On Diophantine square tuples

Katalin Gyarmati

Abstract

Diophantus' problem is generalized by estimating the cardinality of a set of non-zero squares, in which the difference between any two squares is also a square. Such a set with m elements is called a Diophantine square m-tuple. It is proved that there are infinitely many Diophantine square triples. Special cases of this problem are also studied, such as the fact that there is no Diophantine square triple containing only squares of Fibonacci numbers. For a set of integers $\mathcal{A}$, a non-trivial upper bound is given for the number of pairs (a, a') for which $a - a'$ is a square of a Fibonacci number. Some problems and conjectures are also formulated.

1 Introduction

Diophantus of Alexandria, a Greek mathematician, observed that the rational numbers $\frac{1}{16}$, $\frac{33}{16}$, $\frac{17}{4}$, and $\frac{105}{16}$ have the following property: the product of any two of them, increased by 1, is the square of a rational number. Later, Fermat found a set of four positive integers with this property: {1,3,8,120}. Euler found an infinite family of sets of this type, namely

2020 Mathematics Subject Classification: Primary: 11D09, 11D45 and 11B39.

Keywords and phrases: Diophantine equations, squares, Fibonacci numbers.

Research supported by the Hungarian National Research Development and Innovation Funds KKP133819.

$\{a, b, a+b+2r, 4r(r+a)(r+b)\}$, where $ab+1 = r^2$. These examples led Dujella to introduce the following definition:

Definition 1. *A set $A = \{a_1, a_2, \ldots, a_m\} \subset \mathbb{Z}^+$ is called a Diophantine m-tuple if $a_i a_j + 1$ is a perfect square for all $1 \leq i, j \leq m$.*

For a long time, it was a conjecture that there is an absolute bound for the size of a set with this property. The first result of this type was proved by Dujella in 2001: namely, that there is no Diophantine 8-tuple. He later improved this to the nonexistence of Diophantine sixtuples [8]. Since then, many other results have been proved. For example, in 2019, Stoll [27] proved that a certain family of rational Diophantine quadruples can be extended to rational quintuples in only one way. Finally, He, Togbé, and Ziegler [16] proved the following sharpest result:

Theorem A. *There exists no Diophantine quintuple.*

Dujella [9] provided an almost complete reference list of Diophantine m-tuples.

By Theorem A, Diophantus' original problem was essentially solved. However, other similar questions can be asked in this area. For example, Rivat, Sárközy and Stewart [24] proved the following:

Theorem B. *There exists an integer x_0 such that if $x_0 < x \in \mathbb{N}$, $\mathcal{A} \subset \{1, 2, 3, \ldots, x\}$ and for all $a, a' \in \mathcal{A}$, the sum $a + a'$ is a square, then*

$$|\mathcal{A}| < 37 \log x. \tag{1}$$

I do not know how sharp this theorem is. Lagrange [18] and Nicolas [22] found a 6-element set $\mathcal{A}$ satisfying this property:

$$\mathcal{A} = \{-15863902, 17798783, 21126338, 49064546, 82221218, 447422978\}.$$

In the present paper, I consider a related question: what can be said about the size of a set $\mathcal{A}$ of integers such that the difference between any two elements of it is a square? First of all, note that the smallest element can be assumed to be 0, since if the smallest element is a, then subtracting the same a from each element of the set does not change the differences. However, if the smallest element is 0, then the difference $b-0$ for any positive element b of the set is also a positive square, i.e., b itself is a square. So, I am looking for large sets of positive squares, in which the difference between any two elements is also a square. Since this definition closely resembles the original definition of Diophantine m-tuples, I call these sets Diophantine square tuples.

Definition 2. *A set $A = \{a_1^2, a_2^2, \ldots, a_m^2\} \subset \mathbb{Z}^+$ is called a Diophantine square m-tuple if $\left|a_i^2 - a_j^2\right|$ is a non-zero square for all $1 \leq i < j \leq m$.*

The following result is trivial: if a, b, c form a Pythagorean triple $a^2 + b^2 = c^2$, then b^2, c^2 form a Diophantine square pair (in other words, a 2-tuple). Thus, I know that there are infinitely many Diophantine square pairs. The following question is much more interesting: are there any Diophantine square triples, and if so, how many? Using computers to consider the interval $[1, 3000^2]$, I found the following triples of this type:

$(153^2, 185^2, 697^2)$	$(264^2, 520^2, 1105^2)$	$(264^2, 561^2, 1105^2)$
$(306^2, 370^2, 1394^2)$	$(448^2, 952^2, 1073^2)$	$(459^2, 555^2, 2091^2)$
$(495^2, 975^2, 1073^2)$	$(520^2, 533^2, 925^2)$	$(528^2, 1040^2, 2210^2)$
$(528^2, 1122^2, 2210^2)$	$(612^2, 740^2, 2788^2)$	$(644^2, 725^2, 2165^2)$
$(672^2, 680^2, 697^2)$	$(756^2, 765^2, 925^2)$	$(896^2, 1904^2, 2146^2)$
$(952^2, 1073^2, 1105^2)$	$(975^2, 1073^2, 1105^2)$	$(990^2, 1950^2, 2146^2)$
$(1040^2, 1066^2, 1850^2)$	$(1092^2, 2175^2, 2665^2)$	$(1344^2, 1360^2, 1394^2)$
$(1512^2, 1530^2, 1850^2)$	$(1540^2, 2431^2, 2665^2)$	$(1560^2, 1599^2, 2775^2)$
$(1904^2, 2146^2, 2210^2)$	$(1950^2, 2146^2, 2210^2)$	$(2016^2, 2040^2, 2091^2)$
$(2040^2, 2067^2, 2165^2)$	$(2268^2, 2295^2, 2775^2)$	$(2688^2, 2720^2, 2788^2)$

Among these 30 triples, there are 14 whose elements are coprimes. Below, I give a table with n in the first column, the number of Diophantine square triples in the interval $[1, n^2]$ in the second column, the number of such Diophantine square triples whose elements are coprime in the third column, and the proportion of the number of these coprime Diophantine square triples and $n^{1/2}$ in the fourth column (for the triples $(a^2, b^2, c^2) \subseteq [1, n^2]$ I suppose $a < b < c \leq n$).

n	# D. s. triples	# coprime D. s. triples	proportion
200000	4626	232	0.5188
400000	9438	334	0.5281
600000	14306	422	0.5448
800000	19170	468	0.5232
1000000	24030	510	0.51

The table above shows that the number of coprime Diophantine square triples is increasing and that their numbers appear to tend to $(0.5 + \varepsilon)n^{1/2}$. Another interesting fact is that the second and third columns contain only even numbers.

One of the first natural questions is whether there are infinitely many Diophantine square triples whose elements are coprimes. I will prove that the answer to this question is affirmative:

Theorem 1. *There are infinitely many Diophantine square triples* (a^2, b^2, c^2) *such that* $\gcd(a, b, c) = 1$.

Another important question is that of what happens if the size of the Diophantine square tuples increases. Through computer search, I found that there is no Diophantine square quadruple in the interval $[1, 10^{12}]$. Thus, I conjecture the following:

Conjecture 1. *There exists a positive integer n such that there is no Diophantine square n-tuple.*

I have not been able to prove this. Instead, I will prove a result which can be considered as a partial result in this direction. Specifically, I will estimate the size of the Diophantine square tuples in terms of the largest element of the set.

Theorem 2. *For every $\varepsilon > 0$, there exists an integer $x_0 = x_0(\varepsilon)$ such that if $x_0 < x$, $x \in \mathbb{N}$, $\mathcal{A} \subset \{1, 2, 3, \ldots, x\}$, and for all $a, a' \in \mathcal{A}$, $a > a'$, the difference $a - a'$ is a square, then*

$$|\mathcal{A}| < (1 + \varepsilon) \log x.$$

The proof of this theorem will be similar to the proof of Theorem B. I remark that in the case when $a^2 + a'^2$ is always a square in place of $a - a'$, a slightly sharper result can be obtained for the size of the set $\mathcal{A}$. Then in [5, Theorem 5] Bugeaud and I proved that $|\mathcal{A}| \leq 4(\log N)^{1/2}$ (however, the proof of this result uses methods that do not seem applicable to our case). The inverses of these problems were studied by Pintz, Steiger, and Szemerédi [23] and later by Bloom and Maynard [6], who studied sets in which the difference between two elements is never a square.

Probably, Conjecture 1 is a very difficult problem, but perhaps one can make it easier by allowing only certain special subsets of squares instead of all squares. The following statement can be proved easily:

Proposition 1. *There is no Diophantine square triple containing only squares of primes.*

The next question is whether we can replace the squares of primes with other sets. The answer to this question is not so clear. The Fibonacci sequence is a good example of this.

Theorem 3. *There is no Diophantine square triple consisting of squares of Fibonacci numbers.*

Related to Theorem 3, I mention that Fujita and Luca [12] proved that there are no Diophantine quadruples of Fibonacci numbers in the sense of the original Definition 1.

In my first paper on similar problems [14], I extended Diophantus' original problem to higher powers. Later, Bugeaud and Dujella [4] proved, among other things, that if $ac+1$, $ad+1$, $bc+1$ and $bd+1$ are all nth powers, then $n \leq 176$. In [5], Bugeaud and I provided non-trivial estimates for the cardinality of the set $\{(a,a') : a,a' \in A,\ aa'+1 = x^n,\ x \in \mathbb{N}\}$. Furthermore, we extended this result to the case $a-a'$ in [5, p. 1108].

Theorem C. *If $n \geq 3$ and $\mathcal{A} \subseteq \mathbb{Z}$, then at most $\dfrac{|\mathcal{A}|^2}{4}$ pairs (a,a') exist such that $a > a'$ and $a-a'$ is an nth power.*

The proof of Theorem C is a simple consequence of Fermat's theorem. An intriguing question is whether this theorem also holds for squares. I conjecture the following:

Conjecture 2. *There exists a constant $\varepsilon > 0$ such that for all $\mathcal{A} \subseteq \mathbb{Z}$, there exist at most $(1-\varepsilon)\dfrac{|\mathcal{A}|^2}{2}$ pairs (a,a') for which $a-a'$ is a square.*

If Conjecture 1 holds, it is a simple consequence of Turán's theorem [28]. Unfortunately, I was unable to prove Conjecture 2 unconditionally. However, if I consider only the squares of Fibonacci numbers, this conjecture is true and can be improved.

Theorem 4. *If $\mathcal{A} \subseteq \mathbb{Z}$, then at most $|\mathcal{A}|^{3/2} + |\mathcal{A}|$ pairs (a,a') exist such that $a-a'$ is a square of a (positive) Fibonacci number.*

Finally, I note that I can improve Theorem C if I assume certain old conjectures to be true. Euler conjectured the following: for all integers n and k greater than 1, if the sum of n pieces of positive kth powers is itself a kth power, then $n \geq k$. This conjecture was disproved later for $k = 4$ and

5. Lander, Parkin, and Selfridge [19] formulated the following conjecture in 1976: if $\sum_{i=1}^{n} a_i^k = \sum_{j=1}^{m} b_j^k$, where $a_i \neq b_j$ are positive integers for all $1 \leq i \leq n$ and $1 \leq j \leq m$, then $m + n \geq k$. I only need the case of $m = 2$, $n = 2$ in this conjecture:

Conjecture 3. *If $n \geq 5$ is an integer, there are no four positive integers a, b, c, d for which*

$$a^n + b^n = c^n + d^n$$

and $\{a, b\} \neq \{c, d\}$.

Note that the special case $n = 5$ has been conjectured by Erdős. If Conjecture 3 holds, the result of Theorem C can be further improved:

Theorem 5. *If Conjecture 3 holds, then for $n \geq 5$ and $\mathcal{A} \subseteq \mathbb{Z}$, at most $|\mathcal{A}|^{3/2} + |\mathcal{A}|$ pairs (a, a') exist such that $a > a'$ and $a - a'$ is an nth power.*

2 Proofs

Proof of Proposition 1. If (p_1^2, p_2^2, p_3^2) is a Diophantine square triple, where $p_1 < p_2 < p_3$ are different primes, then

$$\sqrt{p_2^2 - p_1^2},\ p_1^2,\ p_2^2$$

is a primitive Pythagorean triple. The primitive Pythagorean triples have a parametric form, and from this, we know that there exist $u, v \in \mathbb{Z}^+$, $u > v$, $u \not\equiv v \pmod 2$, $(u, v) = 1$ such that

$$\sqrt{p_2^2 - p_1^2} = u^2 - v^2, \quad p_1 = 2uv, \quad p_2 = u^2 + v^2 \tag{2}$$

or

$$\sqrt{p_2^2 - p_1^2} = 2uv, \quad p_1 = u^2 - v^2, \quad p_2 = u^2 + v^2. \tag{3}$$

In Case (2), p_1 is an even prime, so $p_1 = 2$, and thus $u = v = 1$, which contradicts $u > v$. In Case (3),

$$p_1 = u^2 - v^2 = (u - v)(u + v),$$

where $1 \leq u - v < u + v$. Since p_1 is prime, we have $u - v = 1$ and $u + v = p_1$. Then $u = \dfrac{p_1 + 1}{2}$ and $v = \dfrac{p_1 - 1}{2}$. Thus

$$\begin{aligned} p_2 = u^2 + v^2 &= \left(\frac{p_1 + 1}{2}\right)^2 + \left(\frac{p_1 - 1}{2}\right)^2 \\ &= \frac{p_1^2 + 1}{2}. \end{aligned}$$

The argument above applies to not only the Pythagorean triple $\sqrt{p_2^2 - p_1^2}, p_1, p_2$ but also the Pythagorean triple $\sqrt{p_3^2 - p_1^2}, p_1, p_3$. Thus, we similarly get

$$p_3 = \frac{p_1^2 + 1}{2}.$$

Then $p_2 = p_3$ which is a contradiction.

Proof of Theorem 1. It is easy to see there exists at least one Diophantine square triple (a_1^2, b_1^2, c_1^2), by giving a simple example, e.g., we may consider the first triple in the table after Definition 2: $a_1 = 153, b_1 = 185, c_1 = 697$ (here all integers are odd). Next, we construct infinitely many Diophantine square triples (a_i, b_i, c_i) by a simple recursion. Assume that for some $i \in \mathbb{N}$, the Diophantine square triple a_i, b_i, c_i is already given. Then, if $a_i + b_i + c_i$ is odd, let

$$\begin{aligned} a_{i+1} &= \left|a_i^2 + b_i^2 - c_i^2\right| \\ b_{i+1} &= \left|a_i^2 - b_i^2 + c_i^2\right| \\ c_{i+1} &= \left|-a_i^2 + b_i^2 + c_i^2\right|. \end{aligned} \tag{4}$$

If $a_i + b_i + c_i$ is even, let

$$\begin{aligned} a_{i+1} &= \frac{1}{2}\left|a_i^2 + b_i^2 - c_i^2\right| \\ b_{i+1} &= \frac{1}{2}\left|a_i^2 - b_i^2 + c_i^2\right| \\ c_{i+1} &= \frac{1}{2}\left|-a_i^2 + b_i^2 + c_i^2\right|. \end{aligned} \tag{5}$$

Clearly, $a_i, b_i, c_i \in \mathbb{N}$.

First, consider the case in which $a_i + b_i + c_i$ is odd. Then,

$$\begin{aligned} b_{i+1}^2 - a_{i+1}^2 &= (a_i^2 - b_i^2 + c_i^2)^2 - (a_i^2 + b_i^2 - c_i^2)^2 \\ &= 4a_i^2(c_i^2 - b_i^2). \end{aligned} \tag{6}$$

Since (a_i, b_i, c_i) is a Diophantine square triple, $c_i^2 - b_i^2$ is a square; thus, $b_{i+1}^2 - a_{i+1}^2$ is a square. Similarly,

$$\begin{aligned} c_{i+1}^2 - a_{i+1}^2 &= 4b_i^2(c_i^2 - a_i^2) \\ c_{i+1}^2 - b_{i+1}^2 &= 4c_i^2(b_i^2 - a_i^2), \end{aligned} \tag{7}$$

and these are also squares. On the other hand, $\gcd(a_{i+1}, b_{i+1}, c_{i+1}) = 1$ since if we denote $\gcd(a_{i+1}, b_{i+1}, c_{i+1})$ by d,

$$\begin{aligned} &d \mid a_i^2 + b_i^2 - c_i^2,\ a_i^2 - b_i^2 + c_i^2,\ -a_i^2 + b_i^2 + c_i^2 \\ &d \mid a_i^2 + b_i^2 - c_i^2 + a_i^2 - b_i^2 + c_i^2 - a_i^2 + b_i^2 + c_i^2 \\ &d \mid a_i^2 + b_i^2 + c_i^2 \\ &d \mid a_i^2 + b_i^2 + c_i^2 - (a_i^2 + b_i^2 - c_i^2) \\ &d \mid 2c_i^2. \end{aligned}$$

Similarly, $d \mid 2a_i^2$ and $d \mid 2b_i^2$. Thus $d \mid 2(a_i^2, b_i^2, c_i^2) = 2$. However, d is odd since $d \mid a_i^2 + b_i^2 + c_i^2$, and $a_i^2 + b_i^2 + c_i^2$ is odd. It follows that $d = 1$. Thus, we have proved that $(a_{i+1}, b_{i+1}, c_{i+1})$ is a Diophantine square tuple, whose elements are coprime.

The case with an even $a_i + b_i + c_i$ is very similar. I will leave the details of the proof of the even case to the reader.

It remains to prove that this recursion gives infinitely many Diophantine square triples. Indeed, let $a_1 = 153$, $b_1 = 185$ and $c_1 = 697$. Then, it is easy to see that the recursion defined above always produces Diophantine square triples such that their elements are odd; thus, we always use recursion (4).

However, then $c_i < c_{i+1}$ since

$$c_{i+1} = \left| -a_i^2 + b_i^2 + c_i^2 \right|,$$

but by $a_i \leq b_i$, $1 < c_i$, and here, $-a_i^2 + b_i^2 + c_i^2$ is positive. Hence

$$c_{i+1} = -a_i^2 + b_i^2 + c_i^2 \geq c_i^2 > c_i.$$

Thus, $c_0 < c_2 < c_3 < \dots$ such that the recursion (4) gives infinitely many Diophantine square triples. This completes the proof of Theorem 1.

Proof of Theorem 2. The main idea of the proof is to use Gallagher's larger sieve [24]. This version of the larger sieve was presented by Erdős, Sárközy and Stewart [10] in 1994.

Lemma 1 (Gallagher's larger sieve)**.** *Suppose that $m, n \in \mathbb{N}$, $\mathcal{A} \subset \{m + 1, m + 2, \dots, m + n\}$ and $\mathcal{B} \subset \mathbb{N}$ is a finite set such that its elements are pairwise coprime. For all $b \in \mathcal{B}$, denote the number of residue classes* mod b *that intersect $\mathcal{A}$ by $\nu(b)$. Then,*

$$|\mathcal{A}| \leq \frac{\sum\limits_{b \in \mathcal{B}} \log b - \log n}{\sum\limits_{b \in \mathcal{B}} \dfrac{\log b}{\nu(b)} - \log n}, \tag{8}$$

provided that the denominator is positive.

We have assumed that for all $a, a' \in \mathcal{A},\ a > a'$, the difference $a - a'$ is always a square. Thus, for all $a, a' \in \mathcal{A},\ a > a'$ and prime p, the difference $a - a'$ is either a quadratic residue $\mod p$ or 0. If -1 is a quadratic residue mod p, then $a' - a$ is also a quadratic residue $\mod p$ or 0. Thus, if $p \equiv 1 \pmod 4$, then for all $a, a' \in \mathcal{A}$, the difference $a - a'$ is either a quadratic residue $\mod p$ or 0 (and the condition $a > a'$ is no longer needed).

We will also need the following lemma of Hanson and Petridis [15, Corollary 1.5].

Lemma 2. *Let p be a prime. If $\mathcal{C} \subset \mathbb{Z}_p$ is a set such that for all $a, a' \in \mathcal{C}$, $a \neq a'$ the difference $a - a'$ is quadratic residue modulo p, then*

$$|\mathcal{C}| \leq \sqrt{p/2} + 1. \tag{9}$$

(Note that this lemma also has a slightly weaker version in which Inequality (9) is replaced by

$$|\mathcal{C}| < \sqrt{p}. \tag{10}$$

The proof of (10) is very simple: if n is a fixed quadratic non-residue, then the differences $a - na'$ with $a, a' \in C$ are all different. This result was slightly improved in [1], which was the best result for a long time, until Hanson and Petridis proved Lemma 2. The application of (10) in the proof of Theorem 2 would lead to the slightly weaker upper bound of

$$|\mathcal{A}| < (2 + o(1)) \log x$$

for $|\mathcal{A}|$.)

Let us return to the proof of Theorem 2. Let $\mathcal{C}$ denote the set of $\mod p$ residue classes that contain an element from $\mathcal{A}$. Then, by Lemma 2, $|\mathcal{C}| \leq \sqrt{p/2} + 1$. Thus, using the notation of Lemma 1, we get $\nu(p) \leq \sqrt{p/2} + 1$. We then use Gallagher's larger sieve. For this, let

$$\mathcal{B} = \{p :\ p \text{ is a prime},\ p \equiv 1 \pmod 4,\ 2 \leq p \leq c(\log x)^2\}.$$

The value of the constant c with $c > 1$ will be fixed later. Using Gallagher's larger sieve yields

$$|\mathcal{A}| \le \frac{\sum\limits_{p\equiv 1 \pmod 4,\ p\le c(\log x)^2} \log p - \log x}{\sum\limits_{p\equiv 1 \pmod 4,\ p\le c(\log x)^2} \dfrac{\log p}{\sqrt{p/2}+1} - \log x}. \tag{11}$$

Now, we estimate the value of the expression on the right. The following sums run over positive primes:

$$\pi(y,4,1) \stackrel{\text{def}}{=} \sum_{p\equiv 1 \pmod 4,\ p\le y} 1$$
$$\theta(y,4,1) \stackrel{\text{def}}{=} \sum_{p\equiv 1 \pmod 4,\ p\le y} \log p$$

We also introduce the following notation:

$p_n(4,1)$ denotes the nth smallest positive prime p with $p \equiv 1 \pmod 4$.

Bennett, Martin, O'Bryan and Rechnitzer [2] proved the following results:

$$\begin{aligned}
\pi(y,4,1) &= (1+o(1))\frac{y}{2\log y} && \text{see Theorem 1.4 in [2]},\\
\theta(y,4,1) &= (1+o(1))\frac{y}{2} && \text{see Corollary 1.7 in [2]},\\
p_n(4,1) &= (1+o(1))2n\log n && \text{see Theorem 1.5 in [2]}.
\end{aligned}$$

It follows from the estimate of $\theta(y,4,1)$ that for the expression in the numerator of the fraction in (11), we have

$$\sum_{p\equiv 1 \pmod 4,\ p\le c(\log x)^2} \log p - \log x = \theta(c(\log x)^2, 4, 1) - \log x$$

$$= (1+o(1))\frac{c}{2}(\log x)^2 - \log x$$

$$= (1+o(1))\frac{c}{2}(\log x)^2. \tag{12}$$

The estimate of the denominator of the fraction (11) is more complicated:

$$\sum_{p\equiv 1 \pmod 4,\ p\le c(\log x)^2} \frac{\log p}{\sqrt{p/2}+1} - \log x$$

$$= \sum_{n=1}^{\pi(c(\log x)^2,4,1)} \frac{\log p_n(4,1)}{\sqrt{p_n(4,1)/2}+1} - \log x$$

$$= (1+o(1)) \sum_{n=1}^{\pi(c(\log x)^2,4,1)} \frac{\log(2n\log n)}{\sqrt{n\log n}} - \log x$$

$$= (1+o(1)) \sum_{n=1}^{\pi(c(\log x)^2,4,1)} \frac{\log n}{\sqrt{n\log n}} - \log x$$

$$= (1+o(1)) \sum_{n=1}^{\pi(c(\log x)^2,4,1)} \frac{\sqrt{\log n}}{\sqrt{n}} - \log x$$

$$= (1+o(1)) \int_{n=1}^{\pi(c(\log x)^2,4,1)} \frac{\sqrt{\log n}}{\sqrt{n}} dn - \log x$$

$$= (1+o(1))2\sqrt{n\log n}\Big]_1^{\pi(c(\log x)^2,4,1)} - \log x$$

$$= (1+o(1))2\sqrt{\pi(c(\log x)^2,4,1)\log\left(\pi(c(\log x)^2,4,1)\right)} - \log x$$

$$= (1+o(1))2\sqrt{\frac{c(\log x)^2}{2\log\left(c(\log x)^2\right)}\log\left(\frac{c(\log x)^2}{2\log\left(c(\log x)^2\right)}\right)} - \log x$$

$$= (1+o(1))2\sqrt{\frac{c(\log x)^2}{4\log\log x}2\log\log x} - \log x$$

$$= (1+o(1))\sqrt{2c}\log x - \log x$$

$$= (1+o(1))(\sqrt{2c}-1)\log x.$$

Using this estimate and (12) in (11) yields the following:

$$|\mathcal{A}| \le (1+o(1))\frac{c}{2\left(\sqrt{2c}-1\right)}\log x \tag{13}$$

if $c > 1$. By choosing $c = 2$ we get

$$|\mathcal{A}| \le (1+\varepsilon)\log x \quad \text{for } x > x_0(\varepsilon),$$

which was to be proved.

Throughout the proofs of Theorems 3 and 4, $F_0 = 0, F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3, F_5 = 5, \ldots$ will denote the Fibonacci numbers.

Proof of Theorem 3. The following lemma is true for Fibonacci numbers:

Lemma 3. *If* $0 < m < n$, $m \equiv n \pmod 2$ *and* $F_n^2 - F_m^2$ *is a square, then*

$$(F_m, F_n) = (F_5, F_7) = (5, 13).$$

In 1979, Bicknell-Johnson [3] began the study of primitive Pythagorean triples in which Fibonacci numbers occur. Indeed, the basic idea in the proof of Lemma 3 is already included in [3]. However, in order to complete the proof of Bicknell-Johnson, we need Lemma 5, which was originally missing and was not proved until much later, in 1998, by McDaniel and Ribenboim [21]. Zhang and Togbé [29] generalized Lemma 3 to higher powers, but only if $\gcd(F_m, F_n) = 1$ holds. Now we will also need the case $\gcd(F_m, F_n) > 1$.

The theorem immediately follows from Lemma 3 since if (F_ℓ^2, F_m^2, F_n^2) is a Diophantine square triple, then based on the pigeon-hole principle, there are two elements of the set $\{\ell, m, n\}$ which are congruent modulo 2, say $m \equiv n \pmod 2$. Then, by applying the lemma, we get $F_m = 5$ and $F_n = 13$ (or vice versa). There are two Pythagorean triples that contain the number 5, namely $(3, 4, 5)$ and $(5, 12, 13)$. Since $\sqrt{|F_\ell^2 - F_m^2|}, F_\ell, F_m$ is a Pythagorean triple, where $F_m = 5$, F_ℓ can only be 3, 4 or 12. Among these numbers, only 3 is a Fibonacci number, so $F_\ell = 3$. However, then $(F_\ell^2, F_m^2, F_n^2) = (3^2, 5^2, 13^2)$ does not form a Diophantine square triple, since $13^2 - 3^2$ is not a square. Thus, in order to prove Theorem 3, we only need to prove Lemma 3.

Proof of Lemma 3. The following identity is due to Ruggles [25]: for all integers n, p, we have

$$F_{n+p}^2 - F_{n-p}^2 = F_{2n}F_{2p}.$$

From this, we immediately get the following:

Lemma 4. *If* $0 < m \leq n$ *and* $0 < m \equiv n \pmod 2$, *then*

$$F_n^2 - F_m^2 = F_{n+m}F_{n-m} \tag{14}$$

By the conditions of Lemma 3, $F_n^2 - F_m^2$ is a square. It is known that for the greatest common divisor of Fibonacci numbers, we have

$$\gcd(F_a, F_b) = F_{\gcd(a,b)}. \tag{15}$$

Indeed, (15) was stated by François Édouard Anatole Lucas in 1876, as Knuth writes in his book [17]. The source where Lucas published (15) is being sought.

By this, if $d \stackrel{\text{def}}{=} \gcd(n-m, n+m)$, then

$$\gcd(F_{n+m}, F_{n-m}) = F_d,$$

and by (14), we get

$$\frac{F_n^2 - F_m^2}{F_d^2} = \frac{F_{n+m}}{F_d} \cdot \frac{F_{n-m}}{F_d}.$$

Here $\frac{F_{n+m}}{F_d}$ and $\frac{F_{n-m}}{F_d}$ are coprime, and their product is a square, so both of them are squares.

McDaniel and Ribenboim [21] proved the following:

Lemma 5. *Assume u, v and y are positive integers such that $u \mid v$ and $\frac{F_v}{F_u} = y^2$. Then, either $u = v$ or $(v, u) \in \{(12, 1), (12, 2), (2, 1), (6, 3)\}$.*

Now, $\frac{F_{n+m}}{F_d}$ is a square, and $d = \gcd(n+m, n-m) \le n - m < n + m$. Thus, by Lemma 5, we get $(n+m, d) \in \{(12, 1), (12, 2), (2, 1), (6, 3)\}$. Then,

$$n, m < n + m \le 12.$$

Using a Python program, it is easy to check that among the first 12 Fibonacci numbers, there are two pairs for which $F_n^2 - F_m^2$ is a square, namely

$$(F_2, F_5) = (3, 5) \quad \text{and} \quad (F_5, F_7) = (5, 13).$$

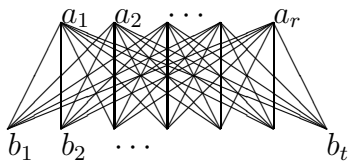
By the condition of the lemma, $n \equiv m \pmod 2$, so the only pair satisfying the lemma is $(F_5, F_7) = (5, 13)$. This completes the proof of Lemma 3.

We note that, similarly to the proof of Lemma 3, if $0 < n \not\equiv m \pmod 2$, then using the identity $F_m^2 + F_n^2 = F_{n-m}F_{n+m}$ (this identity can be found e.g., in [26]), we can prove the following:

Lemma 6. *If $0 < m \le n$ and $m \not\equiv n \pmod 2$, then $F_m^2 + F_n^2$ is never a square.*

Proof of Theorem 4. The proof is based on graph theory. Bugeaud and I proved the following in [5]:

Lemma 7. *Assume that $G(V_1, V_2)$ is a bipartite graph with $|V_1| = n \le |V_2| = m$, and the vertices are labeled by positive real numbers. Suppose that $G(V_1, V_2)$ does not contain a G_0 subgraph $K_{r,t}$*

$G_0 =$ 

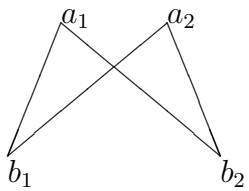
with $a_i < b_j$ for all $1 \le i \le r$, $1 \le j \le t$ (where the a's belong to V_1 and the b's belong to V_2 or vice versa). Then G has at most

$$e(G) \le 2(t-1)^{1/r} m n^{1-1/r} + 2(r-1)m$$

edges.

We only need the special case of this lemma for which $r = t = 2$, but graph G is now a simple graph (i.e., not necessarily a bipartite graph).

Lemma 8. *Assume that $G(V)$ is a graph, in which V denotes the set of the vertices, and in V, the vertices are labeled by positive real numbers. Suppose that $G(V)$ does not contain a G_0 subgraph $K_{2,2}$*

$G_0 =$ 

with $a_i < b_j$ *for all* $1 \leq i \leq 2$, $1 \leq j \leq 2$. *Then* G *has at most*

$$e(G) \leq |V|^{3/2} + |V|$$

edges.

Proof of Lemma 8. The proof of Lemma 8 is based on Lemma 7. In order to use Lemma 7, we need to consider a bipartite graph. Let the graph $G(V)$ be the graph given in Lemma 8, and let us define the bipartite graph $\widetilde{G}(V_1, V_2)$ as follows: $V_1 = V_2 = V$ and $a \in V_1$, $b \in V_2$ are connected by an edge in $\widetilde{G}$ if and only if $a \neq b \in V$ are connected by an edge in G. Then,

$$e(\widetilde{G}) = 2e(G),$$

since if (a, b) is an edge in G, then (a, b) (where $a \in V_1,\ b \in V_2$) and (b, a) (where $b \in V_1,\ a \in V_2$) are edges in $\widetilde{G}$. If G does not contain the subgraph $G_0 = K_{2,2}$ described in Lemma 8, then $\widetilde{G}$ also does not contain the subgraph $G_0 = K_{2,2}$ described in Lemma 7. Thus, by using Lemma 7, we get

$$2e(G) = e(\widetilde{G}) \leq 2|V|^{3/2} + 2|V|,$$

which, when divide by 2, gives us the statement of the lemma.

We are now ready to complete the proof of Theorem 4. Let us define graph $\mathcal{G}$ whose vertices are the elements of the finite set $\mathcal{A} \subset \mathbb{Z}$, and $a, a' \in \mathcal{A}$ are connected by an edge if $|a - a'| = F_n^2$ with $n \geq 1$. We will prove that this graph does not contain the subgraph $G_0 = K_{2,2}$ described in Lemma 8. Indeed, we will assume that this statement is false and that graph G contains the graph $G_0 = K_{2,2}$. We denote the vertices of this $K_{2,2}$ by $a_1, a_2, b_1.b_2$ (where $a_i < b_j$ for all $1 \leq i \leq 2$, $1 \leq j \leq 2$). For symmetry reasons, we can also assume

that $a_1 < a_2$, $b_1 < b_2$. Then,

$$\begin{aligned} b_1 - a_1 &= F_k^2 \\ b_2 - a_2 &= F_\ell^2 \\ b_1 - a_2 &= F_m^2 \\ b_2 - a_1 &= F_n^2, \end{aligned}$$

where $k, \ell, m, n \geq 1$. Then $F_n^2 = b_2 - a_1$ is the greatest among $F_k^2, F_\ell^2, F_m^2, F_n^2$, i.e., $n > k, \ell, m$. Furthermore,

$$F_k^2 + F_\ell^2 = F_m^2 + F_n^2 = b_1 + b_2 - a_1 - a_2.$$

Let $s = \max\{k, \ell\}$ and $t = \min\{k, \ell\}$. Then $t \leq s < n$. Now

$$F_s^2 + F_t^2 = F_m^2 + F_n^2. \tag{16}$$

Next, we distinguish three cases according to $t < s$ or $2 \leq t = s$ or $1 = t = s$.

If $t < s$, then

$$\begin{aligned} F_s^2 + F_t^2 &\leq F_s^2 + F_{s-1}^2 < (F_s + F_{s-1})^2 \\ &= F_{s+1}^2 \leq F_n^2 < F_m^2 + F_n^2, \end{aligned}$$

which is a contradiction.

If $2 \leq t = s$, then it is easy to see that $2F_s^2 < F_{s+1}^2$. Indeed, the following equations are equivalent:

$$\begin{aligned} 2F_s^2 &< F_{s+1}^2 \\ 2F_s^2 &< (F_s + F_{s-1})^2 \\ F_s(F_s - F_{s-1}) &< F_{s-1}(F_s + F_{s-1}) \\ F_s F_{s-2} &< F_{s-1} F_{s+1}. \end{aligned}$$

The last equation holds by $F_s F_{s-2} < F_{s+1} F_{s-2} \leq F_{s+1} F_{s-1}$. Thus,

$$\begin{aligned} F_s^2 + F_t^2 &= 2F_s^2 < F_{s+1}^2 \\ &\leq F_n^2 < F_m^2 + F_n^2, \end{aligned}$$

which is a contradiction.

Finally, if $s = t = 1$, then by (16), $2 = F_m^2 + F_n^2$, so $m = n = 1$, but this contradicts $n > s$.

Thus our assumption always leads to a contradiction, so that our statement is true: the graph G does not contain a $K_{2,2}$ described in Lemma 8. So, by applying Lemma 8, we get

$$\begin{aligned} e(\mathcal{G}) &= |\{(a, a') : \ a, a' \in \mathcal{A}, \ a - a' = F_n^2, \ n \geq 1\}| \\ &\leq |V|^{3/2} + |V| = |\mathcal{A}|^{3/2} + |\mathcal{A}|. \end{aligned}$$

This completes the proof of Theorem 4.

Proof of Theorem 5. Similarly to the proof of Theorem 4, we now define a graph G whose vertices are the elements of $\mathcal{A}$ and show that $a, a' \in \mathcal{A}$ is connected by an edge if $|a - a'|$ is an nth power. This graph does not contain a $K_{2,2}$ described in Lemma 8; if it does contain a $K_{2,2}$ graph, we can denote its vertices by a_1, a_2, b_1, b_2, where $a_i < b_j$ for all $1 \leq i \leq 2$, $1 \leq j \leq 2$. Then

$$\begin{aligned} b_1 - a_1 &= a^n \\ b_2 - a_2 &= b^n \\ b_1 - a_2 &= c^n \\ b_2 - a_1 &= d^n, \end{aligned}$$

where a, b, c, d are positive integers and

$$a^n + b^n = c^n + d^n = b_1 + b_2 - a_1 - a_2.$$

Furthermore, $\{a^n, b^n\} = \{b_1 - a_1, b_2 - a_2\} \neq \{b_1 - a_2, b_2 - a_1\} = \{c^n, d^n\}$, but this contradicts Conjecture 3. Thus, G does not contain a $K_{2,2}$ described in Lemma 8. Using Lemma 8, we get the statement of the theorem.

3 Further problems and conjectures

I have not been able to determine all Diophantine square triples; Theorem 1 states only that there are infinitely many of them. Perhaps, similarly to the case of Pythagorean triples, there is a parametric way to construct all Diophantine square triples. Thus, the following question can be asked:

Problem 1. *Does a parametric system of equations exist that describes all Diophantine square triples?*

Related to the proof of Theorem 1, the following easier question also arises:

Problem 2. *Is there a finite set of coprime Diophantine square triples from which all coprime Diophantine square triples can be obtained using only recursions (4) and (5), or does this statement hold only if it includes further recursions?*

First, I conjectured that there is no Diophantine square triple (a^2, b^2, c^2) whose elements are pairwise coprime. I found this conjecture to be false when I ran a Python program and obtained only one such Diophantine square triple in the $[1, 10^{12}]$ interval, namely

$$(40920^2,\ 41449^2,\ 42601^2).$$

Related to this, I ask the following question:

Problem 3. *Are there infinitely many Diophantine square triples whose elements are pairwise coprime?*

The following question is also important: does Conjecture 1 hold for $n = 4$?

Conjecture 4. *There is no Diophantine square quadruple.*

I checked Conjecture 4 in the the interval $[1, 10^{12}]$ with a Python program and did not find a Diophantine square quadruple (a^2, b^2, c^2, d^2) with $a, b, c, d \leq 10^6$.

In Theorem 3, I proved that squares of Fibonacci numbers cannot form a Diophantine square triple. In place of Fibonacci numbers, we may study other sequences given by linear recursions. For example, we can study whether the following conjecture is true?

Conjecture 5. *There is no binary linear recurrence sequence with integer elements and coefficients that has four elements such that their squares form a Diophantine square quadruple.*

With three elements instead of four, the conjecture is no longer true, since, e.g., in the Diophantine square triple $(153^2, 185^2, 697^2)$ we have $\gcd(153, 185) = 1$. So the linear Diophantine equation $153a + 185b = 697$ can be solved in the range of integers, e.g., $a = 2019, b = -1666$ is a solution. Then the squares of the first three elements of the linear recursion $x_1 = 153, x_2 = 185$ and $x_n = -1666x_{n-1} + 2019x_{n-2}$ give $(153^2, 185^2, 697^2)$, which is a Diophantine square triple.

Of course, if Conjecture 4 is true, so is Conjecture 5 with four elements.

Luca, Fuchs and Szalay studied a slightly different problem in [11] and proved that there exists no a, b, c, for which $ab + 1$, $ac + 1$ and $bc + 1$ are elements of a non-degenerate binary linear recurrence sequence.

Problems similar to Lemmas 3 and 6 were studied by Luca and Patel in [20], where they give all pairs (n, m) for which $F_n + F_m$ or $F_n - F_m$ is a power and $n \equiv m \pmod 2$. Later, Ziegler [30] extended their result to the case $n \not\equiv m \pmod 2$ as well, i.e., he determined all pairs (n, m) for which $F_n + F_m$ or $F_n - F_m$ is a power.

In our case, it is easy to check with a computer program that among the first 1000 positive Fibonacci numbers, there are only two pairs (F_m, F_n) for which $F_n^2 - F_m^2$ is a non-zero square. These pairs are

$$(F_2, F_5) = (3, 5) \quad \text{and} \quad (F_5, F_7) = (5, 13).$$

Furthermore, among the first 1000 positive Fibonacci numbers, there is no pair (F_m, F_n) for which $F_n^2 + F_m^2$ is a square. Based on this, the following conjecture can be formulated, which Bicknell-Johnson already stated in [3], although her paper seems to study only the case $\gcd(F_n, F_m) = 1$:

Conjecture 6. *There are only two pairs (F_m^2, F_n^2) which contain squares of Fibonacci numbers with $0 < m < n$, and such that $F_n^2 - F_m^2$ is a square. These pairs are*

$$(F_2, F_5) = (3, 5) \quad \text{and} \quad (F_5, F_7) = (5, 13).$$

Furthermore, there is no pair (F_m^2, F_n^2) with $0 < m, n$, such that $F_n^2 + F_m^2$ is a square.

In the present paper, the two statements in the conjecture were proved in certain special cases depending on the parity of n and m (see Lemmas 3 and 6), but I have not been able to prove them in general.

An interesting question is whether Theorem 4 can be improved. I think it is very likely that the following holds:

Conjecture 7. *For all positive real number $\varepsilon > 0$, there exists a constant $c(\varepsilon)$ such that if $\mathcal{A} \subseteq \mathbb{Z}$, then at most $c(\varepsilon)|\mathcal{A}|^{1+\varepsilon}$ pairs (a, a') exist, for which $a - a'$ is a square of a non-zero Fibonacci number.*

I wrote the computer programs used in this paper in Python. I also used modern translation systems, such as Google, QuillBot.com along with Magnum Proofreading Services.

Acknowledgments: The author would like to thank András Sárközy and the referees for their valuable advice.

References

[1] C. Bachoc, M. Matolcsi, I. Z. Ruzsa, Squares and difference sets in finite fields. Integers 13 (2013), Paper No. A77, 5 pp.

[2] M. A. Bennett, G. Martin, K. O'Bryant, A. Rechnitzer, Explicit bounds for primes in arithmetic progressions, *Illinois J. Math.* **62** (2018), 427–532.

[3] M. Bicknell-Johnson, Pythagorean triples containing Fibonacci numbers: solutions for $F_n^2 \pm F_k^2 = K^2$, *Fibonacci Quart.* **17**(1) (1979), 1-–12.

[4] Y. Bugeaud, A. Dujella, On a problem of Diophantus for higher powers, *Math. Proc. Cambridge Philos. Soc.* **135** (2003), 1–10.

[5] Y. Bugeaud, K. Gyarmati, On generalizations of a problem of Diophantus, *Illinois J. Math.* **48** (2004), 1105–1115.

[6] T.F. Bloom,, J. Maynard, A new upper bound for sets with no square differences, Compos. Math. 158(8) (2022), 1777–1798.

[7] A. Dujella, Diophantine m-tuples and elliptic curves, *J. Theor. Nombres Bordeaux* **13** (2001) 111–124.

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

[9] A. Dujella, Diophantine m-tuples, References (chronologically), https://web.math.pmf.unizg.hr/~duje/dtuples.html.

[10] P. Erdős, A. Sárközy, C.L. Stewart, On prime factors of subset sums, *J. London Math. Soc.* **49**(2) (1994) 209–218.

[11] C. Fuchs, F. Luca, L. Szalay, Diophantine triples with values in binary recurrences, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. 7* **5** (2008), 579–608.

[12] Y. Fujita, F. Luca, On diophantine quadruples of Fibonacci numbers, *Glas. Mat.* **52**(2) (2017), 221–234.

[13] P. X. Gallagher, A larger sieve, *Acta Arith.* **18** (1971) 77–81.

[14] K. Gyarmati, On a problem of Diophantus, Acta Arith. **97**(1) (2001), 53–65.

[15] B. Hanson, G. Petridis, Refined estimates concerning sumsets contained in the roots of unity, *Proc. London Math. Soc.* **122**(3) (2021), 353–358.

[16] B. He, A. Togbé, V. Ziegler, There is no Diophantine quintuple, *Trans. Amer. Math. Soc.* **371** (2019) 6665–6709.

[17] D. Knuth, The Art of Computer Programming, volume, Fundamental Algorithms, (3rd ed.), Addison-Wesley Professional, 1997.

[18] J. Lagrange, Six entiers dont les sommes deux à deux sont des carrés, *Acta Arith.* **40** (1981) 91–96.

[19] L. J. Lander, T. R. Parkin, J. L. Selfridge, A survey of equal sums of like powers, *Math. Comput.* **21**(99) (1967), 446-–459.

[20] F. Luca, V. Patel, On perfect powers that are sums of two Fibonacci numbers, *J. Number Theory* **189** (2018), 90—96.

[21] W. L. McDaniel, P. Ribenboim, Square-classes in Lucas sequences having odd parameters, *J. Number Theory* **73** (1998), 14-–27.

[22] J.-L. Nicolas, Six nombres dont les sommes deux à deux sont des carrés, *Calculateuren Math. (1975, Limoges)*, *Bull. Soc. Math. France*, Mémoire **49–50** (1977), 141–143.

[23] J. Pintz, W.L. Steiger, E. Szemerédi, On Sets of Natural Numbers Whose Difference Set Contains No Squares, J. Lond. Math. Soc. s2-37(2) (1988), 219–231.

[24] J. Rivat, A. Sárközy and C.L. Stewart, Congruence properties of the Omega-function on sumsets, *Illinois J. Math.* **43**(1) (1999), 1–18.

[25] I. D. Ruggles, Some Fibonacci results using Fibonacci-type sequences **1**(2) (1963), 75–80.

[26] B. Sharpe, On Sums $F_x^2 \pm F_y^2$, Fib. Quart **3**(1) (1965), p. 63.

[27] M. Stoll, Diagonal genus 5 curves, elliptic curves over $\mathbb{Q}(t)$, and rational Diophantine quintuples, *Acta Arith.* **190** (2019), 239–261.

[28] P. Turán, On an extremal problem in graph theory, *Matematikai és Fizikai Lapok* **48** (1941), 436–452.

[29] Z. Zhang, A. Togbé, Perfect powers that are sums of two powers of Fibonacci numbers, *Bull. Aust. Mat. Soc.* **99**(1) (2019), 31-41.

[30] V. Ziegler, Sums of Fibonacci numbers that are perfect powers, *Quaest. Math.* **46**(8) (2023), 1717-1742.

EÖTVÖS LORÁND UNIVERSITY, INSTITUTE OF MATHEMATICS,
H-1117 BUDAPEST, PÁZMÁNY PÉTER SÉTÁNY 1/C,
HUNGARY
E-mail address: katalin.gyarmati@gmail.com