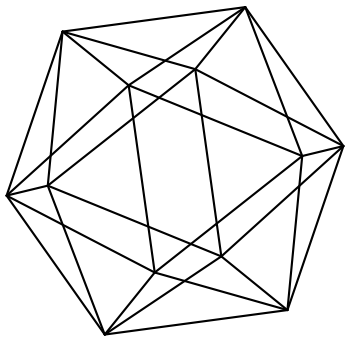
Max-Planck-Institut für Mathematik Bonn

Cyclotomic coefficients: gaps and jumps

by

Oana-Maria Camburu
Emil-Alexandru Ciolan
Florian Luca
Pieter Moree
Igor E. Shparlinski



Max-Planck-Institut für Mathematik
Preprint Series 2015 (32)

Cyclotomic coefficients: gaps and jumps

Oana-Maria Camburu
Emil-Alexandru Ciolan
Florian Luca
Pieter Moree
Igor E. Shparlinski

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

Department of Computer Science
University of Oxford
Oxford OX1 3QD
United Kingdom

Rheinische Friedrich-Wilhelms-Universität
Regina-Pacis-Weg 3
53113 Bonn
Germany

School of Mathematics
University of the Witwatersrand
Private Bag X3
Wits 2050
South Africa

Department of Pure Mathematics
University of New South Wales
Sydney, NSW 2052
Australia

MPIM 15-32

CYCLOTOMIC COEFFICIENTS: GAPS AND JUMPS

OANA-MARIA CAMBURU, EMIL-ALEXANDRU CIOLAN, FLORIAN LUCA, PIETER MOREE, AND IGOR E. SHPARLINSKI

Abstract. We improve several recent results by Hong, Lee, Lee and Park (2012) on gaps and Bzdęga (2014) on jumps amongst the coefficients of cyclotomic polynomials. Besides direct improvements, we also introduce several new techniques that have never been used in this area.

1. Introduction

As usual, for an integer $n \geq 1$, we use $\Phi_n(Z)$ to denote the *nth cyclotomic polynomial*, that is,

$$\Phi_n(Z) = \prod_{\substack{j=0 \\ \gcd(j,n)=1}}^{n-1} (Z - \mathbf{e}_n(j)),$$

where for an integer $m \geq 1$ and a real z, we put

$$\mathbf{e}_m(z) = \exp(2\pi i z/m).$$

Clearly $\deg \Phi_n = \varphi(n)$, where $\varphi(n)$ is the Euler function. Using the above definition one sees that

$$Z^n - 1 = \prod_{d|n} \Phi_d(Z). \tag{1.1}$$

The Möbius inversion formula then yields

$$\Phi_n(Z) = \prod_{d|n} (Z^d - 1)^{\mu(n/d)}, \tag{1.2}$$

where $\mu(n)$ denotes the Möbius function.

We write

$$\Phi_n(Z) = \sum_{k=0}^{\varphi(n)} a_n(k) Z^k.$$

2010 *Mathematics Subject Classification.* 11B83, 11L07, 11N25.

Key words and phrases. Coefficients of cyclotomic polynomials, products of primes, numerical semigroups, double Kloosterman sums.

For $n > 1$ clearly $Z^{\varphi(n)}\Phi_n(1/Z) = \Phi_n(Z)$ and so

$$a_n(k) = a_n(\varphi(n) - k), \qquad 0 \le k \le \varphi(n), \qquad n > 1. \tag{1.3}$$

Recently, there has been a burst of activity in studying the cyclotomic coefficients $a_n(k)$, see, for example, [3, 4, 5, 6, 10, 15, 17, 18, 19, 31, 35] and references therein. Furthermore, in several works *inverse cyclotomic polynomials*

$$\Psi_n(Z) = (Z^n - 1)/\Phi_n(Z)$$

have also been considered, see [7, 21, 22, 23, 29].

The identities $\Phi_{2n}(Z) = \Phi_n(-Z)$, with $n > 1$ odd and $\Phi_{pm}(Z) = \Phi_m(Z^p)$ if $p \mid m$, show that, as far as the study of coefficients is concerned, the complexity of $\Phi_n(Z)$ is determined by its number of distinct odd prime factors. Most of the recent activity concerns the so called *binary* and *ternary* cyclotomic polynomials, which are polynomials $\Phi_n(Z)$ with $n = pq$ and $n = pqr$, respectively, where p, q and r are pairwise distinct odd primes. In particular, the long standing *Beiter conjecture* about coefficients of ternary cyclotomic polynomials has been shown to be wrong (in a very strong sense) by Gallot and Moree [18], see also [3]. It is quite remarkable that the results of [18] are based on seemingly foreign to the problem and deep analytic results such as bounds of Kloosterman sums (see [25, Theorem 11.11]) and a result of Duke, Friedlander and Iwaniec [13] on the distribution of roots of quadratic congruences.

Furthermore, Fouvry [15] has used bounds of exponential sums with reciprocals of primes from [16] in studying the sparsity of binary cyclotomic polynomials and improved a result of Bzdęga [4]. It is quite possible that more recent bounds of Baker [2] and Irving [24] can lead to further progress in this direction.

Here we continue to study ternary cyclotomic and inverse cyclotomic polynomials and using some classical and more recent tools from analytic number theory, we improve several previous results. We also show how to employ a rather nonstandard tool of using numerical semigroups to study the binary cyclotomic polynomials. We give a brief summary of our contributions in Section 5.

We recall that the notations $U = O(V)$, $U \ll V$ and $V \gg U$ are all equivalent to the assertion that the inequality $|U| \le cV$ holds for some constant $c > 0$. If the constant c depends on a parameter, say ε, we indicate this as $U \ll_\varepsilon V$, etc. We also write $U \asymp V$ if $U \ll V \ll U$. For an integer m and a real $M \ge 1$, we write $m \sim M$ to indicate that $m \in [M, 2M]$. If $\mathcal{A}$ is a set of non-negative integers we denote by $\mathcal{A}(x)$ the subset of integers $n \in \mathcal{A}$ with $n \le x$.

As usual, we let $\pi(x)$ denote the number of primes $p \le x$. A few times we use the estimate

$$\pi(x) = \frac{x}{\log x} + O\left(\frac{x}{(\log x)^2}\right). \tag{1.4}$$

We always use the letters ℓ, p, q and r to denote prime numbers, while k, m and n always denote positive integers.

2. Maximum gap

2.1. **Definitions and background.**

Given a polynomial

$$f(Z) = c_1 Z^{e_1} + \cdots + c_t Z^{e_t} \in \mathbb{Z}[Z], \text{with } c_i \neq 0 \text{ and } e_1 < \cdots < e_t,$$

we define the *maximum gap* of f as

$$g(f) = \max_{1 \le i < t} (e_{i+1} - e_i),$$

where we set $g(f) = 0$ when $t = 1$. The study of $g(\Phi_n)$ and $g(\Psi_n)$ has been initiated by Hong, Lee, Lee and Park [22] who have reduced the study of these gaps to the case where n is square-free and odd. They were led to their gap study in an attempt to provide a simple and exact formula for the minimum Miller loop length in the Ate_i pairing arising in elliptic curve cryptography, see [21, 22, 23]. As they write in [23], the crucial idea in finding this exact formula is that it becomes more manageable when it is suitably recast in terms of inverse cyclotomic polynomials and consequently turns into a problem involving the maximum gaps.

It is easy to see [22] that if $p < q$ are odd primes, then

$$g(\Phi_p) = 1, \qquad g(\Psi_p) = 1, \qquad g(\Psi_{pq}) = q - p + 1.$$

The simplest non-trivial case occurs when n is a product of two distinct primes, where by [22, Theorem 1], for two primes $3 \le p < q$ we have

$$g(\Phi_{pq}) = p - 1. \tag{2.1}$$

For ternary n we have by [22, Theorems 2 and 3] the following partial result on $g(\Psi_n)$.

Lemma 2.1. *Put*

$$\mathcal{R}_3 = \{n = pqr \ : \ p < q < r \text{ primes}, \ 4(p-1) > q, \ p^2 > r\}.$$

Let $n = pqr$ with primes $2 < p < q < r$. Then:

- *we have*

$$\max\left\{p - 1, \frac{2n}{p} - \deg \Psi_n\right\} \le g(\Psi_n) < 2n\left(\frac{1}{p} + \frac{1}{q} + \frac{1}{r}\right) - \deg(\Psi_n);$$

- *if $n \notin \mathcal{R}_3$, then $g(\Psi_n) = 2n/p - \deg \Psi_n$.*

Note that $\mathcal{R}_3$ consists of ternary integers only. Although in this work we are mostly interested in the set $\mathcal{R}_3$ that appears in Lemma 2.1, we first make some comments concerning (2.1).

Moree [30] derives (2.1) using a very different technique which is based on numerical semigroups. We come back to this in Section 3.

Here we present yet another short proof of (2.1) communicated to us by Nathan Kaplan. We start by noticing that the nonzero coefficients of $\Phi_{pq}(Z)$ alternate between 1 and -1, see [8, 30]. From (1.2) one easily obtains that $\Phi_p(Z)\Phi_{pq}(Z) = \Phi_p(Z^q)$. Suppose there is a gap of length p or greater, say, for some positive integers $b \ge a + p$ and there are no nonzero coefficients between Z^a and Z^b. One of these coefficients is 1 and the other is -1. Now consider the product $\Phi_p(Z)\Phi_{pq}(Z)$ and examine at the coefficients of the terms with Z^{a+p-1} and Z^b. One of these coefficients is 1 and the other is -1, contradicting the fact that $\Phi_p(Z^q)$ has no coefficient equal to -1 and (2.1) follows.

2.2. **Main result.** We now estimate $\mathcal{R}_3(x)$. We frequently make use of the bound

$$\sum_{p<z} p^k \ll \frac{z^{k+1}}{\log z}, \tag{2.2}$$

which holds for any fixed $k \ge 1$ and real $z \ge 2$ and follows easily from (1.4) by partial summation.

Theorem 2.2. *We have*

$$\#\mathcal{R}_3(x) = \frac{cx}{(\log x)^2} + O\left(\frac{x \log\log x}{(\log x)^3}\right),$$

where $c = (1 + \log 4)\log 4 = 3.30811\ldots$.

Proof. We take $\mathcal{S}(x) = \{n : n \le x/(\log x)^3\}$. Clearly, $\#\mathcal{S}(x) \le x(\log x)^{-3}$. So, from now on, we only consider

$$n \in \mathcal{L}_3(x) = \mathcal{R}_3(x) \setminus \mathcal{S}(x).$$

Now, for $n = pqr \in \mathcal{L}_3(x)$ we have

$$p^3 < pqr \le x \qquad \text{and} \qquad 4p^4 = p^2(4p)p > rqp = n > \frac{x}{(\log x)^3}.$$

Thus,

$$p \in \mathcal{I}(x) = [y(x), x^{1/3}],$$

where

$$y(x) = \frac{x^{1/4}}{\sqrt{2}(\log x)^{3/4}}.$$

In particular, $\log p \asymp \log q \asymp \log r \asymp \log x$.

We now fix a prime $p \in \mathcal{I}(x)$ and consider the interval

$$\mathcal{J}_p = (p, 4(p-1)).$$

If $n = pqr \in \mathcal{L}_3(x)$, then $q \in \mathcal{J}_p$. For fixed p and $q \in \mathcal{J}_p$, we have $r < \min\{p^2, x/pq\}$. We distinguish the following two cases depending of whether $\min\{p^2, x/pq\} = p^2$, and denote by $\mathcal{U}_3(x)$ the set of such $n \in \mathcal{L}_3(x)$, or $\min\{p^2, x/pq\} = x/pq$, and denote by $\mathcal{V}_3(x)$ the set of such $n \in \mathcal{L}_3(x)$.

We estimate the cardinalities of the sets $\mathcal{U}_3(x)$ and $\mathcal{V}_3(x)$ separately.

First, assume that $n = pqr \in \mathcal{U}_3(x)$. Therefore $p^4 < p^2(pq) < x$, so in fact $p < x^{1/4}$. For fixed p and q, the number of primes $r < p^2$ is $O(\pi(p^2)) = O(p^2/\log x)$. Further, the number of choices of $q < 4p$ is $O(\pi(4p)) = O(p/\log x)$. Thus, for fixed $p \in \mathcal{I}(x)$, the number of choices for the pairs (q, r) is

$$O\left(\frac{p^2}{\log x} \times \frac{p}{\log x}\right) = O\left(\frac{p^3}{(\log x)^2}\right).$$

Summing up the above bound over all the choices $p \le x^{1/4}$ and using (2.2) with $k = 3$ we derive

$$\#\mathcal{U}_3(x) \ll \frac{1}{(\log x)^2} \sum_{p < x^{1/4}} p^3 \ll \frac{x}{(\log x)^3}, \tag{2.3}$$

which gets absorbed in the error term of the desired asymptotic formula for $\mathcal{R}_3(x)$.

So we now consider $n = pqr \in \mathcal{V}_3(x)$. In this case, using (1.4) and $\log(x/pq) > \log q \asymp \log x$, we see that the number of choices for the prime $r \in (q, x/pq)$ when p and q are fixed is

$$\pi\left(\frac{x}{pq}\right) - \pi(q) = \frac{x}{pq \log(x/pq)} + O\left(\frac{x}{pq(\log x)^2} + \frac{q}{\log q}\right). \tag{2.4}$$

Note that

$$\log(x/pq) = \log x - \log p - \log q = \log x - 2\log p + O(1)$$

for $q \in \mathcal{J}_p$. Thus, using $(1+z)^{-1} = 1 + O(z)$ and $\log(x/p^2) \asymp \log x$, a simple calculation shows that

$$\frac{1}{\log(x/pq)} = \frac{1}{\log(x/p^2)} + O\left(\frac{1}{(\log x)^2}\right).$$

Hence, using (2.4), we see that the number of choices for r when p and q are fixed is

$$\frac{x}{pq \log(x/p^2)} + O\left(\frac{x}{pq(\log x)^2} + \frac{q}{\log x}\right). \tag{2.5}$$

Now we sum up over $q \in (p, 4(p-1))$ and use the Mertens formula (see [25, Equation (2.15)]) in the form

$$\sum_{\substack{\ell<X\\ \ell \text{ prime}}} \frac{1}{\ell} = \log\log X + \alpha + O\left(\frac{1}{(\log X)^2}\right), \qquad X \to \infty, \tag{2.6}$$

with some constant α to deduce that

$$\begin{aligned}\sum_{p<q<4(p-1)} \frac{1}{q} &= \log\log 4p - \log\log p + O\left(\frac{1}{(\log p)^2}\right)\\ &= \frac{\log 4}{\log p} + O\left(\frac{1}{(\log x)^2}\right).\end{aligned} \tag{2.7}$$

Summing over the choices $q \in \mathcal{J}_p$, we see that the main term in (2.5) contributes

$$\frac{\log 4}{\log p \log(x/p^2)} + O\left(\frac{x}{p(\log x)^3}\right).$$

We now sum up the first error terms in (2.5), getting

$$\sum_{p<q<4(p-1)} \frac{x}{pq(\log x)^2} \ll \frac{x}{p(\log x)^2} \sum_{p<q<4p} \frac{1}{q} \ll \frac{x}{p(\log x)^3},$$

where we have used (2.7), and the second error terms in (2.5) getting

$$\sum_{p<q<4(p-1)} \frac{q}{\log x} \ll \frac{1}{\log x} \sum_{q<4p} q \ll \frac{p^2}{(\log x)^2},$$

where we have used (2.2) with $k = 1$ and $z = 4p$.

Hence, the number of choices for the pair (q, r) when $p \in \mathcal{I}(x)$ is

$$\frac{x \log 4}{p \log p \log(x/p^2)} + O\left(\frac{x}{p(\log x)^3} + \frac{p^2}{(\log x)^2}\right).$$

Now we sum over the primes $p \in \mathcal{I}(x)$. Since $y(x) > x^{1/5}$ for large x, for the total error term we obtain

$$\begin{aligned}\sum_{p\in\mathcal{I}(x)} \left(\frac{x}{p(\log x)^3} + \frac{p^2}{(\log x)^2}\right) &\le \frac{x}{(\log x)^3} \sum_{x^{1/5}<p<x} \frac{1}{p} + \frac{1}{(\log x)^2} \sum_{p<x^{1/3}} p^2\\ &\ll \frac{x}{(\log x)^3}\end{aligned}$$

by (2.6) (applied to the sum of $1/p$) and (2.2) (applied to the sum of p^2). Thus,

$$\#\mathcal{V}_3(x) = \sum_{p\in\mathcal{I}(x)} \frac{x \log 4}{p \log p \log(x/p^2)} + O\left(\frac{x}{(\log x)^3}\right).$$

For the sum of the main term (after we pull out $x\log 4$), using the Stieltjes integral and then partial integration, we obtain

$$\begin{aligned}\sum_{p\in\mathcal{I}(x)}\frac{1}{p\log p\log(x/p^2)} &= \int_{y(x)}^{x^{1/3}}\frac{d\pi(t)}{t\log t\log(x/t^2)}\\ &= \int_{y(x)}^{x^{1/3}}\frac{d\pi(t)}{t\log t\log(x/t^2)} = \frac{\pi(t)}{t\log t\log(x/t^2)}\Big|_{t=y(x)}^{t=x^{1/3}}\\ &\quad - \int_{y(x)}^{x}\pi(t)\frac{d}{dt}\left(\frac{1}{t\log t\log(x/t^2)}\right).\end{aligned}$$

Since $\pi(t)/t = O(1/\log t)$, we see that the first term above is

$$\frac{\pi(t)}{t\log t\log(x/t^2)}\Big|_{t=y(x)}^{t=x^{1/3}} = O\left(\frac{1}{(\log x)^3}\right).$$

We have

$$\begin{aligned}\frac{d}{dt}\left(\frac{1}{t\log t\log(x/t^2)}\right) &= -\frac{\log t\log(x/t^2)+\log(x/t^2)-2\log t}{t^2(\log t)^2(\log x-2\log t)^2}\\ &= -\frac{1}{t^2\log t(\log x-2\log t)}+O\left(\frac{1}{t^2(\log x)^3}\right)\\ &= -\frac{1}{t(\log t)^2(\log x-2\log t)}+O\left(\frac{1}{t\log t(\log x)^3}\right).\end{aligned}$$

Thus, using (1.4), we have

$$\begin{aligned}\int_{y(x)}^{x}\pi(t)\frac{d}{dt}\left(\frac{1}{t\log t\log(x/t^2)}\right)dt &= -\int_{y(x)}^{x^{1/3}}\frac{dt}{t(\log t)^2(\log x-2\log t)}\\ &\quad + O\left(\int_{y(x)}^{x^{1/3}}\frac{dt}{t(\log t)(\log x)^3}\right).\end{aligned}$$

Since

$$\int_{y(x)}^{x^{1/3}}\frac{dt}{t\log t} = \log\log t\Big|_{t=y(x)}^{t=x^{1/3}} \ll 1,$$

we derive

(2.8)
$$\#\mathcal{V}_3(x) = x\log 4\int_{y(x)}^{x^{1/3}}\frac{dt}{t(\log t)^2(\log x-2\log t)}+O\left(\frac{x}{(\log x)^3}\right).$$

Inside the integral, we make the change of variable $t=x^u$, for which $dt = x^u\log x du$. Further, we have $\log t = u\log x$ and we also have

$\log x - 2\log t = (\log x)(1-2u)$. Defining

$$z(x) = \frac{\log y(x)}{\log x} = \frac{1}{4} - \frac{\log(2^{1/2}(\log x)^{3/4})}{\log x} = \frac{1}{4} + O\left(\frac{\log\log x}{\log x}\right),$$

we see that

$$\int_{y(x)}^{x^{1/3}} \frac{dt}{t(\log t)^2(\log x - 2\log t)} = \frac{1}{(\log x)^2}\int_z^{1/3} \frac{du}{u^2(1-2u)}. \tag{2.9}$$

We now write

$$\begin{aligned}\int_z^{1/3} \frac{du}{u^2(1-2u)} &= \int_{1/4}^{1/3} \frac{du}{u^2(1-2u)} + \int_z^{1/4} \frac{du}{u^2(1-2u)} \\ &= \int_{1/4}^{1/3} \frac{du}{u^2(1-2u)} + O\left(\frac{\log\log x}{\log x}\right).\end{aligned} \tag{2.10}$$

One verifies that

$$\frac{d}{du}\left(-2\log(\frac{1}{u}-2) - \frac{1}{u}\right) = \frac{1}{u^2(1-2u)}.$$

Hence,

$$\int_{1/4}^{1/3} \frac{du}{u^2(1-2u)} = -\left(2\log(\frac{1}{u}-2) + \frac{1}{u}\right)\Big|_{u=1/3}^{u=1/4} = 1 + \log 4. \tag{2.11}$$

Substituting (2.10) and (2.11) in (2.9) and recalling (2.8) we obtain

$$\#\mathcal{V}_3(x) = \frac{x(1+\log 4)\log 4}{(\log x)^2} + O\left(\frac{x\log\log x}{(\log x)^3}\right). \tag{2.12}$$

Now, combining (2.3) and (2.12), we conclude the proof. □

Let

$$\mathcal{Q}_3(x) = \{n \le x \ : \ n = pqr \text{ for primes } 2 < p < q < r\}.$$

Using [34, Theorem 4, Section II.6.1] we conclude that

$$\#\mathcal{Q}_3(x) = (1+o(1))\frac{x(\log\log x)^2}{2\log x} \tag{2.13}$$

as $x \to \infty$. Combining this estimate with Theorem 2.2, we immediately derive the following comparison of $\#\mathcal{R}_3(x)$ and $\#\mathcal{Q}_3(x)$.

Corollary 2.3. *As $x \to \infty$ we have*

$$\#\mathcal{R}_3(x) = \frac{2(1+\log 4)\log 4 + o(1)}{(\log x)(\log\log x)^2}\#\mathcal{Q}_3(x).$$

Corollary 2.3 justifies the claim in [22, Remark 1] that $\mathcal{R}_3(x)$ is a sparse set, that is, that $\#\mathcal{R}_3(x) = o(\#\mathcal{Q}_3(x))$ as $x \to \infty$. Indeed, this claim can also be justified with less effort. Namely, it is not difficult to show that

$$\#\mathcal{R}_3(x) \ll \frac{\#\mathcal{Q}_3(x)}{(\log x)(\log\log x)^2}. \tag{2.14}$$

To do so, we note that if $n = pqr \in \mathcal{R}_3(x)$ then

$$p \le (pqr)^{1/3} = n^{1/3} \le x^{1/3}.$$

Hence,

$$\#\mathcal{R}_3(x) \le \sum_{p \le x^{1/3}} \sum_{p<q<4(p-1)} \min\{\pi(\frac{x}{pq}), \pi(p^2)\}.$$

After making some easy estimates one then arrives at (2.14).

Lemma 2.1 also motivates us to study the set $\mathcal{E}_3$ of exceptional ternary integers $n = pqr$, $2 < p < q < r$, for which $g(\Psi_n) \neq 2n/p - \deg \Psi_n$. We now consider a certain subset $\mathcal{E}_4$ of $\mathcal{E}_3$ and estimate $\#\mathcal{E}_4(x)$ (Theorem 2.5). The relevance of this estimate becomes clear in the proof of Theorem 2.6 below.

Lemma 2.4. *Let*

$$\mathcal{E}_3 = \{n = pqr \ : \ 2 < p < q < r \text{ and } g(\Psi_n) \neq 2n/p - \deg \Psi_n.\}.$$

and

$$\mathcal{E}_4 = \{n = pqr \ : \ 2 < p < q < r \text{ and } qr < (q+r)(p-1)\}.$$

We have $\mathcal{E}_4 \subseteq \mathcal{E}_3 \subseteq \mathcal{R}_3 \subseteq \mathcal{Q}_3$.

Proof. Consider the inequality $p - 1 > 2n/p - \deg \Psi_n$. As

$$2n/p - \deg \Psi_n = 2qr - pqr + (p-1)(q-1)(r-1),$$

this is equivalent to

$$qr < (q+r)(p-1). \tag{2.15}$$

By Lemma 2.1, it now follows that $\mathcal{E}_4 \subseteq \mathcal{E}_3 \subseteq \mathcal{R}_3$. The final inclusion is obvious. □

Theorem 2.5. *For the set*

$$\mathcal{E}_4 = \{n = pqr \ : \ 2 < p < q < r \text{ and } qr < (q+r)(p-1)\}$$

we have

$$\frac{x}{(\log x)^3} \ll \#\mathcal{E}_4(x) \ll \frac{x}{(\log x)^3}.$$

Proof. Note that replacing $p > 2$ in the definition of $\mathcal{E}_4$ by $p \ge 2$ does not change the set.

Upper bound. Let $n \in \mathcal{E}_4(x)$. We may assume that $n > x/(\log x)^3$, since the set $\{n \ : \ n \le x/(\log x)^3\}$ has cardinality at most $x/(\log x)^3$. By Lemma 2.4 we have $\mathcal{E}_4(x) \subseteq \mathcal{R}_3(x)$. Thus, writing $n = pqr$ with $p < q < r$, we have

$$4p^4 = p(4p)(p^2) > pqr > \frac{x}{(\log x)^3},$$

therefore $p > x^{1/5}$ for large values of x. The inequality defining $\mathcal{E}_4$ is equivalent to

$$\left(\frac{q}{p-1}\right)\left(\frac{r}{p-1}\right) < \frac{q}{p-1} + \frac{r}{p-1},$$

or

$$\frac{r}{p-1} < \frac{q/(p-1)}{q/(p-1)-1} = \frac{q}{q-p+1}. \tag{2.16}$$

Since $r > q$, we have that $r/(p-1) > q/(p-1)$, giving

$$\frac{q}{p-1} < \frac{q/(p-1)}{q/(p-1)-1},$$

which leads to $q < 2(p-1)$. We put $h = q - (p-1)$, so $h < p-1$. Then (2.16) is equivalent to

$$\frac{r}{p-1} < \frac{q}{q-p+1} = \frac{p-1+h}{h}$$

so

$$r < \frac{(p+h-1)(p-1)}{h}.$$

We also have $r < x/pq$, so we get

$$r < \min\left\{\frac{x}{p(p+h-1)}, \frac{(p+h-1)(p-1)}{h}\right\}. \tag{2.17}$$

We consider both possibilities that may occur on the right hand side of (2.17) separately.

Case 1. Assume that

$$\frac{x}{p(p+h-1)} < \frac{(p+h-1)(p-1)}{h}.$$

Then

$$xh < p(p-1)(p+h-1)^2 < p^2q^2 < 4p^4,$$

so $p > 2^{-1/2}(xh)^{1/4}$. Since $r > p > x^{1/5}$, the number of primes $r < x/pq$ is

$$\pi\left(\frac{x}{pq}\right) = O\left(\frac{x}{pq(\log x)}\right). \tag{2.18}$$

We now fix h and sum up the above bound over $p \in [2^{-1/2}(xh)^{1/4}, x^{1/3}]$ such that p and $p+h-1$ are both primes. Let $\sigma(k)$ denote the sum of all integer positive divisors of $k \geq 1$. By [32, Theorem 7.3], or alternatively by the fundamental lemma of the sieve applied to the sequence $n(n+h-1)$, see [20, Corollary 2.4.1], we see that the counting function of

$$\mathcal{P}_h = \{p : p,\ p+h-1 \text{ are both primes}\}$$

satisfies

$$\#\left(\mathcal{P}_h \cap [1,t]\right)) \ll \frac{t}{(\log t)^2} \prod_{p \mid (h-1)} \left(1+\frac{1}{p}\right) \ll \frac{t}{(\log t)^2} \left(\frac{\sigma(h-1)}{h-1}\right),$$

where we used the observation that

$$\prod_{p \mid (h-1)} \left(1+\frac{1}{p}\right) \leq \sum_{d \mid (h-1)} \frac{1}{d} = \frac{\sigma(h-1)}{h-1}.$$

With the Abel summation formula, we derive that

$$\begin{aligned} \sum_{\substack{p > 2^{-1/2}(xh)^{1/4} \\ p \in \mathcal{P}_h}} \frac{1}{pq} &\leq \sum_{\substack{p > 2^{-1/2}(xh)^{1/4} \\ p \in \mathcal{P}_h}} \frac{1}{p^2} \\ &\ll \frac{1}{(\log x)^2} \left(\frac{\sigma(h-1)}{h-1}\right) \int_{2^{-1/2}(xh)^{1/4}}^{x^{1/3}} \frac{dt}{t^2} \end{aligned}$$

and hence

$$\sum_{\substack{p > 2^{-1/2}(xh)^{1/4} \\ p \in \mathcal{P}_h}} \frac{1}{pq} \ll \frac{1}{x^{1/4}(\log x)^2} \left(\frac{\sigma(h-1)}{h-1}\right) \frac{1}{h^{1/4}}. \tag{2.19}$$

Thus, on combining (2.18) and (2.19), we obtain that, for fixed h, the number of $n = pqr \in \mathcal{E}_4(x)$ in Case 1 is bounded by

$$\sum_{\substack{p > 2^{-1/2}(xh)^{1/4} \\ p \in \mathcal{P}_h}} \pi\left(\frac{x}{pq}\right) = O\left(\frac{x^{3/4}}{(\log x)^3} \left(\frac{\sigma(h-1)}{h-1}\right) \frac{1}{h^{1/4}}\right). \tag{2.20}$$

Note that, by the way it was defined, h is odd and $h \geq 3$. We now sum up over such h and use that

$$\sum_{n \leq x} \frac{\sigma(n)}{n} = \sum_{n \leq x} \sum_{d \mid n} \frac{1}{d} \leq x \sum_{d \leq x} \frac{1}{d^2},$$

to see that

$$\sum_{\substack{3\le h\le t\\ h \text{ odd}}} \frac{\sigma(h-1)}{h-1} = O(t).$$

We use this estimate and the Abel summation formula with the sequence $a_h = \sigma(h-1)/(h-1)$ for $h \ge 3$ and odd (and $a_h = 0$ otherwise) and the function $f(t) = t^{-1/4}$, to get that

$$\sum_{\substack{3\le h\le x^{1/3}\\ h \text{ odd}}} \left(\frac{\sigma(h-1)}{h-1}\right) \frac{1}{h^{1/4}} \ll \int_1^{x^{1/3}} \frac{dt}{t^{1/4}} \ll t^{3/4}\Big|_{t=1}^{t=x^{1/3}} \ll x^{1/4}, \tag{2.21}$$

which, together with (2.20), gives a bound of $O(x/(\log x)^3)$ for the number of $n \in \mathcal{E}_4(x)$ in Case 1.

Case 2. Assume that

$$\frac{x}{p(p+h-1)} \ge \frac{(p+h-1)(p-1)}{h}.$$

Then

$$xh > p(p-1)q^2 \ge q^4/4,$$

giving $q < 2^{1/2}(xh)^{1/4}$. In this case,

$$r \le \frac{q(p-1)}{h} < \frac{4p^2}{h},$$

and the number of such primes is

$$\pi\left(\frac{4p^2}{h}\right) = O\left(\frac{p^2}{h(\log x)}\right),$$

where we used the observation that $p^2/h > p > x^{1/5}$. Again, fixing h and summing up over p, we get a bound of

$$O\left(\frac{1}{h(\log x)} \sum_{\substack{x^{1/5}<p<2^{1/2}(xh)^{1/4}\\ p\in\mathcal{P}_h}} p^2\right). \tag{2.22}$$

Since $p \in \mathcal{P}_h$, it follows, by the Abel summation formula, that

$$\begin{aligned}\sum_{\substack{x^{1/5}\le p<2^{1/2}(xh)^{1/4}\\ p\in\mathcal{P}_h}} p^2 &\ll \frac{1}{(\log x)^2}\left(\frac{\sigma(h-1)}{h-1}\right)\int_{x^{1/5}}^{2^{1/2}(xh)^{1/4}} t^2 dt\\ &\ll \frac{(xh)^{3/4}}{(\log x)^2}\left(\frac{\sigma(h-1)}{h-1}\right).\end{aligned}$$

Inserting this into (2.8), we conclude that, for fixed h, the number of $n \in \mathcal{E}_4(x)$ in Case 2 is

$$O\left(\frac{x^{3/4}}{(\log x)^3}\left(\frac{\sigma(h-1)}{h-1}\right)\frac{1}{h^{1/4}}\right).$$

We now sum again over $h < x^{1/3}$ and use (2.21) to get that the number of $n \in \mathcal{E}_4(x)$ in Case 2 is $O(x/(\log x)^3)$.

Lower bound. Let $\varepsilon > 0$ to be chosen later. Now suppose that there are primes p, q, r such that

$$q \in (p, p(1+\varepsilon)),\ r \in (q, q(1+\varepsilon)). \tag{2.23}$$

Since

$$qr < pq(1+\varepsilon)^2 \qquad \text{and} \qquad (p+q)(p-1) < (q+r)(p-1),$$

the inequality (2.15) holds if

$$1 + \frac{p}{q} - \frac{1}{p} - \frac{1}{q} > 1 + \frac{1}{1+\epsilon} - \frac{1}{p} - \frac{1}{q} > (1+\varepsilon)^2. \tag{2.24}$$

Now we choose any $\varepsilon > 0$ satisfying $1+1/(1+\varepsilon) > (1+\varepsilon)^3$ and observe that there exists $p_0(\varepsilon)$ such that if $p \geq p_0(\varepsilon)$, then inequality (2.24) is satisfied.

Note that if $p^3(1+\varepsilon)^2 \leq x$, then for any choice of $p \geq p_0(\varepsilon)$ and q, r satisfying (2.23), the inequality (2.15) is satisfied and $pqr \leq x$. Putting $c(\varepsilon) = (1+\varepsilon)^{-2/3}$ and noting that by (1.4), we have

$$\pi(x(1+\varepsilon)) - \pi(x) = \frac{(\varepsilon + o(1))\, x}{\log x},$$

it follows that

$$\begin{aligned}
\#\mathcal{E}_3(x) &\geq \sum_{p_0(\varepsilon) < p \leq c(\varepsilon)x^{1/3}} \sum_{p < q < p(1+\varepsilon)} \sum_{q < r < q(1+\varepsilon)} 1 \\
&\gg_\varepsilon \sum_{p_0(\varepsilon) < p \leq c(\varepsilon)x^{1/3}} \sum_{p < q < p(1+\varepsilon)} \frac{q}{\log q} \\
&\gg_\varepsilon \sum_{p_0(\varepsilon) < p \leq c(\varepsilon)x^{1/3}} \frac{p^2}{(\log p)^2} \gg_\varepsilon \int_{p_0(\varepsilon)}^{c(\varepsilon)x^{1/3}} \frac{t^2\,dt}{(\log t)^3} \gg_\varepsilon \frac{x}{(\log x)^3}.
\end{aligned}$$

Taking, e.g., $\varepsilon = 10^{-1}$ completes the proof. □

We have now the ingredients to estimate $\#\mathcal{E}_3(x)$.

Theorem 2.6. *For the set*

$$\mathcal{E}_3 = \{n = pqr : 2 < p < q < r \text{ and } g(\Psi_n) \neq 2n/p - \deg \Psi_n.\}.$$

we have

$$\frac{\#\mathcal{Q}_3(x)}{(\log x)^2(\log\log x)^2} \ll \#\mathcal{E}_3(x) \ll \frac{\#\mathcal{Q}_3(x)}{(\log x)(\log\log x)^2}.$$

Proof. The upper bound follows from Lemma 2.4 and (2.14). The lower bound follows from Lemma 2.4, Theorem 2.5 and the asymptotic formula (2.13) for $\#\mathcal{Q}_3(x)$. □

3. Cyclotomic polynomials and numerical semigroups

3.1. **Numerical semigroups.** A *numerical semigroup* S is a submonoid of $(\mathbb{Z}_{\geq 0}, +)$ with finite complement in $\mathbb{Z}_{\geq 0}$. The nonnegative integers not in S are called *gaps*, and the largest gap is the *Frobenius number*, denoted $F(S)$. A numerical semigroup admits a unique and finite minimal system of generators; its cardinality is called *embedding dimension*, denoted $e(S)$, and its elements *minimal generators*. If $S = \langle a_1, \ldots, a_e\rangle$ is a minimally generated monoid (with a_i positive integers), then S is a numerical semigroup if and only if $\gcd(a_1, \ldots, a_e) = 1$ (for a comprehensive introduction to numerical semigroups, see, for example, [33]).

To a numerical semigroup S we associate $H_S(Z) = \sum_{s\in S} Z^s$, its *Hilbert series*, and $P_S(Z) = (1 - Z)H_S(Z)$, its *semigroup polynomial*. (Note that $P_S(Z)$ is indeed a polynomial, of degree $F(S) + 1$, since all elements larger than $F(S)$ are in S; for the same reason, $H_S(Z)$ is not a polynomial.)

Solely by the definition of H_S and P_S, the following is easy to establish, see [9]:

Lemma 3.1. *Let S be a numerical semigroup and assume that $P_S(Z) = a_0 + a_1Z + \cdots + a_kZ^k$. Then for $s \in \{0, \ldots, k\}$*

$$a_s = \begin{cases} 1 & \text{if } s \in S \text{ and } s - 1 \notin S, \\ -1 & \text{if } s \notin S \text{ and } s - 1 \in S, \\ 0 & \text{otherwise.} \end{cases}$$

It is a folklore result (see, for instance [30]) that, if $S = \langle a, b\rangle$ is a numerical semigroup (that is, $\gcd(a, b) = 1$), then

$$P_S(Z) = \frac{(1 - Z)(1 - Z^{ab})}{(1 - Z^a)(1 - Z^b)}. \tag{3.1}$$

Suppose that $S = \langle p, q\rangle$ with p, q distinct primes. Then by (1.2)

$$P_S(Z) = \Phi_{pq}(Z). \tag{3.2}$$

3.2. **Applications to Φ_{pq}.** The identity (3.2) and Lemma 3.1 allow one to prove results concerning the coefficients of Φ_{pq} by studying certain properties of the numerical semigroup $\langle p, q\rangle$.

Before going further, we give some definitions. Let S be a numerical semigroup.

We say that $D = \{n+1, \ldots, n+k\} \subset \mathbb{Z}_{\geq 0}$ is a *k-gapblock* (or a *gapblock* of size k) if $S \cap (D \cup \{n, n+k+1\}) = \{n, n+k+1\}$. (In the literature this is also named a *k-desert*).

Similarly, $E = \{n+1, \ldots, n+k\} \subset \mathbb{Z}_{\geq 0}$ is a *k-elementblock* (or an *elementblock* of size k) if $S \cap (E \cup \{n, n+k+1\}) = E$.

A key fact is that a numerical semigroup of embedding dimension 2, hence $S = \langle p, q\rangle$ in particular, is *symmetric*, that is, $S \cup (F(S) - S) = \mathbb{Z}$, where $F(S) - S = \{F(S) - s \ : \ s \in S\}$ (see, for example, [30] and [33]). This implies that there is a one to one correspondence between k-gapblock and k-elementblocks in S, which is relevant for our later purposes.

In what follows, the notation $\{0\}_m$ is used to indicate a string $\underbrace{0, \ldots, 0}_{m}$ of m consecutive zeros.

Theorem 3.2. *Let $p < q$ be primes. Then*

(i) *$g(\Phi_{pq}) = p-1$ and the number of maximum gaps equals $2\lfloor q/p \rfloor$;*
(ii) *$\Phi_{pq}(Z)$ contains the sequence of consecutive coefficients of the form $\pm 1, \{0\}_m, \mp 1$ for all $m = 0, 1, \ldots, p-2$ if and only if $q \equiv \pm 1 \pmod{p}$.*

Proof. We prove parts (i) and (ii) separately.

(i). The fact that $g(\Phi_{pq}) = p - 1$ was already proven in [30], so let us only deal with the second claim. Set $q = pk + h$, with $k = \lfloor q/p \rfloor$ and $1 \leq h \leq p-1$. Consider $S = \langle p, q\rangle$ and note that, by Lemma 3.1, a maximum gap in Φ_{pq} corresponds to both a $(p-1)$-gapblock and a $(p-1)$-elementblock in S. Let us see what the elements of S are. First, we have $kp \in S$, for all $k \in \mathbb{Z}_{\geq 0}$. Next, the smallest element in S is $q = pk + h$, which lies in the interval $(pk, p(k+1))$. It is clear now that we have precisely $\lfloor q/p \rfloor$ gapblocks of size $(p-1)$ in S, namely $\{jp+1, \ldots, jp+p-1\}$, for $j = 0, 1, \ldots, k-1$. Since S is symmetric, to any k-gapblock in there corresponds a k-elementblock and conversely. Thus, there are $\lfloor q/p \rfloor$ gapblocks and $\lfloor q/p \rfloor$ elementblocks of size $p-1$ in S, leading to $2\lfloor q/p \rfloor$ maximum gaps in Φ_{pq}.

(ii). Let $S = \langle p, q\rangle$. By (3.2), we have $P_S = \Phi_{pq}$. Note that, in light of Lemma 3.1, the sequence $1, \{0\}_m, -1$ of coefficients of Φ_{pq} corresponds to an $(m+1)$-gapblock in S (similarly, the sequence $-1, \{0\}_m, 1$ corresponds to an $(m+1)$-elementblock in S) and conversely. Since Φ_{pq} is self-reciprocal or equivalently, $S = \langle p, q\rangle$ is symmetric, it is enough to consider only the sequences $1, \{0\}_m, -1$. Therefore, proving part (ii) is equivalent to showing the following:

$$S = \langle p, q\rangle \text{ has gapblocks of sizes } 1, 2, \ldots, p-1 \text{ if and only if } q \equiv \pm 1 \pmod{p}.$$

To prove this claim, we consider both implications separately. First, assume $q \equiv \pm 1 \pmod{p}$. We only treat the case $q \equiv 1 \pmod{p}$, as the other can be dealt with in a similar manner. Let $q = pk + 1$ for some $k \in \mathbb{Z}_{>0}$. Then, for $1 \le m \le p-1$, the intervals $\mathcal{I}_m = [mpk, \ldots, mpk + p)$ are pairwise disjoint. Next, note that if $a, b \in \mathbb{Z}_{\ge 0}$ are so that $mpk \le ap + bq < mpk + p$, then $b \le m$. Conversely, for any such b there is a unique $a \in \mathbb{Z}_{\ge 0}$ such that $mpk \le ap + bq < mpk + p$, since for a fixed $0 \le b \le m$, exactly one number from $\{ap + bq \ : \ a \in \mathbb{Z}_{\ge 0}\}$ lands in $\mathcal{I}_m$. We can write any number

$$mpk + h = (m - h)kp + hq, \quad \text{for } h = 0, 1, \ldots, m,$$

in the form $ap + bq$, with $0 \le b \le m$ and, by the above, no other element of $\mathcal{I}_m$ can be written in this way. Therefore

$$\mathcal{I}_m \cap S = [mpk, mpk + 1, \ldots, mpk + m]$$

and $\{mpk + m + 1, \ldots, mpk + p - 1\}$ is a $(p - m)$-gapblock of S, for $m = 1, 2, \ldots, p - 1$.

By considering the intervals $\mathcal{I}_m = (mpk - p, \ldots, mpk]$ we can give a similar argument in case $q \equiv -1 \pmod{p}$, on taking $k = (q + 1)/p$.

Conversely, since the intervals $\mathcal{J}_k = [pk, \ldots, p(k+1))$ where $k \in \mathbb{Z}_{\ge 0}$, form a partition of $\mathbb{Z}_{\ge 0}$, q must lie in some $\mathcal{J}_k$, $k \ge 1$ (as $p < q$) and $q \ne pk$. We claim that

$$q = pk + 1 \quad \text{or} \quad q = p(k + 1) - 1. \tag{3.3}$$

Assume otherwise. Note that $\mathcal{J}_k \cap S = \{pk, q\}$, because if $s = ap + bq \in \mathcal{J}_k \cap S$ and $s \ne pk$, then $b \ge 1$. If $b \ge 2$, then $s > p + q > p(k + 1)$, and $s \notin \mathcal{J}_k$. Hence, $b = 1$. Next, $s \ne q$ implies $a \ge 1$ and then again $s \ge p + q > p(k + 1)$. Therefore we obtain two gapblocks $\{pk + 1, \ldots, q - 1\}$ and $\{q + 1, \ldots, p(k + 1) - 1\}$, of size at most $p-3$. But now, as $\mathcal{J}_h \cap S = \{ph, p(h + 1) - 1\}$, for all $0 \le h \le k-1$, we have only $(p-1)$-gapblocks before $\mathcal{J}_k$. Also, $q+lp \in \mathcal{J}_{k+l} \cap S$ is different from $p(k+l)+1$ and $p(k+l+1)-1$, hence there can be no $(p-2)$-gapblock

in $\mathcal{J}_{k+l}$, for $l \geq 1$. But then S has no $(p-2)$-gapblock, contradiction. Hence, (3.3) is true and yields $q \equiv \pm 1 \pmod{p}$. □

Let $N_2(x)$ be the number of cyclotomic polynomials Φ_n with $n = pq \leq x$ and $q \equiv \pm 1 \pmod{p}$. By the Dirichlet Theorem on primes in arithmetic progressions, we see that $N_2(x) \to \infty$ as $x \to \infty$. We can make this more precise.

Corollary 3.3. *We have*

$$N_2(x) = C\frac{x}{\log x} + O\left(\frac{x}{(\log x)^2}\right),$$

where

$$C = \frac{1}{2} + \sum_{p \geq 3} \frac{2}{p(p-1)} = -\frac{1}{2} + 2\sum_{k=2}^{\infty} \log \zeta(k) \frac{(\varphi(k) - \mu(k))}{k}.$$

Numerically

$$C = 1.046313338099590255728734919711884 7\ldots.$$

Proof. For a real $z \geq 1$ and integers $k > a \geq 1$, let, as usual, $\pi(z, k, a)$ denote the number of primes $q \leq z$ in the arithmetic progression $q \equiv a \pmod{k}$. We have

$$N_2(x) = \pi\left(\frac{x}{2}\right) + \sum_{3 \leq p \leq x} \left(\pi\left(\frac{x}{p}, p, 1\right) + \pi\left(\frac{x}{p}, p, -1\right)\right). \tag{3.4}$$

We now choose

$$y = \log x \qquad \text{and} \qquad Y = x^{1/3}$$

and for $p < y$, we use the Siegel-Walfisz theorem (see [25, Corollary 5.29]) in the form

$$\pi(z, k, a) = \frac{\pi(z)}{\varphi(k)} + O\left(\frac{z}{\log z)^3}\right).$$

For primes $p \in [y, Y]$, we use the Brun-Titchmarsh theorem (see [25, Theorem 6.6]) in the form

$$\pi(z, k, a) \ll \frac{z}{\varphi(k)\log(2z/k)},$$

which holds for any positive integer $k < z$. Finally, for $p > Y$, we use the trivial bound

$$\pi(z, k, a) \ll \frac{z}{k}.$$

So, collecting the above estimate and bounds we obtain

$$N_2(x) = \pi(x/2) + 2\sum_{3\le p<y}\left(\frac{\pi(x/p)}{p-1} + O\left(\frac{x}{p(\log x)^3}\right)\right) + \sum_{y\le p\le Y}\frac{x}{p^2\log x} + \sum_{Y<p\le x}\frac{x}{p^2},$$

which yields

$$N_2(x) = \pi(x/2) + 2\sum_{3\le p<y}\frac{\pi(x/p)}{p-1} + O\left(\frac{x}{(\log x)^2}\right). \tag{3.5}$$

Now, using (1.4), we derive

$$\begin{aligned}\pi(x/p) &= \frac{x}{p(\log x - \log p)} + O\left(\frac{x}{p(\log x)^2}\right)\\ &= \frac{x}{p\log x} + O\left(\frac{x\log p}{p(\log x)^2}\right).\end{aligned}$$

Hence,

$$\begin{aligned}\sum_{3\le p<y}\frac{\pi(x/p)}{p-1} &= \frac{x}{\log x}\sum_{3\le p<y}\left(\frac{1}{p(p-1)} + O\left(\frac{\log p}{p^2\log x}\right)\right)\\ &= \frac{x}{\log x}\sum_{p\ge 3}\frac{1}{p(p-1)} + O\left(\frac{x}{(\log x)^2}\left(\sum_{p\ge 3}\frac{\log p}{p^2}\right)\right)\\ &= \frac{x}{\log x}\sum_{p\ge 3}\frac{1}{p(p-1)} + O\left(\frac{x}{(\log x)^2}\right).\end{aligned}$$

Substituting this bound into (3.5), we obtain the desired result with C equal to the prime sum given. This sum is not suitable for obtaining C with high numerical precision. In order to achieve that, we express C in terms of values of the zeta-values at integer arguments (which can be approximated with enormous numerical precision very fast). Using the well-known formulas

$$\sum_{p\ge 2}\frac{1}{p^s} = \sum_{n=1}^{\infty}\frac{\mu(n)}{n}\log\zeta(ns) \text{ and } \frac{\varphi(n)}{n} = \sum_{d\mid n}\frac{\mu(d)}{d},$$

we obtain

$$\sum_{p\geq 2}\frac{1}{p(p-1)}=\sum_{m=2}^{\infty}\sum_{p\geq 2}\frac{1}{p^k}=\sum_{m=2}^{\infty}\sum_{n=1}^{\infty}\frac{\mu(n)}{n}\log\zeta(mn)$$
$$=\sum_{k=2}^{\infty}\log\zeta(k)\sum_{n\cdot m=k,\ m\geq 2}\frac{\mu(n)}{n}$$
$$=\sum_{k=2}^{\infty}\log\zeta(k)\frac{(\varphi(k)-\mu(k))}{k}.$$

On noting that the prime sum above equals $C/2+1/4$ the proof of the second equality for C is completed. This identity allows one to evaluate C with hundreds of decimals of precision. □

Remark 3.4. *If we are interested in the number of binary integers $n=pq\leq x$ with $q\equiv\pm 1 \pmod p$, then the above estimate holds with C replaced by $C-1/2$.*

Remark 3.5. *For more details on expressing prime products and prime sums (such as C) in terms of zeta-values, see, for example,* [11, pp. 208-210] *and* [27, 28].

3.3. **A small generalization.** The attentive reader might have noticed that in our proofs of Theorem 3.2 we did not use that p and q are prime, but only that they are coprime. We explore now what happens if we consider $S=\langle a,b\rangle$ with $2\leq a<b$ and a and b coprime. The coprime condition ensures that S is a numerical semigroup. Our earlier proofs then apply to

$$P_S(Z)=\frac{(1-Z)(1-Z^{ab})}{(1-Z^a)(1-Z^b)}=\prod_{d\mid ab,\, d\nmid a, d\nmid b}\Phi_d(Z),$$

where we used (3.1) and (1.1). The polynomial $P_{\langle a,b\rangle}(Z)$ is a *binary inclusion-exclusion polynomial*, see [1].

Theorem 3.6. *Let $2\leq a<b$ be coprime positive integers. Then*

(i) *the maximum gap in $\prod_{d\mid ab,\ d\nmid a,d\nmid b}\Phi_d(Z)$ equals $a-1$ and it occurs precisely $2\lfloor b/a\rfloor$ times;*
(ii) *the polynomial in* (i) *contains the sequence of consecutive coefficients $\pm 1,\{0\}_m,\mp 1$ for all $m=0,1,\ldots,a-2$ if and only if $b\equiv\pm 1 \pmod a$.*

The result that the maximum gap in $\prod_{d\mid ab,\ d\nmid a,d\nmid b}\Phi_d(Z)$ equals $a-1$ already appears in Moree [30].

4. Bound on jumps

4.1. **Definitions and background.** It has been shown by Gallot and Moree [17] that consecutive coefficients of ternary cyclotomic polynomials differ by at most one. Bzdęga [6] gave a very different reproof of this result and initiated the study of the number of "jumps". More precisely, since the sequence of the coefficients $a_n(k)$ is symmetric (1.3), the number of jumps up with $a_n(k) = a_n(k-1)+1$ is the same as the number of jumps down with $a_n(k) = a_n(k-1)-1$. We denote this common number by J_n.

Bzdęga [6] has shown that

$$J_n > n^{1/3}$$

for any ternary cyclotomic polynomial Φ_n and conditionally, on a certain variant of the prime 3-tuple conjecture, that

$$J_n < 15n^{1/3} \tag{4.1}$$

for infinitely many n. As a special case he showed that if $m, 6m-1$ and $12m-1$ are all prime for some $m \geq 7$, then (4.1) holds with $n = m(6m-1)(12m-1)$. Here, we prove an unconditional variant of the upper bound (4.1).

Theorem 4.1. *For infinitely many $n = pqr$ with pairwise distinct odd primes p, q and r, we have*

$$J_n \ll n^{7/8+o(1)}.$$

4.2. **Preparations.** For two integers k and m with $\gcd(k,m) = 1$ we denote by k_m^* the unique integer defined by the conditions

$$kk_m^* \equiv 1 \pmod{m} \qquad \text{and} \qquad 1 \leq k_m^* < m.$$

Our approach depends on showing that for almost all primes p there are many primes ℓ that are not too large and such that ℓ_p^* belongs to a certain interval.

It is natural that we study this problem using bounds of exponential sums involving reciprocals of primes [2, 16, 24].

For a prime p, an integer a and a real number L we define (recall that ℓ and p always denote prime numbers) the double exponential sum

$$S(a; L, P) = \sum_{\ell \sim L} \sum_{p \sim P} \exp(2\pi i a \ell_p^*/p),$$

and formulate a special case of a result of Duke, Friedlander and Iwaniec [14, Theorem 1],

Lemma 4.2. *Let L and P be arbitrary real positive numbers and a an integer satisfying $1 \le |a| \le LP$. Then*

$$|S(a; L, P)| \le \left(L^{1/2} + P^{1/2} + \min\{L, P\}\right)(LP)^{1/2+o(1)}.$$

We now state the *Erdős–Turán inequality* (see [12, 26]), which links the distributional properties of sequences with exponential sums.

To make this precise, for a sequence of N real numbers $\Gamma = (\gamma_n)_{n=1}^N$ of the half-open interval $[0, 1)$, we denote by Δ_Γ its *discrepancy*, that is,

$$\Delta_\Gamma = \sup_{0 \le \alpha \le 1} |T_\Gamma(\alpha) - \alpha N|,$$

where $T_\Gamma(\alpha)$ is the number of points of the sequence Γ in the interval $[0, \alpha]$.

Lemma 4.3. *For any integer $H \ge 1$, the discrepancy Δ_Γ of a sequence $\Gamma = (\gamma_n)_{n=1}^N$ of N real numbers $\gamma_1, \dots, \gamma_N \in [0, 1)$ satisfies the inequality*

$$\Delta_\Gamma \ll \frac{N}{H} + \sum_{a=1}^{H} \frac{1}{a} \left| \sum_{n=1}^{N} \exp(2\pi i a \gamma_n) \right|.$$

Let $D_p(L, P)$ be the discrepancy of the sequence of fractions

$$\frac{\ell_p^*}{p}, \qquad \ell \sim L, \ p \sim P.$$

Combining Lemma 4.2 with Lemma 4.3 (applied with the parameter $H = L$), we obtain

$$D_p(L, P) \ll \frac{(\pi(2L) - \pi(L))(\pi(2P) - \pi(P))}{L} + \left(L^{1/2} + P^{1/2} + \min\{L, P\}\right)(LP)^{1/2+o(1)},$$

and hence the following corollary.

Corollary 4.4. *Let $\delta > 0$ be arbitrary. For any real positive L and P with $L \ge P^\delta$ we have*

$$D_p(L, P) \le \left(L^{1/2} + P^{1/2} + \min\{L, P\}\right)(LP)^{1/2+o(1)},$$

where the implied constant may depend on δ.

4.3. **Proof of Theorem 4.1.**

Proof. We always assume that $L \le P$. Thus, the bound of Corollary 4.4 simplifies as

$$D_p(L, P) \le L^{1/2} P^{1+o(1)} + L^{3/2} P^{1/2+o(1)}.$$

We see from Corollary 4.4 that for any L and P and positive integer $h \le P$ there are

$$\begin{aligned} W(h,P,L) &= h\frac{(\pi(2L)-\pi(L))(\pi(2P)-\pi(P))}{P} \\ &\quad + O\left(L^{1/2}P^{1+o(1)} + L^{3/2}P^{1/2+o(1)}\right) \\ &= (1+o(1))\frac{hL}{\log L \log P} \\ &\quad + O\left(L^{1/2}P^{1+o(1)} + L^{3/2}P^{1/2+o(1)}\right) \end{aligned} \tag{4.2}$$

pairs (ℓ, p) of primes $\ell \sim L$, $p \sim P$ with $\ell_p^* \in [1, hp/P]$.

We now set $L = \lceil P^\rho \rceil$ for some positive constant $\rho < 1$, fix some $\varepsilon > 0$ (sufficiently small compared to ρ) and set

$$h = L^{-1/2}P^{1+\varepsilon} + L^{1/2}P^{1/2+\varepsilon}. \tag{4.3}$$

In particular, we assume that $\varepsilon < \min\{\rho/2, (1-\rho)/2\}$ so that for a sufficiently large P we have $h < P$.

We see from (4.2) that with this choice of h we have

$$W(h,P,L) = (1+o(1))\frac{hL}{\log L \log P} > \pi(2P) - \pi(P).$$

Hence, there is a prime number $p \sim P$ such that for at least two pairs (q,p) and (r,p) counted by $W(h,L,P)$ we have $q_p^*, r_p^* \in [1, hp/P]$.

We define the quantities $\alpha_p(q,r)$ and $\beta_p(q,r)$ as the smallest and the second smallest element in the set $\{q_p^*, p-q_p^*, r_p^*, p-r_p^*\}$. Hence, the above choice of q and r implies that

$$\alpha_p(q,r), \beta_p(q,r) \ll h. \tag{4.4}$$

Now writing

$$qq_p^* = 1 + kp$$

for some integer k, we see that

$$p_q^* = q - k = q - (qq_p^* - 1)/p = q + O(qh/p) = q + O(Lh/P)$$

and similarly

$$p_r^* = r + O(Lh/P).$$

Defining $\alpha_q(p,r)$, $\beta_q(p,r)$, $\alpha_r(p,q)$ and $\beta_r(p,q)$ in full analogy with $\alpha_p(q,r)$ and $\beta_p(q,r)$, we conclude that

$$\alpha_q(p,r), \beta_q(p,r), \alpha_r(p,q), \beta_r(p,q) \ll Lh/P. \tag{4.5}$$

We now see from (4.4) and (4.5) that in the notations of the proof of [6, Theorem 4.1], we have

$$R \ll hL^2, \quad S \ll L^2h^3/P^2, \quad T \ll L^2h^2/P.$$

Hence, we derive

$$R+S+T \ll hL^2 + L^2h^3/P^2 + L^2h^2/P \ll hL^2.$$

Now, by [6, Theorem 4.1], using (4.4) and (4.5), after simple calculations, we derive

$$J_n \ll hL^2 \ll n(h/P). \tag{4.6}$$

Recalling that $L = P^{\rho} + O(1)$, for a sufficiently large P we obtain $n \asymp P^{1+2\rho}$. Since $\varepsilon > 0$ can be arbitrary small, the bound (4.6) implies that

$$J_n \ll n^{1-\vartheta+o(1)},$$

where, recalling (4.3), we have $(1+2\rho)\vartheta = \min\{\rho/2, 1/2-\rho/2\}$. Choosing $\rho = 1/2$ we obtain $\vartheta = 1/8$ and the result follows. □

5. Recapitulation of the results obtained

Here we give a short summary of the results obtained, which are mainly motivated by, and related to, those from [6] and [22].

The paper [22] on maximum gaps in (inverse) cyclotomic polynomials:

- We provide an asymptotic for the quantity $\#\mathcal{R}_3(x)$ (Theorem 2.2), cf. Corollary 2.3. This asymptotic makes Remark 1 in [22] that $\#\mathcal{R}_3(x)$ is small quantitative.
- The right order of magnitude for the number of $n = pqr \le x$ with $p - 1 > 2n/p - \deg\Psi_n$ is provided (Theorem 2.5).
- Estimates for the number of $n = pqr \le x$ with $g(\Psi_n) \ne 2n/p - \deg\Psi_n$ are provided (Theorem 2.6).
- The gap problem for binary cyclotomic polynomials is reformulated in terms of numerical semigroups (Section 3).
- Using this approach the gap structure is studied more in detail (Theorem 3.2).
- An easy reproof of $g(\Phi_{pq}) = p - 1$ due to Nathan Kaplan is given (Section 2.1).
- The Φ_{pq} with $q \equiv \pm 1 \pmod p$ are shown to be special (Theorem 3.2) and their number for $pq \le x$ quantified (Corollary 3.3).
- The gap structure for binary inclusion-exclusion polynomials is considered (Section 3.3).

The paper [6] on jumps of cyclotomic polynomials:

- We use bounds of double Kloosterman sums over primes to show that $J_n \ll n^{7/8+o(1)}$ for infinitely many ternary n (Theorem 4.1).

Acknowledgements

The authors would like to thank Nathan Kaplan for his permission to present his elegant proof of (2.1) (communicated by e-mail to the third author). Further, Alessandro Languasco and Alessandro Zaccagnini for e-mail correspondence regarding the numerical evaluation of the constant C that appears in Corollary 3.3.

The first author worked on cyclotomic coefficient gaps while carrying an internship at the Max Planck Institute for Mathematics in August 2010. The second author worked on this paper at the Max Planck Institute for Mathematics during a one-week visit in January 2014, the third author in April 2015 and the fifth author during a visit July-December 2013. All authors gratefully acknowledge the support, the hospitality and the excellent conditions for collaboration at the Max Planck Institute for Mathematics. The project was continued whilst the authors were working at other institutions. The fifth author was also supported in part by the ARC Grants DP130100237 and DPDP140100118.

References

[1] G. Bachman, 'On ternary inclusion-exclusion polynomials', *Integers*, **10** (2010) A48 623–638.
[2] R.C. Baker, 'Kloosterman sums with prime variable', *Acta Arith.*, **156** (2012), 351–372.
[3] B. Bzdęga, 'Bounds on ternary cyclotomic coefficients', *Acta Arith.*, **144** (2010), 5–16.
[4] B. Bzdęga, 'Sparse binary cyclotomic polynomials', *J. Number Theory*, **132** (2012), 410–413.
[5] B. Bzdęga, 'On the height of cyclotomic polynomials', *Acta Arith.*, **152** (2012), 349–359.
[6] B. Bzdęga, 'Jumps of ternary cyclotomic coefficients', *Acta Arith.*, **163** (2014), 203–213.
[7] B. Bzdęga, 'On a certain family of inverse ternary cyclotomic polynomials', *J. Number Theory* **141** (2014), 1–12.
[8] L. Carlitz, The number of terms in the cyclotomic polynomial $F_{pq}(x)$, *Amer. Math. Monthly*, **73** (1966), 979–981.
[9] A. Ciolan, P.A. García-Sánchez and P. Moree, 'Cyclotomic numerical semigroups', *Preprint*, 2014 (see http://arxiv.org/abs/1409.5614).
[10] C. Cobeli, Y. Gallot, P. Moree and A. Zaharescu, 'Sister Beiter and Kloosterman: A tale of cyclotomic coefficients and modular inverses', *Indag. Math.*, **24** (2013), 915–929.
[11] H. Cohen, Number theory. Vol. II. Analytic and modern tools. Graduate Texts in Mathematics **240**, Springer, New York, 2007.
[12] M. Drmota and R.F. Tichy, *Sequences, discrepancies and applications*, Springer-Verlag, Berlin, 1997.
[13] W. Duke, J.B. Friedlander and H. Iwaniec, 'Equidistribution of roots of a quadratic congruence to prime moduli', *Ann. Math.*, **141** (1995), 423–441.

[14] W. Duke, J. B. Friedlander and H. Iwaniec, 'Bilinear forms with Kloosterman fractions', *Invent. Math. J.*, **128** (1997), 23–43.
[15] É. Fouvry, 'On binary cyclotomic polynomials', *Algebra Number Theory*, **7** (2013), 1207–1223.
[16] É. Fouvry and I.E. Shparlinski, 'On a ternary quadratic form over primes', *Acta Arith.*, **150** (2011), 285–314.
[17] Y. Gallot and P. Moree, 'Neighboring ternary cyclotomic coefficients differ by at most one', *J. Ramanujan Math. Soc.*, **24** (2009), 235–248.
[18] Y. Gallot and P. Moree, 'Ternary cyclotomic polynomials having a large coefficient', *J. Reine Angew. Math.*, **632** (2009), 105–125.
[19] Y. Gallot, P. Moree and R. Wilms, 'The family of ternary cyclotomic polynomials with one free prime', *Involve*, **4** (2011), 317–341.
[20] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, London, 1974.
[21] H. Hong, E. Lee and H.-S. Lee, 'Explicit formula for optimal ate pairing over cyclotomic family of elliptic curves', *Finite Fields Appl.*, **34** (2015), 45–74.
[22] H. Hong, E. Lee, H.-S. Lee and C.-M. Park, 'Maximum gap in (inverse) cyclotomic polynomial', *J. Number Theory*, **132** (2012), 2297–2315.
[23] H. Hong, E. Lee, H.-S. Lee and C.-M. Park, 'Simple and exact formula for minimum loop length in Ate$_i$ pairing based on Brezing-Weng curves', *Des. Codes Cryptogr.*, **67** (2013), 271–292.
[24] A.J. Irving, 'Average bounds for Kloosterman sums over primes', *Funct. Approx. Comment. Math.*, **51** (2014), 221–235.
[25] H. Iwaniec and E. Kowalski, *Analytic number theory*, American Mathematical Society, Providence, RI, 2004.
[26] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience, New York-London-Sydney, 1974.
[27] M. Mazur and B.V. Petrenko, 'Representations of analytic functions as infinite products and their application to numerical computations', *Ramanujan J.* **34** (2014), 129–141.
[28] P. Moree, 'Approximation of singular series and automata', *Manuscr. Math.*, **101** (2000), 385–399.
[29] P. Moree, 'Inverse cyclotomic polynomials', *J. Number Theory*, **129** (2009), 667–680.
[30] P. Moree, 'Numerical semigroups, cyclotomic polynomials and Bernoulli numbers', *Amer. Math. Monthly* **121** (2014), 890–902.
[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] M.B. Nathanson, Additive number theory. The classical bases, Graduate Texts in Mathematics **164**, Springer-Verlag, New York, 1996.
[33] J.C. Rosales and P.A. García-Sánchez, 'Numerical semigroups', *Developments in Mathematics*, **20**, Springer, New York, 2009.
[34] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge University Press, 1995.
[35] J. Zhao and X. Zhang, 'Coefficients of ternary cyclotomic polynomials', *J. Number Theory*, **130** (2010), 2223–2237.

Department of Computer Science, University of Oxford, Oxford OX1 3QD, United Kingdom
E-mail address: oana-maria.camburu@cs.ox.ac.uk

Rheinische Friedrich-Wilhelms-Universität Bonn, Regina-Pacis-Weg 3, D-53113 Bonn, Germany
E-mail address: ciolan@uni-bonn.de

School of Mathematics, University of the Witwatersrand, Private Bag X3, Wits 2050, South Africa
E-mail address: florian.luca@wits.ac.za

Max-Planck-Institut für Mathematik, Vivatsgasse 7, D-53111 Bonn, Germany
E-mail address: moree@mpim-bonn.mpg.de

Department of Pure Mathematics, University of New South Wales, Sydney, NSW 2052, Australia
E-mail address: igor.shparlinski@unsw.edu.au