LONGER GAPS BETWEEN VALUES OF BINARY QUADRATIC FORMS

RAINER DIETMANN, CHRISTIAN ELSHOLTZ, ALEXANDER KALMYNIN, SERGEI KONYAGIN, AND JAMES MAYNARD

Abstract. We prove new lower bounds on large gaps between integers which are sums of two squares, or are represented by *any* binary quadratic form of discriminant D, improving results of Richards. Let $s_1, s_2, \dots$ be the sequence of positive integers, arranged in increasing order, that are representable by *any* binary quadratic form of fixed discriminant D, then

$$\limsup_{n\to\infty} \frac{s_{n+1}-s_n}{\log s_n} \gg \frac{|D|}{\varphi(|D|)\log|D|},$$

improving a lower bound of $\frac{1}{|D|}$ of Richards. In the special case of sums of two squares, we improve Richards's bound of $1/4$ to $\frac{390}{449} = 0.868\ldots$.

We also generalize Richards's result in another direction: If d is composite we show that there exist constants C_d such that for all integer values of x none of the values $p_d(x) = C_d + x^d$ is a sum of two squares. Let d be a prime. For all $k \in \mathbb{N}$ there exists a smallest positive integer y_k such that none of the integers $y_k + j^d, 1 \le j \le k$, is a sum of two squares. Moreover,

$$\limsup_{k\to\infty} \frac{k}{\log y_k} \gg \frac{1}{\sqrt{\log d}}.$$

1. Introduction

Let

$$\mathcal{S} = (s_1, s_2, s_3, s_4, s_5, s_6, \dots) = (0, 1, 2, 4, 5, 8, \dots)$$

denote the sequence of integers, in increasing order, which can be written as a sum of two squares of integers. The question of the size of large gaps between these integers was investigated by Turán and Erdős [9], Warlimont [24] and Richards [22]; see also [2], [23], [14], [1], [18] and [3] for related or more recent work by Bambah and Chowla, Shiu, Hooley, Balog and Wooley, Maynard, and Bonfoh and Enyi. Erdős writes that Turán proved that infinitely often

$$s_{n+1} - s_n \gg \frac{\log s_n}{\log\log s_n}$$

holds, which Erdős [9] improved to

$$s_{n+1} - s_n \gg \frac{\log s_n}{\sqrt{\log\log s_n}}. \tag{1}$$

In fact, Erdős's result was a bit more general, and Warlimont [24] independently obtained the same estimate (1), again in a more general context for sequences with

2020 *Mathematics Subject Classification.* 11N25, 11N37 (Primary), 11K65 (Secondary).

Key words and phrases. Numbers represented by binary quadratic forms; sums of two squares; large gaps.

hypotheses slightly different from [9]. In a very short and elegant paper Richards [22] improved this further to

$$\limsup_{n\to\infty} \frac{s_{n+1}-s_n}{\log s_n} \geq \frac{1}{4}. \tag{2}$$

In fact, the result he obtained was again more general: Fix a fundamental discriminant D and denote by $s_1, s_2, \ldots$ the integers, in increasing order, representable by *any* binary quadratic form of discriminant D. Then

$$\limsup_{n\to\infty} \frac{s_{n+1}-s_n}{\log s_n} \geq \frac{1}{|D|}. \tag{3}$$

The special case $D=-4$ corresponds to sums of two squares and recovers (2).

Apparently, Richards's record has not been broken since 1982. In this paper we obtain the following improvements to (2) and (3).

Theorem 1. *Let $s_1 < s_2 < \ldots$ be the sequence of positive integers that are sums of two squares. Then*

$$\limsup_{n\to\infty} \frac{s_{n+1}-s_n}{\log s_n} \geq \frac{390}{449} = 0.868\ldots$$

Equivalently, for $X \to +\infty$ we have

$$g(X) \geq \left(\frac{390}{449} + o(1)\right) \log X,$$

where

$$g(X) = \max_{s_{n+1}\leq X} (s_{n+1}-s_n).$$

Remark 1. The following table records some numerics on $g(X)$.

X	$g(X)$
10^6	35
10^7	50
10^8	60
10^9	74
10^{10}	105
10^{11}	107

One might wonder what the true order of magnitude of $g(X)$ is. The Cramér random model would predict that $g(X)$ is of order of magnitude $(\log X)^{3/2}$, see appendix C.

Theorem 2. *Let $D \neq 1$ be a fundamental discriminant, i.e. $D \equiv 1 \pmod 4$ and D being squarefree, or $D \equiv 0 \pmod 4$, $\frac{D}{4}$ being squarefree and $\frac{D}{4} \equiv 2 \pmod 4$ or $\frac{D}{4} \equiv 3 \pmod 4$. Further, let $(s_1, s_2, \ldots)$ be the sequence of positive integers, in increasing order, that are representable by* any *binary quadratic form of discriminant D, and let φ denote Euler's totient function. Then the following two estimates hold:*

A)

$$\limsup_{n\to\infty} \frac{s_{n+1}-s_n}{\log s_n} \geq \frac{|D|-1}{2|D|(1+\log\varphi(|D|))},$$

B)

$$\limsup_{n\to\infty} \frac{s_{n+1}-s_n}{\log s_n} \geq \frac{|D|}{2\varphi(|D|)(\log|D| + O((\log\log|D|)^3))}.$$

This is a significant improvement over the bound $1/|D|$ by Richards and now the dependence on $|D|$ has become very mild. Note that the result in A) is asymptotically weaker for large $|D|$ with many prime factors, but is completely explicit without any O-term.

Our approach for proving Theorem 1 largely follows that of Richards, but introduces two new key ideas: a *modular refinement*, and in addition a *probabilistic refinement*, whereas Richards' construction is completely deterministic; his argument used a sieving construction, which amounted to using a variant of the following proposition with $\Phi(x)$ of the form $e^{(4+o(1))x}$.

Proposition 1. *Let Φ be a continuous increasing function, such that for every $Y > 1$ there are integers $a(Y)$ and $P(Y)$ satisfying:*

a) $a(Y) \le P(Y) \le \Phi(Y)/2$ *and* $P(Y) > Y$
b) $P(Y)$ *has no prime factors of the form* $4k+1$
c) *For every integer* $1 \le j \le Y$, *at least one of the two conditions is satisfied:*

condition (I) *There is an odd prime* $p \equiv 3 \pmod 4$ *and an odd integer* k *such that* p^{k+1} *divides* $P(Y)$, p^k *divides* $a(Y)+j$ *but* p^{k+1} *does not divide* $a(Y)+j$.

condition (II) *There is an integer* k *such that* 2^{k+2} *divides* $P(Y)$ *and* $a(Y)+j \equiv 3\times 2^k \pmod{2^{k+2}}$.

Then for all $x \ge 1$ we have

$$g(x) \ge \Phi^{-1}(x),$$

where Φ^{-1} is the inverse function of Φ.

Richards [22] chose the following number to be $P(Y)$ in his construction:

$$P(Y) = \prod_{\substack{p \le 4Y \\ p \equiv 3 \pmod 4}} p^{\beta_p+1}, \tag{4}$$

where $\beta_p = [\log(4Y)/\log p]$. In this case $a(Y) \in [1, P(Y)]$ is the solution of the congruence

$$4a(Y) \equiv -1 \pmod{P(Y)}. \tag{5}$$

From (5) we obtain

$$4(a(Y)+j) \equiv 4j-1 \pmod{P(Y)} \quad (1 \le j \le Y).$$

Hence there exist a prime p with $p \equiv 3 \pmod 4$ and an odd integer α such that $4j-1$ is divisible by p^α, but is not divisible by $p^{\alpha+1}$. As $\alpha \le \beta_p$, and $P(Y)$ is divisible by p^{β_p+1}, it follows that condition (I) of Proposition 1 is satisfied. By the prime number theorem in arithmetic progressions (see for example formula (17.1) in [16]), we have

$$P(Y) < (4Y)^{2(1+\varepsilon)(2Y/\log 4Y)} = \exp((1+\varepsilon)4Y),$$

for sufficiently large $Y \ge Y(\varepsilon)$, so that (after redefining ε) $\exp((4+\varepsilon)Y)$ is an admissible choice of $\Phi(Y)$ for arbitrarily small $\varepsilon > 0$. From Proposition 1 we then recover (2); note that so far we have not used condition (II) in Proposition 1. The following theorem is the improved version of Proposition 1, with the probabilistic refinement included.

Theorem 3. *Assume that for Y large enough there are $P(Y)$ and $a(Y)$ satisfying the assumptions of Proposition 1. Suppose that $\log \Phi(Y) = o(Y \log Y)$ and that small prime factors make a small contribution to the size of $P(Y)$: we assume that*

$$\lim_{\varepsilon \to 0} \limsup_{Y \to +\infty} \frac{\log P(Y, \varepsilon)}{\log P(Y)} = 0, \tag{6}$$

where

$$P(Y, \varepsilon) = \prod_{\substack{p^k || P(Y) \\ p \le \varepsilon Y}} p^k.$$

Further assume that for all $\varepsilon > 0$ and all sufficiently large Y the inequality $Y \le \Phi(Y)^\varepsilon$ holds. Then for any $\varepsilon > 0$ and large enough X we have

$$g(X) \ge \Phi^{-1}(X^{2-\varepsilon}).$$

Remark 2. It would be natural to try to incorporate some of the techniques behind recent improvements for showing large gaps between primes [19, 11, 10] and other sieved sets [12] to try to gain a further improvement. Unfortunately neither of these approaches seems to give an improvement in the situation of gaps between sums of two squares. Both techniques followed a sieving set-up, where one first would sieve by 'small primes' (following earlier approaches), and then, having sieved by small primes, follows a more complicated sieving procedure for 'large primes' to sieve out many residue classes for these primes. In the Richards' setup, having sieved by primes $p < Y^{1/2}$ one is left with those $j < Y$ such that $4j - 1$ is the product of a prime $p_1 \equiv 3 \pmod 4$ and other primes all congruent to 1 (mod 4). Richards' setup proceeds by removing those j for which $4j - 1$ is a multiple of p for each large prime p. It is easy to see that this sieving is perfectly disjoint, and removes roughly as many elements as possible for each large prime. Therefore there does not appear to be any scope to refine this 'large prime' sieving, and so the newer techniques offer no improvement in our setup.

Remark 3. In our proof of Theorem 2 we only used the modular, not the probabilistic refinement. However, the proof of Theorem 3 does not use any properties of the set of primes of the form $4n + 3$ outside of the statement of Lemma 7 for $D = -4$. This means that for general D our probabilistic refinement essentially allows us to replace all conditions of the form $p^\alpha \,||\, n + a(Y)$ for some odd α by $p \,||\, n + a(Y)$ for all $p \ge \delta Y$ and arbitrarily small positive δ. Hence, the bounds of Theorem 2 can be doubled by our methods.

Finally, we study the distribution of the set $\mathcal{S}$ along consecutive values in polynomial sequences. It is obvious that all values of $p(x) = x^2 + 1$ are in $\mathcal{S}$, and that all values of polynomials such as $4x + 3$ or $2x^3 + 2x + 3$ are always congruent to 3 mod 4 and hence not in $\mathcal{S}$. Here we concentrate on the following class of polynomials $p_d(x) = C_d + x^d$, where C_d is a suitable constant.

We first study in Theorem 4 composite values of d in more detail, and show that for these values of d one can find a constant C_d such that *all* values $C_d + x^d, x \in \mathbb{Z}$ are not in $\mathcal{S}$. There is no congruence obstruction that would render the results in part B) and C) as trivial. We are not aware of a comparable result of this type in the literature, even though, with hindsight, it has an elementary proof. For the proof we took some inspiration from work of Jagy and Kaplansky [17] on Waring's

problem with mixed exponents. The method can apparently not be adapted to the case of prime values d. We then state a result (Theorem 5), which generalizes Richards's result (corresponding to $d = 1$) to all prime values d.

Theorem 4. A) *If $d \geq 4$ is even, then for no $x \in \mathbb{Z}$ the value $p_d(x) = 6 + x^d$ is in $\mathcal{S}$.*

B) *Let d be odd and composite, and let q be the least prime factor of d. Then for no $x \in \mathbb{Z}$ the value $p_d(x) = (c_q q)^q + x^d$ is in $\mathcal{S}$, where $c_q = 2$, when $q \equiv 3 \pmod 4$, and $c_q = 6$, when $q \equiv 1 \pmod 4$.*

C) *If $d \in \{2, 3\}$, then for all $C_d \in \mathbb{Z}$ there exists $x \in \mathbb{Z}$ such that $p(x) = C_d + x^d$ is in $\mathcal{S}$.*

Conjecture 1. (1) *If d is prime, then for all $C_d \in \mathbb{Z}$ there exists $x \in \mathbb{Z}$ such that $p(x) = C_d + x^d$ is in $\mathcal{S}$.*

(2) *If d is prime or $d = 1$, then there are sequences $C_d(1), C_d(2), \ldots$ and $x_1, x_2, \ldots$, both tending to infinity, such that $p(x) = C_d(i) + x^d$ is not in $\mathcal{S}$ for all $0 \leq x \leq x_i$, such that*

$$\limsup_{i\to\infty} \frac{x_i}{\log C_d(i)} = \infty.$$

In regard to the conjecture we remark that there can exist very long chains without an element in $\mathcal{S}$. We list several polynomials with long chains: The values of the polynomial $p_1(x) = 11^{23} + x^{23}$ are not in $\mathcal{S}$, when $-10 \leq x \leq 222$. This example has quite small coefficients. Searching a bit more, one finds that $p_2(x) = 25190^{11} + x^{11}$ is $\mathcal{S}$-free in the interval $[-209, 1229]$, and $p_3(x) = 642006^3 + x^3$ in $[-81, 717]$. In Appendix C we discuss a probabilistic model suggesting that $\log C_1(i)$ can be chosen of order of magnitude $x_i^{2/3}$.

Theorem 5. *Let $d \geq 3$ be an odd prime. For all $k \in \mathbb{N}$ there exists a smallest positive integer y_k such that none of the integers $y_k + j^d, 1 \leq j \leq k$, is a sum of two squares. The following estimates hold:*

$$\limsup_{k\to\infty} \frac{k}{\log y_k} \geq \frac{d-1}{4d\left(1 - \frac{1}{d} + \sum_{\text{odd } m \in \mathcal{S}\cap[1,d-1]} \frac{1}{m}\right)},$$

where for $d = 3, 5, 7, 11, 13, 17$ one can take $\frac{1}{10}, \frac{1}{9}, \frac{5}{48}, \frac{225}{2198}, \frac{135}{1307}, \frac{585}{5791}$ as a lower bound for the lim sup. *If $\varepsilon > 0$ and d is sufficiently large in terms of ε, then*

$$\limsup_{k\to\infty} \frac{k}{\log y_k} \geq \frac{d-1}{4d(1 - \frac{1}{d} + (C_R + \varepsilon)\sqrt{\log d})},$$

where

$$C_R = \frac{1}{\sqrt{2}} \prod_{p\equiv 3 \pmod 4} \left(1 - \frac{1}{p^2}\right)^{-1/2} \approx 0.7642\ldots \tag{7}$$

is the Landau-Ramanujan constant. Moreover, when $d \geq 17$, then

$$\limsup_{k\to\infty} \frac{k}{\log y_k} > \frac{1}{60\sqrt{\log d}}.$$

Remark 4. Note that we refrained from introducing our modular and probabilistic refinements of Theorem 1 in order to keep the statement clean as in this case our focus was more on the qualitative side of the result.

Notation: We say that a prime power p^α *exactly divides* an integer n if p^α divides n but $p^{\alpha+1}$ does not, and use the notation $p^\alpha \,||\, n$. The symbol $\left(\frac{m}{n}\right)$ always denotes the Kronecker symbol of m over n, and we use the usual notation φ, μ and γ for the Euler totient function, Möbius function and the Euler-Mascheroni constant, respectively. Also, as the proof of Theorem 1 in section 4 requires quite heavy notation, we have a summary of abbreviations in appendix A and some examples for the sets of residue classes introduced in appendix B.

Acknowledgments: All authors want to thank Igor Shparlinski for helpful comments regarding the exposition of this paper, and Pieter Moree for drawing our attention to Lemma 9 in [20].
C.E. and R.D. would like to express thanks to the TU Graz, Royal Holloway and the ETH Zürich for support and favourable working conditions, and R.D. and J.M. want to express the same thanks to the University of Oxford.
C.E. was partially supported by the Austrian Science Fund (FWF): W1230 and a joint FWF-ANR grant (FWF: I 4945-N and ANR-20-CE91-0006).
A.K. was partially supported by the HSE University Basic Research Program, the "Russian Young Mathematics" contest, the Simons foundation and the foundation for Advancement of Theoretical Physics and Mathematics "BASIS".
The work of S. K. was performed at the Steklov International Mathematical Center and supported by the Ministry of Science and Higher Education of the Russian Federation (agreement no. 075-15-2019-1614).
J.M. was supported by a Royal Society Wolfson Merit Award and ERC grant agreement No 851318.

2. Outline of the proof of Theorem 1

As the details of the proof of Theorem 1 are quite technical, let us briefly outline the underlying ideas and structure of the proofs.

The key observation for the modular refinement is as follows: We concluded that $4j-1$ is exactly divisible by some prime power p^α with $p \equiv 3 \pmod 4$ and odd α, say $4j-1 = p^\alpha r$. If $4j-1 < 4Y$ is composite, then $p < \frac{4}{5}Y$. Primes p with $\frac{4}{5}Y \le p \le 4Y$ included in the product (4) therefore are only used once in the argument, namely when $p = 4j-1$. Now integers in $I = \{a(Y)+1, \dots, a(Y)+Y\}$ which are congruent to 3 (mod 4) are obviously not sums of two squares. Hence, additionally assuming that $a(Y) \equiv 0 \pmod 4$, we conclude that for $j \equiv 3 \pmod 4$ we trivially know that $a(Y)+j$ is not a sum of two squares. As $4j-1 \equiv 11 \pmod{16}$ in this case, we deduce that primes $p \equiv 11 \pmod{16}$, with $4Y/5 \le p \le 4Y$, are not needed in the product (4). In other words, for fixed Y one can use a smaller P, or for a given size of $a(Y)$, one can find a larger Y than in Richards's argument. This basic approach just described in fact would replace the $\frac{1}{4}$ in (2) by

$$\left(2 \times \frac{1}{2}\left(\frac{4}{5} + \frac{3}{4}\left(4 - \frac{4}{5}\right)\right)\right)^{-1} = \frac{5}{16} = 0.3125,$$

but it can be further refined in two ways: First, considering higher powers of 2 one finds a larger proportion of residue classes for j one can dispense with; for example the residue classes $j \equiv 3, 6, 7$ modulo 8 immediately rule out that $a(Y)+j$ is a sum of two squares, provided that $a(Y) \equiv 0 \pmod 8$. Therefore the primes $p = 4j-1 \ge \frac{4}{5}Y$ with $p \equiv 11$, 23, or 27 (mod 32) are not needed. Secondly, also smaller primes p can be considered; for example primes p with $\frac{4}{9}Y \le p < \frac{4}{5}Y$

can only occur in the argument if either $4j-1=p$ or $4j-1=5p$. One residue class one can disregard here is for example $p\equiv 11 \pmod{32}$, as for such p both p and $5p$ are modulo 32 in the set of residue classes 11, 23 and 27 which were ruled out above for $4j-1$. This optimization, making use of powers of 2, is our first refinement, the modular one, and is encoded in condition (II) of Proposition 1. Properly implementing it as in section 4 would yield a result with

$$\frac{195}{449}=\frac{1}{2}\times\frac{390}{449}$$

in place of $\frac{390}{449}$ in Theorem 1.

A similar idea is used to prove Theorem 2. Here instead of $4j-1$ the progression $|D|j+r$ shows up, for some r with $(\frac{D}{r})=-1$. Whereas Richards [22] uses all primes p with $(\frac{D}{p})=-1$, it turns out that for large primes p used only once, namely when $|D|j+r=p$ (i.e. $L/\ell_2<p\le L$ in the notation of section 7), only p with $p\equiv r \pmod{|D|}$ need to be considered. Similarly, for somewhat smaller primes p used only twice, only two residue classes modulo $|D|$ have to be covered, and so on. This leads to a considerably smaller product P. An even more elaborate analysis along the same lines as described before for the modular refinement could probably lead to a further slight improvement in Theorem 2. For the sake of simplicity of exposition, though, we again opted not to implement a further refinement here.

The second key observation, the probabilistic refinement, is that when $p\mid a(Y)+j$ for some large prime $p\equiv 3 \pmod 4$ that divides $P(Y)$, then it is rather unlikely that $p^2\mid a(Y)+j$. Using a probabilistic construction instead of the deterministic one so far, one can reduce the required exponent β_p for large primes p to 1. Theorem 3 indeed shows that under certain additional restrictions on $P(Y)$ and $\Phi(Y)$ the conditions (I) and (II) of Proposition 1 imply the lower bound $g(X)\ge\Phi^{-1}(X^{2-o(1)})$. In this way we can improve the result that comes out of the modular approach by another factor 2, with the probabilistic approach yielding $2\times\frac{195}{449}=\frac{390}{449}$ in Theorem 1.

3. The modular refinement I: Proof of proposition 1

The interval $I=[1+a(Y);Y+a(Y)]$ lies inside the interval $[1,\Phi(Y)]$ for all Y as $Y+a(Y)\le 2P(Y)\le\Phi(Y)$. Let us show that this interval does not contain any elements of the set $\mathcal{S}$. Indeed, if $n\in I$ then for some $j\le Y$ one has $n=j+a(Y)$. Now, if condition I of Proposition 1 holds for j then n is not a sum of two squares, because for some prime number p with $p\equiv 3 \pmod 4$ there is an odd number k, a positive integer a_j which is not divisible by p and a positive integer b_j such that $P(Y)=p^{k+1}b_j$ and $n=a(Y)+j\equiv p^k a_j \pmod{P(Y)}$. From this we obtain

$$n=p^k a_j+cP(Y)=p^k(a_j+b_j cp)$$

for some integer c. Therefore, some prime p congruent to 3 modulo 4 has an odd exponent in the prime factorization of n and thus n is not in $\mathcal{S}$.

On the other hand, if for j the second condition holds then an analogous argument shows that n is equal to $2^k(4u-1)$ for some positive integer u, but numbers of this form are not sums of two squares.

Hence, for all large enough Y the interval $[1,\Phi(Y)]$ contains a subinterval of length Y that does not intersect with $\mathcal{S}$. Consequently,

$$g(\Phi(Y))\ge Y.$$

Choosing $Y = \Phi^{-1}(x)$ we get the desired result.

4. The modular refinement II: Two-adic preparations

For all $\ell \in \mathbb{N}$ with $\ell \geq 2$ define

$$S_\ell = \{2^a b \in \{1, \dots, 2^\ell\} : a \leq \ell - 2, b \equiv 3 \pmod 4\}. \tag{8}$$

Clearly

$$S_\ell \subset S_{\ell+1} \tag{9}$$

for all $\ell \geq 2$.

The following lemma is immediate.

Lemma 1. *For $\ell \geq 2$ we have*

$$\#S_\ell = 2^{\ell-1} - 1.$$

In the following it is convenient to use the projection

$$\pi_\ell : \mathbb{Z} \to \{1, \dots, 2^\ell\}$$
$$x \mapsto n \in \{1, \dots, 2^\ell\} \text{ such that } x \equiv n \pmod{2^\ell}.$$

Define the map τ by

$$\tau : \mathbb{Z} \to \mathbb{Z}; \quad j \mapsto 4j - 1,$$

and for $\ell \in \mathbb{N}$, $\ell \geq 2$ define

$$T_{\ell+2} = \tau(S_\ell) \subset \{1, \dots, 2^{\ell+2}\}.$$

Again, one observes that

$$T_\ell \subset T_{\ell+1} \tag{10}$$

for all $\ell \geq 4$.

Lemma 2. *Let $\ell \geq 2$ and $s \in S_\ell$ with $s \neq 3 \times 2^{\ell-2}$. Then we have $\pi_\ell(s + 2^{\ell-1}) \in S_\ell$.*

Proof. As $s \in S_\ell$, $s \neq 3 \times 2^{\ell-2}$, it follows that $s \equiv 2^a b \pmod{2^\ell}$ where $a \leq \ell - 3$ and $b \equiv 3 \pmod 4$. Hence $s + 2^{\ell-1} \equiv 2^a (b + 2^{\ell-1-a}) \pmod{2^\ell}$, where $b + 2^{\ell-1-a} \equiv b \equiv 3 \pmod 4$, whence $\pi_\ell(s + 2^{\ell-1}) \in S_\ell$. □

Corollary 1. *Let $\ell \geq 2$ and $t \in T_{\ell+2}$ with $t \neq 3 \times 2^\ell - 1$. Then $\pi_{\ell+2}(t + 2^{\ell+1}) \in T_{\ell+2}$.*

Lemma 3. *Let $\ell \geq 3$. Then $3 \times 2^\ell - 1 \notin T_{\ell+1}$.*

Proof. Let $x = 3 \times 2^\ell - 1$. Since $T_{\ell+1} \subset \{1, \dots, 2^{\ell+1}\}$ and $x > 2^{\ell+1}$, clearly $x \notin T_{\ell+1}$. □

Let

$$U_3 = \{3\} \subset S_3, \tag{11}$$

and for $\ell \geq 4$, recursively define

$$U_\ell = U_{\ell-1} \cup \{u + 2^{\ell-1} : u \in U_{\ell-1}\} \cup \{3 \times 2^{\ell-2}\}. \tag{12}$$

It follows by induction from (8), (9), Lemma 2, (11) and (12) that $U_\ell \subset S_\ell$ for all $\ell \geq 3$. Moreover, for $\ell \geq 2$ define

$$V_\ell = \{s \in S_\ell : \pi_{\ell+2}(5\tau(s)) \in T_{\ell+2}\}.$$

In a similar way, define

$$W_5 = \{24\} \subset U_5,$$

and, for $\ell \geq 6$,

$$W_\ell = W_{\ell-1} \cup \{u + 2^{\ell-1} : u \in W_{\ell-1}\} \cup \{3 \times 2^{\ell-2}\}.$$

As above one shows that

$$W_\ell \subset U_\ell$$

for all $\ell \geq 5$. Further, for $\ell \geq 2$ let

$$R_\ell = \{s \in S_\ell : \pi_{\ell+2}(5\tau(s)) \in T_{\ell+2} \text{ and } \pi_{\ell+2}(9\tau(s)) \in T_{\ell+2}\}.$$

Note that

$$R_\ell \subset V_\ell$$

for all $\ell \geq 5$.

Lemma 4. *Let $\ell \geq 3$. Then $\#U_\ell = 2^{\ell-2} - 1$ and $U_\ell \subset V_\ell$.*

Proof. We prove the lemma by induction on ℓ. For $\ell = 3$, it is immediate to check that

$$U_3 = V_3 = \{3\},$$

whence

$$\#U_3 = 1 = 2^{\ell-2} - 1,$$

and no element in U_3 is divisible by $2^{\ell-1}$. Now let $\ell \geq 4$, and suppose that $\#U_{\ell-1} = 2^{\ell-3} - 1$, no element in $U_{\ell-1}$ is divisible by $2^{\ell-2}$ and $U_{\ell-1} \subset V_{\ell-1}$. The three sets on the right hand side of (12) are disjoint as $U_{\ell-1} \subset \{1, \dots, 2^{\ell-1}\}$, $\{u + 2^{\ell-1} : u \in U_{\ell-1}\} \subset \{2^{\ell-1} + 1, \dots, 2^\ell\}$, and no element in $U_{\ell-1}$ is divisible by $2^{\ell-2}$. Hence

$$\#U_\ell = 2\#U_{\ell-1} + 1 = 2^{\ell-2} - 1$$

and no element in U_ℓ is divisible by $2^{\ell-1}$. To prove $U_\ell \subset V_\ell$, let $s \in U_\ell$.

Case I: $s = 3 \times 2^{\ell-2}$. Then

$$\begin{aligned}\frac{1}{4}\left(5\tau(s) - 3 \times 2^{\ell+2} + 1\right) &= \frac{1}{4}\left(5 \times (4 \times 3 \times 2^{\ell-2} - 1) - 3 \times 2^{\ell+2} + 1\right) \\ &= 3 \times 2^{\ell-2} - 1,\end{aligned}$$

hence

$$5\tau(s) \equiv 4 \times (3 \times 2^{\ell-2} - 1) - 1 \pmod{2^{\ell+2}}.$$

Now $3 \times 2^{\ell-2} - 1 \equiv 3 \pmod 4$, so $3 \times 2^{\ell-2} - 1 \in S_\ell$, whence $4 \times (3 \times 2^{\ell-2} - 1) - 1 \in T_{\ell+2}$, so $\pi_{\ell+2}(5\tau(s)) \in T_{\ell+2}$.

Case II: $s \neq 3 \times 2^{\ell-2}$. Then by definition of U_ℓ, we have $\pi_{\ell-1}(s) \in U_{\ell-1}$, so $\pi_{\ell+1}(5\tau(s)) \in T_{\ell+1}$ by our inductive assumption $U_{\ell-1} \subset V_{\ell-1}$. Hence $\pi_{\ell+2}(5\tau(s)) \in T_{\ell+1}$ or $\pi_{\ell+2}(5\tau(s)) \in \{u + 2^{\ell+1} : u \in T_{\ell+1}\}$. If $\pi_{\ell+2}(5\tau(s)) \in T_{\ell+1}$ then by (10) we immediately obtain $\pi_{\ell+2}(5\tau(s)) \in T_{\ell+2}$ as required, whereas if $\pi_{\ell+2}(5\tau(s)) \in \{u + 2^{\ell+1} : u \in T_{\ell+1}\}$ then (10), Lemma 3 and Corollary 1 again yield $\pi_{\ell+2}(5\tau(s)) \in T_{\ell+2}$. □

Lemma 5. *Let $\ell \geq 5$. Then $\#W_\ell = 2^{\ell-4} - 1$ and $W_\ell \subset R_\ell$.*

Proof. We use a similar strategy as in the proof of Lemma 4; the proof for $\#W_\ell = 2^{\ell-4} - 1$ is completely analogous, so let us focus on the second part $W_\ell \subset R_\ell$. The case $\ell = 5$ is immediately checked directly. Now suppose that $\ell \geq 6$ and

$W_{\ell-1} \subset R_{\ell-1}$. For $s \neq 3 \times 2^{\ell-2}$ we argue in exactly the same way as in the proof of Lemma 4. Therefore let us only discuss the case $s = 3 \times 2^{\ell-2}$. Here

$$\begin{aligned}\frac{1}{4}\left(9\tau(s) - 6 \times 2^{\ell+2} + 1\right) &= \frac{1}{4}\left(9 \times (4 \times 3 \times 2^{\ell-2} - 1) - 6 \times 2^{\ell+2} + 1\right)\\ &= 2 \times (3 \times 2^{\ell-3} - 1),\end{aligned}$$

hence

$$9\tau(s) \equiv 4 \times 2 \times (3 \times 2^{\ell-3} - 1) - 1 \pmod{2^{\ell+2}}.$$

Now $2 \times (3 \times 2^{\ell-3} - 1) \equiv 6 \pmod 8$, so $2 \times (3 \times 2^{\ell-3} - 1) \in S_\ell$, thus $4 \times 2 \times (3 \times 2^{\ell-3} - 1) - 1 \in T_{\ell+2}$ and $\pi_{\ell+2}(9\tau(s)) \in T_{\ell+2}$; from the proof of Lemma 4 we already know that $\pi_{\ell+2}(5\tau(s)) \in T_{\ell+2}$. □

We now follow the idea of Richards [22] already explained in the introduction. Let $\varepsilon > 0$. Then, in terms of ε, fix a sufficiently large positive integer $\ell \geq 5$ and a sufficiently large positive integer Y, and let the sets S_ℓ, $T_{\ell+2}$, U_ℓ, V_ℓ, W_ℓ, R_ℓ be defined as above. In the following it is convenient to define

$$A := T_{\ell+2}, \quad B := \pi_{\ell+2}(\tau(U_\ell)), \quad C := \pi_{\ell+2}(\tau(W_\ell)).$$

Note that

$$C \subset B \subset A$$

since $W_\ell \subset U_\ell \subset S_\ell$. By Lemma 1, Lemma 4 and Lemma 5 we have

$$\begin{aligned}&\#A = \#T_{\ell+2} = \#S_\ell = 2^{\ell-1} - 1, \quad \#B = \#U_\ell = 2^{\ell-2} - 1,\\ &\#C = \#W_\ell = 2^{\ell-4} - 1.\end{aligned}$$

Hence if ℓ is chosen sufficiently large in terms of ε, then

$$\frac{\#A}{\varphi(2^{\ell+2})} \geq \frac{1}{4}(1-\varepsilon); \quad \frac{\#B}{\varphi(2^{\ell+2})} \geq \frac{1}{8}(1-\varepsilon); \quad \frac{\#C}{\varphi(2^{\ell+2})} \geq \frac{1}{32}(1-\varepsilon). \tag{13}$$

Now for each prime $p \leq 4Y$ define

$$\beta_p = \max_{p^m \leq 4Y} m,$$

let

$$X_1 = (1+\varepsilon)\frac{4}{13}Y, \quad X_2 = (1+\varepsilon)\frac{4}{9}Y, \quad X_3 = (1+\varepsilon)\frac{4}{5}Y,$$

and let

$$\begin{aligned}P(Y) =& 2^\ell \prod_{\substack{p_1 \leq X_1:\\ p_1 \equiv 3 \pmod 4}} p_1^{\beta_{p_1}+1} \prod_{\substack{X_1 < p_2 \leq X_2:\\ p_2 \equiv 3 \pmod 4,\\ \pi_{\ell+2}(p_2) \notin C}} p_2^{\beta_{p_2}+1}\\ &\times \prod_{\substack{X_2 < p_3 \leq X_3:\\ p_3 \equiv 3 \pmod 4,\\ \pi_{\ell+2}(p_3) \notin B}} p_3^{\beta_{p_3}+1} \prod_{\substack{X_3 < p_4 \leq 4Y:\\ p_4 \equiv 3 \pmod 4,\\ \pi_{\ell+2}(p_4) \notin A}} p_4^{\beta_{p_4}+1},\end{aligned}$$

where $p_1, \ldots, p_4$ denote prime numbers, varying over the respective intervals. Then by the prime number theorem in arithmetic progressions (see for example formula (17.1) in [16]), using the upper bound $p^{\beta_p+1} \leq (4Y)^2$, the lower bounds (13) and the fact that all elements in A are congruent to 3 modulo 4, we obtain

$$P(Y) \leq (4Y)^{2(1+\varepsilon)Y\alpha/\log(4Y)},$$

where

$$\begin{aligned}\alpha &= \frac{1}{2}\left(\frac{4}{13}+\left(\frac{4}{9}-\frac{4}{13}\right)\frac{15}{16}+\left(\frac{4}{5}-\frac{4}{9}\right)\frac{3}{4}+\left(4-\frac{4}{5}\right)\frac{1}{2}\right)\\ &= \frac{1}{2}\times\frac{449}{195}.\end{aligned}$$

Hence

$$P(Y) \le \exp\left((1+\varepsilon)\frac{449}{195}Y\right). \tag{14}$$

Now use the Chinese Remainder Theorem to find $a(Y) \in \{1,\dots,P(Y)\}$ such that

(i) $2^\ell \mid y$,
(ii) if $p \equiv 3 \pmod 4$ and $p \le X_1$, then $4a(Y) \equiv -1 \pmod{p^{\beta_p+1}}$,
(iii) if $p \equiv 3 \pmod 4$, $X_1 < p \le X_2$ and $\pi_{\ell+2}(p) \notin C$, then $4a(Y) \equiv -1 \pmod{p^{\beta_p+1}}$,
(iv) if $p \equiv 3 \pmod 4$, $X_2 < p \le X_3$ and $\pi_{\ell+2}(p) \notin B$, then $4a(Y) \equiv -1 \pmod{p^{\beta_p+1}}$,
(v) if $p \equiv 3 \pmod 4$, $X_3 < p \le 4Y$ and $\pi_{\ell+2}(p) \notin A$, then $4a(Y) \equiv -1 \pmod{p^{\beta_p+1}}$.

We claim that all the numbers $a(Y)+1,\dots,a(Y)+Y$ satisfy condition (I) or condition (II) of Proposition 1. To settle the claim, let $1 \le j \le Y$. If $\pi_\ell(j) \in S_\ell$, then by property (i) above $a(Y)+j$ satisfies condition (II) of Proposition 1, so we can assume that $\pi_\ell(j) \notin S_\ell$, whence $\pi_{\ell+2}(\tau(j)) \notin A$. Now $\tau(j) \equiv 3 \pmod 4$, so $\tau(j)$ must be divisible by a prime p with $p \equiv 3 \pmod 4$ where $3 \le p \le 4Y-1$ and $p^\gamma \mid\mid \tau(j)$ for odd $\gamma \le \beta_p$.

Case I: $p \le X_1$. Then by property (ii) above $4a(Y) \equiv -1 \pmod{p^{\beta_p+1}}$.

Case II: $X_3 < p \le 4Y$. Then $p \equiv 3 \pmod 4$ and $p \mid \tau(j)$ imply that $\tau(j) = p$, so $\pi_{\ell+2}(\tau(j)) = \pi_{\ell+2}(p) \notin A$ and by property (v) above, we have $4a(Y) \equiv -1 \pmod{p^{\beta_p+1}}$.

Case III: $X_2 < p \le X_3$. Then $p \equiv 3 \pmod 4$ and $p \mid \tau(j)$ imply that $\tau(j) = p$ or $\tau(j) = 5p$. As above, if $\tau(j) = p$ we conclude that $\pi_{\ell+2}(p) \notin A$, whereas if $\tau(j) = 5p$ we obtain $\pi_{\ell+2}(5p) \notin A$. Writing $p = \tau(s)$ for some positive integer s, we then find that $\pi_\ell(s) \notin V_\ell$, whence by Lemma 4 also $\pi_\ell(s) \notin U_\ell$, hence $\pi_{\ell+2}(p) \notin B$. As $B \subset A$, we get $\pi_{\ell+2}(p) \notin B$ regardless of whether $\tau(j) = p$ or $\tau(j) = 5p$, so by property (iv) again $4a(Y) \equiv -1 \pmod{p^{\beta_p+1}}$.

Case IV: $X_1 < p \le X_2$. Then $p \equiv 3 \pmod 4$ and $p \mid \tau(j)$ imply that $\tau(j) = p$ or $\tau(j) = 5p$ or $\tau(j) = 9p$. If $\tau(j) = p$, then $\pi_{\ell+2}(p) = \pi_{\ell+2}(\tau(j)) \notin A$. Next, if $\tau(j) = 5p$, then $\pi_{\ell+2}(\tau(j)) = \pi_{\ell+2}(5p) \notin A$, hence as above $\pi_{\ell+2}(p) \notin B$ by Lemma 4. Finally, if $\tau(j) = 9p$, then $\pi_{\ell+2}(\tau(j)) = \pi_{\ell+2}(9p) \notin A$, hence as above $\pi_{\ell+2}(p) \notin C$ by Lemma 5. As $C \subset B \subset A$, we obtain $\pi_{\ell+2}(p) \notin C$ regardless whether $\tau(j) = p$, $\tau(j) = 5p$ or $\tau(j) = 9p$, whence $4a(Y) \equiv -1 \pmod{p^{\beta_p+1}}$ by property (iii) above.

In all cases,

$$4(a(Y)+j) \equiv 4j-1 \pmod{p^{\beta_p+1}},$$

so $p^\gamma \mid\mid (a(Y)+j)$. Since $p \equiv 3 \pmod 4$ and γ is odd, $a(Y)+j$ satisfies condition (I) of Proposition 1.

Remark 5. One might wonder if the study of further iterations leading to analogous sets $D, E, \dots$ would reduce the size of P even more. As far as we see this is not the case because $C \cap 13C = \emptyset$, but this does not exclude other refinements.

5. The probabilistic refinement: Proof of Theorem 3

Let us choose $\delta = \delta(\varepsilon) < 1$ so that for large enough Y we have $\log P(Y, \delta) < \varepsilon \log P(Y)$. Set

$$\mathcal{P}(Y) = P(Y, \delta) \prod_{\substack{p \mid P(Y) \\ p > \delta Y}} p^{\gamma_p},$$

where $\gamma_p = 1$ if there is a positive integer $j \le Y$ with $p^{k+1} \mid P(Y)$ and

$$a(Y) + j \equiv p^k a_j \pmod{P(Y)}$$

for some odd k with a_j coprime to p. Let $\gamma_p = 0$ otherwise. Here $a(Y)$ is the same as in Proposition 1.

Note that $\mathcal{P}(Y) \le P(Y)^{1/2+\varepsilon/2}$. Indeed, every exponent γ_p is at most half the exponent of $p > \delta Y$ in the factorization of $P(Y)$. Therefore

$$\mathcal{P}(Y) \le P(Y, \delta) \sqrt{\frac{P(Y)}{P(Y, \delta)}} = \sqrt{P(Y, \delta) P(Y)} < P(Y)^{1/2+\varepsilon/2}$$

by the choice of δ. Now choose a positive integer $a_0(Y)$ such that the congruence $a_0(Y) \equiv a(Y) \pmod{\mathcal{P}(Y)}$ and the inequalities $0 < a_0(Y) \le \mathcal{P}(Y)$ hold. Define the family of intervals

$$I_n = [1 + a_0(Y) + n\mathcal{P}(Y); Y + a_0(Y) + n\mathcal{P}(Y)],$$

where the variable n takes integer values with $0 \le n \le \delta Y$. Let us show that at least one of the constructed intervals does not contain any sum of two squares. Indeed, assume that $m \in I_n$ is an element of $\mathcal{S}$. As m lies in I_n, for some $j \le Y$ the equality $m = a_0(Y) + j + n\mathcal{P}(Y)$ holds. By the definition of $a(Y)$ and $P(Y)$, at least one of the following conditions holds:

(I) There are positive integers k and m with $2^{k+2} \mid P(Y)$ and

$$a(Y) + j \equiv 2^k(4m - 1) \pmod{P(Y)}.$$

In this case we also have

$$a_0(Y) + j \equiv 2^k(4m - 1) \pmod{\mathcal{P}(Y)}$$

and $2^{k+2} \mid \mathcal{P}(Y)$, therefore m cannot be the sum of two squares, which is a contradiction.

(II) There are an odd prime p and an odd positive integer k with $p^{k+1} \mid P(Y)$ and

$$a(Y) + j \equiv p^k a_j \pmod{P(Y)}$$

for some a_j that is not divisible by p. If $p \le \delta Y$ then these congruences and divisibilities remain true for $a_0(Y)$ and $\mathcal{P}(Y)$, which once again leads us to a contradiction. If, on the contrary, $p > \delta Y$, then necessarily $\gamma_p = 1$ and hence

$$m = a_0(Y) + j + n\mathcal{P}(Y) \equiv 0 \pmod{p}.$$

As $p \equiv 3 \pmod 4$ and m is the sum of two squares, we have $p^2 \mid m$.

Notice now, that for fixed $j \le Y$ and $p > \delta Y$ there exists at most one n such that $a_0(Y) + j + n\mathcal{P}(Y)$ is divisible by p^2. Indeed, otherwise for two distinct $0 \le n_1, n_2 \le \delta Y$ we have

$$a_0(Y) + j + n_1\mathcal{P}(Y) \equiv a_0(Y) + j + n_2\mathcal{P}(Y) \pmod{p^2}.$$

As $p^2 \nmid \mathcal{P}(Y)$ we obtain $n_1 \equiv n_2 \pmod{p}$, which is impossible because

$$0 < |n_1 - n_2| \le \delta Y < p.$$

Furthermore, for a fixed prime $p > \delta Y$ there are at most $1 + 1/\delta \le 2/\delta$ numbers $j \le Y$ with $a_0(Y) + j \equiv 0 \pmod{p}$, as any two numbers having this property are congruent modulo p. Thus, the number of n for which I_n contains a sum of two squares is at most the number of all exceptional pairs (j, p), i.e. at most $2F/\delta$, where F is the number of prime factors $p > \delta Y$ of $P(Y)$. Clearly, $P(Y) \ge (\delta Y)^F$ so

$$F \le \log P(Y)/\log(\delta Y) \ll \log P(Y)/\log Y \le \log \Phi(Y)/\log Y = o(Y)$$

due to the conditions of Theorem 3. Therefore, all but $o(Y)$ of the intervals I_n do not intersect $\mathcal{S}$. In particular, for all large enough Y there is at least one interval with this property.

Next, all the resulting intervals lie inside the interval $[1, Y\Phi(Y)^{1/2+\varepsilon/2}]$ because

$$Y + a_0(Y) + \delta Y\mathcal{P}(Y) \le Y + \mathcal{P}(Y) + \delta Y\mathcal{P}(Y) \le YP(Y)^{1/2+\varepsilon/2} \le Y\Phi(Y)^{1/2+\varepsilon/2}.$$

By the conditions of Theorem 3 for all large enough Y we have $Y \le \Phi(Y)^{\varepsilon/2}$. Consequently, for all large Y the inequality $Y\Phi(Y)^{1/2+\varepsilon/2} \le \Phi(Y)^{1/2+\varepsilon}$ is true, which means that the interval $[1, \Phi(Y)^{1/2+\varepsilon}]$ contains a subinterval of length Y that does not contain sums of two squares. Choosing $Y = \Phi^{-1}(X^{2/(1+2\varepsilon)})$, we obtain the estimate $g(X) \ge \Phi^{-1}(X^{2/(1+2\varepsilon)})$. As ε was an arbitrary positive real number, this concludes our proof.

6. Proof of Theorem 1

In this section we will prove Theorem 1. By (14), there are $P(Y)$ and $a(Y)$ which satisfy the conditions of Proposition 1 and the relation

$$\log P(Y) = (449/195 + o(1))Y.$$

Furthermore, all the prime factors of $P(Y)$ are at most $4Y$ and all the exponents of p in the factorization do not exceed

$$\beta_p + 1 = [\log(4Y)/\log p] + 1 \le 2\beta_p.$$

Therefore, if $p^\alpha \mid\mid P(Y)$ then $p^\alpha \le 16Y^2$. It follows that small primes make a small contribution in $P(Y)$. Indeed, if $\varepsilon > 0$ then

$$P(Y, \varepsilon) = \prod_{\substack{p^k \mid\mid P(Y) \\ p \le \varepsilon Y}} p^k \le \prod_{\substack{p^k \mid\mid P(Y) \\ p \le \varepsilon Y}} 16Y^2 \le (4Y)^{2\pi(\varepsilon Y)}.$$

Consequently,

$$\log P(Y, \varepsilon) \le 2\pi(\varepsilon Y)\log(4Y) \sim 2\frac{\varepsilon Y}{\log \varepsilon Y}\log Y \sim 2\varepsilon Y.$$

As we also have $\log P(Y) \gg Y$, we finally get

$$\lim_{\varepsilon\to 0}\limsup_{Y\to+\infty}\frac{\log P(Y,\varepsilon)}{\log P(Y)}\le\lim_{\varepsilon\to 0}\frac{2\varepsilon}{c}=0,$$

for some $c>0$. Thus, the assumption (6) of Theorem 3 is satisfied. Choosing $\Phi(Y)=\exp((449/195+\varepsilon)Y)$, we observe that also the assumptions $\log\Phi(Y)=o(Y\log Y)$ and $Y\le\Phi(Y)^{\varepsilon}$ are satisfied, so from Theorem 3 for arbitrary $\varepsilon_1>0$ we get

$$\begin{aligned}g(X)&\ge\Phi^{-1}(X^{2-\varepsilon_1})\\&=(449/195+\varepsilon)^{-1}(2-\varepsilon_1)\log X\\&\ge(390/449-2\varepsilon-195/449\varepsilon_1)\log X.\end{aligned}$$

Small enough values of ε and ε_1 give the desired result.

7. Longer gaps between numbers representable by binary quadratic forms of discriminant D: Proof of Theorem 2

Again, we follow the idea of Richards [22]: By our assumptions on D in Theorem 2, we can choose and fix a positive integer $r\in\{1,\dots,|D|\}$ such that the Kronecker symbol (D/r) has value

$$\left(\frac{D}{r}\right)=-1.$$

Also note that necessarily $|D|\ge 3$. The following three well known results are provided for easy later reference.

Lemma 6. *For fixed $m\in\mathbb{Z}\backslash\{0\}$ with $m\equiv 0\pmod 4$ or $m\equiv 1\pmod 4$, the Kronecker symbol $(\frac{m}{\cdot})$ is periodic of period dividing $|m|$, i.e. for all $k,n\in\mathbb{Z}$ where $n\ne 0$ and $n+km\ne 0$ we have*

$$\left(\frac{m}{n+km}\right)=\left(\frac{m}{n}\right).$$

Proof. This is Theorem 2.29 in [5]. □

Lemma 7. *Let $D\ne 1$ be a fundamental discriminant, $n\in\mathbb{N}$ and p be a prime such that p^{α} exactly divides n for odd α and with $(\frac{D}{p})=-1$. Then n is not representable by* any *binary quadratic form of discriminant D.*

Proof. Let $R_D(n)$ be the total number of representations of n by any binary quadratic form of discriminant D. Then by Theorem 3 in [25, §8], we have

$$R_D(n)=\sum_{\ell\mid n}\left(\frac{D}{\ell}\right). \tag{15}$$

Since $(\frac{D}{\ell})$ is multiplicative in ℓ, $R_D(n)$ is multiplicative in n as well. Now α is odd and $(\frac{D}{p})=-1$, whence $R_D(p^{\alpha})=0$. By multiplicativity, as p^{α} exactly divides n, also $R_D(n)=0$, so indeed n is not represented by any binary quadratic form of discriminant D. □

Lemma 8. *Let $n\ge 3$ be an integer. Then*

$$\sum_{d\mid n}\frac{\log d}{d}\ll(\log\log n)^2.$$

Proof. As shown on page 208 of [15], we have

$$\sum_{d|n} \frac{\log d}{d} \ll (\log\log n)\sigma_{-1}(n),$$

where

$$\sigma_{-1}(n) = \sum_{d|n} \frac{1}{d} = \frac{\sigma(n)}{n}$$

and

$$\sigma(n) = \sum_{d|n} d \ll n \log\log n$$

due to Gronwall's theorem (see for example Theorem 323 in [13]). □

Next, let

$$t = \varphi(|D|),$$

and let $\ell_1 = 1, \dots, \ell_t = |D| - 1$ be the coprime residue classes modulo $|D|$, ordered by size. Further, for $i \in \mathbb{N}$ define

$$T_i = \{x \in (\mathbb{Z}/|D|\mathbb{Z})^* : \ell_j x \equiv r \pmod{|D|} \text{ for some } j \le i\}, \tag{16}$$

and let π be the projection

$$\pi : \mathbb{Z} \to \mathbb{Z}/|D|\mathbb{Z}.$$

Further, fix $\varepsilon > 0$ and in terms of ε and $|D|$ fix a sufficiently large positive integer k. Moreover, let

$$L = |D|(k+1),$$

and for prime p let

$$\beta_p = \max_{p^m \le L} m.$$

Finally, define

$$P = \prod_{\substack{p_t \le L/\ell_t: \\ (p_t, D)=1}} p_t^{\beta_{p_t}+1} \prod_{i=1}^{t-1} \prod_{\substack{L/\ell_{i+1} < p_i \le L/\ell_i: \\ \pi(p_i) \in T_i}} p_i^{\beta_{p_i}+1}, \tag{17}$$

where $p_1, \dots, p_t$ denote prime numbers, varying over the respective intervals. Using the observations

$$p^{\beta_p+1} \le L^2$$

and

$$\#T_i = i \quad (i \le t)$$

together with the prime number theorem in arithmetic progressions, from (17) we obtain

$$P \le L^{2(1+\varepsilon)\alpha},$$

where

$$\begin{aligned}
\alpha &= \frac{L/\ell_t}{\log L/\ell_t} + \frac{L}{t} \sum_{i=1}^{t-1} \frac{i(1/\ell_i - 1/\ell_{i+1})}{\log(L/\ell_i - L/\ell_{i+1})} \\
&\le \frac{L/\ell_t}{\log L/\ell_t} + \frac{L}{t} \frac{1}{\log L/\ell_t^2} \sum_{i=1}^{t-1} i \left(\frac{1}{\ell_i} - \frac{1}{\ell_{i+1}}\right) \\
&\le \frac{L/\ell_t}{\log L/\ell_t^2} R(|D|),
\end{aligned}$$

say; here

$$R(|D|) = \sum_{i=1}^{t-1} \frac{i}{\ell_i} - \sum_{i=1}^{t-1} \frac{i}{\ell_{i+1}} + \frac{t}{\ell_t} = \sum_{i=1}^{t} \frac{1}{\ell_i}.$$

To estimate $R(|D|)$, we rewrite the sum and use Lemma 8 to obtain

$$\begin{aligned} R(|D|) &= \sum_{\substack{n \le |D| \\ (n,D)=1}} \frac{1}{n} = \sum_{n \le |D|} \frac{1}{n} \sum_{d|(n,|D|)} \mu(d) = \sum_{d||D|} \frac{\mu(d)}{d} \sum_{k \le |D|/d} \frac{1}{k} \\ &= \sum_{d||D|} \frac{\mu(d)}{d} \left(\log |D| + \gamma - \log d + O(d/|D|)\right) \\ &= (\log |D| + \gamma) \sum_{d||D|} \frac{\mu(d)}{d} + O\left(\sum_{d||D|} \frac{\log d}{d}\right) \\ &= \frac{\varphi(|D|)}{|D|} \log |D| + O((\log\log |D|)^2). \end{aligned}$$

Let

$$\delta = \frac{|D|}{\ell_t} = \frac{|D|}{|D| - 1}.$$

With

$$\lim_{k \to +\infty} \frac{\log L}{\log(L/\ell_t^2)} = 1$$

we obtain, for k sufficiently large in terms of $|D|$ and $\varepsilon > 0$,

$$\begin{aligned} P &\le \exp\left(2(1+2\varepsilon)\frac{L}{\ell_t}\left(\frac{t}{|D|}\log|D| + O((\log\log|D|)^2)\right)\right) \\ &\le \exp\left(2(1+2\varepsilon)\delta(k+1)\left(\frac{t}{|D|}\log|D| + O((\log\log|D|)^2)\right)\right), \end{aligned}$$

and from this

$$\begin{aligned} \frac{k+1}{\log P} &\ge \frac{1}{2(1+2\varepsilon)\delta(\frac{t}{|D|}\log|D| + O((\log\log|D|)^2))} \\ &= \frac{(|D|-1)/\varphi(|D|)}{2(1+2\varepsilon)(\log|D| + O((\log\log|D|)^3))}. \end{aligned} \tag{18}$$

Now choose $y \in \{1, \dots, P\}$ such that

$$|D|y \equiv r \pmod{P}$$

which is possible as $(D, P) = 1$ by definition (17) of P. We claim that none of the numbers $y+1, \dots, y+k$ can be represented by any binary quadratic form of discriminant D, which together with $y \le P$ and (18) proves the theorem. To settle the claim, fix $j \in \{1, \dots, k\}$. Now

$$|D|(y+j) \equiv |D|j + r \pmod{P}. \tag{19}$$

Since

$$-1 = \left(\frac{D}{r}\right) = \left(\frac{D}{|D|j+r}\right)$$

by Lemma 6, we conclude that $|D|j+r$ must be divisible by a prime p with $(D/p) = -1$ to an odd power γ at most β_p: Writing

$$|D|j + r = p^\gamma \ell$$

where ℓ is a certain positive integer coprime to D and p, we find that $|D|j + r \leq |D|(k + 1) = L$, so $\gamma \leq \beta_p$. If $\gamma \geq 3$, then $p \leq L^{1/3}$. As $L/\ell_t \geq L^{1/3}$ for sufficiently large k (in terms of D), by (17) we can then assume that p^{β_p+1} divides P. If $\gamma = 1$, then $|D|j + r = p\ell$, so $p \leq L$. Moreover, if $L/\ell_{i+1} < p \leq L/\ell_i$, then $\ell \leq \ell_i$, so $\pi(p) \in T_i$ by definition (16) of T_i, whence p^{β_p+1} again divides P by (17) as well as in case $p \leq L/\ell_t$, once more by definition (17). Using (19), we conclude that p divides $y + j$ to an odd power, so as $(D/p) = -1$ by Lemma 7 the number $y + j$ indeed cannot be represented by any binary quadratic form of discriminant D. This proves part B) of Theorem 2. For part A), which is without any O-term, we use $\ell_i \geq i$ for all $i \in \mathbb{N}$ to obtain the alternative upper bound

$$R(|D|) = \sum_{i=1}^{t} \frac{1}{\ell_i} \leq \sum_{i=1}^{t} \frac{1}{i} \leq (1 + \log t).$$

Now as above we obtain, for k sufficiently large in terms of $|D|$ and ε,

$$P \leq \exp\left(2(1 + 2\varepsilon)\delta(k + 1)(1 + \log t)\right),$$

and from this

$$\frac{k+1}{\log P} \geq \frac{|D| - 1}{2(1 + 2\varepsilon)|D|(1 + \log \varphi(|D|)}.$$

We now proceed as above to complete the proof of part A).

8. Sums of two squares in the sequence $C_d + x^d$: Proof of Theorem 4

Proof. A) If x is odd, then $p_d(x) \equiv 2 + 1 = 3 \pmod 4$, hence $p_d(x) \notin \mathcal{S}$. If x is even, then $p_d(x) \equiv 6 + 0 = 6 \pmod 8$, hence $p_d(x) \notin \mathcal{S}$.

B) Let d be odd and a composite number $d = qd_2$, where q is the least prime factor of d. Then $p_d(x) = N = (c_q q)^q + x^d$. Suppose that $N \in \mathcal{S}$.

Case I: Let $q \equiv 3 \pmod 4$. As $c_q = 2$, it is not possible that x is even, as otherwise $N/2^q \equiv 3 \pmod 4$, and $N \notin \mathcal{S}$. If $x \equiv 3 \pmod 4$, then $N \equiv 3 \pmod 4$, and again $N \notin \mathcal{S}$. Therefore $x \equiv 1 \pmod 4$. One can factorize N as follows:

$$N = (c_q q)^q + x^d = (c_q q + x^{d_2}) \left(\sum_{i=1}^{q} (-1)^{i+1} (c_q q)^{i-1} x^{d-id_2} \right) = N_1 N_2, \tag{20}$$

say. Note that $N_1 = c_q q + x^{d_2} \equiv 2q + 1 \equiv 3 \pmod 4$, as $x \equiv 1 \pmod 4$ and $q \equiv 3 \pmod 4$. Further, $N \in \mathcal{S}$. Hence there exists a prime $t \equiv 3 \pmod 4$ which must divide both factors N_1 and N_2. Now $c_q q \equiv -x^{d_2} \pmod t$. This gives

$$\begin{aligned} N_2 &= x^{d-d_2} - (c_q q)x^{d-2d_2} \pm \cdots + (c_q q)^{q-1} \\ &\equiv x^{d-d_2} + x^{d-d_2} + \cdots + x^{d-d_2} \equiv q x^{d-d_2} \pmod t. \end{aligned} \tag{21}$$

We distinguish whether $t \mid q$ or $t \nmid q$. Let us first assume that $t \mid q$, then also $t \mid x$, as $t \mid c_q q + x^{d_2}$. Now $d_2 > 1$, so t exactly divides $N_1 = c_q q + x^{d_2}$. On the other hand t exactly divides N_2 to an even power, as with $c_q = 2$ each of the summands $(2q)^{i-1} x^{d-id_2}$ is divisible by q^{q-1}, and all but one

summand are even divisible by q^q. Hence t^{q-1} exactly divides N_2, and t^q exactly divides N. Since $q \equiv t \equiv 3 \pmod 4$, this contradicts $N \in \mathcal{S}$.

Hence $t \nmid q$, so from (21) we obtain that $t \mid x$ and hence $t \mid 2q$, so $t = 2$, contradicting $t \equiv 3 \pmod 4$.

Case II: Let $q \equiv 1 \pmod 4$. As $c_q = 6$, it is not possible that x is even, as otherwise $N/2^q \equiv 3 \pmod 4$, and $N \notin \mathcal{S}$. If $x \equiv 3 \pmod 4$, then $N \equiv 3 \pmod 4$, and again $N \notin \mathcal{S}$. Therefore $x \equiv 1 \pmod 4$. One can factorize N in the same way as in (20). Note that the first factor satisfies $N_1 = c_q q + x^{d_2} \equiv 6q+1 \equiv 3 \pmod 4$, as $x \equiv 1 \pmod 4$ and $q \equiv 1 \pmod 4$. As in Case I, this factor N_1 contains a prime divisor $t \equiv 3 \pmod 4$ with odd multiplicity, which must also divide the second factor N_2. Further, as in Case 1) we have $c_q q \equiv -x^{d_2} \pmod t$ from which we obtain (21). In this case $t \nmid q$ is obvious, as $t \equiv 3 \bmod 4$ and $q \equiv 1 \pmod 4$ are distinct primes. It follows from (21) that $t \mid x$ and hence also $t \mid 6q$. Therefore $t = 3$. But then the first factor $N_1 = 6q + x^{d_2}$ shows that $3 \mid x$. With $d_2 > 1$ we see again that $t = 3$ exactly divides N_1, whereas in the second factor N_2 each term $(c_q q)^{i-1} x^{d-id_2}$ is divisible by an even power of 3, and exactly one summand is divisible by 3 with the minimum exponent $q-1$. Hence again N_2 is exactly divisible by 3 with even exponent, and so $N \notin \mathcal{S}$.

C) It is known that every integer $N \in \mathbb{Z}$ can be written in the form $N = x^2 + y^2 + z^3$, with $z \in \mathbb{Z}$. Indeed, Elkies, Kaplansky and Adler [7] showed the existence of a finite set of congruences covering all integers:

$$\begin{aligned}
2x+1 &= (x^3 - 3x^2 + x)^2 + (x^2 - x - 1)^2 - (x^2 - 2x)^3 \\
4x+2 &= (2x^3 - 2x^2 - x)^2 + (2x^3 - 4x^2 - x + 1)^2 - (2x^2 - 2x - 1)^3 \\
8x+4 &= (x^3 + x + 2)^2 + (x^2 - 2x - 1)^2 - (x^2 + 1)^3 \\
16x+8 &= (2x^3 - 8x^2 + 4x + 2)^2 + (2x^3 - 4x^2 - 2)^2 - (2x^2 - 4x)^3 \\
16x &= (x^3 + 7x - 2)^2 + (x^2 + 2x + 11)^2 - (x^2 + 5)^3
\end{aligned}$$

Now, for every $N \in \mathbb{Z}$ there exists some $z \in \mathbb{Z}$ such that $N - z^3 = N + (-z)^3$ is a sum of two squares. An easier argument holds for the exponent $d = 2$. Every odd integer is a difference of two consecutive squares: $2k+1 = (k+1)^2 - k^2 = y^2 - z^2$. With $x = 0$ or 1 we see that every integer can be written as $N = x^2 + y^2 - z^2$ and hence that for every $N \in \mathbb{Z}$ there exists a $z \in \mathbb{Z}$ such that $N + z^2 = x^2 + y^2$.

□

9. Long gaps between sums of two squares in sparse sequences: Proof of Theorem 5

In order to prove Theorem 5, we need the following auxiliary results.

Lemma 9. *Let d be an odd prime as in the statement of Theorem 5, and let j be a positive integer. Then*

$$\gcd\left(\frac{(4dj)^d - 1}{4dj - 1}, 4dj - 1\right) = 1.$$

Proof. Let $a = 4dj$. Then the claim is

$$\gcd\left(\frac{a^d - 1}{a - 1}, a - 1\right) = 1.$$

As $a^i \equiv 1 \pmod{a-1}$, when $a > 2$, one also has:

$$\frac{a^d-1}{a-1} = \sum_{i=0}^{d-1} a^i \equiv \sum_{i=0}^{d-1} 1 \equiv d \pmod{a-1}.$$

Applied with $a = 4dj$, this gives

$$\gcd\left(\frac{(4dj)^d-1}{4dj-1}, 4dj-1\right) = (d, 4dj-1) = 1$$

and the claim follows. □

Lemma 10. *Let $\mathcal{S}(x)$ denote the counting function of $\mathcal{S}\backslash\{0\}$. For $x \geq 2$ we have*

$$\left|\mathcal{S}(x) - C_R\frac{x}{\sqrt{\log x}}\right| \leq 9.62\frac{x}{\log x},$$

where C_R is the Landau-Ramanujan constant (see equation (7)).

Proof. This is Lemma 9 a) in [20]. □

Lemma 11. *Let $C_1 = 9.62$. For $x \geq 2$ we have*

$$\sum_{1\leq m\leq x, m\in\mathcal{S}} \frac{1}{m} \leq 2C_R\sqrt{\log x} + C_1\log\log x + 1 - 2C_R\sqrt{\log 2} - C_1\log\log 2 + \frac{C_R}{\sqrt{\log x}} + \frac{C_1}{\log x}.$$

Proof. By Lemma 10 the inequality

$$\mathcal{S}(x) \leq C_R\frac{x}{\sqrt{\log x}} + C_1\frac{x}{\log x}$$

holds. Applying partial summation with $a_m = 1$ if $m \in \mathcal{S}$, and $a_m = 0$ otherwise, we obtain

$$\sum_{2\leq m\leq x} \frac{a_m}{m} = \frac{1}{x}\sum_{2\leq m\leq x} a_m + \int_2^x \left(\sum_{m\leq u} a_m\right)\frac{1}{u^2}\,du.$$

It follows for $x \geq 2$ that

$$\begin{aligned}\sum_{1\leq m\leq x, m\in\mathcal{S}} \frac{1}{m} &\leq 1 + \frac{C_R}{\sqrt{\log x}} + \frac{C_1}{\log x} + \int_2^x \left(\frac{C_R}{u\sqrt{\log u}} + \frac{C_1}{u\log u}\right)du \\ &\leq 2C_R\sqrt{\log x} + C_1\log\log x - 1 + \frac{C_R}{\sqrt{\log x}} + \frac{C_1}{\log x} \\ &\quad - 2C_R\sqrt{\log 2} - C_1\log\log 2.\end{aligned}$$

□

Lemma 12. *Let $\varepsilon > 0$ and let d be sufficiently large in terms of ε. Then*

$$\sum_{m\leq d, m\ odd, m\in\mathcal{S}} \frac{1}{m} \leq (C_R+\varepsilon)\sqrt{\log d}. \tag{22}$$

Proof. It clearly follows from Lemma 10, for sufficiently large $d \geq d_\varepsilon$, that

$$\sum_{1\leq m\leq d, m\in\mathcal{S}} \frac{1}{m} \leq (2C_R+\varepsilon)\sqrt{\log d}$$

holds. As the sum in our estimate only runs over odd values m we observe: if $m \in \mathcal{S}$ is odd, then $2m, 4m, \dots$ is also in $\mathcal{S}$. The sum over a dyadic interval gives a bounded contribution, hence the major contribution comes from small $m \le x/2^i$, for some large i. Hence the "odd" contribution is almost half the contribution of all such power-of-2-multiples of m, also counted in the above sum:

$$\frac{1}{m} \le \frac{1}{m} + \frac{1}{2m} + \cdots + \frac{1}{2^i m} = \frac{2}{m} - \frac{1}{2^i m}.$$

Hence for $d \ge d_\varepsilon$ we obtain (22). □

Lemma 13. *For $d \ge 17$ we have*

$$\sum_{1 \le m \le d, m \in \mathcal{S}} \frac{1}{m} \le 13\sqrt{\log d}. \tag{23}$$

Proof. Let C_R be the Landau-Ramanujan constant (see equation (7)) and let $C_1 = 9.62$. Using the inequality

$$\log\log x \le \frac{2}{e}\sqrt{\log x} \quad (x \ge 3),$$

which is easy to verify by calculus, we obtain

$$2C_R\sqrt{\log d} + C_1 \log\log d < 8.66\sqrt{\log d}.$$

Further, observing for $d \ge 17$ the inequality

$$1 + \frac{C_R}{\sqrt{\log d}} + \frac{C_1}{\log d} - 2C_R\sqrt{\log 2} - C_1 \log\log 2 < 4.242\sqrt{\log d}$$

then via Lemma 11 confirms (23). □

Proof of Theorem 5. For convenience, let us introduce the function $f_d(j) = j^d$. Recall that $d \ge 3$ and fix the gap size k. For each prime $p \le 4kd, p \equiv 3 \pmod 4$ let $\beta = \beta_p$ be the highest power with $p^\beta \le 4kd$. Let

$$P = \prod_{\substack{p \le 4k \\ p \equiv 3 \pmod 4, (p,d)=1}} p^{\beta_p + 1} \prod_{\substack{m \le d, (m,4d)=1 \\ m \in \mathcal{S}}} \prod_{\substack{p \le 4kd/m \\ pm \equiv -1 \pmod{4d}}} p^{\beta_p + 1}.$$

Define $y \in \{1, \dots, P\}$ by the congruence $(4d)^d y \equiv -1 \pmod P$. We show below that none of the integers in

$$I = \{y + f_d(1), \dots, y + f_d(k)\} \tag{24}$$

is the sum of two squares. To estimate the size of P, by the prime number theorem in arithmetic progressions, we have

$$\prod_{\substack{p \le 4k \\ p \equiv 3 \pmod 4, (p,d)=1}} p^{\beta_p + 1} \le (4k)^{2(1+\varepsilon)2k/\log(4k)} \le \exp(4(1+\varepsilon)k).$$

Similarly, for any integer m with

$$0 < m \le d, (m, 4d) = 1$$

we obtain

$$\prod_{\substack{p\le 4kd/m\\ pm\equiv -1 \pmod{4d}}} p^{\beta_p+1} \le (4kd)^{2(1+\varepsilon)4kd/(m\varphi(4d)\log(4kd/m))} = \exp\left(\frac{8kd(1+\varepsilon)\log(4kd)}{\varphi(4d)m\log(4kd/m)}\right).$$

Now

$$\frac{\log(4kd)}{\log(4kd/m)} = 1 + O(1/\log k) = 1 + o(1)$$

as $k \to +\infty$. Noting that $\varphi(4d) = 2\varphi(d) = 2(d-1)$ as d is an odd prime, we obtain

$$\log P \le 4(1+\varepsilon)k + \sum_{\text{odd } m\in\mathcal{S}\cap[1,d-1]} \frac{4kd(1+\varepsilon)(1+o(1))}{(d-1)m} \le \frac{4kd(1+\varepsilon)(1+o(1))\left(1-\frac{1}{d}+\sum_{\text{odd } m\in\mathcal{S}\cap[1,d-1]}\frac{1}{m}\right)}{d-1}.$$

Consequently,

$$\frac{k}{\log P} \ge \frac{d-1}{4d(1+\varepsilon)(1+o(1))\left(1-\frac{1}{d}+\sum_{\text{odd } m\in\mathcal{S}\cap[1,d-1]}\frac{1}{m}\right)},$$

so using $y \le P$ we find that

$$\limsup_{k\to\infty}\frac{k}{\log y} \ge \frac{d-1}{4d\left(1-\frac{1}{d}+\sum_{\text{odd } m\in\mathcal{S}\cap[1,d-1]}\frac{1}{m}\right)} \tag{25}$$

as required in the first part of Theorem 5; the explicit numerical values for $d = 3, 5, 7, 11, 13, 17$ follow from a straightforward computation.

From (25) and Lemma 12 we obtain the second statement in Theorem 5.

Finally, (25) and Lemma 13 yield

$$\limsup_{k\to\infty}\frac{k}{\log y_k} \ge \frac{d-1}{4d\left(1+\sum_{m\in\mathcal{S}\cap[1,d-1]}\frac{1}{m}\right)} \ge \frac{d-1}{4d(1+13\sqrt{\log d})} \ge \frac{1}{60\sqrt{\log d}},$$

when $d \ge 17$. Clearly, for all particular regions of d better constants can be achieved.

To prove our claim from above, that none of the integers in (24) is a sum of two squares, note that for $1 \le j \le k$

$$(4d)^d(y + f_d(j)) \equiv f_d(4dj) - 1 \pmod{P}.$$

Now $a = 4dj - 1$ is a divisor of $f_d(4dj) - 1$ and further, as $a \equiv 3 \pmod 4$ and $(a, d) = 1$, the number a is exactly divisible by some prime power p^α with

$$p \equiv 3 \pmod 4, \quad p \ne d \quad \text{and} \quad \alpha \text{ odd}. \tag{26}$$

By Lemma 9 the codivisor $\frac{(4dj)^d-1}{4dj-1}$ is coprime to a, which implies that $f_d(4dj) - 1$ is exactly divisible by this prime power p^α. Assume that p^α is the smallest prime power dividing $4dj - 1$ with the condition in (26). As $\alpha \le \beta_p$, if $p \le 4k$, then P is divisible by $p^{\alpha+1} \mid p^{\beta_p+1}$, so that $y + f_d(j)$ is also exactly divisible by p^α, and is therefore not the sum of two squares. On the other hand, if $p > 4k$, then for sufficiently large k the number $4dj-1$ cannot be divisible by any other prime power $q^\nu \equiv 3 \pmod 4$, because in this case due to the minimality of p^α, α odd, we obtain

$q^\nu p \geq p^2 > 16k^2$ and thus $4dj-1 > 16k^2 > 4dk-1$, contradicting $j \leq k$. Therefore, the number $m = (4dj-1)/p$ is a sum of two squares and also $\frac{4dj-1}{p} < \frac{4dk}{4k} = d$, so that $pm = 4dj - 1 \equiv -1 \pmod{4d}$ and p^{β_p+1} divides P. Now $\alpha \leq \beta_p$, hence $y + f_d(j)$ is exactly divisible by p^α and is not a sum of two squares. □

10. Concluding remarks

Remark 6. One may wonder about another connection to primes. It is known that every interval of type $[X, X + X^{0.525}]$, where X is sufficiently large, contains at least $\gg \frac{X^{0.525}}{\log X}$ many primes.

Let $\varepsilon > 0$, and $n \geq n_\varepsilon$ be sufficiently large. By a binomial estimate, see [21], the interval $[2^n - 2^{0.525n}, 2^n]$ must contain primes where the proportion of binary digits being '1' is, for any $\varepsilon > 0$, larger than $0.7375 - \varepsilon$. Similarly, the interval $[2^n, 2^n + 2^{0.525n}]$ contains many primes with at least $(0.7375 - \varepsilon)n$ many binary digits being '0'.

In view of $2^n + 2^i$ it is clear that $\mathcal{S}$ contains elements with very few binary 1-digits only. Interestingly, with another identity one can also achieve this for sums of two squares with only very few '0' digits: The integers of the form

$$\begin{aligned} N &= \left(\sum_{i=0}^{2^n+1} 2^i\right) - \left(2^{2^n} + 2\right) = 3(2^{2^n} - 1) \\ &= 3(2^{2^1} - 1)(2^{2^1} + 1)(2^{2^2} + 1) \cdots (2^{2^{n-1}} + 1) \\ &= 9(2^2 + 1)(2^4 + 1) \cdots (2^{2^{n-1}} + 1) \end{aligned}$$

have exactly two binary digits being zero. Here we repeatedly used the binomial formula $2^{2i} - 1 = (2^i + 1)(2^i - 1)$. Moreover as a product of integers

$$9, 2^2 + 1, 2^4 + 1, \ldots,$$

where each factor is a sum of two squares, and as this property is multiplicative, N is itself a sum of two squares. Hence the problem of binary digits in the set $\mathcal{S}$ is considerably easier than in the set of primes, due to explicit identities.

For the quadratic form $3x^2 + y^2$ this also works with exactly three '0' digits. We use several times the identity

$$(2^{3^i} - 1)(2^{2\times 3^i} + 2^{3^i} + 1) = 2^{3^{i+1}} - 1.$$

Now

$$\begin{aligned} N &= 7(2^{3^n} - 1) \\ &= 7(2^3 - 1)(2^6 + 2^3 + 1) \times (2^{18} + 2^9 + 1) \times \cdots \times (2^{2\times 3^{n-1}} + 2^{3^{n-1}} + 1) \\ &= \left(\sum_{i=0}^{3^n+2} 2^i\right) - \left(2^1 + 2^2 + 2^{3^n}\right). \end{aligned}$$

Note that N is a product of $49 = 3 \times 4^2 + 1^2$ and integers of the form

$$2^{2\times 3^i} + 2^{3^i} + 1 = 3x^2 + (x+1)^2 = 4x^2 + 2x + 1$$

with $x = 2^{3^i - 1}$. A product of two integers of the form $3x^2 + y^2$ is again of this type, in view of the identity

$$(3a^2 + b^2)(3c^2 + d^2) = 3(bc + ad)^2 + (3ac - bd)^2.$$

Hence N is of type $3x^2 + y^2$ and has only three of its binary digits being 0.

Remark 7. It is trivial that short gaps exist between integers in $\mathcal{S}$, such as $3 = (n^2 + 4) - (n^2 + 1)$. Brüdern and Dietmann [4] showed that each positive integer occurs as the gap between pairs in $\mathcal{S}$ infinitely often, and they also study triples of integers in $\mathcal{S}$. In view of the Green-Tao theorem on primes in arithmetic progressions it is clear that $\mathcal{S}$ contains arbitrarily long arithmetic progressions. Good upper bounds on the length of progressions in $\mathcal{S}$ in terms of the size of the gap have recently been given in [8].

References

[1] Antal Balog and Trevor D. Wooley. Sums of two squares in short intervals. *Canad. J. Math.*, 52(4):673–694, 2000.

[2] R. P. Bambah and S. Chowla. On numbers which can be expressed as a sum of two squares. *Proc. Nat. Inst. Sci. India*, 13:101–103, 1947.

[3] Ahmed Bonfoh and Cyril D. Enyi. The Cahn-Hilliard equation as limit of a conserved phase-field system. *Asymptot. Anal.*, 101(3):97–148, 2017.

[4] Jörg Brüdern and Rainer Dietmann. On the gaps between values of binary quadratic forms. *Proc. Edinb. Math. Soc. (2)*, 54(1):25–32, 2011.

[5] Henri Cohen. *Number theory. Vol. I. Tools and Diophantine equations*, volume 239 of *Graduate Texts in Mathematics*. Springer, New York, 2007.

[6] H. Cramér. On the order of magnitude of the difference between consecutive prime numbers. *Acta Arith.*, 2(1):23–46, 1936.

[7] Noam Elkies, I. Kaplansky, and A. Adler. Sums of two squares and a cube, (problem 10426 by Elkies and Kaplansky, solution by Adler). *The American Mathematical Monthly*, 104 (6):574, 1997.

[8] Christian Elsholtz and Christopher Frei. Arithmetic progressions in binary quadratic forms and norm forms. *Bull. Lond. Math. Soc.*, 51(4):595–602, 2019.

[9] P. Erdös. Some problems and results in elementary number theory. *Publ. Math. Debrecen*, 2:103–109, 1951.

[10] Kevin Ford, Ben Green, Sergei Konyagin, James Maynard, and Terence Tao. Long gaps between primes. *J. Amer. Math. Soc.*, 31(1):65–105, 2018.

[11] Kevin Ford, Ben Green, Sergei Konyagin, and Terence Tao. Large gaps between consecutive prime numbers. *Ann. of Math. (2)*, 183(3):935–974, 2016.

[12] Kevin Ford, Sergei Konyagin, James Maynard, Carl Pomerance, and Terence Tao. Long gaps in sieved sets. *J. Eur. Math. Soc. (JEMS)*, 23(2):667–700, 2021.

[13] G. H. Hardy and E. M. Wright. *An introduction to the theory of numbers*. Oxford University Press, Oxford, sixth edition, 2008. Revised by D. R. Heath-Brown and J. H. Silverman, With a foreword by Andrew Wiles.

[14] Christopher Hooley. On the intervals between numbers that are sums of two squares. IV. *J. Reine Angew. Math.*, 452:79–109, 1994.

[15] A. E. Ingham. Some Asymptotic Formulae in the Theory of Numbers. *J. London Math. Soc.*, 2(3):202–208, 1927.

[16] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[17] William C. Jagy and Irving Kaplansky. Sums of squares, cubes, and higher powers. *Experiment. Math.*, 4(3):169–173, 1995.

[18] James Maynard. Sums of two squares in short intervals. In *Analytic number theory*, pages 253–273. Springer, Cham, 2015.

[19] James Maynard. Large gaps between primes. *Ann. of Math. (2)*, 183(3):915–933, 2016.

[20] Pieter Moree and Herman J. J. te Riele. The hexagonal versus the square lattice. *Math. Comp.*, 73(245):451–473, 2004.

[21] Eric Naslund. The tail distribution of the sum of digits of prime numbers. *Unif. Distrib. Theory*, 10(1):63–68, 2015.

[22] Ian Richards. On the gaps between numbers which are sums of two squares. *Adv. in Math.*, 46(1):1–2, 1982.

[23] Peter Shiu. The gaps between sums of two squares. *Math. Gaz.*, 97(539):256–262, 2013.
[24] Richard Warlimont. Über einen Satz von P. Erdős. *Arch. Math. (Basel)*, 27(2):164–168, 1976.
[25] D. B. Zagier. *Zetafunktionen und quadratische Körper*. Springer-Verlag, Berlin-New York, 1981. Eine Einführung in die höhere Zahlentheorie. [An introduction to higher number theory], Hochschultext. [University Text].

Appendix A. Table of abbreviations

In this appendix we briefly collect some notation used in the proof of Theorem 1 in section 4. First recall the maps

$$\pi_\ell : \mathbb{Z} \to \{1, \dots, 2^\ell\}, \quad x \mapsto x \pmod{2^\ell}$$

and

$$\tau : \mathbb{Z} \to \mathbb{Z}; \quad j \mapsto 4j - 1.$$

The sets S_ℓ, T_ℓ, U_ℓ, V_ℓ, W_ℓ and R_ℓ are then defined by

$$\begin{aligned}
S_l &= \{2^a b \in \{1, \dots, 2^\ell\} : a \le \ell - 2, b \equiv 3 \pmod 4\} \quad (\ell \ge 2),\\
T_{\ell+2} &= \pi_{\ell+2}(\tau(S_\ell)) \quad (\ell \ge 2),\\
U_3 &= \{3\},\\
U_\ell &= U_{\ell-1} \cup \{u + 2^{\ell-1} : u \in U_{\ell-1}\} \cup \{3 \times 2^{\ell-2}\} \quad (\ell \ge 4),\\
V_\ell &= \{s \in S_\ell : \pi_{\ell+2}(5\tau(s)) \in T_{\ell+2}\} \quad (\ell \ge 2),\\
W_5 &= \{24\},\\
W_\ell &= W_{\ell-1} \cup \{u + 2^{\ell-1} : u \in W_{\ell-1}\} \cup \{3 \times 2^{\ell-2}\} \quad (\ell \ge 6),\\
R_\ell &= \{s \in S_\ell : \pi_{\ell+2}(5\tau(s)) \in T_{\ell+2} \text{ and } \pi_{\ell+2}(9\tau(s)) \in T_{\ell+2}\} \quad (\ell \ge 2).
\end{aligned}$$

Note the inclusions

$$\begin{aligned}
R_\ell &\subset V_\ell \subset S_\ell \quad (\ell \ge 5),\\
W_\ell &\subset U_\ell \subset S_\ell \quad (\ell \ge 5),\\
U_\ell &\subset V_\ell \quad (\ell \ge 3),\\
W_\ell &\subset R_\ell \quad (\ell \ge 5).
\end{aligned}$$

Appendix B. Some examples for the sets of residue classes introduced in section 4

We have

$$\begin{aligned}
S_2 &= \{3\},\\
S_3 &= \{3,6,7\},\\
S_4 &= \{3,6,7,11,12,14,15\},\\
S_5 &= \{3,6,7,11,12,14,15,19,22,23,24,27,28,30,31\},\\
T_4 &= \{11\},\\
T_5 &= \{11,23,27\},\\
T_6 &= \{11,23,27,43,47,55,59\},\\
T_7 &= \{11,23,27,43,47,55,59,75,87,91,95,107,111,119,123\},\\
U_3 &= \{3\} = V_3 \subset S_3,\\
U_4 &= \{3,11,12\} = V_4 \subset S_4,\\
U_5 &= \{3,11,12,19,24,27,28\} = V_5 \subset S_5,\\
\pi_5(\tau(U_3)) &= \{11\} \subset T_5,\\
\pi_6(\tau(U_4)) &= \{11,43,47\} \subset T_6,\\
\pi_7(\tau(U_5)) &= \{11,43,47,75,95,107,111\} \subset T_7,\\
W_5 &= \{24\} = R_5,\\
\pi_7(\tau(W_5)) &= \{95\} \subset \pi_7(\tau(U_5)) \subset T_7.
\end{aligned}$$

Appendix C. Cramér's model, adapted to sums of two squares

In this section we briefly discuss the adaptation of Cramér's model (see [6]) for primes to the set $\mathcal{S}$. The Cramér random model for the set $\mathcal{S}$ of numbers representable as the sum of two squares suggests that distributional properties of $\mathcal{S}$ should be similar to that of a random set $\mathcal{R}$ where each integer $k > 1$ is included in $\mathcal{R}$ independently with probability $C_R/\sqrt{\log k}$. Here C_R is the Landau-Ramanujan constant

$$C_R = \frac{1}{\sqrt{2}} \prod_{p \equiv 3 \pmod 4} \left(1 - \frac{1}{p^2}\right)^{-1/2} \approx 0.7642\ldots$$

(already introduced in (7)), chosen such that the size of $\mathcal{S} \cap \{1,\ldots,x\}$ is asymptotically that of $\mathcal{R} \cap \{1,\ldots,x\}$ with probability 1.

Given $c > 0$, let $f(k) := \lfloor c(\log k)^{3/2}/C_R \rfloor$ and let $\mathcal{E}_k$ denote the random event that

$$\mathcal{R} \cap \{k+1,\ldots,k+f(k)\} = \emptyset.$$

We then see that

$$\mathbb{P}(\mathcal{E}_k) = \prod_{j=1}^{f(k)} \Big(1 - \frac{C_R}{\sqrt{\log(k+j)}}\Big) = \exp\Big(-\frac{C_R(1+o(1))f(k)}{\sqrt{\log k}}\Big) = k^{-c+o(1)}.$$

If $c > 1$ then $\sum_k \mathbb{P}(\mathcal{E}_k) < \infty$, and so the Borel-Cantelli lemma implies that almost surely only finitely many of the events $\mathcal{E}_k$ occur. On the other hand, we see that the events $\mathcal{E}_{\lceil k(\log k)^2 \rceil}$ are independent for k large enough (the underlying sets are disjoint) and if $c < 1$ then $\sum_k \mathbb{P}(\mathcal{E}_{\lceil k(\log k)^2 \rceil}) = \infty$. Therefore by the Borel-Cantelli

lemma if $c < 1$ then almost surely infinitely many of the events occur. Therefore we find that if $\mathcal{R} = \{r_1, r_2, \dots\}$ with $r_1 < r_2 < \dots$ then with probability 1 we have

$$\limsup_{k\to\infty} \frac{r_{k+1} - r_k}{(\log r_k)^{3/2}} = \frac{1}{C_R}.$$

In particular, the Cramér random model would predict that the maximal gap between elements of $\mathcal{S} \cap \{1, \dots, x\}$ should be of size roughly $(\log x)^{3/2}$. (It is likely the maximal gap would differ slightly from the precise prediction $(1/C_R + o(1))(\log x)^{3/2}$ from this model because the model does not account for divisibility effects caused by small primes.)

Department of Mathematics, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, United Kingdom
Email address: rainer.dietmann@rhul.ac.uk

Institute of Analysis and Number Theory, Graz University of Technology, Kopernikusgasse 24/II, Graz, A-8010 Graz, Austria
Email address: elsholtz@math.tugraz.at

National Research University Higher School of Economics, Russian Federation, 6 Usacheva Str.,Moscow, Russia, 119048
Email address: alkalb1995cd@mail.ru

Steklov Institute of Mathematics, 8 Gubkin Str., Moscow, Russia, 119991
Email address: konyagin@mi-ras.ru

Mathematical Institute, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, England
Email address: james.alexander.maynard@gmail.com