Diophantine m-tuples for primes

Andrej Dujella and Florian Luca

Abstract

In this paper, we show that if p is a prime and if $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$ is a set of positive integers with the property that $a_i a_j + p$ is a perfect square for all $1 \leq i < j \leq m$, then $m < 3 \cdot 2^{168}$. More generally, when p is replaced by a squarefree integer n, the inequality $m \leq f(\omega(n))$ holds with some function f, where $\omega(n)$ is the number of prime divisors of n. We also give upper bounds for m when p is replaced by an arbitrary integer which hold on a set of n of asymptotic density one.

1 Introduction

Let n be any nonzero integer. A *Diophantine m-tuple with the property $D(n)$* is a set of m positive integers $\{a_1, \ldots, a_m\}$ such that $a_i a_j + n$ is a perfect square for all $1 \leq i < j \leq m$. Diophantus found the quadruple $\{1, 33, 68, 105\}$ with the property $D(256)$, while the first Diophantine quadruple with the property $D(1)$, the set $\{1, 3, 8, 120\}$, was found by Fermat (see [4]). Baker and Davenport [1] proved that this Fermat set cannot be extended to a Diophantine quintuple. The first author proved recently that there does not exist a Diophantine sextuple with the property $D(1)$, and that there are only finitely many such quintuples (see [7]). On the other hand, there are examples of Diophantine sextuples, e.g. $\{99, 315, 9920, 32768, 44460, 19534284\}$ with the property $D(2985984)$, found by Gibbs [9].

The question is what can be said about the size of sets with the property $D(n)$ for $n \neq 1$. Let

$$M_n = \sup\{|\mathcal{S}| : \mathcal{S} \text{ has the property } D(n)\}.$$

[0] *2000 Mathematics Subject Classification* 11D45.

Considering congruences modulo 4, it is easy to see that $M_n = 3$ if $n \equiv 2 \pmod 4$ (see [2]). On the other hand, if n is not congruent to 2 modulo 4 and $n \notin \{-4, -3, -1, 3, 5, 8, 12, 20\}$, then $M_n \geq 4$ (see [5]).

Since the number of integer points on the elliptic curve

$$y^2 = (a_1x + n)(a_2x + n)(a_3x + n)$$

is finite, we conclude that there does not exist an infinite set with the property $D(n)$. Furthermore, the hyperelliptic curve

$$y^2 = (a_1x + n)(a_2x + n)(a_3x + n)(a_4x + n)(a_5x + n)$$

has genus $g = 2$. Caporaso, Harris and Mazur [3], proved that the Lang conjecture on varieties of general type implies that for $g \geq 2$ the number $B(g, \mathbb{K}) = \max_{\mathcal{C}} |\mathcal{C}(\mathbb{K})|$ is finite. Here, $\mathcal{C}$ runs over all curves of genus g over a number field $\mathbb{K}$, and $\mathcal{C}(\mathbb{K})$ denotes the set of all $\mathbb{K}$-rational points on $\mathcal{C}$. However, even the question whether $B(2, \mathbb{Q}) < \infty$ is still open. Since $M_n \leq 5 + B(2, \mathbb{Q})$ (by [11], we also have $M_n \leq 4 + B(4, \mathbb{Q})$), we see that the Lang conjecture implies that there exist an absolute constant C such that $M_n \leq C$ for all nonzero integers n. However, at present, the best known upper bound for M_n has the form $M_n \leq C \log |n|$ (see [6, 8]).

The main result of this paper consists in an absolute upper bound for the size of sets with properties $D(p)$ and $D(-p)$, where p is a prime.

Throughout the paper, the letter p will always denote a prime number. For a nonzero integer n, we write $\omega(n)$ and $P(n)$ for the number of prime divisors and the largest prime factor of n, respectively, with the convention that $P(\pm 1) = 1$. As usual, $\pi(x)$ denotes the number of primes $p \leq x$. We use the Vinogradov symbols $\ll$ and $\gg$, as well as the Landau symbols O and o, with their usual meanings.

Acknowledgments. The authors would like to thank the referee for valuable comments. This paper was written during a visit of the second author at the University of Zagreb in October of 2004. He warmly thanks this University for its hospitality. Both authors were partly supported by the Croatian Ministry of Science, Education and Sport Grant 0037110.

2 Results

The first result of this paper is an absolute upper bound on the size m of a Diophantine m-uple with the property $D(\pm p)$ which holds for all primes p.

Theorem 2.1. *There exists an absolute constant C such that any Diophantine m-tuple with the property $D(p)$ or $D(-p)$, where p is a prime, satisfies $m < C$. Furthermore, C can be chosen to be $3 \cdot 2^{168}$.*

We next give a more general result, namely an upper bound on the size m of a Diophantine m-tuple with the property $D(\pm n)$, where n is squarefree, which depends only on the number of prime divisors $\omega(n)$ of the positive integer n.

Theorem 2.2. *There exists a function $f : \mathbb{N} \longrightarrow \mathbb{N}$ such that if n is any squarefree positive integer, then any Diophantine m-tuple with the property $D(n)$ or $D(-n)$ satisfies $m < f(\omega(n))$, where $\omega(n)$ is the number of distinct prime divisors of n.*

We finally present an upper bound on m which is valid for most positive integers n. Let us call a Diophantine m-tuple $\mathcal{A}$ with the property $D(n)$ *reduced* if $\gcd(a, n) = 1$ holds for all $a \in \mathcal{A}$.

Theorem 2.3. *For every $\varepsilon > 0$, the set of positive integers n with the property that there exists a Diophantine m-tuple with the property $D(n)$ or $D(-n)$ and with $m > (1 + \varepsilon) \log \log n$, is of asymptotic density zero. Furthermore, if only reduced Diophantine m-tuples are considered, then the same result holds with $(1+\varepsilon) \log \log n$ replaced by any increasing function $f(n)$ such that $\lim_{x \to \infty} f(x) = +\infty$.*

3 The proof of Theorem 2.1

We shall analyze in detail the case of the Diophantine tuples with the property $D(p)$, and we shall only point out the minor differences in the argument for the case of the Diophantine tuples with the property $D(-p)$.

We start with a short outline of the methods used in the proof. In Section 3.1 we show that in order to prove Theorem 2.1, it suffices to establish a gap principle of the form $a_{i+\ell} > p^{\gamma} a_i$ for the elements of a Diophantine tuple.

In the first step, in Section 3.2, a factorization is assigned to any triple with the property $D(p)$. A combination of p-adic and archimedean estimates suffice to determine that the second factor appearing in the factorization should be divisible by p^2, since divisibility of the first factor leads to a quick conclusion of the proof.

In the second step, we obtain in Lemma 3.1 a congruence condition associated with solutions and deduce from this a set of new polynomial congruences. By elimination of variables, one gets a new polynomial of bounded height which must satisfy a congruence modulo p for rather small values of the variables if the gap principle we want is not fulfilled. The smallness of the variables yields the key fact that the congruence in question must be in fact an equation.

The final step consists in deducing from this the existence of a polynomial of bounded degree M in five variables which vanishes on a set which is a cartesian product $\mathcal{M}^5$, with $\mathcal{M}$ a set of cardinality larger than M, which is impossible. This establishes the gap principle.

3.1 Reductions of the original problem

Let $\mathcal{A} = \{a_1, \ldots, a_m\}$ be a Diophantine m-tuple with the property $D(p)$. Since the main result from [8] shows that if $\mathcal{A}$ is any Diophantine m-tuple with the property $D(n)$ (n any nonzero integer), then $m \leq 16 \log |n|$ for $|n| > 400$, and $m \leq 31$ for $|n| \leq 400$, it follows that from now on we may assume that $p > 2^{2^{76}}$, and that $C > 2^{80} \log 2$.

We assume that $a_1 < a_2 < \cdots < a_m$. Furthermore, whenever we write $a_i a_j + p = x_{ij}^2$, we make the convention that $x_{ij} > 0$. We note that there exists at most one element of $\mathcal{A}$ which is a multiple of p. Indeed, for if two such elements, say a_i and a_j exist, then reducing the equation $a_i a_j + p = x_{ij}^2$ modulo p^2 we get $p \equiv x_{ij}^2 \pmod{p^2}$, which is a contradiction. Eliminating such an element from $\mathcal{A}$, it follows that we may assume that p does not divide a for any $a \in \mathcal{A}$. We now note that $a_3 > p^{1/4}$. Indeed, note that if we write

$$a_1 a_3 + p = x_{13}^2 \qquad \text{and} \qquad a_2 a_3 + p = x_{23}^2,$$

then $x_{23} > x_{13} > p^{1/2}$. In particular,

$$a_3^2 > a_3(a_2 - a_1) = x_{23}^2 - x_{13}^2 = (x_{23} - x_{13})(x_{23} + x_{13}) \geq 2p^{1/2}.$$

Hence, $a_3 > p^{1/4}$. When p is replaced by $-p$, we then have

$$a_2^2 > a_1 a_2 = p + x_{12}^2 > p,$$

therefore $a_2 > p^{1/2}$, which is even a better inequality. Thus, eliminating the smallest two elements, if needed, we may assume that $a > p^{1/4}$ holds for all

$a \in \mathcal{A}$. Furthermore, a result from [6] shows that there exist at most 21 elements $a \in \mathcal{A}$ such that $a > p^3$. Eliminating those elements too, we may assume that $p^{1/4} < a < p^3$ holds for all $a \in \mathcal{A}$. We now note that if there exist constants γ and ℓ such that the inequality $a_{i+\ell} > p^\gamma a_i$ holds for all $i \in \{1, \ldots, m-\ell\}$, then, by induction on i, the inequality

$$a_i > p^{1/4+\lfloor i/\ell \rfloor \gamma}$$

holds for $i = 1, \ldots, m$. Since $a_m < p^3$, we get the inequality

$$\left\lfloor \frac{m}{\ell} \right\rfloor \gamma + \frac{1}{4} < 3.$$

Hence,

$$m < \frac{(11+4\gamma)\ell}{4\gamma}. \tag{1}$$

The above argument together with the fact that we first eliminated at most 24 elements from $\mathcal{A}$, shows that one may take

$$C = \max\left\{ \left\lceil 2^{80} \log 2 \right\rceil, \frac{(11+4\gamma)\ell}{4\gamma} + 24 \right\}. \tag{2}$$

Thus, it suffices to find such constants γ and ℓ.

3.2 A factorization

To any triple with the property $D(p)$, a factorization can be assigned. We will show that p^2 divides exactly one of the factors appearing in the factorization. The case when the smaller factor is divisible by p^2 leads easily to desired inequality of the form $a_{i+\ell} > p^\gamma a_i$. The case when the larger factor is divisible by p^2 is much more involved, and will be considered later.

Let $a < b < c$ be any three elements in $\mathcal{A}$. We write

$$ab + p = x^2, \qquad bc + p = y^2 \qquad \text{and} \qquad ac + p = z^2,$$

where x, y and z are positive integers. Then,

$$\begin{aligned}
(xyz)^2 &= (ab+p)(bc+p)(ac+p) \\
&= (abc)^2 + pabc(a+b+c) + p^2(ab+bc+ac) + p^3 \\
&= (abc + p(a+b+c)/2)^2 \\
&\quad + p^2\left(ab+ac+bc-(a+b+c)^2/4\right) + p^3,
\end{aligned}$$

therefore

$$\begin{aligned}
&(2xyz - 2abc - p(a+b+c))\,(2xyz + 2abc + p(a+b+c)) \\
&\quad = p^2\left(2ab + 2bc + 2ac - a^2 - b^2 - c^2 + 4p\right) \\
&\quad = p^2\left(4(ab+p) - (a+b-c)^2\right) \\
&\quad = p^2\,(2x - a - b + c)\,(2x + a + b - c)\,. \qquad (3)
\end{aligned}$$

We note that if both sides of the above expression are zero then $c = a + b + 2x$, and x depends only on a and b. Thus, given $a < b$, there exists at most one value for $c > b$ such that both sides of the above equation vanish. (In the case $D(-p)$ there is an additional possibility for c; namely, $c = a + b - 2x$.) From now on, we assume that both sides of the above equation are nonzero. We further note that the greatest common divisor of the two factors appearing in the left hand side of the above identity (3) is not a multiple of p. Indeed, for if it were, then reducing these expressions modulo p we would get $2xyz - 2abc \equiv 0 \pmod{p}$ and $2xyz + 2abc \equiv 0 \pmod{p}$. Subtracting these congruences we get $4abc \equiv 0 \pmod{p}$, which is impossible because p is odd and p does not divide any member of $\mathcal{A}$. From the above remarks and equation (3), we conclude that p^2 divides one of the two factors from the left hand side of equation (3). If p^2 divides the smaller factor, then $|2xyz - 2abc - p(a+b+c)| \geq p^2$, therefore equation (3) implies that

$$\begin{aligned}
4abc &< 2xyz + 2abc + p(a+b+c) \\
&\leq |2ab + 2bc + 2ac - a^2 - b^2 - c^2 + 4p| \\
&< 6c^2 + 4p < 2\max\{6c^2, 4p\}. \qquad (4)
\end{aligned}$$

If $4p > 6c^2$, then the above expression is bounded by $8p$. However,

$$2xyz + 2abc + p(a+b+c) > 2xyz + p(a+b+c) > 3p^{5/4}.$$

Comparing these inequalities, we get $8p > 3p^{5/4}$, or $p^{1/4} < 8/3$, which contradicts our assumptions on p. Thus, $6c^2 > 4p$, and the above inequality becomes

$$4abc < 12c^2,$$

which leads to $c > ab/3 > p^{1/4}b/3 > p^{1/8}b$. The last inequality above follows from the fact that $p > 3^8$. When p is replaced by $-p$, the small factor is again considered to be

$$2xyz - (2abc - p(a+b+c)).$$

Furthermore, note that since $a > p^{1/2}$ in this case, the large factor is bounded below by

$$\begin{aligned} 2xyz + 2abc - p(a+b+c) &> 4abc - p(a+b+c) \\ &> 4abc - 3pc \\ &= abc\left(4 - \frac{3p}{ab}\right) \\ &> abc, \end{aligned}$$

because $p < ab$. Thus, the analogue of inequality (4) is now

$$abc < 2xyz + 2abc - p(a+b+c) < 2\max\{6c^2, 4p\} = 12c^2,$$

which leads to the same inequality $c > ab/12 > p^{1/2}b/12 > p^{1/8}b$, because $p > 12^{4/3}$.

We now let ℓ be some positive integer to be determined later and let $a_i < a_{i+1} < \cdots < a_{i+\ell}$ be a sequence of length $\ell + 1$ of consecutive elements of $\mathcal{A}$. We let $b_1 = a_i,\ b_2 = a_{i+1}$ and write $b_1b_2 + p = y_{12}^2$. If $b_1 + b_2 \pm 2y_{12} = a_j$ for some $j \in \{i+2, \ldots, i+\ell\}$, we then eliminate the element a_j from our sequence ($b_1 + b_2 - 2y_{12} = a_j$ is possible only in the $D(-p)$ case). Let $b_3 > b_2$ be the smallest element of this sequence (note that $b_3 = a_{i+2}$ or $b_3 = a_{i+3}$ or $b_3 = a_{i+4}$). We put $b_1b_3 + p = y_{13}^2$ and $b_2b_3 + p = y_{23}^2$. If either one of $b_1 + b_3 + 2y_{13}$ and $b_2 + b_3 + 2y_{23}$ belong to our sequence, we eliminate those elements too. Continuing in this way, if $\ell \geq 16$, we can then select $b_1 < b_2 < b_3 < b_4 < b_5$ such that $b_i + b_j \pm 2y_{ij} \neq b_k$ for any $i < j < k$ in $\{1, 2, \ldots, 5\}$. Note that $b_5 \leq a_{i+16}$ (in the case $D(p)$ we have $b_5 \leq a_{i+10}$). For each subset $\{i, j, k\}$ of three elements of $\{1, \ldots, 5\}$, we compute

$$2y_{ij}y_{ik}y_{jk} + \varepsilon(2b_ib_jb_k + p(b_i + b_j + b_k)), \qquad \text{where} \qquad \varepsilon \in \{\pm 1\}.$$

From the preceding remarks, none of these numbers is zero, and for each such subset $\{i, j, k\}$ there exists only one $\varepsilon \in \{\pm 1\}$ such that the above number is a multiple of p^2. If this value for ε equals -1 for at least one of these subsets, then, by the preceding argument, we have $a_{i+\ell} \geq a_{i+10} \geq b_k > p^{1/8}b_1 = p^{1/8}a_i$, which is the desired inequality with $\gamma = 1/8$ and $\ell = 16$.

3.3 Auxiliary polynomials and variable elimination

From now on, we assume that p^2 divides the above expression with $\varepsilon = 1$ for all such subsets $\{i, j, k\}$. As we have already mentioned, the analysis of this case is much more involved than the preceding one.

Using a lemma on congruence properties of Diophantine quadruples, we will assign to any quadruple with the property $D(p)$ an integer Λ. In that way, to any quintuple with the property $D(p)$ we can assign a quintuple of such integers $(\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5)$. The Elimination Theory will be used to construct a nonzero polynomial Q such that $Q(\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5) \equiv 0 \pmod{p}$. It is important that we have also bounds for the degree and height of Q. This will allow us to conclude that $Q(\Lambda_1, \Lambda_2, \Lambda_3, \Lambda_4, \Lambda_5) = 0$, under the assumption that all the Λ_i's are small.

Lemma 3.1. *Let $\mathcal{A} = \{c_1, c_2, c_3, c_4\}$ be a Diophantine quadruple with the property $D(p)$, where p is a prime. Let $c_i c_j + p = z_{ij}^2$ for $1 \leq i < j \leq 4$. Assume that p does not divide any of the c_i's and that the congruence*

$$z_{ij} z_{jk} z_{ik} \equiv -c_i c_j c_k \pmod{p}$$

holds for all subsets with three elements $\{i, j, k\} \subset \{1, \ldots, 4\}$. Then, $z_{ij} z_{k\ell} \equiv z_{i\ell} z_{jk} \pmod{p}$ holds for all permutations (i, j, k, ℓ) of $\{1, 2, 3, 4\}$. Here, we make the convention that $z_{ji} = z_{ij}$ if $i < j$. A similar statement holds for Diophantine quadruples with the property $D(-p)$.

Proof. We prove the lemma only for Diophantine quadruples with the property $D(p)$, since the proof for the case when p is replaced by $-p$ is entirely similar.

Since p does not divide any of the c_i's, it follows that p does not divide any of the z_{ij}'s either. We use the notation (a, b, c, d) for an arbitrary permutation of (c_1, c_2, c_3, c_4). We write

$$ab + p = x_1^2, \qquad ac + p = x_2^2, \qquad bc + p = x_3^2,$$

$$ad + p = x_4^2, \qquad bd + p = x_5^2, \qquad cd + p = x_6^2.$$

Of course, since this is an arbitrary permutation, it suffices to prove the particular congruence $x_1 x_6 \equiv x_2 x_5 \pmod{p}$. We multiply the congruences

$$x_1 x_2 x_3 \equiv -abc \pmod{p} \qquad \text{and} \qquad x_1 x_4 x_5 \equiv -abd \pmod{p},$$

and we obtain

$$x_1^2 x_2 x_3 x_4 x_5 \equiv a^2 b^2 cd \pmod{p}.$$

Since $x_1^2 \equiv ab \pmod{p}$, and all elements $a,\ b,\ c$ and d are invertible modulo p, we get $x_2x_3x_4x_5 \equiv abcd \pmod{p}$. In the same way, multiplying the congruences

$$x_1x_2x_3 \equiv -abc \pmod{p} \qquad \text{and} \qquad x_2x_4x_6 \equiv -acd \pmod{p},$$

we obtain

$$x_1x_2^2x_3x_4x_6 \equiv a^2bc^2d \pmod{p},$$

and since $x_2^2 \equiv ac \pmod{p}$, we arrive at the congruence $x_1x_3x_4x_6 \equiv abcd \pmod{p}$. Hence, $x_1x_3x_4x_6 \equiv x_2x_3x_4x_5 \pmod{p}$, leading to $x_1x_6 \equiv x_2x_5 \pmod{p}$. □

In what follows, we let $\{c_1, c_2, c_3, c_4\}$ be a Diophantine m-tuple with the property $D(p)$ satisfying the hypotheses of Lemma 3.1. There are precisely three sets consisting each of a pair of opposite sides in any quadrilateral of vertices $a,\ b,\ c$ and d (here, by side we mean any edge connecting two of the four points); namely, they are $\{ab, cd\},\ \{ac, bd\}$ and $\{ad, bc\}$. For each two of the above three subsets of pairs, say $\{t_1, t_2\}$ and $\{t_3, t_4\}$, where t_i are edges connecting two of the four points, we look at the expression $|(t_1 + p)(t_2 + p) - (t_3 + p)(t_4 + p)|$. We assume that $a < b < c < d$.

When $\{t_1, t_2\} = \{ab, cd\}$ and $\{t_3, t_4\} = \{ac, bd\}$, we get

$$\begin{aligned} p(d-a)(c-b) &= (ab+p)(cd+p) - (ac+p)(bd+p) \\ &= (x_1x_6)^2 - (x_2x_5)^2 \\ &= (x_1x_6 - x_2x_5)(x_1x_6 + x_2x_5). \end{aligned}$$

By Lemma 3.1, it follows that if we write

$$\lambda_1 = \frac{(x_1x_6 - x_2x_5)}{p}, \tag{5}$$

then λ_1 is a positive integer and

$$d^2 > (d-a)(c-b) = \lambda_1(x_1x_6 + x_2x_5) > 2\lambda_1(abcd)^{1/2} > \lambda_1 a^2.$$

Hence, $d > \lambda_1^{1/2}a$. In particular, if δ is some constant such that the inequality $\lambda_1 \geq p^\delta$ holds, then

$$a_{i+\ell} \geq d > \lambda_1^{1/2}a \geq p^{\delta/2}a, \tag{6}$$

which is the desired inequality with $\gamma = \min\{1/8, \delta/2\}$. From now on, we assume that $\lambda_1 < p^\delta$.

In the same manner, we conclude that we may assume that $\lambda_2 < p^\delta$ and $\lambda_3 < p^\delta$, where λ_2 and λ_3 are positive integers defined by

$$\lambda_2 = \frac{(x_1x_6 - x_3x_4)}{p}, \tag{7}$$

$$\lambda_3 = \frac{(x_2x_5 - x_3x_4)}{p}. \tag{8}$$

We now exploit the fact that all λ_i for $i = 1, 2, 3$ are small. Rewriting equation (5) as

$$x_1x_6 = x_2x_5 + p\lambda_1,$$

and squaring both sides of it, we get

$$(ab + p)(cd + p) = p^2\lambda_1^2 + 2p\lambda_1\sqrt{(ac + p)(bd + p)} + (ac + p)(bd + p).$$

Performing the obvious cancellations in both sides and simplifying by a factor of p, we arrive at

$$(d - a)(c - b) - p\lambda_1^2 = 2\lambda_1\sqrt{(ac + p)(bd + p)}.$$

Squaring both sides of the above relation and reducing the resulting equation modulo p, we get

$$(d - a)^2(c - b)^2 \equiv 4\lambda_1^2abcd \pmod{p}.$$

Performing the same manipulations with λ_2 and λ_3 and multiplying the three congruences obtained in this way, we get

$$((d - c)(d - b)(d - a)(c - b)(c - a)(b - a))^2 \equiv 2^6\Lambda(abcd)^3 \pmod{p},$$

where $\Lambda = (\lambda_1\lambda_2\lambda_3)^2$. We now write $P = P(X, Y, Z, T, \Lambda)$ for the polynomial in $\mathbb{Z}[X, Y, Z, T, \Lambda]$ given by

$$P = ((T - Z)(T - Y)(T - X)(Z - Y)(Z - X)(Y - X))^2 - 2^6\Lambda(XYZT)^3. \tag{9}$$

Note that the above polynomial, as a polynomial in X, Y, Z and T with coefficients in $\mathbb{Z}[\Lambda]$, is homogeneous, symmetric, and of degree $D = 12$.

We now return to our instance in which $\{b_1, \dots, b_5\}$ is a subset of cardinality five of $\mathcal{A}$ such that every subset of it with four elements satisfies the hypotheses of Lemma 3.1. For every $i \in \{1, \dots, 5\}$, we write $\mathbf{b}_i$ for the quadruple $\{b_j : j \neq i\}$ and Λ_i for its corresponding Λ. Furthermore, we define $P_i(X_1, X_2, X_3, X_4, X_5, \Lambda) \in \mathbb{Z}[X_1, X_2, X_3, X_4, X_5, \Lambda]$ as

$$P_i(X_1, X_2, X_3, X_4, X_5, \Lambda) = P(X_1, \dots, \hat{X}_i, \dots, \Lambda),$$

where by $\hat{X}_i$ we mean that the variable X_i has been eliminated (note that P_i depends only on four variables of the type X_j). We then get that the five positive integers b_i for $i = 1, \dots, 5$ satisfy the system of five congruences

$$P_i(b_1, \dots, b_5, \Lambda_i) \equiv 0 \pmod{p} \qquad \text{for } i = 1, \dots, 5.$$

Since we have 5 homogeneous polynomial relations in the indeterminates $b_1, \dots, b_5$, by *variable elimination*, we get a relation of the form

$$Q(\Lambda_1, \dots, \Lambda_5) \equiv 0 \pmod{p}. \tag{10}$$

The polynomial Q is called the integral resultant for the system $P_i = 0$, $i = 1, \dots, 5$. From the Elimination Theory (see [13, Section V.1] and [14, Section 9]), it follows that Q is not the constant zero polynomial.

We now find upper bounds for the *total degree* $\deg(Q)$ of Q and its *height* $h(Q)$, which we define as the sum of the absolute values of all its coefficients.

Assume that

$$f(X) = E_0 + E_1 X + \cdots + E_n X^n \qquad \text{and} \qquad g(X) = F_0 + F_1 + \cdots + F_m X^m$$

are polynomials of degrees $\max\{m, n\} \leq D$. Then their resultant $\mathrm{Res}_X(f, g)$ is a polynomial in $\mathbb{Z}[E_0, \dots, E_n, F_0, \dots, F_m]$ of total degree not exceeding $2D$. If furthermore our polynomials $f(X)$ and $g(X)$ shown above have the property that each one of the coefficients $a_0, \dots, a_m, b_0, \dots, b_n$ are polynomials in $\mathbb{Z}[Y, Z, T, \dots]$ such that the degrees and heights in the sense mentioned above of both f and g as polynomials in $\mathbb{Z}[X, Y, Z, T, \dots]$ fulfill the inequalities $\max\{\deg(f), \deg(g)\} \leq D$ and $\max\{h(f), h(g)\} \leq H$, then it is easy to see that their resultant $\mathrm{Res}_X(f, g)$ as a polynomial in $\mathbb{Z}[Y, Z, \dots]$ has total degree at most $2D^2$ and height $\leq (2D)! H^{2D} < (2DH)^{2D}$. Define the two sequences of upper bounds for degrees and heights, say $(D_k)_{k \geq 0}$ and $(H_k)_{k \geq 0}$,

respectively, as $D_0 = 12$ (the total degree of P), $H_0 = 2^{12}$ (an obvious upper bound for the height of P),

$$D_{k+1} = 2D_k^2 \qquad \text{and} \qquad H_{k+1} = (2D_kH_k)^{2D_k} \tag{11}$$

for $k \geq 1$. Then Q is obtained from the P_i's (which have the same degrees and heights as P) by taking successive resultants. For example, the first step, say eliminating X_1, means computing the three polynomials $U_j = \mathrm{Res}_{X_1}(P_2, P_j)$ for $j = 3, 4, 5$, thus reducing the problem from the five homogeneous polynomials P_i for $i = 1, \ldots, 5$, in the five variables $X_1, \ldots, X_5$, to the four homogeneous polynomials P_1, U_3, U_4, U_5, in the four variables $X_2, \ldots, X_5$. Inductively, it is easy to see that $\deg(Q) \leq D_4$ and $h(Q) \leq H_4$, where the numbers D_4 and H_4 can be computed using the above recurrence (11).

We now show that

$$H_k \leq 2^{D_k(k+1)}. \tag{12}$$

The above inequality is certainly true for $k = 0$. Assuming it to be true at some $k \geq 0$, and using the inequality $2D_k \leq 2^{D_k}$, we obtain

$$H_{k+1} = (2D_kH_k)^{2D_k} \leq (2D_k)^{2D_k} \cdot 2^{2D_k^2(k+1)} \leq 2^{2D_k^2} \cdot 2^{2D_k^2(k+1)} = 2^{D_{k+1}(k+2)}.$$

As for D_k, one proves by induction on k that the formula

$$D_k = 2^{2^k-1} D_0^{2^k} \tag{13}$$

holds for all $k \geq 0$. Evaluating at $k = 4$, $D_0 = 12$ first in formula (13) and then in inequality (12), we get

$$\deg(Q) \leq D_4 = 2^{15} \cdot 12^{16} < 2^{73} \qquad \text{and} \qquad h(Q) \leq H_4 < 2^{2^{75}}.$$

These bounds will alow us to show that in (10) the congruence can be replaced by the equality.

Since $\max\{\Lambda_1, \ldots, \Lambda_5\} < p^{6\delta}$, we get that

$$|Q(\Lambda_1, \ldots, \Lambda_5)| \leq h(Q)\,(\max\{\Lambda_i\})^{\deg(Q)} < (2p^{6\delta})^{2^{75}} < p^{2^{79}\delta}, \tag{14}$$

where the last inequality holds when $p^{10\delta} > 2$. We now choose $\delta = 2^{-79}$ and conclude for $p > 2^{2^{76}}$ we have that

$$|Q(\Lambda_1, \ldots, \Lambda_5)| < p. \tag{15}$$

Comparing (10) with (15), we conclude that the only possibility is

$$Q(\Lambda_1, \ldots, \Lambda_5) = 0. \tag{16}$$

3.4 Another auxiliary polynomial

In what follows, we study the equation (16). Namely, we show that there exists a polynomial $R(X,Y,Z,T,W) \in \mathbb{Z}[X,Y,Z,T,W]$, which is not the constant zero polynomial, such that if $\Lambda_1, \ldots, \Lambda_5$ satisfy (16), then $(\Lambda_1, \ldots, \Lambda_5)$ arise from a quintuple $(b_1, \ldots, b_5)$ (via the formulas $\Lambda_i = (\lambda_{1i}\lambda_{2i}\lambda_{3i})^2$ where $\lambda_{1i}, \lambda_{2i}$ and λ_{3i} are given by formulas (5), (7) and (8), respectively, when $(a,b,c,d) = (b_1, \ldots, \hat{b_i}, \ldots, b_5)$), such that

$$R\left(\frac{b_1}{\sqrt{p}}, \frac{b_2}{\sqrt{p}}, \frac{b_3}{\sqrt{p}}, \frac{b_4}{\sqrt{p}}, \frac{b_5}{\sqrt{p}}\right) = 0. \tag{17}$$

We also give an upper bound for the total degree of R.

Fix again some Diophantine quadruple $\{a,b,c,d\}$ with the property $D(p)$. With the substitution $(x,y,z,t) = (a/\sqrt{p}, b/\sqrt{p}, c/\sqrt{p}, d/\sqrt{p})$, formulas (5), (7) and (8) show that we have

$$\lambda_1 = \sqrt{(xy+1)(zt+1)} - \sqrt{(xz+1)(yt+1)}, \tag{18}$$

$$\lambda_2 = \sqrt{(xy+1)(zt+1)} - \sqrt{(xt+1)(yz+1)}, \tag{19}$$

and

$$\lambda_3 = \sqrt{(xt+1)(yz+1)} - \sqrt{(xz+1)(yt+1)}. \tag{20}$$

Clearly, $\lambda_1, \lambda_2, \lambda_3$ are algebraic functions belonging to the unique extension $\mathbb{K}$ of $\mathbb{F} = \mathbb{Q}(x,y,z,t)$ of degree 2^6 which contains all the expressions $e_{\{u,v\}} = \sqrt{uv+1}$ for all subsets of two elements $\{u,v\} \subseteq \{x,y,z,t\}$. Now we have that

$$\Lambda = (\lambda_1\lambda_2\lambda_3)^2$$

is a function belonging to the same field, which is also symmetric in the four variables x,y,z,t. Furthermore, with respect to the canonical basis

$$\{f_1, \ldots, f_{2^6}\} = \left\{ \prod_{\substack{I_1,\ldots,I_k \subset \{x,y,z,t\} \\ \#I_j=2,\ I_j \text{ distinct}}} \prod_{j=1}^{k} e_{I_j} : k = 0, \ldots, 6 \right\}$$

of $\mathbb{K}$ over $\mathbb{F}$ (here an empty product is taken to be 1), we have that

$$\Lambda = \sum_{m=1}^{2^6} g_m(x,y,z,t) f_m, \tag{21}$$

where $g_m(x,y,z,t) \in \mathbb{F}$ are polynomials of degrees at most 12. Let now $(b_1,\ldots,b_5)$ be a quintuple for which relation (16) is fulfilled. Write $z_i = b_i/\sqrt{p}$ for all $i = 1,\ldots,5$, and represent

$$\Lambda_i = \Lambda(z_1,\ldots,\hat{z_i},\ldots,z_5) \tag{22}$$

for all $i = 1,\ldots,5$, in the form (21). If we insert all the five relations (22) above into (16), we get a relation of the type

$$\phi(z_1,\ldots,z_5) = 0, \tag{23}$$

where

$$\phi(z_1,\ldots,z_5) = \sum_{n=1}^{2^{10}} h_n(z_1,\ldots,z_5)e_n, \tag{24}$$

$$\{e_1,\ldots,e_{2^{10}}\} = \left\{ \prod_{\substack{I_1,\ldots,I_k \subset \{z_1,z_2,z_3,z_4,z_5\} \\ \#I_j=2,\ I_j \text{ distinct}}} \prod_{j=1}^{k} e_{I_j} : k = 0,\ldots,10 \right\}$$

is the canonical basis of the smallest extension of $\mathbb{L} = \mathbb{Q}(z_1,\ldots,z_5)$ containing all the functions $e_{\{u,v\}}$ for $\{u,v\} \subseteq \{z_1,\ldots,z_5\}$, and $h_n(z_1,\ldots,z_5)$ are polynomials of degrees at most $12 \cdot \deg(Q) \leq 12 \cdot 2^{73} < 2^{77}$. We shall later show that this relation is nontrivial (i.e., it is not constant zero). Assuming that we have proved this, the above relation (23) leads to a polynomial relation

$$R_1(z_1,\ldots,z_5) = 0$$

of total degree at most $2^{10} \cdot 2^{77} = 2^{87}$. Clearly, the polynomial R_1 can be regarded as being obtained from the minimal polynomial of $\phi(z_1,\ldots,z_5)$ over $\mathbb{L}$. This polynomial R_1 is *almost* the polynomial R. To choose R, we note that when we have selected $\{b_1,\ldots,b_5\}$ out of $\{a_i,\ldots,a_{i+\ell}\}$, we chose not to consider any instance of the type $\{a_i,a_j,a_k\}$ such that $a_k = a_i + a_j \pm 2x_{ij}$. However, if $\{a,b,c\}$ is a Diophantine triple having the property that

$$c = a + b \pm 2\sqrt{ab \pm p},$$

we then get the relation

$$\frac{1}{4}\left(\frac{c}{\sqrt{p}} - \frac{a}{\sqrt{p}} - \frac{b}{\sqrt{p}}\right)^2 - \frac{a}{\sqrt{p}}\frac{b}{\sqrt{p}} \pm 1 = 0,$$

which is a polynomial relation of the type

$$U_{\pm}(a/\sqrt{p}, b/\sqrt{p}, c/\sqrt{p}) = 0,$$

where

$$U_{\pm}(x, y, z) = \frac{1}{4}(z - x - y)^2 - xy \pm 1$$

is a polynomial of degree 2. We now take

$$\begin{aligned} R(z_1, \ldots, z_5) &= R_1(z_1, \ldots, z_5) \prod_{1 \le i < j \le 5} (z_i - z_j) \\ &\times \prod_{\varepsilon \in \{\pm\}} \prod_{I \subset \{1,\ldots,5\},\ \#I=3} U_{\varepsilon}(z_i, z_j, z_k : \{i, j, k\} = I), \quad (25) \end{aligned}$$

and we note that this is a nonzero polynomial of degree $\le \deg(R_1)+20+40 < 2^{88}$.

Before proceeding further, we show that R_1 (hence, R), is not the zero polynomial. Assuming that it is, it follows that $\phi(z_1, \ldots, z_5)$ is constant zero. Since Q is not constant zero, and ϕ is the image of $Q(\Lambda_1, \ldots, \Lambda_5)$ via the algebraic map

$$(\Lambda_1, \ldots, \Lambda_5) = (\Lambda_i(z_1, \ldots, z_5))_{i=1,\ldots,5}$$

which is differentiable (in fact, of class $\mathcal{C}^{\infty}$) in the real open set $\mathcal{B} = \{z_i > 1 : i = 1, \ldots, 5\} \subset \mathbb{R}^5$, it suffices to show that the Jacobian of this map is not constant zero in $\mathcal{B}$. Here, we could not find a theoretical argument to this effect, so we simply computed the specialization of this Jacobian in $(z_1, \ldots, z_5) = (2, 3, 4, 5, 6)$ using Maple and obtained a nonzero value for it. In fact, using 100 digits precision, we obtained that this Jacobian is $-0.1226252714 \cdot 10^{-30}$. When p is replaced by $-p$ (i.e., when all the $+1$'s are replaced by -1's in the formulas (18), (19) and (20)), we obtained that the value of this Jacobian is $0.2933578498 \cdot 10^{-28}$. This shows that indeed ϕ is not constant zero as an algebraic element over $\mathbb{L}$.

3.5 The end of the proof

We will finish the proof of Theorem 2.1 by proving the non-vanishing of polynomial R in an appropriate quintuple. This will imply the non-vanishing of Q in the corresponding quintuple of Λ's, contradicting the already proved

property of Q. The obtained contradiction leads to a lower bound for one of the λ's (from the definition of Λ), and this implies the desired inequality of the form $a_{i+\ell} > p^{\gamma} a_i$.

We recall the following elementary result about the non-vanishing of polynomials with more than one indeterminate.

Lemma 3.2. *Let M be a positive integer, and let R be a nonzero polynomial with complex coefficients in $t \geq 1$ variables $X_1, \ldots, X_t$ of total degree $< M$. Let $\mathcal{M}$ be any subset of $\mathbb{C}$ of cardinality $\geq M$. Then, there exists $\mathbf{x} = (x_1, \ldots, x_t) \in \mathcal{M}^t$ such that $R(\mathbf{x}) \neq 0$.*

Proof. We use induction on t. When $t = 1$, the assertion of the Lemma 3.2 follows from the fact that a nonzero polynomial $R(x)$ of degree $< M$ cannot have a number of roots $\geq M$, and therefore there exists an element $x \in \mathcal{M}$ such that $R(x) \neq 0$.

Assume now that $t \geq 2$ and that the assertion of Lemma 3.2 holds for all nonzero polynomials with less than t variables. Let R be a polynomial in $\mathbb{C}[X_1, \ldots, X_t]$ which depends on all t variables (i.e., $\partial R/\partial X_i$ is not the zero polynomial for any $i = 1, \ldots, t$), and write it as

$$R = X_t^d R_d + X_t^{d-1} R_{d-1} + \cdots + R_0,$$

where $d < M$, and R_i are polynomials in the variables $X_1, \ldots, X_{t-1}$ for $i = 0, \ldots, d$, with $R_d \neq 0$, and of degree $< M$. By the induction hypothesis, there exist $x_i \in \mathcal{M}$ for $i = 1, \ldots, t-1$, such that $R_d(x_1, \ldots, x_{t-1}) \neq 0$. Specializing R in $(X_1, \ldots, X_{t-1}) = (x_1, \ldots, x_{t-1})$, and letting the tth variable X_t free, we obtain a nonzero polynomial in one variable X_t of degree $< M$, and therefore there exists a choice of an element $x_t \in \mathcal{M}$ so that this last polynomial in the variable X_t does not vanish when evaluated in x_t.

□

By assuming that the polynomial R appearing in the above Lemma 3.2 is a multiple of all the linear polynomials $X_i - X_j$ for all $i \leq j$ in $\{1, \ldots, t\}$, we may insure that every point $\mathbf{x} \in \mathcal{M}^t$ realizing $R(\mathbf{x}) \neq 0$ has distinct coordinates.

We now take $t = 5$, R to be the polynomial appearing at (25), $\ell = M = 2^{88}$, and $\mathcal{M} = \{a_i, \ldots, a_{i+\ell}\}$. By Lemma 3.2, there exists $(a, b, c, d, e) \in \mathcal{M}^5$ such that $R(a, b, c, d, e) \neq 0$. Since $z_i - z_j$ divides $R(z_1, \ldots, z_5)$ for all $i \neq j$

in $\{1,\dots,5\}$, it follows that $\{a,b,c,d,e\}=\{a_{i_1},\dots,a_{i_5}\}$, where $i\le i_1<\dots<i_5\le i+\ell$. Furthermore, since $U_\epsilon(z_i,z_j,z_k)$ divides $R(z_1,\dots,z_5)$ for all subsets with three elements $\{i,j,k\}$ of $\{1,\dots,5\}$ and both $\varepsilon\in\{\pm1\}$, it follows that $a_{i_u}\ne a_{i_v}+a_{i_w}\pm 2\sqrt{a_{i_v}a_{i_w}\pm p}$ holds for all triples (u,v,w) of distinct indices in $\{1,\dots,5\}$. Finally, since $R_1(a_{i_1}/\sqrt{p},\dots,a_{i_5}/\sqrt{p})\ne 0$, it follows that $Q(\Lambda_{i_1},\dots,\Lambda_{i_5})\ne 0$, and this contradicts (16). Note that we have obtained the equality (16), under the assumption that the corresponding λ's are small. Therefore, we conclude that the inequality $\lambda_{\kappa i_j}\ge p^\delta$ holds for some $\kappa\in\{1,2,3\}$, $j\in\{1,\dots,5\}$. But then $a_{i_5}>p^{\delta/2}$ (see (6)), and if we take $\gamma=\delta/2=2^{-80}$ and $\ell=2^{88}$, then the desired inequality

$$a_{i+\ell}\ge a_{i_5}>p^\gamma a_{i_1}\ge p^\gamma a_i$$

does hold. Finally, inequality (2) tells us that we can take

$$C=24+\ell+11\ell\gamma^{-1}4^{-1}<3\ell\gamma^{-1}=3\cdot 2^{168},$$

and Theorem 2.1 is completely proved.

4 The Proof of Theorem 2.2

We treat again only the case of the positive squarefree integer n, and we shall point out the minor modifications required to deal with the case of the negative integer n. Put

$$M_n=\max\{|\mathcal{A}|:\mathcal{A}\text{ is a Diophantine tuple with the property }D(n)\},$$

and we wish to show that the inequality $M_n\le f(\omega(n))$ holds with some function f. If $p|n$, then there can be at most one element $a\in\mathcal{A}$ such that $p|a$. Indeed, if two such occur, call them a and b, we then get that $ab+n=x^2$ holds with some integer x such that $p|x$. Reducing this equation modulo p^2, we get $n\equiv 0\pmod{p^2}$, which is impossible because n is squarefree. Thus, eliminating at most $\omega(n)$ elements from $\mathcal{A}$, we may assume that the remaining ones are all coprime to n.

As in the case when n was prime, we may assume that $n>2^{2^{76}}$, and therefore that $f(\omega(n))>2^{80}\log 2$.

Again as in the case when n was a prime, at most two of its elements are $<n^{1/4}$. Eliminating those ones, we may assume that the inequality $a>n^{1/4}$ holds for all $a\in\mathcal{A}$.

Let again $\{a, b, c\}$ be some triple of elements of $\mathcal{A}$ with $a < b < c$ and write again

$$ab + n = x^2, \qquad bc + n = y^2, \qquad \text{and} \qquad ac + n = z^2$$

with some positive integers $x,\ y$ and z. We write again equation (3), which in this instance is

$$\begin{aligned}(2xyz - 2abc - n(a+b+c))\,(2xyz + 2abc + n(a+b+c)) \\ = n^2\,(2x - a - b + c)\,(2x + a + b - c)\,. \end{aligned} \tag{26}$$

If both sides of the above equation (26) are zero, then $c = a + b + 2x$ (or $c = a + b - 2x$ if n is negative). Therefore c is (almost) uniquely determined in terms of a and b. As in the proof of Theorem 2.1, we shall ignore such triples. Furthermore, as in the proof of Theorem 2.1, since each of $a,\ b$ and c is coprime to n, it follows that the greatest common divisor of the two factors appearing on the first line of the above array of equations (26) is coprime to n if n is odd and to $n/2$ if n is even. We write n_1 for the largest odd divisor of n.

We now write

$$d_s = \gcd\left(2xyz - 2abc - n(a+b+c), n_1^2\right), \tag{27}$$

and

$$d_l = \gcd\left(2xyz + 2abc + n(a+b+c), n_1^2\right), \tag{28}$$

and remark that $d_s d_l = n_1^2 \geq n^2/4$.

We will first consider the case when the "small factor" is divisible by some "large divisor" of n^2. More precisely, we assume that $d_s \geq n^{8/5}$. In this case, we get that

$$2\max\{4xyz, 3nc\} \geq |(2xyz - (2abc + n(a+b+c))| \geq n^{8/5}. \tag{29}$$

We distinguish the following situations.

Case 1. $3nc > 4xyz$.

In this case, $nc > xyz > abc$, therefore $ab < n$. In particular, $a < n^{1/2}$. Since also $6nc = 2\max\{4xyz, 3nc\} > n^{8/5}$, we get

$$c > n^{3/5}/6 > n^{1/10}a/6 > n^{1/11}a,$$

where the last inequality above follows from the fact that $n > 6^{110}$. Hence, in this case we get the desired inequality with $\ell = 2$ and $\gamma = 1/11$.

From now on, we have $4xyz \geq 3nc$. Clearly,

$$xyz \leq 2\sqrt{2}\max\{\sqrt{ab}, \sqrt{n}\}\max\{\sqrt{bc}, \sqrt{n}\}\max\{\sqrt{ac}, \sqrt{n}\}.$$

If $ab < ac < bc < n$, then $xyz \leq 2\sqrt{2}n^{3/2}$, and now inequality (29) leads to the inequality

$$16\sqrt{2}n^{3/2} \geq 8xyz = 2\max\{4xyz, 3nc\} > n^{8/5}.$$

In turn, the above inequality implies $n < (16\sqrt{2})^{10} = 2^{45}$, contradicting our assumption on n. Thus, $bc > n$.

Case 2. *$ab < n$ and $ac < n$.*

In this case, $a^2 < ab < n$, therefore $a < n^{1/2}$. Furthermore, the inequality (29) leads to

$$16\sqrt{2}nc > 16\sqrt{2}n\sqrt{bc} \geq 2\max\{4xyz, 3nc\} > n^{8/5},$$

which in turn implies

$$c > n^{3/5}/(16\sqrt{2}) > n^{1/10}a/(16\sqrt{2}) > n^{1/11}a,$$

which is the same inequality as the one from Case 1. The last inequalities hold for the range $n > (16\sqrt{2})^{110} = 2^{495}$, which is our case.

Case 3. *$ab < n$ but $ac \geq n$.*

In this case, $a^2 < ab < n$, therefore $a < n^{1/2}$. Furthermore, the inequality (29) becomes

$$16\sqrt{2}nc > 16\sqrt{2}\sqrt{n}\sqrt{abc} = 16\sqrt{2}\sqrt{n}\sqrt{ac}\sqrt{bc} \geq 2\max\{4xyz, 3nc\} \geq n^{8/5},$$

so we reached again the same inequality as the one from Case 2.

Case 4. *$ab \geq n$.*

In this case, the desired inequality is

$$16\sqrt{2}c^3 > 16\sqrt{2}abc > 3\max\{4xyz, 3nc\} > n^{8/5},$$

leading to $c > n^{8/15}/2^{3/2}$. We now have, by (26),

$$\begin{aligned} & n^{8/5}(2(xyz+abc)+n(a+b+c)) \\ & \quad \le |2xyz-2abc-n(a+b+c)||2xyz+2abc+n(a+b+c)| \\ & \quad \le n^2|2(ab+ac+bc)-a^2-b^2-c^2-4n| \\ & \quad \le n^2 \max\{4n, 6c^2\}. \end{aligned}$$

If $6c^2 < 4n$, then $c < n^{1/2}$, which together with the fact that $c > n^{8/15}/2^{3/2}$ leads to the conclusion that $n < (2^{3/2})^{30} = 2^{45}$, contradicting our assumption on n. Hence, we must have the inequality

$$4abc < 2xyz + 2abc + n(a+b+c) < 6n^{2/5}c^2,$$

leading to $c > (2/3)abn^{-2/5}$. Since $b > n^{1/2}$ (because $ab \ge n$), we have $c > an^{1/10}(2/3) > n^{1/11}a$, where the last inequality follows because $n > (3/2)^{110}$.

Hence, in all the above four cases we reached the conclusion that the inequality $c > n^{1/11}a$ holds when $d_s > n^{8/5}$. The case in which n is replaced by $-n$ can be analyzed with similar arguments.

We now look at the case $d_s < n^{8/5}$. In this case, $d_l > n^2/(4d_s) > n^{2/5}/4 > n^{1/5}$ because $n > 4^5$. Since every prime in d_l appears at the exponent exactly 2, it follows that the largest prime $p|d_l$ is odd and fulfills $p > n^{\frac{1}{10\omega(n)}}$.

Let $\{p_1, \ldots, p_{t(n)}\}$ be the set of all the distinct prime factors of n dividing some divisor of n_1 larger than $n^{1/10}$ (in the case of the above triple, such a divisor can be taken to be $\sqrt{d_l}$). We also put

$$u(n) = \max\left\{\frac{\log n}{\log p_i} : i = 1, \ldots, t(n)\right\}. \tag{30}$$

Clearly, $u(n) \le 10\omega(n)$. We note that $t(n) \le \omega(n)$.

To continue, we need to introduce the following combinatorial definition of the *Ramsey like number* $\mathcal{R}(A, B)$, where A and B are arbitrary positive integers. For an arbitrary finite set $\mathcal{S}$, let us refer to an arbitrary subset with three elements of $\mathcal{S}$ as to a *triangle*.

Definition 4.1. *Given A and B, we write $\mathcal{R}(A, B)$ for the smallest positive integer S such that every finite set $\mathcal{S}$ of cardinality $\#\mathcal{S} \ge S$, and for every coloring of the set of all its triangles with at most A colors, there exists a subset $\mathcal{S}_1$ of $\mathcal{S}$ of cardinality at least B such that all its triangles have the same color.*

We may recognize $\mathcal{R}(A,B)$ as the classical generalized Ramsey number $R(\underbrace{B,\ldots,B}_{A \text{ times}};3)$. Here, $R(m_1,\ldots,m_k;q)$ is the minimum positive integer R such that if $\mathcal{S}$ is a set of cardinality R and all its q-element subsets are colored with k colors, then there exists $i \in \{1,\ldots,k\}$ and a subset $\mathcal{S}_i$ of $\mathcal{S}$ of cardinality at least m_i such that all its q-element subsets have color i.

It is a deep theorem due to Ramsey that $\mathcal{R}(A,B)$ exists (see, for example, Theorem 1 on page 3 of [10]). We shall discuss upper bounds for it in the proof of Theorem 2.3.

For our instance, we choose $A = t(n)$ and $B = 2^{88}$, which is an upper bound on the degree of the polynomial $R(z_1,\ldots,z_5)$ appearing towards the end of the proof of Theorem 2.1. We let $\ell = \mathcal{R}(A,B)$ and let $\mathcal{S} = \{a_i,\ldots,a_{i+\ell}\}$. We assume that every triangle in $\mathcal{S}$ is either *regular*; i.e., if $a<b<c$ is the triangle then $c = a+b\pm 2\sqrt{ab+n}$, or that it has $d_s < n^{8/5}$. In case $d_s < n^{8/5}$, we color this triangle by assigning to it the color p_i for $i \in \{1,\ldots,t(n)\}$, where p_i is the maximal prime factor of d_l (which is larger than $n^{\frac{1}{u(n)}}$). In the regular case, we simply assign to $\{a,b,c\}$ an arbitrary color from the set $\{p_1,\ldots,p_{t(n)}\}$. From the definition of $\mathcal{R}(A,B)$ (more precisely, from the fact that it exists), it follows that we can select a subset of it $\{a_j : j \in J\}$, where $J \subset \{i,\ldots,i+\ell\}$ has at least $B = 2^{88}$ elements, and such that furthermore there exists a prime number $p \in \{p_1,\ldots,p_{t(n)}\}$ such that if $\{j_1,j_2,j_3\} \in J$ is an arbitrary triangle with $j_1<j_2<j_3$, then either $\{a_{j_1},a_{j_2},a_{j_3}\}$ is regular, or $a_{j_1}a_{j_2}a_{j_3} \equiv -x_{j_1j_2}x_{j_1j_3}x_{j_2j_3} \pmod p$.

At this stage, the argument from the end of the proof of Theorem 2.1 based on Lemma 3.2 shows that there exists a quintuple of elements among $\{a_j : j \in J\}$, let us call it $\{b_1,\ldots,b_5\}$, such that $R(b_1/\sqrt{p},\ldots,b_5/\sqrt{p}) \neq 0$. In particular, $Q(\Lambda_1,\ldots,\Lambda_5) \neq 0$, where the numbers Λ_i are the parameters constructed in the proof of Theorem 2.1 with respect to the quintuple $\{b_1,\ldots,b_5\}$. As in the proof of Theorem 2.1, all the b_i's are distinct and no triangle among them is regular. We now write δ for the nonnegative real number such that

$$\delta = \max\left\{\frac{\log \lambda_{ji}}{\log p} : j=1,2,3;\ i=1,\ldots,5\right\},$$

where $\Lambda_i = (\lambda_{1i}\lambda_{2i}\lambda_{3i})^2$, and the λ_{ji}'s are defined through relations (5)- (8) for $j=1,2,3$ for the quadruple $\{b_1,\ldots,\hat{b_i},\ldots,b_5\}$. The divisibility condition

(10) together with inequality (14) now show that

$$p \leq |Q(\Lambda_1, \ldots, \Lambda_5)| \leq h(Q) \left(\max\{\Lambda_i\}\right)^{\deg(Q)} < (2p^{6\delta})^{2^{75}} < p^{2^{79}\delta}, \qquad (31)$$

provided that $p^{10\delta} > 2$. We will justify this inequality later. The above inequality shows that $\delta \geq 2^{-79}$. It remains to justify that $p > 2^{0.1\delta^{-1}}$. For this, it suffices that $p > 2^{2^{76}}$. However, since $p > n^{\frac{1}{u(n)}} > n^{\frac{1}{10\omega(n)}}$, it suffices that $n > 2^{2^{76} \cdot 10\omega(n)}$, which is implied by

$$\log n > 2^{80}\omega(n). \qquad (32)$$

By the trivial estimate $\omega(n)! \leq n$ and Stirling's formula, we get that

$$\log(n) \geq \omega(n) \log(\omega(n)/e). \qquad (33)$$

If $\omega(n)/e > e^{2^{80}}$, then inequality (32) is implied by (33). If on the other hand $\omega(n)/e < e^{2^{80}}$, then $\omega(n) < 2^{2^{81}}$, $2^{80}\omega(n) < 2^{2^{82}}$, therefore inequality (32) is valid for $n > e^{2^{2^{82}}}$. Thus, by the result from [8], we need that $f(\omega(n))$ is larger than $16 \cdot 2^{2^{82}}$, and in particular we may assume that it is larger than $2^{2^{83}}$.

The conclusion is that with our values for ℓ and δ, we insured that the inequality

$$a_{i+\ell} \geq b_5 \geq p^{\delta/2} b_1 \geq p^{\delta/2} a_i \geq n^{\frac{2^{-80}}{u(n)}} a_i = n^{\gamma} a_i$$

holds, where one can take $\gamma = 2^{-80}/u(n)$. Since at any rate $a_m < n^3$, it follows that we are entitled to apply inequality (2) and get that

$$\begin{aligned} M_n &< \max\left\{2^{2^{83}},\ 23 + \omega(n) + \ell\left(1 + 11/4\gamma^{-1}\right)\right\} \\ &< \max\left\{2^{2^{83}},\ 23 + \omega(n) + \mathcal{R}(t(n), 2^{88})(1 + 11/4 \cdot 2^{80}u(n))\right\} \\ &< \max\left\{2^{2^{83}},\ \omega(n) + 2^{82}u(n)\mathcal{R}(t(n), 2^{88})\right\}. \qquad (34) \end{aligned}$$

Finally, since $t(n) \leq \omega(n)$ and $u(n) \leq 10\omega(n)$, we get the inequality

$$M_n < \max\left\{2^{2^{83}},\ 2^{86}\omega(n)\mathcal{R}(\omega(n), 2^{88})\right\},$$

which is a bound depending only on $\omega(n)$.

5 The Proof of Theorem 2.3

Since this result addresses only most positive integers, throughout this section we shall make extensive use of the symbols O, o, $\ll$ and $\gg$ recalled at the end of Section 1.

Let X be a large positive real number and let $n \in [1, X]$. We let $\mathcal{E}_1(X) = \{n : n \leq X/\log X\}$. Since $\#\mathcal{E}_1(X) = O(X/\log X) = o(X)$, we may assume that $n \notin \mathcal{E}_1(X)$. Write

$$n = s(n)n_1,$$

where $\gcd(s(n), n_1) = 1$, $s(n)$ is squarefull except for the prime 2; i.e., if $p|s(n)$ and $p > 2$ then $p^2|n$, and n_1 is odd and squarefree. It is wellknown that if we write $S(X)$ for the counting function of the set of squarefull numbers $m \leq X$, then

$$S(X) = \beta\sqrt{X} + O(X^{1/3}), \qquad \beta = \frac{\zeta(3/2)}{\zeta(3)} \approx 2.1732, \tag{35}$$

where ζ is the Riemann zeta function (see, for example, Theorem 14.4 of [12]). We now let

$$\mathcal{E}_2(X) = \{n \notin \mathcal{E}_1(X) : s(n) > \log n\}.$$

Since $n > X/\log X$, we get that $s(n) > 1/2 \log X$ holds for all X sufficiently large. To estimate $\#\mathcal{E}_2(X)$, we fix a number m in the interval $[1/2 \log X, X]$ such that either m or $2m$ is squarefull and count the number of positive integers $n \leq X$ which are multiples of m. The number of such integers is

$$\leq \left\lfloor \frac{X}{m} \right\rfloor \leq \frac{X}{m}.$$

Thus,

$$\#\mathcal{E}_2(X) \leq X \sum_{\substack{m > 1/2 \log X \\ m \in S(X)}} \frac{1}{m} + X \sum_{\substack{m > 1/2 \log X \\ m \in S(X/2)}} \frac{1}{2m} \ll \frac{X}{\sqrt{\log X}}, \tag{36}$$

where estimate (36) above follows from (35) by partial summation. From now on, we work only with those positive integers $n \leq X$ not belonging to $\mathcal{E}_1(X) \cup \mathcal{E}_2(X)$.

By the Túran-Kubilius estimate

$$\sum_{n \le X} |\omega(n) - \log\log X|^2 = O(X \log\log X)$$

(see, for example, [16]), it follows that the set $\mathcal{E}_3(X)$ of positive integers $n \le X$ not in $\mathcal{E}_1(X) \cup \mathcal{E}_2(X)$ which have the property that $|\omega(n) - \log\log X| \ge (\log\log X)^{2/3}$ has cardinality $O(X/(\log\log X)^{1/3}) = o(X)$. Thus, from now on, we work only with the integers $n \le X$ not in $\mathcal{E}_1(X) \cup \mathcal{E}_2(X) \cup \mathcal{E}_3(X)$. It is easy to see that all such integers n have the property that $|\omega(n) - \log\log n| < 2(\log\log n)^{2/3}$ once X is large enough.

We now let $\mathcal{A}$ be an arbitrary Diophantine m-tuple with the property $D(\pm n)$. By the arguments from the beginning of the proof of Theorem 2.2, there are at most $24 + \omega(n_1)$ elements a of $\mathcal{A}$ such that either $a \notin (n^{1/4}, n^3)$, or a is a multiple of some prime p dividing n_1. Clearly, for us,

$$\begin{aligned} \omega(n) &\ge \omega(n_1) \ge \omega(n) - \omega(s(n)) \ge \omega(n) + O\left(\frac{\log(s(n))}{\log\log s(n)}\right) \\ &= \omega(n) + o(\log\log n), \end{aligned}$$

therefore $\omega(n_1) = (1 + o(1))\omega(n) = (1 + o(1))\log\log n$. From now on, we eliminate such elements from $\mathcal{A}$.

We now let $\{a, b, c\}$ be an arbitrary triple in $\mathcal{A}$ with $a < b < c$ which is not regular. We follow the arguments from the proof of Theorem 2.2, where d_s and d_l are again given by formulas (27) and (28), respectively, except that $n_1 = n/s(n) > n/\log n$. As in the proof of Theorem 2.2, if $d_s > n^{8/5}$, then $c > n^{1/11}a$ once n (hence, X) is sufficiently large. We now consider the case when $d_s < n^{8/5}$, for which $d_l > n_1^2/d_s > n^{2/5}/(\log n)^2 > n^{1/5}$ once n is sufficiently large. Since d_l is also a perfect square (in fact, the square of a squarefree number), it follows that $\sqrt{d_l} > n^{1/10}$ is a divisor of n. As in Theorem 2.2, we put $\{p_1, \ldots, p_{t(n)}\}$ for a set consisting of some prime factors of n as large as possible such that every divisor $d > n^{1/10}$ of n_1 is a multiple of one of the p_i's for $i = 1, \ldots, t(n)$. Such set is provided by

$$\{p_1, \ldots, p_{t(n)}\} = \{P(d) : d | n \text{ and } d > n^{1/10}\}. \tag{37}$$

We also put $u(n)$ as in (30). The argument from the proof of Theorem 2.2 shows that if n is large (i.e., if X is large), then the inequality

$$M_n \le 23 + \omega_1(n) + 2^{88} u(n) \mathcal{R}(t(n), 2^{88}) \tag{38}$$

holds for all our $n \leq X$ not in $\mathcal{E}_1(X) \cup \mathcal{E}_2(X) \cup \mathcal{E}_3(X)$. Furthermore, the term $\omega(n_1)$ from the right hand side of the above inequality (38) can be eliminated once the Diophantine m-tuple $\mathcal{A}$ is reduced (i.e., one no longer has to "eliminate" the eventual $a \in \mathcal{A}$ such that $p|a$ for some prime factor p of n_1 since $\mathcal{A}$ no longer contains such elements).

Unfortunately, the only inequality known for our number $R(t, 2^{88})$ is

$$R(t, 2^{88}) = \underbrace{\exp \exp \ldots \exp}_{O(t) \text{ times}}(O(1)). \tag{39}$$

Indeed, the above inequality follows from what is known about the classical generalized Ramsey numbers (see [10, pp. 90-91]). However, regardless of the size of $R(t, 2^{88})$, the full conclusion of Theorem 2.3 will now follow at once from inequality (38), and the following purely probabilistic result, which might be of independent interest.

Lemma 5.1. *Let $g : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{\geq 0}$ be any increasing function with the property that $\lim_{x \to \infty} g(x) = \infty$. Then the inequality*

$$\max\{t(n), u(n)\} < g(n)$$

holds for almost all positive integers n.

Proof. We let X be a large positive real number and put $u = g(X)$. We may assume that $u < \log \log X$ for large X. We put $Y = X^{1/u}$. We assume again that $n > X/\log X$. We put $\mathcal{E}_4(X) = \{n \leq X : P(n) \leq Y\}$. By standard results from the distribution of smooth numbers (see, for example, Section III.5.4 in [15]), we know that

$$\#\mathcal{E}_4(X) = \Psi(X; Y) = \frac{X}{\exp(u(1 + o(1)) \log u)} = o(X).$$

We put $\mathcal{E}_5(X)$ to be the set of $n \leq X$ not in $\cup_{i=1}^4 \mathcal{E}_i(X)$ such that $p^2|n$ for some $p > Y$. Then

$$\#\mathcal{E}_5(X) \leq \sum_{Y < p} \left\lfloor \frac{X}{p^2} \right\rfloor \leq X \sum_{p > Y} \frac{1}{p^2} \ll \frac{X}{Y \log Y} = o(X).$$

We now let $n \leq X$ not in $\cup_{i=1}^5 \mathcal{E}_i(X)$. We write $\mathcal{E}_6(X)$ for the set of such n such that there exists a divisor $d|n$ with $d > n^{1/10}$ and $P(d) \leq Y$. Since

$n > X/\log X$, we get that $d > n^{1/10} > X^{1/11}$. For $t \in [X^{1/11}, X]$, we note that $u(t) = \log t/\log Y \geq u/11$. Thus, for these values of t we have

$$\Psi(t;Y) = \frac{t}{\exp(u(t)(1+o(1))\log(u(t)))} > \frac{t}{\exp(u(\log u)/12)} \quad \text{for } t \in [Y, X],$$

once X is sufficiently large. Hence, by partial summation, we get that

$$\begin{aligned} \mathcal{S}_1 &= \sum_{\substack{X^{1/11}<n<X \\ P(n)<Y}} \frac{1}{n} \ll \int_{X^{1/11}}^{X} \frac{1}{t} d\Psi(t;Y) \\ &\ll \frac{1}{\exp(u(\log u)/12)} \int_{X^{1/11}}^{X} \frac{dt}{t} \\ &\ll \frac{\log X}{\exp(u(\log u)/12)}. \end{aligned} \tag{40}$$

To bound $\mathcal{E}_6(X)$, we let $n \in \mathcal{E}_6(X)$. Then $n = Pab$, where $P = P(n) > Y$, $b > X^{1/11}$, $P(b) \leq Y$, and a has the property that all its prime factors are $> Y$. Fixing $a,\ b$, we have $P \leq X/ab$, therefore the number of choices for P is

$$\leq \pi\left(\frac{X}{ab}\right) \leq \frac{X}{ab\log(X/ab)} \leq \frac{Xu}{ab\log X},$$

where in the last inequality above we used the fact that $X/ab \geq P \geq Y = X^{1/u}$. Summing up over a and b, we get

$$\#\mathcal{E}_6(X) \leq \frac{X}{u}\log X \left(\sum_{\substack{X^{1/11}<a<X \\ P(a)<Y}} \frac{1}{n}\right)\left(\sum_{\substack{b<X \\ p|b \Rightarrow p>Y}} \frac{1}{b}\right).$$

The first sum above is the sum $\mathcal{S}_1$ which is bounded in (40). For the second one, note that the b's we are considering are squarefree (because $b|n$ and $n \notin \mathcal{E}_5(X)$). Thus,

$$\mathcal{S}_2 = \sum_{\substack{b<X \\ p|b \Rightarrow p>Y}} \frac{1}{b} \leq \sum_{k\geq 0} \sum_{\substack{b<X \\ p|b \Rightarrow p>Y \\ \omega(b)=k}} \frac{1}{b} \leq \sum_{k\geq 0} \frac{1}{k!}\left(\sum_{Y<p<X} \frac{1}{p}\right)^k.$$

Since

$$z = \sum_{Y<p<X} \frac{1}{p} = \log\log X - \log\log Y + o(1) = \log u + o(1) < 1 + \log u,$$

we get that

$$\mathcal{S}_2 \leq \sum_{k\geq 0} \frac{z^k}{k!} = \exp(z) \ll \exp(\log u) = u. \tag{41}$$

From inequalities (40) and (41), we get that

$$\#\mathcal{E}_6(X) \ll \frac{xu^2}{\exp(u(\log u)/12)} = o(X).$$

Hence, all $n \leq X$ but for $o(X)$ of them have the property that if $p \in \mathcal{P}(n) = \{P(d) : d|n,\ d > n^{1/10}\}$, then

$$u(n) = \max\left\{\frac{\log n}{\log p} : p \in \mathcal{P}\right\} \leq u = g(X).$$

As for the size of the number $t(n)$, let $K = \lfloor 2e\log u \rfloor$. Let $\mathcal{E}_7(X)$ be the set of $n \leq X$ not in $\cup_{i=1}^{6}\mathcal{E}_i(X)$, such that $t(n) \geq K$. Let $P_1, \ldots, P_K$ be any K distinct primes factors of n than Y. Then $n = P_1 \ldots P_K m$, where m is some integer. Clearly, m can be chosen in at most $X/(P_1 \ldots P_K)$ ways. Summing this up over all the possible choices of distinct primes $P_i \in (Y, X)$ for $i = 1, \ldots, K$, we get

$$\begin{aligned}
\#\mathcal{E}_7(X) &\leq X \sum_{Y<P_1<P_2<\cdots<P_K<X} \frac{X}{P_1 \ldots P_K} < X \cdot \frac{z^K}{(K)!} \\
&\leq X\left(\frac{2e(u+1)}{K}\right)^K = \frac{X}{u^{2e(1+o(1))\log 2}} = o(X).
\end{aligned}$$

Hence, $t(n) \leq K \leq 2e\log u < u = g(X)$ holds for all $n \leq X$ with $o(X)$ exceptions. □

References

[1] A. Baker and H. Davenport, *The equations* $3x^2-2=y^2$ *and* $8x^2-7=z^2$, Quart. J. Math. Oxford Ser. (2) **20** (1969), 129–137.

[2] E. Brown, *Sets in which* $xy+k$ *is always a square*, Math. Comp. **45** (1985), 613–620.

[3] L. Caporaso, J. Harris and B. Mazur, *Uniformity of rational points*, J. Amer. Math. Soc. **10** (1997), 1–35.

[4] L. E. Dickson, *History of the Theory of Numbers, Vol. 2*, Chelsea, 1966, pp. 513–520.

[5] A. Dujella, *Generalization of a problem of Diophantus*, Acta Arith. **65** (1993), 15–27.

[6] A. Dujella, *On the size of Diophantine m-tuples*, Math. Proc. Cambridge Phil. Soc. **132** (2002), 23–33.

[7] A. Dujella, *There are only finitely many Diophantine quintuples*, J. Reine Angew. Math. **566** (2004), 183–214.

[8] A. Dujella, *Bounds for the size of sets with the property* $D(n)$, Glas. Mat. Ser. III **39** (2004), 199–205.

[9] P. Gibbs, *Some rational Diophantine sextuples*, `math.NT/9902081`, preprint.

[10] R. L. Graham, B. L. Rothschild, J. H. Spencer, *Ramsey Theory*, John Wiley, New York, 1980.

[11] E. Herrmann, A. Pethő and H. G. Zimmer, *On Fermat's quadruple equations*, Abh. Math. Sem. Univ. Hamburg **69** (1999), 283–291.

[12] A. Ivić, *The Riemann Zeta-Function, Theory and Applications*, Dover Publications, Mineola, New York, 2003.

[13] W. M. Schmidt, *Equations over Finite Fields: An Elementary Approach*, Springer-Verlag, Berlin, 1976.

[14] A. Schinzel, *Selected Topics on Polynomials*, University of Michigan Press, Ann Arbor, 1982.

[15] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge University Press, 1995.

[16] P. Túran, *On a Theorem of Hardy and Ramanujan*, J. London Math. Soc. **9** (1934), 274–276.

Andrej Dujella
Department of Mathematics, University of Zagreb, Bijenička cesta 30,
10000 Zagreb, Croatia
duje@math.hr

Florian Luca
Instituto de Matemáticas, Universidad Nacional Autónoma de México,
C.P. 58180, Morelia, Michoacán, México
fluca@matmor.unam.mx