $\operatorname{Aut} E[\ell] \simeq \operatorname{GL}_2(\mathbb{F}_\ell)$ having a nonzero fixed vector is

$$\frac{\ell^3 - 2\ell}{(\ell^2-1)(\ell^2-\ell)} = \frac{1}{\ell} + O\left(\frac{1}{\ell^2}\right).$$

The sum of this over ℓ diverges, so as $C \to \infty$, the fraction of elements of $\prod_{\ell < C, \ell \notin L'} \operatorname{Aut} E[\ell]$ having fewer than t components with a nonzero fixed vector tends to 0. Applying the Chebotarev Density Theorem (see Théorème 1 of [Ser81] for a version using natural density) and letting $C \to \infty$, we obtain the result. □

Proposition 9.4. *The set T_2^a has natural density* 0.

Proof. Because of Lemma 9.3, it suffices to show that the upper natural density of

$$T_2^{a,t} := \{\, p \in T_2^a : \omega(\#E(\mathbb{F}_p)) \geq t \,\}$$

tends to 0 as $t \to \infty$.

Suppose $p = p_\ell \in T_2^{a,t}$. By Lemma 8.2, $\ell \mid \#E(\mathbb{F}_p)$. By definition of $T_2^{a,t}$, the integer $\#E(\mathbb{F}_p)$ is divisible by at least $t-1$ other primes, so $2^{t-1}\ell \leq \#E(\mathbb{F}_p)$. There exists a degree-2 map $E \to \mathbb{P}^1$ over $\mathbb{F}_p$, so $\#E(\mathbb{F}_p) \leq 2(p+1) \leq 4p$. Combining the previous two sentences yields $\ell \leq 2^{3-t}p$. Since every element of $T_2^{a,t}$ is p_ℓ for some ℓ, we have

$$\#\{\, p \in T_2^{a,t} : p \leq X \,\} \leq \pi(2^{3-t}X) = (2^{3-t} + o(1))\,\pi(X)$$

as $X \to \infty$. Thus by definition, the upper natural density of $T_2^{a,t}$ is at most 2^{3-t}. This goes to 0 as $t \to \infty$. □

Thus T_1 and T_2 have natural density 0.

10. Proof of Theorem 1.3

By Proposition 5.2, $E'(\mathbb{Z}[S^{-1}])$ differs from $\{\, \pm\ell_i P : i \geq 1 \,\}$ by at most a finite set. Since $y(\pm\ell_i P)$ is within $1/10$ of $\pm i$, any bounded subset of $\mathbb{R}^2$ contains at most finitely many points of $E'(\mathbb{Z}[S^{-1}])$. Part (1) of Theorem 1.3 follows.

We next construct a diophantine model A of the positive integers over $\mathbb{Z}[S^{-1}]$. The set of nonzero elements of $\mathbb{Z}[S^{-1}]$ is diophantine (see Theorem 4.2 of [Shl94]), and we can represent elements of $\mathbb{Q}$ as fractions of elements of $\mathbb{Z}[S^{-1}]$ with nonzero denominator. Therefore equations over $\mathbb{Q}$ can be rewritten as systems of equations over $\mathbb{Z}[S^{-1}]$, and there is no harm in using them in our diophantine definitions. In particular, we may use the predicate $x \geq y$, since it can be encoded as $(\exists z_1, z_2, z_3, z_4 \in \mathbb{Q})(x = y + z_1^2 + z_2^2 + z_3^2 + z_4^2)$.

For $i \in \mathbb{Z}_{>0}$, define $y_i := y(\ell_i P)$. Let $A = \{y_1, y_2, \dots\}$. Then A is diophantine over $\mathbb{Z}[S^{-1}]$, because it consists of the nonnegative elements of the set of y-coordinates of $E'(\mathbb{Z}[S^{-1}])$ minus a finite set. We have a bijection $\mathbb{Z}_{>0} \to A$ taking i to y_i.

It remains to show that the graphs of addition and multiplication on $\mathbb{Z}_{>0}$ correspond to diophantine subsets of A^3. We know $|y_i - i| \leq 1/(10i) \leq 1/10$, so the idea is that the addition on $\mathbb{Z}_{>0}$ should correspond to the operation of adding elements of A and then rounding to the nearest element of A. A similar idea will work for squaring, and we will get multiplication from addition and squaring.

Lemma 10.1. *Let $m, n, q \in \mathbb{Z}_{>0}$. Then*

(1) *$m + n = q$ if and only if $|y_m + y_n - y_q| \leq 3/10$.*
(2) *$m^2 = n$ if and only if $|y_m^2 - y_n| \leq 4/10$.*

Proof.

(1) The quantity $y_m + y_n - y_q$ differs from the integer $m + n - q$ by at most $1/10 + 1/10 + 1/10$.

(2) The quantity $y_m^2 - y_n$ differs from the integer $m^2 - n$ by at most

$$|y_m^2 - m^2| + |y_n - n| \leq \left| \left(m + \frac{1}{10m} \right)^2 - m^2 \right| + \frac{1}{10} \leq \frac{4}{10}. \qquad \square$$

Lemma 10.1 shows that the two predicates $m + n = q$ and $m^2 = n$ on $\mathbb{Z}_{>0}$ correspond to diophantine predicates on A. Building with these, we can show the same for $mn = q$, since

$$mn = q \quad \iff \quad (m + n)^2 = m^2 + n^2 + q + q.$$

This completes the proof of part (2) of Theorem 1.3. Part (3) follows from (2) and Matijasevič's Theorem.

Acknowledgements

I thank Thanases Pheidas and the referee for helpful comments, in particular for simplifying the diophantine definition of multiplication at the end. I thank also Gunther Cornelissen and Graham Everest for suggesting references for Section 3 and Alexandra Shlapentokh for suggesting some improvements in the exposition.

References

[Aya92] Mohamed Ayad, *Points S-entiers des courbes elliptiques*, Manuscripta Math. **76** (1992), no. 3-4, 305–324. MR **93i:**11064

[CZ00] Gunther Cornelissen and Karim Zahidi, *Topology of Diophantine sets: remarks on Mazur's conjectures*, Hilbert's tenth problem: relations with arithmetic and algebraic geometry (Ghent, 1999), Amer. Math. Soc., Providence, RI, 2000, pp. 253–260. MR **2001m:**11217

[DLPVG00] Jan Denef, Leonard Lipshitz, Thanases Pheidas, and Jan Van Geel (eds.), *Hilbert's tenth problem: relations with arithmetic and algebraic geometry*, American Mathematical Society, Providence, RI, 2000, Papers from the workshop held at Ghent University, Ghent, November 2–5, 1999. MR **2001g:**00018

[DPR61] Martin Davis, Hilary Putnam, and Julia Robinson, *The decision problem for exponential diophantine equations*, Ann. of Math. (2) **74** (1961), 425–436. MR **24:**A3061

[Eve02] Graham Everest, *Zsigmondy's theorem for elliptic curves*, preprint, 11 October 2002.

[KR92] Ki Hang Kim and Fred W. Roush, *An approach to rational Diophantine undecidability*, Proceedings of Asian Mathematical Conference, 1990 (Hong Kong, 1990) (River Edge, NJ), World Sci. Publishing, 1992, pp. 242–248.

[Mat70] Yuri V. Matijasevič, *The Diophantineness of enumerable sets*, Dokl. Akad. Nauk SSSR **191** (1970), 279–282. MR **41:**3390

[Maz92] Barry Mazur, *The topology of rational points*, Experiment. Math. **1** (1992), no. 1, 35–45. MR **93j:**14020

[Maz95] Barry Mazur, *Speculations about the topology of rational points: an update*, Astérisque (1995), no. 228, 4, 165–182, Columbia University Number Theory Seminar (New York, 1992). MR **96c:**11068

[Ser72] Jean-Pierre Serre, *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques*, Invent. Math. **15** (1972), no. 4, 259–331. MR **52:**8126

[Ser73] Jean-Pierre Serre, *A course in arithmetic*, Springer-Verlag, New York, 1973, Translated from the French, Graduate Texts in Mathematics, No. 7. MR **49:**8956

[Ser81] Jean-Pierre Serre, *Quelques applications du théorème de densité de Chebotarev*, Inst. Hautes Études Sci. Publ. Math. (1981), no. 54, 323–401. MR **83k:**12011

[Ser97] Jean-Pierre Serre, *Lectures on the Mordell-Weil theorem*, third ed., Friedr. Vieweg & Sohn, Braunschweig, 1997, Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt, With a foreword by Brown and Serre. MR **2000m:**11049

[Shl94] Alexandra Shlapentokh, *Diophantine classes of holomorphy rings of global fields*, J. Algebra **169** (1994), no. 1, 139–175. MR **95h:**12007

[Shl97] Alexandra Shlapentokh, *Diophantine definability over some rings of algebraic numbers with infinite number of primes allowed in the denominator*, Invent. Math. **129** (1997), no. 3, 489–507. MR **98h:**11163

[Shl00] Alexandra Shlapentokh, *Defining integrality at prime sets of high density in number fields*, Duke Math. J. **101** (2000), no. 1, 117–134. MR **2001a:**11200

[Shl02] Alexandra Shlapentokh, *Diophantine definability and decidability in large subrings of totally real number fields and their totally complex extensions of degree 2*, J. Number Theory **95** (2002), no. 2, 227–252. MR **2003h:**03068

[Shl03] Alexandra Shlapentokh, *A ring version of Mazur's conjecture on topology of rational points*, Internat. Math. Res. Notices (2003), no. 7, 411–422.

[Sil88] Joseph H. Silverman, *Wieferich's criterion and the abc-conjecture*, J. Number Theory **30** (1988), no. 2, 226–237. MR **89m:**11027

[Sil92] Joseph H. Silverman, *The arithmetic of elliptic curves*, Springer-Verlag, New York, 1992, Corrected reprint of the 1986 original. MR **95m:**11054

[Vin54] I. M. Vinogradov, *The method of trigonometrical sums in the theory of numbers*, Interscience Publishers, London and New York, 1954, Translated, revised and annotated by K. F. Roth and Anne Davenport. MR **15:**941b

[War48] Morgan Ward, *Memoir on elliptic divisibility sequences*, Amer. J. Math. **70** (1948), 31–74. MR **9:**332j

Department of Mathematics, University of California, Berkeley, California 94720-3840

E-mail address: poonen@math.berkeley.edu