Mathematika 67 (2021) 214–234 doi:10.1112/mtk.12069

CYCLOTOMIC POLYNOMIALS WITH PRESCRIBED HEIGHT AND PRIME NUMBER THEORY

ALEXANDRE KOSYAK, PIETER MOREE, EFTHYMIOS SOFOS AND BIN ZHANG

Abstract. Given any positive integer n, let $A(n)$ denote the height of the nth cyclotomic polynomial, that is its maximum coefficient in absolute value. It is well known that $A(n)$ is unbounded. We conjecture that every natural number can arise as value of $A(n)$ and prove this assuming that for every pair of consecutive primes p and p' with $p \geqslant 127$, we have $p' - p < \sqrt{p} + 1$. We also conjecture that every natural number occurs as the maximum coefficient of some cyclotomic polynomial and show that this is true if Andrica's conjecture holds, that is, that $\sqrt{p'} - \sqrt{p} < 1$ always holds. This is the first time, as far as the authors know that a connection between prime gaps and cyclotomic polynomials is uncovered. Using a result of Heath–Brown on prime gaps, we show unconditionally that every natural number $m \leqslant x$ occurs as $A(n)$ value with at most $O_\epsilon(x^{3/5+\epsilon})$ exceptions. On the Lindelöf Hypothesis, we show there are at most $O_\epsilon(x^{1/2+\epsilon})$ exceptions and study them further by using deep work of Bombieri–Friedlander–Iwaniec on the distribution of primes in arithmetic progressions beyond the square-root barrier.

§1. *Introduction.* Let $n \geqslant 1$ be an integer. The nth cyclotomic polynomial

$$\Phi_n(x) = \sum_{j=0}^{\varphi(n)} a_n(j)x^j,$$

is a polynomial of degree $\varphi(n)$, with φ Euler's totient function. For $j > \varphi(n)$, we put $a_n(j) = 0$. The coefficients $a_n(j)$ are usually very small. In the 19th century, mathematicians even thought that they are always 0 or ± 1. The first counterexample to this claim occurs at $n = 105$: indeed, $a_{105}(7) = -2$. The number 105 is the smallest ternary number (see Definition 1) and these will play a major role in this article. Issai Schur proved that every negative even number occurs as a cyclotomic coefficient. Emma Lehmer [28] reproduced his unpublished proof. Schur's argument is easily adapted to show that *every* integer occurs as a cyclotomic coefficient; see Suzuki [34] or Moree and Hommersom [30, Proposition 5]. Let $m \geqslant 1$ be given. Ji, Li and Moree [23] adapted Schur's argument and proved that

$$\{a_{mn}(j) : n \geqslant 1,\ j \geqslant 0\} = \mathbb{Z}. \tag{1}$$

Fintzen [11] determined the set of all cyclotomic coefficients $a_n(j)$ with j and n in prescribed arithmetic progression, thus generalizing (1).

We put

$$A(n) = \max_{k \geqslant 0} |a_n(k)|, \quad \mathcal{A} = \cup_{n \in \mathbb{N}} A(n), \quad A\{n\} = \{a_n(k) : k \geqslant 0\},$$

in particular $A(n)$ is the height of the cyclotomic polynomial Φ_n.

Received 5 November 2019, published online 6 January 2021.
MSC (2020): 11B83, 11C08 (primary).
The authors, except the fourth, are or were supported by the Max Planck Institute for Mathematics and thankful for this. The fourth author is supported by the National Natural Science Foundation of China (Grant No. 11801303), project ZR2019QA016 supported by the Shandong Provincial Natural Science Foundation and a project funded by the China Postdoctoral Science Foundation (Grant No. 2018M640617).

It is a classical result that if n has at most two distinct odd prime factors, then $A(n) = 1$, cf. Lam and Leung [26]. The first non-trivial case arises where n has precisely three distinct odd prime divisors and thus is of the form $n = p^e q^f r^g$, with $2 < p < q < r$ prime numbers. It is easy to deduce that $A\{p^e q^f r^g\} = A\{pqr\}$ using elementary properties of cyclotomic polynomials (as given, e.g., in [30, Lemma 2]). It thus suffices to consider only the case where $e = f = g = 1$ and so $n = pqr$. This motivates the following definition.

Definition 1. A cyclotomic polynomial $\Phi_n(x)$ is said to be *ternary* if $n = pqr$, with $2 < p < q < r$ primes. In this case, we call the integer $n = pqr$ *ternary*. We put $\mathcal{A}_t = \{A(n) : n \text{ is ternary}\}$.

Note that $\mathcal{A}_t \subseteq \mathcal{A}$. In this article, we address the nature of the sets $\mathcal{A}$, $\mathcal{A}_t$ and $\mathcal{A}_{opt}$ (see Definition 2).

CONJECTURE 1. We have $\mathcal{A} = \mathbb{N}$, that is for any given natural number m, there is a cyclotomic polynomial having height m.

CONJECTURE 2. We have $\mathcal{A}_t = \mathbb{N}$, that is for any given natural number m, there is a ternary n such that Φ_n has height m.

The argument of Schur cannot be adapted to resolve Conjecture 1, as it allows one to control only the coefficients in a tail of a polynomial that quickly becomes very large if we want to show that some larger number occurs as a coefficient, and typically will have much larger coefficients than the coefficient constructed. Instead, we will make use of various properties of ternary cyclotomic polynomials. This class of cyclotomic polynomials has been intensively studied as it is the simplest one where the coefficients display non-trivial behavior. For these, we still have $\{a_n(j) : n \text{ is ternary},\ j \geqslant 0\} = \mathbb{Z}$, as a consequence of the following result.

THEOREM 1 [3]. *For every odd prime p, there exists an infinite family of polynomials Φ_{pqr} such that $A\{pqr\} = [-(p-1)/2, (p+1)/2] \cap \mathbb{Z}$ and another one such that $A\{pqr\} = [-(p+1)/2, (p-1)/2] \cap \mathbb{Z}$.*

If n is ternary, then $A\{n\}$ consists of consecutive integers. Moreover, we have $|a_n(j+1) - a_n(j)| \leqslant 1$ for $j \geqslant 0$; see [15]. Note that for each of the members Φ_{pqr} of the two families, the cardinality of $A\{pqr\}$ is $p + 1$. This is not always the case for arbitrary ternary n and even best possible in the sense that $\#A\{pqr\} \leqslant p + 1$ for arbitrary ternary n (by [2, Corollary 3]).

Definition 2. If the cardinality of $A\{pqr\}$ is exactly $p + 1$, we say that Φ_{pqr} is ternary optimal and call $n = pqr$ optimal. We denote the set of all $A(n)$ with n optimal by $\mathcal{A}_{opt}$.

Note that the bound for the size of $\mathcal{A}\{pqr\}$ depends only on the smallest prime factor, p. Similarly, it has been known since the 19th century that $A(pqr) \leqslant p - 1$.

We expect the following to be true regarding ternary optimal polynomials.

CONJECTURE 3. We have $\mathcal{A}_{opt} = \mathbb{N}\backslash\{1, 5\}$.

We will see that this conjecture is closely related to the following prime number conjecture we propose (with p_n the nth prime number).

CONJECTURE 4. Let $n \geqslant 31$ (and so $p_n \geqslant 127$). Then

$$p_{n+1} - p_n < \sqrt{p_n} + 1. \tag{2}$$

Although prime gaps $d_n := p_{n+1} - p_n$ have been studied *in extenso* in the literature, we have not come across this particular conjecture. It is a bit stronger than Andrica's conjecture (see [35] for some numerics).

CONJECTURE 5 (Andrica's conjecture). For $n \geqslant 1$, $p_{n+1} - p_n < \sqrt{p_n} + \sqrt{p_{n+1}}$, or equivalently $\sqrt{p_{n+1}} - \sqrt{p_n} < 1$, or equivalently $p_{n+1} - p_n < 2\sqrt{p_n} + 1$.

Both conjectures seem to be far out of reach, as under RH the best result is due to Cramér [9] who showed in 1920 that $d_n = O(\sqrt{p_n}\log p_n)$. More explicitly, Carneiro *et al.* [8] showed under RH that $d_n \leqslant \frac{22}{25}\sqrt{p_n}\log p_n$ for every $p_n > 3$.

There is a whole range of conjectures on gaps between consecutive primes. The most famous one is Legendre's that there is a prime between consecutive squares is a bit weaker, but, for example, Firoozbakht's conjecture that $p_n^{1/n}$ is a strictly decreasing function of n is much stronger. Firoozbakht's conjecture implies that $d_n < (\log p_n)^2 - \log p_n + 1$ for all n sufficiently large (see [33]), contradicting a heuristic model; see Banks *et al.* [5], suggesting that given any $\epsilon > 0$ there are infinitely many n such that $d_n > (2e^{-\gamma} - \epsilon)(\log p_n)^2$, with γ Euler's constant.

This is in line with the famous conjecture that

$$0 < \liminf_{x\to\infty} \frac{\max\{d_n : p_n \leqslant x\}}{(\log x)^2} \leqslant \limsup_{x\to\infty} \frac{\max\{d_n : p_n \leqslant x\}}{(\log x)^2} < \infty,$$

stated in 1936 by Cramér [10], who also provided heuristic arguments in support of it. His conjecture implies that $d_n = O((\log p_n)^2)$, which if true, clearly shows that the claimed bound in Conjecture 4 holds for all sufficiently large n. Further work on d_n can be found in [5, 13, 18].

We denote the set of natural numbers $\leqslant h$ by $\mathbb{N}_h$.

THEOREM 2. *Let h be an integer such that* (2) *holds for* $127 \leqslant p_n < 2h$. *Then*

$$\mathbb{N}_h \subseteq \mathcal{A}_t \subseteq \mathcal{A}, \quad \mathbb{N}_h \backslash \{1,5\} \subseteq \mathcal{A}_{opt}.$$

Moreover, $1, 5 \notin \mathcal{A}_{opt}$.

COROLLARY 1. *If Conjecture* 4 *is true, then so are Conjectures* 1–3.

Theorem 2 is in essence a consequence of a result of Moree and Roşu [31] (Theorem 6) generalizing Theorem 1, as we shall see in § 2.

A lot of numerical work on large gaps has been done (see the website [32]). This can be used to infer that the inequality (2) holds whenever $127 \leqslant p_n \leqslant 2 \cdot 2^{63} \approx 1.8 \cdot 10^{19}$; see [35]. This in combination with Theorem 2 leads to the following proposition.

PROPOSITION 1. *Every integer up to $9 \cdot 10^{18}$ occurs as the height of some ternary cyclotomic polynomial.*

The following theorem is the main result of our paper. Its proof rests on combining Lemma 3 b, the key lemma used to prove Theorem 2, with deep work by Heath–Brown [**21**] and Yu [**37**] on gaps between primes.

THEOREM 3. *Almost all positive integers occur as the height of an optimal ternary cyclotomic polynomial. Specifically, for any fixed $\epsilon > 0$, the number of positive integers $\leqslant x$ that do not occur as a height of an optimal ternary cyclotomic polynomial is $\ll_\epsilon x^{3/5+\epsilon}$. Under the Lindelöf Hypothesis, this number is $\ll_\epsilon x^{1/2+\epsilon}$.*

(Readers unfamiliar with the Lindelöf Hypothesis are referred to the paragraph in § 3 before the statement of Lemma 7.)

In addition to Conjecture 4, there are two further prime number conjectures of relevance for the topic at hand: Conjecture 5, that we have not come across in the literature, and Andrica's conjecture (Conjecture 5).

CONJECTURE 6. Let $h > 1$ be odd. There exists a prime $p \geqslant 2h - 1$, such that $1 + (h - 1)p$ is a prime too.

The widely believed Bateman–Horn conjecture [**1**] implies that given an odd $h > 1$, there are infinitely many primes p such that $1 + (h - 1)p$ is a prime too, and thus Conjecture 6 is a weaker version of this.

THEOREM 4. *If Conjecture 6 holds true, then $\mathcal{A}_t$ contains all odd natural numbers. Unconditionally $\mathcal{A}_t$ contains a positive fraction of all odd natural numbers.*

The first assertion is a consequence of work of Gallot *et al.* [**17**] and involves ternary cyclotomic polynomials that are not optimal. The second makes use of deep work of Bombieri *et al.* [**6**] on the level of distribution of primes in arithmetic progressions with fixed residue and varying moduli. The level of distribution that is needed here goes beyond the square root barrier (that is studied in the Bombieri–Vinogradov theorem, for example) and this is due to the condition $p \geqslant 2h - 1$ in Conjecture 6; see Remark 3 for more details. As far as we know, this is the first time that this kind of level of distribution is used in the subject of cyclotomic coefficients (see § 4 for the details). We would like to point out though that Fouvry [**14**] has used the classical Bombieri–Vinogradov theorem in a rather different way and context, namely, for studying the number of nonzero coefficients of cyclotomic polynomials Φ_n with n having two distinct prime factors.

In the final section, we consider cyclotomic polynomials with prescribed maximum or minimum coefficient. We will prove the following result.

THEOREM 5. *Andrica's conjecture implies that every natural number occurs as the maximum coefficient of some cyclotomic polynomial.*

That prime numbers play such an important role in our approach is a consequence of working with ternary cyclotomic polynomials. One would want to work with Φ_n with n having at least four prime factors, however, this leads to a loss of control over the behavior of the coefficients in general and the maximum in particular.

§2. *More on ternary cyclotomic polynomials.* Given any $m \geqslant 1$, Moree and Roşu [**31**] constructed infinite families of ternary optimal Φ_{pqr} such that $A(pqr) = (p+1)/2 + m$, provided that p is large enough in terms of m. This result, Theorem 6, allows one to show that for $p \geqslant 11$ there are cyclotomic polynomials having heights $(p+1)/2 + 1, \ldots, (p+1)/2 + k$, with k an integer close to $\sqrt{p}/2$. If the gaps between consecutive primes are always small enough, these heights cover all integers large enough and this would allow one to prove Conjecture 2. If large prime gaps do occur, then we are led to study the total length of prime gaps large enough up to x (cf. $E(x)$ in Lemma 5). Conveniently for us a good upper bound for this was recently obtained by Heath–Brown [**21**]; see Lemma 6.

The remainder of this section is devoted to deriving consequences of Theorem 6 and proving Theorem 2.

THEOREM 6 [**31**, Theorem 1.1]. *Let $p \geqslant 4m^2 + 2m + 3$ be a prime, with $m \geqslant 1$ any integer. Then there exists an infinite sequence of prime pairs $\{(q_j, r_j)\}_{j=1}^{\infty}$ with $q_j < q_{j+1}$, $pq_j < r_j$, such that*

$$A\{pq_jr_j\} = \left\{-\frac{(p-1)}{2} + m, \ldots, \frac{p+1}{2} + m\right\}.$$

We note that the two families of Theorem 1 are also infinite in the sense of this theorem. Thus Theorem 6 also holds for $m = 0$.

Put

$$\mathcal{R} = \left\{\frac{p+1}{2} + m : p \text{ is a prime}, \ m \geqslant 0, \ 4m^2 + 2m + 3 \leqslant p\right\}. \tag{3}$$

LEMMA 1. *We have $\mathcal{R} \subseteq \mathcal{A}_{opt}$.*

Proof. For the elements of $\mathcal{R}$ with $m = 0$, this follows from Theorem 1, for those with $m \geqslant 1$ it follows from Theorem 6. □

LEMMA 2. *If $p_{n+1} - p_n < \sqrt{p_n} + 1$ holds for $127 \leqslant p_n < 2h$ with h an integer, then we have $\mathbb{N}_h \backslash \{1, 5, 63\} \subseteq \mathcal{R}$.*

The proof is a consequence of part (a) of the following lemma and the computational observation that 1,5 and 63 are the only natural numbers < 64 that are not in $\mathcal{R}$.

By $\lfloor r \rfloor$, we denote the entire part of a real number r.

LEMMA 3. *Let $n \geqslant 5$.*

(a) *If $p_{n+1} - p_n < \sqrt{p_n} + 1$, then $I_n \cap \mathbb{N} \subset \mathcal{R}$, where $I_n := [\frac{p_n+1}{2}, \frac{p_{n+1}-1}{2}]$.*

(b) *If $p_{n+1} - p_n \geqslant \sqrt{p_n} + 1$, then there are at most*

$$\lfloor (p_{n+1} - p_n - \sqrt{p_n} + 1)/2 \rfloor \tag{4}$$

integers in the interval I_n that are not in $\mathcal{R}$.

Proof. The assumption on n implies $p_n \geqslant 11$. Put $z_n = (\sqrt{p_n} - 1)/2$. Note that $4z_n^2 + 2z_n + 3 = p_n - \sqrt{p_n} + 3 < p_n$. As $4x^2 + 2x + 3$ is increasing for $x \geqslant 0$, the inequality $4x^2 + 2x + 3 < p_n$ is satisfied for every real number $0 \leqslant x \leqslant z_n$. In particular it is satisfied for

$x = m_n$, with m_n the unique integer in the interval $[z_n - 1, z_n]$. Thus $m_n \geqslant (\sqrt{p_n} - 3)/2$ and $4m_n^2 + 2m_n + 3 \leqslant p_n$. It follows that

$$\left[\frac{p_n+1}{2}, \frac{p_n+1}{2} + m_n\right] \cap \mathbb{N} \subseteq \mathcal{R}.$$

As $(p_{n+1}+1)/2$ is clearly in $\mathcal{R}$, part (a) follows if we can show that the final number $(p_n + 1)/2 + m_n$ is at least $(p_{n+1} - 1)/2$. Since both numbers are integers, we can express this as $(p_n + 1)/2 + m_n > (p_{n+1} - 3)/2$. The validity of this inequality is obvious, since

$$\frac{p_n+1}{2} + m_n \geqslant \frac{p_n+1}{2} + \frac{\sqrt{p_n}-3}{2} > \frac{p_{n+1}-3}{2},$$

where the second inequality is a consequence of our assumption $d_n < \sqrt{p_n} + 1$.

Part (b) follows on noting that the number of integers of $\mathcal{R}$ that are not in I_n is bounded above by $d_n/2 - 1 - m_n$, which we see is bounded above by the integer in (4) on using $m_n \geqslant (\sqrt{p_n} - 3)/2$. □

Since we believe that (2) holds for all $p_n \geqslant 127$, Lemma 2 leads us to make the following conjecture.

CONJECTURE 7. We have $\mathcal{R} = \mathbb{N}\backslash\{1, 5, 63\}$.

The numbers 1,5 and 63 are special in our story.

LEMMA 4. *The integers* 1 *and* 5 *are in* $\mathcal{A}_t \subseteq \mathcal{A}$, *but not in* $\mathcal{A}_{opt}$. *The integer* 63 *is in* $\mathcal{A}_{opt} \subset \mathcal{A}_t \subseteq \mathcal{A}$, *but not in* $\mathcal{R}$.

Proof. If pqr is optimal, then $A(pqr) \geqslant (p+1)/2 \geqslant 2$ and so $1 \notin \mathcal{A}_{opt}$. It is also easy to see that there is no optimal pqr such that $A(pqr) = 5$. If such an optimal pqr would exist, then as $A(pqr) \leqslant 3$ for $p \leqslant 5$ and $A(pqr) \geqslant 6$ for $p \geqslant 11$ (for an optimal pqr), this would force $p = 7$ and $A\{7qr\} = [-5, 2] \cap \mathbb{Z}$ or $A\{7qr\} = [-2, 5] \cap \mathbb{Z}$, contradicting the result of Zhao and Zhang [38] that $A\{7qr\} \subseteq [-4, 4] \cap \mathbb{Z}$.

The number 63 is in $\mathcal{A}_{opt}$. This follows on applying Theorem 3.1 of [31]. The obvious approach is to consider the largest prime p such that $(p+1)/2 < 63$, which is $p = 113$, and take $l = 11$ (here and below we use the notation of Theorem 3.1). For this combination, the result does not apply, unfortunately. However, it does for $p = 109$ and $l = 15$, in which case we obtain $A\{109 \cdot 6803 \cdot 12084113\} = [-46, \ldots, 63] \cap \mathbb{Z}$ (with $q = 6803$, $\rho = 2870$, $\sigma = 62$, $s = 46$, $\tau = 18$, $w = 45$, $r_1 = 12\,084\,113$). □

Proof of Theorem 2. This follows on combining Lemmas 1, 2 and 4. □

§3. *Proof of Theorem* 3. The goal of this section is to prove Theorem 3. The quantity of central interest, $N(x)$, is defined below.

Definition 3. The number of integers $\leqslant x$ that does not occur as a height of an optimal ternary cyclotomic polynomial is denoted by $N(x)$.

LEMMA 5. *We have* $N(x) \leqslant E(2x)/2 + O(1)$, *where*

$$E(x) = \sum_{\substack{p_n \leqslant x \\ p_{n+1}-p_n \geqslant \sqrt{p_n}+1}} (p_{n+1} - p_n - \sqrt{p_n} + 1).$$

Proof. By Lemma 1, it suffices to bound above the number of integers $\leqslant x$ that are not in $\mathcal{R}$. By Lemma 3(b) this cardinality, in turn, is bounded above by $E(2x)/2 + O(1)$. □

If Cramér's conjecture, $d_n = O((\log p_n)^2)$ holds true, then this lemma implies $N(x) = O(1)$.

Heath–Brown [**21**] recently proved the following result which gives an upper bound for $E(x)$.

LEMMA 6 (Heath–Brown). *We have*

$$\sum_{\substack{p_n \leqslant x \\ p_{n+1}-p_n \geqslant \sqrt{p_n}}} (p_{n+1} - p_n) \ll_\epsilon x^{3/5+\epsilon}.$$

Proof of the conditional bound of Theorem 3. This follows on combining the latter upper bound for $E(x)$ with Lemma 5. □

In order to complete the proof of Theorem 3, we need to improve the exponent $3/5$ in Lemma 6 to $1/2$, conditionally on the Lindelöf Hypothesis. The Lindelöf Hypothesis states that for all fixed $\epsilon > 0$, we have

$$\zeta(1/2 + it) = O_\epsilon(t^\epsilon),\ t \in \mathbb{R},\ t > 1,$$

where as usual ζ denotes the Riemann zeta function. It is well known that the Riemann Hypothesis implies the Lindelöf Hypothesis, but not vice versa. There is a large body of work concerning the Lindelöf Hypothesis (see, e.g., the recent work of Bourgain [**7**]), however, it is still open.

We will make use of the following result of Yu [**37**].

LEMMA 7 (Yu). *Fix any* $\epsilon > 0$. *Under the Lindelöf Hypothesis, we have*

$$\sum_{p_n \leqslant x} (p_{n+1} - p_n)^2 \ll_\epsilon x^{1+\epsilon}.$$

From it one can easily derive a conditional improvement of Lemma 6.

LEMMA 8. *Assume the Lindelöf Hypothesis and fix any* $\epsilon > 0$. *Then we have*

$$\sum_{\substack{p_n \leqslant x \\ p_{n+1}-p_n \geqslant \sqrt{p_n}}} (p_{n+1} - p_n) \ll_\epsilon x^{1/2+\epsilon}.$$

Proof. Using dyadic division of the interval $[1, x]$, we obtain

$$\sum_{\substack{p_n \leqslant x \\ d_n \geqslant \sqrt{p_n}}} d_n \ll (\log x) \max_{1 \leqslant y \leqslant x} \sum_{\substack{y < p_n \leqslant 2y \\ d_n \geqslant \sqrt{p_n}}} d_n \leqslant (\log x) \max_{1 \leqslant y \leqslant x} \sum_{\substack{y < p_n \leqslant 2y \\ d_n \geqslant \sqrt{p_n}}} \frac{d_n^2}{\sqrt{p_n}},$$

which by Lemma 7 is at most

$$(\log x) \max_{1\leqslant y\leqslant x} \frac{1}{\sqrt{y}} \sum_{y<p_n\leqslant 2y} d_n^2 \ll_\epsilon x^{1/2+\epsilon}. \qquad \square$$

Proof of the conditional bound of Theorem 3. This follows on combining Lemma 5 with Lemma 8. □

Remark 1. Although it is not required for the applications in the present paper, one can prove slightly stronger variants of Lemmas 6 and 8, namely, that for every fixed $C > 0$ and $\epsilon > 0$, we have

$$\sum_{\substack{p_n\leqslant x \\ p_{n+1}-p_n\geqslant C\sqrt{p_n}}} (p_{n+1} - p_n) \ll_{C,\epsilon} x^{\alpha+\epsilon},$$

with $\alpha = 3/5$ (unconditionally) and $\alpha = 1/2$ under the Lindelöf Hypothesis (for details, see [25]).

§4. *A special case of the Bateman–Horn conjecture on average.* The goal of this section is to prove Theorem 4. Although the unconditional statement in Theorem 4 is surpassed by the unconditional statement in Theorem 3, the proof of Theorem 4 is, in a way, "orthogonal" to the one of Theorem 3; it thus has the potential of working in variations of the problem where the method behind Theorem 3 would fail. Interestingly, like our prime gap criterion, it rests on a variation (implicit in Lemma 9) of a certain very well-studied problem involving prime numbers. Both prime number questions are, however, quite different. Lemma 9 allows one to show that many odd heights occur among the ternary cyclotomic polynomials in a way different from Theorem 6.

LEMMA 9. *Let $h > 1$ be odd. If there exists a prime $p \geqslant 2h - 1$, such that the integer $q := 1 + (h - 1)p$ is a prime too, then $A(pqr) = h$ for some prime $r > q$. For r, one can take any prime $r_1 > q$ satisfying $r_1(p + q)/2 \equiv 1 \pmod{pq}$.*

Proof. Define

$$M(p; q) = \max_{r>q}\{A(pqr) : 2 < p < q < r\}. \tag{5}$$

Gallot et al. [17, Theorem 43] showed that if $q \equiv 1 \pmod p$, then

$$M(p; q) = \min\left\{\frac{q-1}{p} + 1, \frac{p+1}{2}\right\}.$$

The conditions on p and h ensure that $M(p; q) = h$. By [17, Lemma 24], it follows that $A(pqr_1) \geqslant h$. This in combination with $M(p; q) = h$ shows that $A(pqr_1) = h$. □

In case $p \geqslant 2h + 1$, the ternary cyclotomic polynomials from Lemma 9 are not optimal. We demonstrate this in the case $h = 63$ (with $p = 131$ and $q = 8123$).

Example 1. Using the latter result and [17, Lemma 24], we find that

$$A(131 \cdot 8123 \cdot 25497973) = \frac{8123 - 1}{131} + 1 = 63$$

and $a_{131\cdot 8123\cdot 25497973}(13459462019674) = -63$.

We define the set $G \subset \mathbb{N}$ as follows:

$G := \{m \in \mathbb{N} : \exists\, p \in (4m, 32m)$ such that $1 + 2mp$ is prime$\}$.

We would like to point out that the requirement $p < 32m$ is not necessary for the proof of Theorem 4, but is needed in the proof of Theorem 9.

In the remaining part of this section, we show that the density of G among all integers is positive, that is, that there exists $c_0 > 0$ such that

$$\liminf_{M\to+\infty} \frac{\#\{m \in G \cap [1, M]\}}{M} \geqslant c_0. \tag{6}$$

For any natural number m and any real number x, we define

$\pi_m(x) := \#\{p \in [\frac{x}{2}, x) : 1 + 2mp$ is prime$\}$.

Further, for any $x \geqslant 0$, we define

$$G(x) := \{m \in \mathbb{N} \cap [1, x/4) : \exists p \in (4m, x] \text{ such that } 1 + 2mp \text{ is prime}\}.$$

LEMMA 10. *For all $x, M \in \mathbb{R}$ with $x > 8M$ and $M \geqslant 1$, we have*

$$\#\{m \in G(x) \cap (M/4, M]\} \sum_{1\leqslant m\leqslant M} \pi_m(x)^2 \geqslant \left(\sum_{M/4<m\leqslant M} \pi_m(x)\right)^2.$$

Proof. Put

$$u_m(x) = \begin{cases} 1, & \text{if } \pi_m(x) \neq 0; \\ 0, & \text{otherwise.} \end{cases}$$

Fix $x > 8M$. By Cauchy's inequality, we have

$$\sum_{M/4<m\leqslant M} \pi_m(x) = \sum_{M/4<m\leqslant M} \pi_m(x) u_m(x)$$

$$\leqslant \#\{M/4 < m \leqslant M : \pi_m(x) > 0\}^{1/2} \left(\sum_{1\leqslant m\leqslant M} \pi_m(x)^2\right)^{1/2}.$$

If $m \leqslant M$ and $p \geqslant x/2$, then $4m \leqslant 4M < x/2 \leqslant p$, hence

$$\#M/4 < m \leqslant M : \pi_m(x) > 0\} \leqslant \#\{m \in G(x) \cap (M/4, M]\},$$

concluding the proof. □

We would like to estimate the sums $\sum_{M/4<m\leqslant M} \pi_m(x)$ and $\sum_{1\leqslant m\leqslant M} \pi_m(x)^2$ appearing above. An upper bound, say A, for $\sum_{1\leqslant m\leqslant M} \pi_m(x)^2$ is easily obtained by using standard sieve results. Now if we could derive a lower bound B for $\sum_{M/4<m\leqslant M} \pi_m(x)$, then by Lemma 10 we get

$$\#\{m \in G(x) \cap (M/4, M]\} \geqslant \frac{B^2}{A}.$$

Unfortunately the condition $x > 8M$ makes it difficult to obtain a good lower bound for $\sum_{M/4<m\leqslant M} \pi_m(x)$. We overcome this by using deep work of Bombieri *et al.* regarding the level of distribution of primes in arithmetic progressions with fixed residue and varying moduli.

We start with estimating $\sum_{1\leqslant m\leqslant M} \pi_m(x)^2$, for which we need the following lemma, which is obtained on putting $b = k = l = 1$ in [19, Theorem 3.12].

LEMMA 11. *Let a be a positive even integer. Then for all $x > 1$, we have, uniformly in a, that*

$$\#\{p \leqslant x : ap+1 \text{ is prime}\} \leqslant \frac{8C_2 x}{(\log x)^2} \prod_{\substack{p|a \\ p>2}} \left(\frac{p-1}{p-2}\right) \left\{1 + O\left(\frac{\log\log x}{\log x}\right)\right\},$$

where

$$C_2 = \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right)$$

is the twin prime constant.

Remark 2. Hardy and Littlewood conjectured, based on heuristic reasoning, that asymptotically

$$\#\{p \leqslant x : p+2 \text{ is prime}\} \sim 2C_2 \frac{x}{(\log x)^2}.$$

A similar heuristic reasoning leads to the conjecture that asymptotically

$$\#\{p \leqslant x : ap+1 \text{ is prime}\} \sim C_2 \prod_{\substack{p|a \\ p>2}} \left(\frac{p-1}{p-2}\right) \frac{x}{(\log x)^2}.$$

Both conjectures are special cases of the Bateman–Horn conjecture, cf. [1].

LEMMA 12. *Let x, M be any two positive real numbers. Then*

$$\sum_{1 \leqslant m \leqslant M} \pi_m(x)^2 \leqslant 64 C_1 C_2^2 M \frac{x^2}{(\log x)^4} \left\{1 + O\left(\frac{\log\log x}{\log x} + \frac{1}{\sqrt{M}}\right)\right\},$$

where

$$C_1 := \prod_{p>2} \left(1 + \frac{2}{p(p-2)} + \frac{1}{p(p-2)^2}\right),$$

C_2 is the twin prime constant, and the implied constant is absolute.

Proof. By Lemma 11 with $a = 2m$, we get

$$\pi_m(x)^2 \leqslant 8^2 C_2^2 \frac{x^2}{(\log x)^4} \prod_{\substack{p|2m \\ p>2}} \left(\frac{p-1}{p-2}\right)^2 \left\{1 + O\left(\frac{\log\log x}{\log x}\right)\right\},$$

therefore, we conclude that $\sum_{1 \leqslant m \leqslant M} \pi_m(x)^2$ is at most

$$8^2 C_2^2 \frac{x^2}{(\log x)^4} \left\{1 + O\left(\frac{\log\log x}{\log x}\right)\right\} \sum_{1 \leqslant m \leqslant M} \prod_{\substack{p|m \\ p>2}} \left(\frac{p-1}{p-2}\right)^2.$$

We define the multiplicative function f via

$$f(p^e) := \chi_{p>2}(p) \chi_{e=1}(e) \left(\frac{2}{p-2} + \frac{1}{(p-2)^2}\right), \quad (e \in \mathbb{N},\ p \text{ prime}).$$

with χ the characteristic function. One can easily verify that

$$\prod_{\substack{p|k\\p>2}}\left(\frac{p-1}{p-2}\right)^2=\sum_{d|k}f(d)=\sum_{\substack{d|k\\2\nmid d}}f(d)$$

for all non-zero integers k. This shows that

$$\begin{aligned}\sum_{1\leqslant m\leqslant M}\prod_{\substack{p|2m\\p>2}}\left(\frac{p-1}{p-2}\right)^2&=\sum_{\substack{1\leqslant d\leqslant M\\2\nmid d}}f(d)\sum_{\substack{1\leqslant m\leqslant M\\d|2m}}1\\&=M\sum_{1\leqslant d\leqslant M}\frac{f(d)}{d}+O\left(\sum_{1\leqslant d\leqslant M}f(d)\right),\end{aligned}$$

where we used several times that $f(d)=0$ if d is even. Noting that $f(p)\leqslant C/p$ for some absolute constant $C>0$ yields the bound

$$f(d)\leqslant\mu(d)^2\frac{C^{\omega(d)}}{d}\ll\frac{1}{\sqrt{d}},\quad(d\in\mathbb{N}),$$

which can be used to obtain

$$\sum_{1\leqslant d\leqslant M}\frac{f(d)}{d}=\sum_{d=1}^{\infty}\frac{f(d)}{d}+O\left(\sum_{d>M}\frac{1}{d^{3/2}}\right)=C_1+O\left(\frac{1}{\sqrt{M}}\right)$$

and

$$\sum_{1\leqslant d\leqslant M}f(d)\ll\sum_{1\leqslant d\leqslant M}\frac{1}{\sqrt{d}}\ll\sqrt{M}.$$

Putting everything together it follows that

$$\sum_{1\leqslant m\leqslant M}\prod_{\substack{p|2m\\p>2}}\left(\frac{p-1}{p-2}\right)^2=C_1M+O(\sqrt{M}),$$

which is sufficient for our purposes. □

We now proceed to evaluate the sum $\sum_{1\leqslant m\leqslant M}\pi_m(x)$ appearing in Lemma 10. Writing $n=1+2mp$, we see that it equals

$$\begin{aligned}\sum_{M/4<m\leqslant M}\sum_{\substack{x/2\leqslant p<x\\1+2mp\text{ prime}}}1&=\sum_{x/2\leqslant p<x}\sum_{\substack{M/4<m\leqslant M\\1+2mp\text{ prime}}}1\\&=\sum_{x/2\leqslant p<x}\#\{n\text{ prime}:1+Mp/2<n\leqslant1+2Mp,\ n\equiv1(\text{mod }p)\}\\&\geqslant\frac{1}{\log(1+2Mx)}\sum_{x/2\leqslant p<x}\sum_{\substack{n\text{ prime}\\1+Mp/2<n\leqslant1+2Mp\\n\equiv1(\text{mod }p)}}\log n,\end{aligned}\tag{7}$$

where we used that $\log n\leqslant\log(1+2Mp)\leqslant\log(1+2Mx)$.

Remark 3. One now recognizes the argument in the latter sum as a counting function of primes in an arithmetic progression of varying modulus p, as p runs through $[x/2, x)$. We would now use the Bombieri–Vinogradov theorem; however, the size of the primes n is of the order of magnitude

$$1 + 2Mp \approx 2Mx,$$

since the moduli p have typical size x. Thus, owing to the condition $x > 8M$, we are counting primes in a progression whose modulus exceeds the square-root of the size of the primes. Therefore, the Bombieri–Vinogradov theorem cannot be applied in our case. To be more precise, it can only be applied when the moduli are bounded by $\sqrt{z}/(\log z)^A$, where $A > 0$ and z is the length of the interval $(0, z]$ we are counting primes in. This means that we need

$$p \leqslant \frac{\sqrt{2Mp}}{(\log(2Mp))^A},$$

for some fixed $A > 0$, and this can only happen when $x = o(M)$. To deal with this problem, we shall need a special case (Lemma 13), of the work of Bombieri, Friedlander and Iwaniec [6].

As usual, let

$$\theta(x; q, a) := \sum_{\substack{p \leqslant x \\ p \equiv a (\mathrm{mod}\ q)}} \log p, \quad \psi(x; q, a) := \sum_{\substack{n \leqslant x \\ n \equiv a (\mathrm{mod}\ q)}} \Lambda(n),$$

with Λ the von Mangoldt function.

LEMMA 13 (Bombieri–Friedlander–Iwaniec [6]). *For any $t \geqslant y \geqslant 3$, we have*

$$\sum_{\sqrt{ty}/2 \leqslant q < \sqrt{ty}} \left| \psi(t; q, 1) - \frac{t}{\phi(q)} \right| \ll t \left(\frac{\log y}{\log t} \right)^2 (\log \log t)^B,$$

where B is an absolute constant and the implied constant is absolute.

This estimate is obtained on setting $a = 1$, $x = t$ and $Q = \sqrt{xy}$ in [6, Main Theorem, p. 363].

LEMMA 14. *For any $t \geqslant y \geqslant 3$ with $y \leqslant t^{1/20}$, we have*

$$\sum_{\substack{q \text{ prime} \\ \sqrt{ty}/2 \leqslant q < \sqrt{ty}}} \left| \theta(t; q, 1) - \frac{t}{\phi(q)} \right| \ll t \left(\frac{\log y}{\log t} \right)^2 (\log \log t)^B,$$

where B is an absolute constant and the implied constant is absolute.

Proof. Clearly

$$\psi(t; q, 1) = \theta(t; q, 1) + \sum_{k=2}^{\infty} \sum_{\substack{p \leqslant t^{1/k} \\ p^k \equiv 1 (\mathrm{mod}\ q)}} \log p.$$

The inner sum vanishes if $t^{1/k} < 2$, therefore only the integers $k \leqslant (\log t)/\log 2$ need to be taken into account. The contribution of all such integers with $k \geqslant 3$ is $\ll t^{1/3} \log t$, since the sum over p is $\ll t^{1/k}$ by the prime number theorem. The steps so far are the standard arguments that one performs when moving from asymptotics for ψ to asymptotics for θ; however, in our case, owing to the level of distribution being comparable to the square root of the length of the interval, the term $k = 2$ cannot be controlled with the classical arguments. Instead, we use the bound

$$\frac{1}{\log t} \sum_{\substack{p \leqslant \sqrt{t} \\ p^2 \equiv 1 \pmod q}} \log p \leqslant \sum_{\substack{m \leqslant \sqrt{t} \\ m^2 \equiv 1 \pmod q}} 1 = \sum_{\substack{m \leqslant \sqrt{t} \\ m \equiv -1 \pmod q}} 1 + \sum_{\substack{m \leqslant \sqrt{t} \\ m \equiv 1 \pmod q}} 1,$$

where we used the fact that q is prime. Each of the sums in the right side is trivially $\ll \sqrt{t}/q + 1$ and therefore

$$\sum_{\substack{p \leqslant \sqrt{t} \\ p^2 \equiv 1 \pmod q}} \log p \ll (\log t)\left(\frac{\sqrt{t}}{q} + 1\right).$$

We thus find that

$$\psi(t; q, 1) = \theta(t; q, 1) + O\left(t^{1/3}(\log t) + \frac{\sqrt{t}}{q} \log t\right).$$

This shows that the sum over q in the statement of this lemma is

$$\ll \sum_{\sqrt{ty}/2 \leqslant q < \sqrt{ty}} \left| \psi(t; q, 1) - \frac{t}{\phi(q)} \right| + \sum_{\sqrt{ty}/2 \leqslant q < \sqrt{ty}} \left(t^{1/3}(\log t) + \frac{\sqrt{t}}{q} \log t \right).$$

The first sum can be bounded by Lemma 13. Noting that $\sum_{x/2 < q \leqslant x} 1/q = O(1)$, cf. (8), we see that the second sum is

$$\ll \sqrt{ty}\, t^{1/3}(\log t) + \sqrt{t} \log t,$$

which is $\ll t^{19/20} \ll t(\log t)^{-2}$, as $y \leqslant t^{1/20}$. □

LEMMA 15. *Let* $\psi : (1, \infty) \to (4, \infty)$ *be any function satisfying* $\psi(M) \leqslant \log M$. *For any* $M > 1$, *we let* $x = M\psi(M)$ *and have*

$$\sum_{M/4 < m \leqslant M} \pi_m(x) \geqslant \frac{Mx}{2\log(Mx)} \frac{\log 2}{\log x} \left\{ 1 + O\left(\frac{(\log \log x)^{B+2}}{\log x} \right) \right\},$$

where B *is the absolute constant from Lemma* 14.

Proof. The condition $p \in [x/2, x)$ in the definition of $\pi_m(x)$ ensures that the interval $(1 + Mx/2, 1 + Mx]$ is contained in the interval $(1 + Mp/2, 1 + 2Mp]$. Therefore, by (7) we see that the sum in our lemma is at least

$$\frac{1}{\log(1 + 2Mx)} \sum_{x/2 \leqslant p < x} \sum_{\substack{n \text{ prime} \\ 1 + Mx/2 < n \leqslant 1 + Mx \\ n \equiv 1 \pmod p}} \log n.$$

Using Lemma 14 with $t = Mx$ and $y = \psi(M)$ shows that this is

$$\frac{(1 + Mx) - (1 + M\frac{x}{2})}{\log(1 + 2Mx)} \sum_{x/2 \leqslant p < x} \frac{1}{p - 1} + O\left(\frac{Mx}{\log(Mx)} \left(\frac{\log \psi(M)}{\log x} \right)^2 (\log \log x)^B \right).$$

Using the standard estimate

$$\sum_{p \leqslant x} \frac{1}{p-1} = \log\log x + C' + O\left(\frac{1}{(\log x)^2}\right),$$

we obtain

$$\sum_{x/2 < p \leqslant x} \frac{1}{p-1} = \frac{\log 2}{\log x}\left\{1 + O\left(\frac{1}{\log x}\right)\right\}. \tag{8}$$

It follows that the main term is as claimed in our lemma. Furthermore, on using the bound $\log \psi(M) \ll \log\log M \ll \log\log x$, we see that the error term is

$$\ll \frac{Mx}{\log(Mx)} \frac{(\log\log x)^{B+2}}{(\log x)^2},$$

as required. □

Proof of Theorem 4. The first assertion is a corollary of Lemma 9.

The inequalities obtained in Lemmas 12 and 15 with $\psi(M) = 9$ in combination with the inequality in Lemma 10 give rise, on choosing $x = 1 + 8M$, to the inequality

$$\#\{m \in G(1+8M) \cap (M/4, M]\}\, 64 C_1 C_2^2 M \frac{x^2}{(\log x)^4} \geqslant \left(\frac{Mx}{\log(Mx)} \frac{\log 2}{2\log x}\right)^2 (1 + o(1)).$$

In particular, the estimate $\log(Mx) \leqslant 2\log x$ yields

$$\#\{m \in G(1+8M) \cap (M/4, M]\} \geqslant c' M(1 + o(1)),$$

where

$$c' = \frac{(\log 2)^2}{1024 C_1 C_2^2} > 0.$$

Suppose $m \in G(1+8M) \cap (M/4, M]$. Note that since $M/4 < m$, we have

$$p \leqslant x = 1 + 8M < 1 + 32m,$$

and hence $p < 32m$, therefore, the set $G(1+8M)$ is contained in G. We conclude that (6) holds with $c_0 = c'$. It follows that a positive proportion of all integers m have the property that there exists a prime $p > 4m$ with also $1 + 2mp$ being a prime. By Lemma 9, we have $1 + 2m \in \mathcal{A}_t$ for each of those m, and it thus follows that unconditionally $\mathcal{A}_t$ contains a positive fraction of all odd natural numbers. □

Remark 4. The proof actually yields that a positive proportion of all integers m have the property that there exists a prime $4m < p < 32m$ with also $1 + 2mp$ being a prime. This is what we will use in the proof of Theorem 7.

§5. *Some related issues.*

5.1. *Estimating the smallest n for which $A(n) = h$.* *Definition* 4. Given a natural number h, let n_h be the smallest ternary integer, if it exists, such that $A(n_h) = h$.

The entries in the column $k/\varphi(pqr)$ in Table 1 suggest the following question.

Table 1: Ternary examples with prescribed height

Height	p	q	r	k	sign	diff.	$\frac{k}{\phi(pqr)}$	$\frac{\log(pqr)}{\log h}$
1	3	7	11	0	+	2	0	
2	3	5	7	7	−	**3**	0.146	6.714
3	5	7	11	119	−	**5**	0.496	5.418
4	11	13	17	677	−	7	0.353	5.623
5	11	13	19	1008	−	9	0.467	4.913
6	13	23	29	2499	−	10	0.338	5.060
7	17	19	53	6013	+	14	0.402	5.009
8	17	31	37	5596	−	14	0.324	4.750
9	17	47	53	14 538	−	**17**	0.379	4.848
10	17	29	41	4801	−	**17**	0.267	4.305
11	23	37	61	20 375	−	16	0.428	4.527
12	23	37	41	14 471	+	21	0.456	4.209
13	31	59	73	58 333	−	25	0.465	4.601
14	37	53	61	52 286	+	27	0.465	4.430
15	37	47	61	45 939	−	29	0.462	4.273
16	41	79	97	133 844	−	30	0.446	4.565
17	41	43	53	38 240	+	33	0.437	4.039
18	61	97	103	178 013	−	34	0.302	4.608
19	43	83	89	101 051	−	33	0.333	4.302
20	47	83	131	235 842	+	37	0.481	4.387
21	47	101	109	217 278	−	41	0.437	4.321
22	53	83	89	165 453	−	44	0.441	4.166
23	43	71	109	108 355	+	**43**	0.341	4.055
24	53	103	109	189 160	−	42	0.330	4.183
25	61	79	97	224 640	−	47	0.500	4.055
26	41	71	97	96 529	−	**41**	0.359	3.852
27	61	109	113	332 589	−	54	0.458	4.105
28	53	89	131	186 685	−	**53**	0.314	4.001
29	83	109	139	552 035	−	58	0.452	4.170
30	67	131	137	389 139	−	52	0.333	4.116
31	83	107	113	444 435	+	61	0.456	4.024
32	79	149	163	881 529	+	63	0.471	4.174
33	73	103	113	389 314	+	61	0.473	3.904
34	71	109	113	409 320	−	60	0.483	3.879
35	83	103	139	544 198	−	69	0.471	3.934
36	127	149	151	1 246 462	−	72	0.445	4.148
37	71	101	239	671 716	+	67	0.403	3.975
38	127	137	409	3 355 658	−	75	0.479	4.337
39	83	149	157	941 094	+	76	0.497	3.952
40	79	233	239	1 624 556	+	**79**	0.377	4.146

QUESTION 1. Let $h > 1$ be an integer. Does there exist an absolute constant $0 < c \leqslant 1/2$ such that if $|a_n(k)| = h$, then $k > c\,\varphi(n_h)$?

A further question is to relate the size of n_h to h. See the final column of Table 1 for some numerical data. The 19th century estimate $A(pqr) \leqslant p - 1$ implies $n_h \gg h^3$.

CONJECTURE 8. There are constants E_1 and E_2 such that $h^{E_1} \ll n_h \ll h^{E_2}$ and $E_1 \geqslant 3$.

Theorem 7 shows that for a positive fraction of integers h the upper bound in the conjecture holds true. Its formulation involves Linnik's constant L.

Definition 5. Let $r \geqslant 0$ be an arbitrary fixed real number. For coprime integers a and d, let $p_r(a, d)$ denote the smallest prime $> d^r$ in the progression $a \pmod{d}$.

Linnik proved in 1944 that there exist positive constants C and L such that $p_0(a, d) \leqslant C\, d^L$. The constant L is known as *Linnik's constant*. Xylouris [36] proved that $L \leqslant 5$, heavily relying on a fundamental paper by Heath–Brown [20], who obtained $L \leqslant 5.5$. On GRH, Lamzouri *et al.* [27] showed that $p(a, d) \leqslant (\varphi(d) \log d)^2$ for $d > 3$.

The following result generalizes Linnik's theorem.

LEMMA 16. *There exists some absolute constant C such that $p_r(a, d) \ll d^{r+C}$, where the implied constant is also absolute.*

Proof. We use Corollary 18.8 of the book of Iwaniec and Kowalski [22]. It states that there exists an explicit effectively computable constant $L_1 > 0$ such that for all sufficiently large d and all $x \geqslant d^{L_1}$, we have

$$\psi(x; d, a) \gg \frac{x}{\varphi(d)\sqrt{d}},$$

where the implied constant is absolute. Since $\varphi(d) \leqslant d$, this implies

$$\psi(x; d, a) \gg \frac{x}{d^{3/2}}.$$

For all $x > d^6$, we have $\sqrt{x} \leqslant x^{3/4} d^{-3/2}$ and hence,

$$\psi(x; d, a) - \theta(x; d, a) \leqslant \psi(x) - \theta(x) \ll \sqrt{x} \leqslant \frac{x^{3/4}}{d^{3/2}}.$$

Therefore, if $x > d^{L_1+6}$, we deduce that

$$\theta(x; d, a) \gg \frac{x}{d^{3/2}},$$

where the implied constant is absolute. To conclude our proof, we note that $p_r(a, d)$ is bounded by any real number $x > d^r$ which satisfies

$$\theta(x; d, a) > \theta(d^r; d, a).$$

Since $\theta(d^r; d, a) \leqslant \theta(d^r) < 2d^r$ by the prime number theorem, it suffices to find the least $x > d^r$ for which $\theta(x; d, a) \geqslant 2d^r$. Clearly, this holds as long as $x > d^{L_1+6}$ and $x\, d^{-3/2} > Cd^r$ for some large constant C. For both of these properties to hold, it is sufficient that $x \gg d^{r+6+L_1}$, from which we infer that

$$p_r(a, d) \ll d^{r+6+L_1},$$

with an absolute implied constant. □

The next result makes some progress toward Conjecture 8. It requires only Linnik's theorem for its proof. Under GRH, the estimate holds with $L = 2$.

THEOREM 7. *Let $\epsilon > 0$. Let n_h be the smallest ternary integer, if it exists, such that $A(n_h) = h$. There exists a constant $c_\epsilon > 0$ such that $n_h < c_\epsilon\, h^{3(L+1+\epsilon)}$ for a positive proportion of the odd natural numbers h.*

Proof. Let m be an integer such that there exists a prime $4m < p < 32m$ with also $q := 1 + 2mp$ being a prime. For any such m, we will show that $h := 1 + 2m \in \mathcal{A}_t$ and construct a ternary n such that $A(n) = h$ and n satisfies the required upper bound. Since, as we have seen in the proof of Theorem 4 (cf. Remark 4), there is a positive proportion of such m, the result follows.

We let $0 < r_1 < pq$ be the unique solution of $r_1(p+q)/2 \equiv 1 \pmod{pq}$. If r_1 is even, we put $r = p_0(r_1, pq)$. Note that $r > pq$. If r_1 is odd, we let s be the smallest prime not dividing $r_1 + pq$. Let $\delta > 0$ be arbitrary. Since the product of the primes not exceeding x is of size $e^{(1+o(1))x}$, we conclude that $s < (pq)^\delta$ for all m large enough. Observe that $r_1 + pq$ and spq are coprime. We put $r = p_0(r_1 + pq, spq)$. Note that $r > q$. By Linnik's theorem, we have $r \leqslant C(pq)^{(1+\delta)L}$. By Lemma 9, we have $A(pqr) = h$. Since $pqr > n_h$ and

$$pqr = O(m \cdot m^2 \cdot (m^{3(1+\delta)})^L) = O(h \cdot h^2 \cdot (h^{3(1+\delta)})^L) = O(h^{3(L+1+\epsilon)}),$$

with $\epsilon = \delta L$, the proof is completed. $\square$

The next result can be seen as a supplement to Theorem 3. The proof requires Lemma 16 and a more precise version of Theorem 6 that is too long to be formulated here.

THEOREM 8. *Let t_h be the smallest optimal ternary integer, if it exists, such that $A(t_h) = h$. There exist positive constants c and T such that $t_h < c\,h^T$ for all $h \leqslant x$ with at most $\ll_\epsilon x^{3/5+\epsilon}$ exceptions.*

Proof. We will use [31, Theorem 3.1], the full version of Theorem 6. As Theorem 6 is used in the proof of Theorem 3, we get the same number of possible exceptions $h \leqslant x$. In terms of the m of Theorem 6, we have $l = 2m - 1$, with $l \leqslant \sqrt{p}$. We take $h = (p + l + 2)/2$. The prime q indicated in the theorem is bounded above by $p_2(a, p)$, with a an appropriate residue class. The prime r has to exceed pq and be in an appropriate residue class modulo pq. By Lemma 16, we have $pq \leqslant pp_2(a, p) \ll p^{T_1}$ for some constant T_1. Thus by Lemma 16 again, r is $\ll p^{T_2}$ for some constant T_2. Thus $pqr \ll p^{T_1+T_2}$. The result then follows with $T = T_1 + T_2$ on noticing that $p = O(h)$. $\square$

5.2. *Prescribed maximum or minimum coefficient.* So far we focused on possible heights of cyclotomic polynomials. Instead one can ask for possible maxima and minima. In this section, we will argue why the following conjecture is reasonable.

CONJECTURE 9. Each non-zero integer occurs either as the maximum or as the minimum coefficient of some cyclotomic polynomial.

Definition 6. We denote the maximum and minimum coefficients of Φ_n by $A^+(n)$, respectively, $A^-(n)$. We put $\mathcal{A}_t^+ = \{A^+(n) : n \text{ is ternary}\}$ and define $\mathcal{A}_t^-$ analogously. We denote by $\mathcal{A}_{opt}^+$ the set of all $A^+(n)$, with n optimal and define $\mathcal{A}_{opt}^-$ analogously.

Remark 5. Using the elementary identity $\Phi_n(1) = e^{\Lambda(n)}$ (valid for $n > 1$), we infer that

$$A^-(n) = \begin{cases} 1 & \text{if } n = p^k \text{ for some prime } p \text{ and } k \geqslant 1; \\ < 0 & \text{otherwise.} \end{cases}$$

Since our arguments rest on properties of ternary cyclotomic polynomials, the next result due to Kaplan makes it plausible that asking which maximal coefficients can occur is in essence the same as asking which possible minimum coefficients can occur.

PROPOSITION 2 (Implicit in Kaplan [24], explicit in Bachman and Moree [4]). *If $r, s > pq$, then*

$$A\{pqr\} = \begin{cases} A\{pqs\} & \text{if } s \equiv r \pmod{pq}; \\ -A\{pqs\} & \text{if } s \equiv -r \pmod{pq}. \end{cases}$$

This proposition can be used to prove the following lemma (recall that $M(p;q)$ is defined in (5)).

LEMMA 17. *If $A(pqr) = M(p;q)$, then there exist primes r_1 and r_2 such that $A^+(pqr_1) = M(p;q)$ and $A^-(pqr_2) = -M(p;q)$.*

Proof. The integers in $[-M(p;q), M(p;q)] \cap \mathbb{Z}$ are precisely those that appear in Φ_{pqr} as r ranges over the primes exceeding q; see Gallot, Moree and Wilms [17, Proposition 1]. $\square$

In the proof of Theorem 4 exclusively heights are considered that are of the form $M(p;q)$. This observation together with Lemma 17 then leads to a proof of the following variant of Theorem 4.

THEOREM 9. *If Conjecture 6 holds true, then $\mathcal{A}_t^- \cup \mathcal{A}_t^+$ contains all odd integers. Unconditionally both $\mathcal{A}_t^-$ and $\mathcal{A}_t^+$ contain a positive fraction of all odd integers.*

In our proof of Theorem 2, we actually show that $\mathcal{R} \subseteq \mathcal{A}_{opt}^+$ (recall that $\mathcal{R}$ is defined in (3)). The optimal ternary cyclotomic polynomials Φ_{pqr} used come from Theorem 6 and satisfy $r > pq$. This allows one then to invoke Proposition 2 and conclude that $-\mathcal{R} \subseteq \mathcal{A}_{opt}^-$.

The following result is analogous to Theorem 3. The proof of that result (given in § 3) rests on bounding above the integers $\leqslant x$ that are not in $\mathcal{R}$. Likewise the proof of Theorem 10 rests on bounding above the integers in $[-x, x]$ that are not in $-\mathcal{R} \cup \mathcal{R}$.

THEOREM 10. *The set $\mathcal{A}_{opt}^- \cup \mathcal{A}_{opt}^+$ contains almost all integers. Specifically, for any fixed $\epsilon > 0$, the number of integers with absolute value $\leqslant x$ that do not occur in $\mathcal{A}_{opt}^- \cup \mathcal{A}_{opt}^+$ is $\ll_\epsilon x^{3/5+\epsilon}$. Under the Lindelöf Hypothesis, this number is $\ll_\epsilon x^{1/2+\epsilon}$.*

Finally, we will derive a variant of Theorem 2, namely Lemma 18.

We put

$$\mathcal{R}^{\pm} = \left\{ \frac{p-1}{2} - m : p \text{ is a prime},\ m \geqslant 0,\ 4m^2 + 2m + 3 \leqslant p \right\}$$
$$\cup \left\{ \frac{p-1}{2} + m : p \text{ is a prime},\ m \geqslant 0,\ 4m^2 + 2m + 3 \leqslant p \right\}.$$

We saw that $\mathcal{R} \subseteq \mathcal{A}_{opt}^+$ and $-\mathcal{R} \subseteq \mathcal{A}_{opt}^-$. However, more is true.

LEMMA 18. *We have $\mathcal{R}^{\pm} \subseteq \mathcal{A}_{opt}^+$ and $-\mathcal{R}^{\pm} \subseteq \mathcal{A}_{opt}^-$.*

Proof. For the elements of $\mathcal{R}^{\pm}$ with $m = 0$, this follows from Theorem 1, for those with $m \geqslant 1$ it follows from Theorem 6 in combination with Proposition 2. □

Taking $p = 3, 11, 127$ and $m = 0$ we see that $\{1, 5, 63\}$ are in $\mathcal{R}^{\pm}$. This in combination with Conjecture 7 and Lemma 18 leads to the following conjecture.

CONJECTURE 10. We have $\mathcal{R}^{\pm} = \mathbb{N}$, $\mathcal{A}^{+}_{opt} = \mathbb{N}$ and $\mathcal{A}^{-}_{opt} = -\mathbb{N}$.

5.3. *Connection with Andrica's conjecture.* The aim of this subsection is to prove Theorem 5. Our proof is a consequence of the following lemma that is analogous to Lemma 3.

LEMMA 19. *Let* $n \geqslant 5$ *and* $I_n := [\frac{p_n+1}{2}, \frac{p_{n+1}-1}{2}]$.
(a) *If* $p_{n+1} - p_n < \sqrt{p_n} + \sqrt{p_{n+1}}$, *then* $I_n \cap \mathbb{N} \subseteq \mathcal{R}^{\pm}$.
(b) *If* $p_{n+1} - p_n < \sqrt{p_n} + \sqrt{p_{n+1}}$ *holds for* $11 \leqslant p_n < 2h$ *with* h *an integer, then we have* $\mathbb{N}_h \subseteq \mathcal{R}^{\pm}$.

Proof. We let the integer m_n be as in the proof of Lemma 3 and recall that $m_n \geqslant (\sqrt{p_n} - 3)/2$. Part (a) follows if we can show that the final number $(p_n + 1)/2 + m_n$ is at least $(p_{n+1} - 1)/2 - m_{n+1} - 1$. Since both numbers are integers, it suffices to require that

$$\frac{p_n + 1}{2} + m_n > \frac{p_{n+1} - 1}{2} - m_{n+1} - 2.$$

This is equivalent with $d_n/2 < m_n + m_{n+1} + 3$. Now our assumption on d_n implies

$$d_n/2 < (\sqrt{p_n} - 3)/2 + (\sqrt{p_{n+1}} - 3)/2 + 3 \leqslant m_n + m_{n+1} + 3,$$

as wanted.
(b) This is a consequence of part (a) and the observation that 1,2,3,4 and 5 are in $\mathcal{R}^{\pm}$. □

Proof of Theorem 5. A consequence of Lemma 19 part (b) and the observation that we also have $p' - p < \sqrt{p} + \sqrt{p'}$ for $p \leqslant 11$. □

§6. *Ternary cyclotomic polynomials of small height.* Table 1 gives the minimum ternary integer $n = pqr$ with $p < q < r$ such that $A(n) = m$ for the numbers $m = 1, \ldots, 40$. The integer k has the property that $a_{pqr}(k) = \pm m$, with the sign coming from the sixth column. The seventh column records the difference between the largest and smallest coefficient and is in bold if this is optimal, that is, if the difference equals p (compare Definition 2). The second-to-last column gives the relative position of k in Φ_{pqr}. The final column gives, for $h > 1$, the exponent e such that $pqr = h^e$.

The heights h in Table 1 satisfy $h \leqslant 2p/3$ with equality only in case $h = 2$. This is consistent with the generalized Sister Beiter conjecture due to Gallot and Moree [16].

Acknowledgements. The authors thank Danilo Bazzanella, Adrian Dudek, Tomás Oliveira e Silva, Alberto Perelli and Tim Trudgian for helpful email correspondence. Olivier Ramaré kindly provided us with a high accuracy evaluation of C_1. We thank the referee for excellent remarks that helped to improve the exposition of this paper and questions that led to the addition of Section 5.

In case an integer h is not in $\mathcal{R}$ (defined in (3)), still the work of Moree and Roşu [31] offers some hope to show that h occurs as a height (as we saw in case $h = 63$). To make this more precise involves understanding the distribution of inverses modulo primes. We thank Cristian Cobeli for sharing some observations and numerical experiments on this.

The first author is a novice in number theory and is very grateful to Pieter Moree for introducing him to the field.

Open access funding enabled and organized by Projekt DEAL.

References

1. S. L. Aletheia-Zomlefer, L. Fukshansky and S. R. Garcia, The Bateman-Horn conjecture: heuristics, history and applications. *Expos. Math.* **38** (2020), 430–479.
2. G. Bachman, On the coefficients of ternary cyclotomic polynomials. *J. Number Theory* **100** (2003), 104–116.
3. G. Bachman, Ternary cyclotomic polynomials with an optimally large set of coefficients. *Proc. Amer. Math. Soc.* **132** (2004), 1943–1950.
4. G. Bachman and P. Moree, On a class of ternary inclusion-exclusion polynomials. *Integers* **11** (2011), A8.
5. W. Banks, K. Ford and T. Tao, Large prime gaps and probabilistic models. *Preprint*, arXiv:1908.08613, 2019.
6. E. Bombieri, J. B. Friedlander and H. Iwaniec, Primes in arithmetic progressions to large moduli. II. *Math. Ann.* **277** (1987), 361–393.
7. J. Bourgain, Decoupling, exponential sums and the Riemann zeta function. *J. Amer. Math. Soc.* **30** (2017), 205–224.
8. E. Carneiro, M. B. Milinovich and K. Soundararajan, Fourier optimization and prime gaps. *Comment. Math. Helv.* **94** (2019), 533–568.
9. H. Cramér, Some theorems concerning prime numbers. *Arkiv f. Math. Astr. Fys.* **15** (1920), 1–33. [Collected Works **1**, 85–91, Springer, Berlin-Heidelberg, 1994.]
10. H. Cramér, On the order of magnitude of the difference between consecutive primes. *Acta Arith.* **2** (1936), 23–46. [Collected Works **2**, 871–894, Springer, Berlin–Heidelberg, 1994.]
11. J. Fintzen, Cyclotomic polynomial coefficients $a(n, k)$ with n and k in prescribed residue classes. *J. Number Theory* **131** (2011), 1852–1863.
12. K. Ford, B. Green, S. Konyagin, J. Maynard and T. Tao, Large gaps between consecutive prime numbers. *Ann. of Math.* (2) **183** (2016), 935–974.
13. K. Ford, B. Green, S. Konyagin, J. Maynard and T. Tao, Long gaps between primes. *J. Amer. Math. Soc.* **31** (2018), 65–105.
14. É. Fouvry, On binary cyclotomic polynomials. *Algebra Number Theory* **7** (2013), 1207–1223.
15. Y. Gallot and P. Moree, Neighboring ternary cyclotomic coefficients differ by at most one. *J. Ramanujan Math. Soc.* **24** (2009), 235–248.
16. Y. Gallot and P. Moree, Ternary cyclotomic polynomials having a large coefficient. *J. reine angew. Math.* **632** (2009), 105–125.
17. Y. Gallot, P. Moree and R. Wilms, The family of ternary cyclotomic polynomials with one free prime. *Involve* **4** (2011), 317–341.
18. A. Granville, Harald Cramér and the distribution of prime numbers. *Scand. Actuar. J.* **1** (1995), 12–28.
19. H. Halberstam and H.-E. Richert, *Sieve Methods* (London Mathematical Society Monographs **4**), Academic Press (London–New York, 1974).
20. D. R. Heath-Brown, Zero-free regions for Dirichlet L-functions, and the least prime in an arithmetic progression. *Proc. Lond. Math. Soc.* **64** (1992), 265–338.
21. D. R. Heath-Brown, The differences between consecutive primes. V. *Int. Math. Res. Not. IMRN*, to appear, https://academic.oup.com/imrn/article-abstract/doi/10.1093/imrn/rnz295/5676434.
22. H. Iwaniec and E. Kowalski, *Analytic Number Theory* (American Mathematical Society Colloquium Publications **53**), American Mathematical Society (Providence, RI, 2004).
23. C.-G. Ji, W.-P. Li and P. Moree, Values of coefficients of cyclotomic polynomials II. *Discrete Math.* **309** (2009), 1720–1723.
24. N. Kaplan, Flat cyclotomic polynomials of order three. *J. Number Theory* **127** (2007), 118–126.
25. A. Kosyak, P. Moree, E. Sofos and B. Zhang, Cyclotomic polynomials with prescribed height and prime number theory, Max Planck Institute for Mathematics. *Preprint*, MPIM2019-59.
26. T. Y. Lam and K. H. Leung, On the cyclotomic polynomial $\Phi_{pq}(X)$. *Amer. Math. Monthly* **103** (1996), 562–564.
27. Y. Lamzouri, X. Li and K. Soundararajan, Conditional bounds for the least quadratic non-residue and related problems. *Math. Comp.* **84** (2015), 2391–2412.
28. E. Lehmer, On the magnitude of the coefficients of the cyclotomic polynomials. *Bull. Amer. Math. Soc.* **42** (1936), 389–392.
29. A. Migotti, Zur Theorie der Kreisteilungsgleichung. *S.-B. der Math.-Naturwiss. Classe der Kaiserlichen Akademie der Wissenschaften, Wien,* (2) **87** (1883), 7–14.
30. P. Moree and H. Hommersom, Value distribution of Ramanujan sums and cyclotomic polynomial coefficients, https://arxiv.org/abs/math/0307352.
31. P. Moree and E. Roşu, Non-Beiter ternary cyclotomic polynomials with an optimally large set of coefficients. *Int. J. Number Theory* **8** (2012), 1883–1902.
32. T. R. Nicely, First occurrence prime gaps. http://www.trnicely.net/gaps/gaplist.html.
33. Z.-W. Sun, On a sequence involving sums of primes. *Bull. Aust. Math. Soc.* **88** (2013), 197–205.
34. J. Suzuki, On coefficients of cyclotomic polynomials. *Proc. Japan Acad. Ser. A Math. Sci.* **63** (1987), 279–280.

35. M. Visser, Strong version of Andrica's conjecture. *Int. Math. Forum* **14** (2019), 181–188.
36. T. Xylouris, *Über die Nullstellen der Dirichletschen L-Funktionen und die kleinste Primzahl in einer arithmetischen Progression*, Bonner Mathematische Schriften **404**, Universität Bonn, Mathematisches Institut (Bonn, 2011).
37. G. Yu, The differences between consecutive primes. *Bull. Lond. Math. Soc.* **28** (1996), 242–248.
38. J. Zhao and X. Zhang, Coefficients of ternary cyclotomic polynomials. *J. Number Theory* **130** (2010), 2223–2237.

Alexandre Kosyak,
Institute of Mathematics,
Ukrainian National Academy of Sciences,
3 Tereshchenkivs'ka Str.,
Kyiv, 01024,
Ukraine
Email: kosyak02@gmail.com

Pieter Moree,
Max-Planck-Institut für Mathematik,
Vivatsgasse 7 Bonn, D-53111,
Germany
Email: moree@mpim-bonn.mpg

Efthymios Sofos,
School of Mathematics and Statistics,
University of Glasgow,
University Place Glasgow, G12 8SQ,
United Kingdom
Email: efthymios.sofos@glasgow.ac.uk

Bin Zhang,
School of Mathematical Sciences,
Qufu Normal University,
57 Jingxuan West Road Qufu, 273165,
P. R. China
Email: zhangbin100902025@163.com

Mathematika is owned by University College London and published by the London Mathematical Society. All surplus income from the publication of *Mathematika* is returned to mathematicians and mathematics research via the Society's research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.