ACTA ARITHMETICA
143.3 (2010)

On a polynomial conjecture of Pál Turán

by

Pradipto Banerjee and Michael Filaseta (Columbia, SC)

1. Introduction. P. Turán (see [11]) posed the problem of showing that there is an absolute constant C such that, if

$$f(x) = \sum_{j=0}^{r} a_j x^j \in \mathbb{Z}[x],$$

then there is a $w(x) = \sum_{j=0}^{r} b_j x^j \in \mathbb{Z}[x]$ with $\sum_{j=0}^{r} |b_j| \le C$ such that $f(x) + w(x)$ is irreducible over $\mathbb{Q}$. The problem remains open, though it has been verified by A. Bérczes and L. Hajdu [1, 2], with $C = 4$, for all $f(x)$ with $r \le 24$. If we allow for the possibility that $\deg w > r$, then the problem was resolved in general by A. Schinzel in [13]. As a consequence of his work, we have

Theorem 1 (Schinzel, 1970). *For every $f(x) = \sum_{j=0}^{r} a_j x^j \in \mathbb{Z}[x]$, there exist infinitely many polynomials $w(x) = \sum_{j=0}^{s} b_j x^j \in \mathbb{Z}[x]$ with $\sum_{j=0}^{s} |b_j| \le 3$ and $f(x) + w(x)$ irreducible. One of these is such that*

$$s < \exp((5r+7)(\|f\|^2 + 3)),$$

where $\|f\| = \sqrt{\sum_{j=0}^{r} a_j^2}$.

In this paper, we establish a version of Schinzel's theorem where the dependence on the degree of $f(x)$ in the bound for s is improved from exponential to linear. The dependence on $\|f\|^2$ however remains exponential. Specifically, we prove the following theorem:

Theorem 2. *For every $f(x) = \sum_{j=0}^{r} a_j x^j \in \mathbb{Z}[x]$, there exist infinitely many polynomials $w(x) = \sum_{j=0}^{s} b_j x^j \in \mathbb{Z}[x]$ such that $\sum_{j=0}^{s} |b_j| \le 3$ and $f(x) + w(x)$ is irreducible. One of these is such that*

$$s \le 8 \max\{r+3, n_0\} \exp((\log 5)(8\|f\|^2 + 9)),$$

where n_0 is an effectively computable absolute constant.

2010 *Mathematics Subject Classification*: Primary 11C08; Secondary 11R09, 12E05.
Key words and phrases: irreducible polynomial, Turán's problem.

DOI: 10.4064/aa143-3-4

Similarly to Schinzel's work, we consider

$$w(x) = x^m - x^n \quad \text{or} \quad w(x) = x^m - x^n \pm 1,$$

where

$$m \in (M, 2M], \qquad n \in (N, 2N].$$

To obtain our results, we take

$$N \geq \max\{r+3, n_0\} \quad \text{and} \quad M = 4N\exp((\log 5)(8\|f\|^2+9)). \tag{1}$$

The improvement will follow largely due to estimates given in [6], which can be viewed as replacing the role of [12] in [13].

The role of n_0 is simply to allow us to take N and M large in our arguments. We will not concern ourselves with justifying that n_0 is effectively computable though that it is will be clear from our arguments.

We note that the estimates in [6] give considerably more easily a similar result to Theorem 2 if one is willing to settle for 5 as an upper bound on $\sum_{j=0}^{s}|b_j|$. Getting the bound 3 is somewhat more involved. The bound of 3 may be best possible depending on whether a certain covering system exists, which is associated with a long outstanding problem posed by P. Erdős and J. Selfridge (see [5] and [11]).

2. Preliminaries. Define

$$g(x) = \begin{cases} f(x) & \text{if } f(0) \neq 0 \neq f(1), \\ f(x)-1 & \text{if } f(0)=0,\ f(1)\neq 1 \text{ or } f(0)\neq 1,\ f(1)=0, \\ f(x)+1 & \text{if } f(0)=0,\ f(1)=1 \text{ or } f(0)=1,\ f(1)=0. \end{cases}$$

This choice of $g(x)$ implies that 0 and 1 are not roots of

$$G(x) = G_{m,n}(x) = x^m - x^n + g(x).$$

We prove Theorem 2 by showing that for sufficiently large values of M and N as in (1), there exists a pair $(m,n) \in (M,2M]\times(N,2N]$ such that $G_{m,n}(x)$ is irreducible. We consider the pairs $(m,n) \in (M,2M]\times(N,2N]$ where $m-n$ is prime. Since $M \geq 4N$ implies that $p = m-n > M-2N \geq 2N$ so that $p > 2N \geq n$, we easily deduce that $\gcd(m,n) = 1$ with m and n as indicated.

If a polynomial $h(x) \in \mathbb{C}[x]$ satisfies $h(x) = \pm x^{\deg h}h(1/x)$, we refer to it as *reciprocal*. The terminology is partly due to the fact that $h(x)$ is reciprocal if and only if for every root α of $h(x)$, we have $\alpha \neq 0$ and $1/\alpha$ is also a root of $h(x)$. Every $H(x) \in \mathbb{Z}[x]$ can be factored uniquely in $\mathbb{Z}[x]$ as $H_1(x)H_2(x)$ where $H_1(x)$ has a positive leading coefficient, every irreducible factor of $H_1(x)$ is reciprocal, every irreducible factor of $H_2(x)$ is not reciprocal, and the content (that is, the greatest common divisor of the coefficients) of $H_2(x)$ is 1. We refer to $H_2(x)$ as the *non-reciprocal part of* $H(x)$. We use $N(H(x))$ to denote the non-reciprocal part of $H(x)$ (so $N(H(x)) = H_2(x)$).

The following is an immediate consequence of a theorem of K. Ford, S. Konyagin and the second author [6]. Recall that, in the introduction, we noted this result will replace the use of a similar theorem of Schinzel in [12] for establishing Theorem 1.

LEMMA 1. *If* $m \geq 2n\exp((\log 5)(8\|f\|^2+9))$, *then* $N(G_{m,n})$ *is irreducible unless at least one of the following holds:*

(i) $x^n - g(x)$ *is a* p*th power for some prime* $p \mid m$.
(ii) $-x^n + g(x)$ *is* 4 *times a* 4*th power in* $\mathbb{Z}[x]$ *and* $4 \mid m$.

Recall that $n > N > r = \deg g$. Since $\gcd(m,n) = 1$, we deduce that $x^n - g(x)$ is not a pth power for any prime p dividing m. Also, the leading coefficient of $-x^n + g(x)$ is not divisible by 4. We deduce (i) and (ii) above do not hold, and hence the non-reciprocal part of $G_{m,n}(x)$ is irreducible provided m and n satisfy the inequality stated in the lemma.

For sufficiently large values of M depending on N, we now see that $N(G_{m,n})$ is irreducible for $(m,n) \in (M,2M] \times (N,2N]$. We will show that for M and N sufficiently large, there exists at least one pair $(m,n) \in (M,2M]\times(N,2N]$ such that $G_{m,n}(x)$ has no irreducible reciprocal factor. As a consequence, we deduce that for sufficiently large values of M and N, there exists a pair $(m,n) \in (M,2M] \times (N,2N]$ such that $G_{m,n}(x) = N(G_{m,n})$ and hence $G_{m,n}(x)$ is irreducible.

Most of the remainder of our arguments concerns obtaining results about irreducible reciprocal polynomials that can divide $G(x) = x^m - x^n + g(x)$ where $(m,n) \in (M,2M] \times (N,2N]$. As we will see momentarily, it will be helpful to separate discussions about cyclotomic factors of $G(x)$, which are necessarily reciprocal, from information on reciprocal irreducible non-cyclotomic factors of $G(x)$. We begin, however, by addressing both types of reciprocal factors.

LEMMA 2. *Fix a positive integer* $n \in (N,2N]$. *There is a non-zero polynomial* $V(x) = V_n(x)$ *of degree* $\leq 4N$ *such that for every* $m \in (M,2M]$, *each reciprocal divisor of* $x^m - x^n + g(x)$ *is also a divisor of* $V(x)$.

Proof. For a non-zero polynomial $F(x)$, we set $\widetilde{F}(x) = x^{\deg F}F(1/x)$. Observe that if a reciprocal polynomial divides $F(x)$, then it also divides $\widetilde{F}(x)$. Let $v(x) = x^n - g(x)$ and $G(x) = x^m - x^n + g(x)$. Suppose a reciprocal polynomial $h(x)$ divides $G(x)$. Then $h(x)$ also divides $\widetilde{G}(x)$. Thus, $h(x)$ divides the polynomial

$$\begin{aligned}\widetilde{v}(x)G(x) + x^n\widetilde{G}(x) &= \widetilde{v}(x)(x^m - v(x)) + x^{n+m}\left(\frac{1}{x^m} - v\left(\frac{1}{x}\right)\right)\\ &= \widetilde{v}(x)x^m - v(x)\widetilde{v}(x) + x^n - x^m\widetilde{v}(x) = -v(x)\widetilde{v}(x) + x^n.\end{aligned}$$

This last expression is non-zero; in fact, the coefficient of x^n in $v(x)\widetilde{v}(x)$ is $\|v\|^2 \geq 2$ so that the coefficient of x^n in $-v(x)\widetilde{v}(x) + x^n$ is non-zero. As the degree of $-v(x)\widetilde{v}(x) + x^n$ is $\leq 2n \leq 4N$, the lemma follows by taking this polynomial to be $V(x)$. ∎

Fix $n \in (N, 2N]$. Recalling that $M \geq 4N$, a simple application of the Prime Number Theorem implies that the number of $m \in (M, 2M]$ with $m - n$ prime is asymptotic to $M/\log M$. More precisely, given $\varepsilon > 0$, if N is sufficiently large, $n \in (N, 2N]$ and $M \geq 4N$, then the number of primes in $(M - n, 2M - n]$ is in the interval

$$((1 - \varepsilon)M/\log M, (1 + \varepsilon)M/\log M).$$

We deduce that the number of pairs $(m, n) \in (M, 2M] \times (N, 2N]$ with $m - n$ prime is asymptotic to $MN/\log M$. Thus, it is enough to show that for sufficiently large values of M and N satisfying (1), the number of pairs $(m, n) \in (M, 2M] \times (N, 2N]$ for which $m - n$ is prime and $G_{m,n}(x)$ has a reciprocal irreducible factor is less than $\rho MN/\log M$ for some constant $\rho < 1$.

Now, we consider pairs (m, n) where $G_{m,n}(x)$ has an irreducible reciprocal non-cyclotomic factor. The following lemma gives a bound for the number of such pairs.

LEMMA 3. *The number of pairs $(m, n) \in (M, 2M] \times (N, 2N]$ such that $G_{m,n}(x)$ has a reciprocal irreducible non-cyclotomic factor is $\leq N \log^3 N$.*

Proof. Let $h(x)$ be a reciprocal irreducible non-cyclotomic polynomial in $\mathbb{Z}[x]$. We begin by showing that $h(x)$ divides at most one polynomial $G_{m,n}(x)$. Assume there are two pairs (m, n) and (m', n') of positive integers, with $(m, n) \neq (m', n')$, such that $h(x)$ divides both $G_{m,n}(x)$ and $G_{m',n'}(x)$. Then $h(x)$ also divides

$$G_{m,n}(x) - G_{m',n'}(x) = x^m - x^n - x^{m'} + x^{n'}.$$

As $h(x)$ is reciprocal, $h(x) \neq x$. We deduce that there are integers a, b and c with $a > b > c > 0$ and numbers $\varepsilon_j \in \{1, -1\}$ for $1 \leq j \leq 3$ such that $h(x)$ divides

$$F(x) = x^a + \varepsilon_1 x^b + \varepsilon_2 x^c + \varepsilon_3.$$

As $h(x)$ is reciprocal, it also divides

$$\widetilde{F}(x) = \varepsilon_3 x^a + \varepsilon_2 x^{a-c} + \varepsilon_1 x^{a-b} + 1.$$

Thus, $h(x)$ is a factor of the polynomial

$$\begin{aligned} F(x) - \varepsilon_3 \widetilde{F}(x) &= \varepsilon_1 x^b + \varepsilon_2 x^c - \varepsilon_2\varepsilon_3 x^{a-c} - \varepsilon_1\varepsilon_3 x^{a-b} \\ &= (\varepsilon_1 x^c - \varepsilon_2\varepsilon_3 x^{a-b})(x^{b-c} + \varepsilon_1\varepsilon_2). \end{aligned}$$

If neither of the factors on the right is identically 0, then we obtain a contradiction since $h(x)$ is non-cyclotomic. Therefore, either $a = b + c$ and $\varepsilon_1\varepsilon_2\varepsilon_3 = 1$, or $b = c$ and $\varepsilon_1\varepsilon_2 = -1$. In the first of these two cases,

$$F(x) = x^{b+c} + \varepsilon_1 x^b + \varepsilon_2 x^c + \varepsilon_1\varepsilon_2 = (x^b + \varepsilon_2)(x^c + \varepsilon_1);$$

in the second, $F(x) = x^a + \varepsilon_3$. In either case, we again have a contradiction to $h(x)$ being non-cylcotomic. Hence, $h(x)$ divides at most one $G_{m,n}(x)$.

If m and n are positive integers, now shown to be unique, such that $h(x)$ divides $G_{m,n}(x)$, then Lemma 2 implies that $h(x)$ divides a polynomial $V_n(x)$ of degree $\le 4N$. Since $h(x)$ is a divisor of $G_{m,n}(x)$, we may and do suppose that $h(x)$ is monic. Observe that as $n \in (N, 2N]$ varies, there are N polynomials $V_n(x)$. To obtain our lemma, it therefore suffices to show that $d = \deg h(x)$ necessarily satisfies $d \ge N/\log^3 N$, as then, for each $n \in (N, 2N]$, there are at most $4\log^3 N$ possibilities for $h(x)$.

We establish that $d \ge N/\log^3 N$ by appealing to a result of E. Dobrowolski [4] on the maximum size of the absolute value of a root of a non-cyclotomic polynomial. Dobrowolski's result can be considered as an improvement on a classical result of L. Kronecker [8] that a monic irreducible non-cyclotomic polynomial in $\mathbb{Z}[x]$ must have a root with modulus > 1. Observe that the sum of the absolute values of the coefficients of $x^n - g(x)$ is bounded above by $\|f\|^2 + 2$. Hence, if α is a root of $G_{m,n}(x)$ with $|\alpha| > 1$, then

$$|\alpha|^m = |\alpha^m - G_{m,n}(\alpha)| = |\alpha^n - g(\alpha)| \le |\alpha|^n(\|f\|^2 + 2).$$

From (1),

$$m - n \ge m - 2N \ge 2N\exp((\log 5)(8\|f\|^2 + 9)) > N(\|f\|^2 + 2).$$

The maximum value of $x^{1/x}$ for $x > 0$ is $e^{1/e} < 2$. Hence, we deduce that

$$|\alpha| < 2^{1/N}. \tag{2}$$

We use that N is large. As $h(x)$ divides $G_{m,n}(x)$, the roots of $h(x)$ are roots of $G_{m,n}(x)$. We now take α to be the root of $h(x)$ with $|\alpha|$ maximal. Recall that $h(x)$ is monic. Observe that for any D, there are finitely many monic non-cyclotomic polynomials of degree $\le D$ having each root < 2 in absolute value. By Kronecker's theorem mentioned above, each of these has a root exceeding 1 in absolute value. Since N is large, with D fixed, we may suppose that the roots for each of these polynomials are not all $< 2^{1/N}$. Hence, we may consider d large (larger than any fixed number). Dobrowolski's result then implies that

$$|\alpha| > 1 + \frac{1}{d}\left(\frac{\log\log d}{\log d}\right)^3.$$

Combining this with (2) yields $d \ge N/\log^3 N$, and therefore the lemma follows. ∎

The significance of the above lemma is clear. We want to show that there are $< \rho MN/\log M$ pairs $(m,n) \in (M,2M] \times (N,2N]$ for which $G_{m,n}(x)$ has a reciprocal factor. The above lemma indicates that there are $\leq N\log^3 N$ such pairs when we restrict to irreducible reciprocal factors that are not cyclotomic. Recall that Theorem 2 has an n_0 in the bound for s, allowing us to consider M large. Since $N\log^3 N$ is very small compared to $MN/\log M$ for $M > N$ large, the contribution of pairs (m,n) for which $G_{m,n}(x)$ has an irreducible reciprocal non-cyclotomic factor is insignificant.

We turn now to estimating pairs (m,n) for which $G_{m,n}(x)$ has a cyclotomic factor. As usual, we denote the lth cyclotomic polynomial by $\Phi_l(x)$. Recall $\deg \Phi_l = \phi(l)$ where ϕ is Euler's ϕ-function. By Lemma 2, one has $\phi(l) \leq 4N$. Schinzel and the second author [7] have shown that if a polynomial in $\mathbb{Z}[x]$ has a cyclotomic factor, then it also has a cyclotomic factor $\Phi_l(x)$ such that every prime divisor of l is less than or equal to the number of terms of the polynomial. Recall $N \geq \deg f + 3$. Thus, if $G_{m,n}(x)$ is divisible by a cyclotomic polynomial for some $(m,n) \in (M,2M] \times (N,2N]$, then it is divisible by some $\Phi_l(x)$ with l having each prime factor $\leq N$ and such that $\phi(l) \leq 4N$. The condition $\phi(l) \leq 4N$ by itself implies

$$l \leq c_1 N \log\log N \tag{3}$$

for some constant c_1. In fact, [10, formula (3.42)] easily implies one may take $c_1 = 8$ for N large. Hence, estimating the number of $G_{m,n}(x)$, with $(m,n) \in (M,2M] \times (N,2N]$, that are divisible by a cyclotomic polynomial corresponds to estimating the number of such $G_{m,n}(x)$ divisible by some $\Phi_l(x)$ with each prime factor of l being $\leq N$ and with (3) holding. We now consider only such l.

Next, we show that if $\Phi_l(x)$ divides $G_{m,n}(x)$, then $\Phi_l(x)$ does not divide $g(x)$. Observe that if $\Phi_l(x)$ divides both $G_{m,n}(x)$ and $g(x)$, then it divides $x^m - x^n$ and, hence, $x^{m-n} - 1$. Recall $G_{m,n}(1) \neq 0$. Also, $m-n$ is a prime. Hence, we deduce $l = m - n$. But $M \geq 4N$ and $n \leq 2N$ imply that l is a prime $\geq 2N$, contrary to the fact that we are only considering l that have each prime divisor $\leq N$.

Denoting the largest prime factor of l by $P(l)$, we are left with estimating the size of the set

$$\begin{aligned}\mathcal{S} = \{&(m,n) \in (M,2M] \times (N,2N] : m-n \text{ prime}, \exists\, l \in \mathbb{Z} \text{ satisfying} \\ &l \geq 2,\ l \leq c_1 N\log\log N,\ P(l) \leq N,\ \Phi_l(x) \,|\, G_{m,n}(x) \text{ and } \Phi_l(x) \nmid g(x)\}.\end{aligned}$$

Recall that M and N are large and the specific conditions on M and N imply $m - n \geq 2N$ above so that, in particular, $m-n$ is a large prime. Also, given the previous paragraph, when we discuss l with $\Phi_l(x) \,|\, G_{m,n}(x)$, the condition $\Phi_l(x) \nmid g(x)$ necessarily follows. We are interested in establishing

that

$$|\mathcal{S}| < \rho MN/\log M \quad \text{for some fixed } \rho < 1.$$

The following lemma will provide us with an upper bound for the number of pairs (m,n) modulo l such that $G_{m,n}(x)$ is divisible by $\Phi_l(x)$ for a given l as above. Note that if $\Phi_l(x)$ divides both $G_{m,n}(x)$ and $G_{m',n'}(x)$, then

$$\zeta_l^m - \zeta_l^n = \zeta_l^{m'} - \zeta_l^{n'} \neq 0 \quad \text{where } \zeta_l = e^{2\pi i/l}.$$

The result of the lemma can be found in [13] (see also the main result in [14] and Theorem 7 of [3]); we include the details to keep the paper more self-contained.

LEMMA 4. *Fix integers a and b and a positive integer l with $\zeta_l^a - \zeta_l^b \neq 0$. If l is odd or $l = 2$, then m and n are positive integers such that $\zeta_l^m - \zeta_l^n = \zeta_l^a - \zeta_l^b$ if and only if $m \equiv a \pmod{l}$ and $n \equiv b \pmod{l}$. If l is even and $l > 2$, then m and n are positive integers such that $\zeta_l^m - \zeta_l^n = \zeta_l^a - \zeta_l^b$ if and only if either $m \equiv a \pmod{l}$ and $n \equiv b \pmod{l}$, or $m \equiv b + l/2 \pmod{l}$ and $n \equiv a + l/2 \pmod{l}$.*

Proof. One can verify directly that the "if" parts of the statements in Lemma 4 hold. We concern ourselves therefore simply with the "only if" parts of the statements.

Let $\zeta = \zeta_l$ where $l > 2$. We denote the real part of a complex number z and the imaginary part of z by $\Re(z)$ and $\Im(z)$, respectively. First, we show that if u and v are real numbers with $\zeta^u - \zeta^v \neq 0$ and $\Re(\zeta^u - \zeta^v) = 0$, then ζ^u and ζ^v are complex conjugates. We have

$$\Re(\zeta^u - \zeta^v) = 0 \Rightarrow \Re(\zeta^u) = \Re(\zeta^v).$$

Since ζ^u and ζ^v are on the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ and both have the same real part, we deduce that $\Im(\zeta^u) = \pm\Im(\zeta^v)$. If $\Im(\zeta^u) = \Im(\zeta^v)$, then $\zeta^u = \zeta^v$ contradicting that $\zeta^u - \zeta^v \neq 0$. Hence, $\Im(\zeta^u) = -\Im(\zeta^v)$, establishing that ζ^u and ζ^v are complex conjugates.

Let $\alpha = (a+b)/2$. From $\zeta^m - \zeta^n = \zeta^a - \zeta^b \neq 0$, we obtain

$$\zeta^{m-\alpha} - \zeta^{n-\alpha} = \zeta^{(a-b)/2} - \zeta^{(b-a)/2} \neq 0.$$

The middle expression $\zeta^{(a-b)/2} - \zeta^{(b-a)/2}$ is a difference of two conjugates and hence has real part 0. Therefore, $\Re(\zeta^{m-\alpha} - \zeta^{n-\alpha}) = 0$. The previous paragraph implies that $\zeta^{m-\alpha}$ and $\zeta^{n-\alpha}$ are complex conjugates. Setting

$$A = \zeta^{m-\alpha} - \zeta^{n-\alpha} = \zeta^{(a-b)/2} - \zeta^{(b-a)/2},$$

we see that the two (not necessarily distinct) roots of $z^2 - Az - 1$ are, on the one hand, $\zeta^{(a-b)/2}$ and $-\zeta^{(b-a)/2}$ and, on the other hand, $\zeta^{m-\alpha}$ and $-\zeta^{n-\alpha}$. Hence, either $\zeta^{(a-b)/2} = \zeta^{m-\alpha}$ or $-\zeta^{(b-a)/2} = \zeta^{m-\alpha}$. Rewriting these, we see that either $\zeta^a = \zeta^m$ or $-\zeta^b = \zeta^m$. In the former case, we also have $\zeta^b = \zeta^n$ so that $m \equiv a \pmod{l}$ and $n \equiv b \pmod{l}$. In the latter case, $\zeta^{m-b} = -1$

and there are no solutions if l is odd. If l is even, then, since also $\zeta^{n-a} = -1$ in this case, we obtain

$$m \equiv b + l/2 \pmod{l} \quad \text{and} \quad n \equiv a + l/2 \pmod{l}.$$

The above verifies the lemma in every case except $l = 2$. In this case, we have $\zeta = -1$. Since $\zeta^a - \zeta^b \neq 0$, either $\zeta^a - \zeta^b = 2$ or $\zeta^a - \zeta^b = -2$. If $\zeta^a - \zeta^b = 2$, then $\zeta^m - \zeta^n = 2$ so that $m \equiv 0 \pmod 2$ and $n \equiv 1 \pmod 2$. Thus, there is only one solution in m and n modulo 2. A similar analysis works in the case $\zeta^a - \zeta^b = -2$, completing the proof. ∎

We note that in the case that l is even and > 2, similar to the case $l = 2$ above, it is possible that there is only one pair (m, n) modulo l such that $\zeta_l^m - \zeta_l^n = \zeta_l^a - \zeta_l^b$ in Lemma 4. This occurs precisely when $a \equiv b + l/2 \pmod l$. The argument for $l = 2$ above is simply justifying that this congruence must hold in this case.

Before proceeding, we show that Lemma 4 implies that we can reduce our consideration of l satisfying (3) to l satisfying a stronger inequality. For example, we show that we can take

$$l \leq N^{3/4}, \tag{4}$$

where the particular bound on the right is not so important but our arguments later will depend on having $l \ll N^\theta$ for some $\theta < 1$. The same estimate from [10] that led to (3) implies even more directly that

$$\phi(l) \geq \frac{l}{1.8 \log\log l} \quad \text{for } l \text{ sufficiently large.}$$

In particular, for N large and $l > N^{3/4}$ satisfying (3), we have

$$\phi(l) \geq \frac{l}{2 \log\log N} > \frac{N^{3/4}}{2 \log\log N}.$$

We deduce that there can be at most $8N^{1/4} \log\log N$ different $\Phi_l(x)$ with $l > N^{3/4}$ that can divide a polynomial of degree $\leq 4N$. By Lemma 2, there are at most $8N^{5/4} \log\log N$ different cyclotomic polynomials $\Phi_l(x)$ with $l > N^{3/4}$ that can divide a $G_{m,n}(x)$ with $(m, n) \in (M, 2M] \times (N, 2N]$. By Lemma 4, the total number of such pairs (m, n) for which such a $\Phi_l(x)$ divides $G_{m,n}(x)$ is bounded by

$$8N^{5/4} \log\log N \cdot 2\left(\left\lfloor \frac{N}{N^{3/4}} \right\rfloor + 1\right) \cdot \left(\left\lfloor \frac{M}{N^{3/4}} \right\rfloor + 1\right) < MN^{0.8},$$

for N large. We wish to compare the above bound to $MN/\log M$, but we do not necessarily know that $\log M$ is small compared to a power of N. So we apply the above estimate only in the case that $M \leq N^2$. For such M, we easily deduce that the total number of pairs $(m, n) \in (M, 2M] \times (N, 2N]$

for which $\Phi_l(x)$ divides $G_{m,n}(x)$ for some $l > N^{3/4}$ is small compared to $MN/\log M$.

For $M > N^2$, we do a similar estimate but also use a version of the Brun–Titchmarsh inequality obtained by H. Montgomery and R. Vaughan [9] (though a weaker version would suffice). For each fixed $l > N^{3/4}$ that satisfies also (3), we appeal to Lemma 4 to deduce that for a fixed $n \in (N, 2N]$ with N large, there are at most

$$\frac{4M}{\phi(l)\log(2M/l)} \le 4M \cdot \frac{2\log\log N}{N^{3/4}} \cdot \frac{3}{\log M} \le \frac{M}{N^{0.7}\log M}$$

integers $m \in (M, 2M]$ such that $m - n$ is prime and $\Phi_l(x)$ divides $G_{m,n}(x)$. Therefore, as n and l vary, we deduce that the total number of such pairs (m, n) for which $G_{m,n}(x)$ is divisible by some $\Phi_l(x)$ with $l > N^{3/4}$ is

$$\le 8N^{5/4}\log\log N \cdot \frac{M}{N^{0.7}\log M} \cdot 2\left(\left\lfloor \frac{N}{N^{3/4}} \right\rfloor + 1\right) < \frac{MN^{0.9}}{\log M}.$$

For $M > N^2$, we deduce again that the total number of pairs $(m, n) \in (M, 2M] \times (N, 2N]$ for which $\Phi_l(x)$ divides $G_{m,n}(x)$ for some $l > N^{3/4}$ is small compared to $MN/\log M$. Hence, it suffices to show that the number of $(m, n) \in (M, 2M] \times (N, 2N]$ for which $G_{m,n}(x)$ is divisible by some $\Phi_l(x)$ with l satisfying (4) and with $m - n$ prime is $< \rho MN/\log M$ for some $\rho < 1$.

Suppose $\Phi_l(x)$ divides some $G_{a,b}(x)$ with $0 \le a, b \le l - 1$. We count the number of $(m, n) \in (M, 2M] \times (N, 2N]$ with $m - n$ prime, $m \equiv a \pmod l$ and $n \equiv b \pmod l$. Observe that for each such (m, n), $\Phi_l(x)$ divides $G_{m,n}(x)$. Furthermore, by Lemma 4, each l corresponds to either one or possibly two different (a, b) depending on whether l is odd or 2, or l is an even number > 2.

Fix $n \in (N, 2N]$ with $n \equiv b \pmod l$. The number of m in $(M, 2M]$ such that $m \equiv a \pmod l$ and $m - n$ is prime is equal to the number of primes p in $(M - n, 2M - n]$ such that $p \equiv a - b \pmod l$. We consider two possibilities depending on whether l is small or large. Fix L arbitrarily; in the end, we will choose $L = 20000$. For $l \le L$, by an asymptotic form of Dirichlet's theorem, we deduce that the number of m in $(M, 2M]$ as above is either asymptotic to $M/(\phi(l)\log M)$ or, in the case that $\gcd(a - b, l) \ne 1$, identically 0 (for M large). For $l > L$, we appeal again to the Brun–Titchmarsh inequality as established in [9] to deduce that the number of such m is bounded by $2M/(\phi(l)\log(M/l))$. From (4), $M/l \ge M/N^{3/4} \ge M^{1/4}$. Since there are at most $\lfloor N/l \rfloor + 1$ different $n \equiv b \pmod l$, we deduce that for every $\varepsilon > 0$ and for N and, hence, M sufficiently large (depending on ε), the number of $(m, n) \in (M, 2M] \times (N, 2N]$ with $m - n$ prime, $m \equiv a \pmod l$ and $n \equiv b \pmod l$ is bounded above by

$$B_l = B_l(M,N,\varepsilon) = \begin{cases} \dfrac{(1+\varepsilon)MN}{l\phi(l)\log M} + \dfrac{(1+\varepsilon)M}{\phi(l)\log M} & \text{if } l \le L, \\ \dfrac{8(1+\varepsilon)MN}{l\phi(l)\log M} + \dfrac{8(1+\varepsilon)M}{\phi(l)\log M} & \text{if } l > L. \end{cases}$$

Recall that we want to show that the total number of pairs $(m,n) \in (M,2M] \times (N,2N]$ that arise with $G_{m,n}(x)$ having a cyclotomic factor is $< \rho MN/\log M$ for some constant $\rho < 1$. To bound the total number of such pairs, we sum the various bounds B_l above, recalling however that we need to double the bound above in the case of even $l > 2$ since in this case $\Phi_l(x)$ might divide $G_{a,b}(x)$ for two different pairs (a,b) with $0 \le a,b \le l-1$. Recalling (4), it suffices to show that

$$(5) \qquad \sum_{l=2}^{L} \frac{1}{l\phi(l)} + \sum_{l=2}^{L/2} \frac{1}{2l\phi(2l)} + \sum_{L<l\le N^{3/4}} \frac{16}{l\phi(l)} + \sum_{2\le l\le N^{3/4}} \frac{16}{N\phi(l)} < 0.99.$$

Note that the first three sums above correspond to estimates based on the first expressions appearing in the bounds given by B_l and the last sum above incorporates all estimates based on the second remaining expressions appearing in the bounds given by B_l. For convenience, we have made some over-estimates here that will not affect our discussion.

The naive idea of simply taking L sufficiently large, computing the first two sums in (5) directly and estimating the remaining sums fails. One can easily check that the sum of $1/(l\phi(l))$ by itself over say $2 \le l \le 1000$ exceeds 1.2. Sieving could be used to take into account over-counting of pairs (m,n) divisible by more than one $\Phi_l(x)$, and the authors initially approached the problem with this in mind. But we found in the end that taking advantage of extra information on pairs (m,n) for which $G_{m,n}(x)$ can be divisible by $\Phi_l(x)$ for small values of l was sufficient for improving on this naive idea enough to obtain the results we wanted. So we turn to obtaining such extra information to complete the proof with the goal of revising the first two sums in (5) to account for some over-counting that is currently done there.

3. Refining the estimates. Although we will revise (5) before obtaining our results, it is convenient here to note that the last two sums in this inequality can be handled without difficulty. We will make use of the same estimates later for our revised form of (5).

We begin with the last sum in (5). For $X > 0$, we have

$$\sum_{X<l\le 2X} \frac{1}{\phi(l)} = \sum_{X<l\le 2X} \frac{1}{l\prod_{p\mid l}(1-1/p)} = \sum_{X<l\le 2X} \frac{\prod_{p\mid l}(1+1/p)}{l\prod_{p\mid l}(1-1/p^2)}$$

$$\le \zeta(2) \sum_{X<l\le 2X} \frac{1}{l} \prod_{p|l} \Big(1+\frac{1}{p}\Big) = \zeta(2) \sum_{X<l\le 2X} \frac{1}{l} \sum_{d|l} \frac{\mu^2(d)}{d}$$

$$\le \frac{\zeta(2)}{X} \sum_{X<l\le 2X} \sum_{d|l} \frac{\mu^2(d)}{d} = \frac{\zeta(2)}{X} \sum_{d\le 2X} \sum_{\substack{X<l\le 2X \\ d|l}} \frac{\mu^2(d)}{d}$$

$$\le \frac{\zeta(2)}{X} \sum_{d\le 2X} \frac{2X}{d} \cdot \frac{\mu^2(d)}{d} \le 2\zeta(2) \sum_{d=1}^{\infty} \frac{\mu^2(d)}{d^2}$$

$$= 2\zeta(2) \prod_{p} \Big(1+\frac{1}{p^2}\Big) = \frac{2\zeta(2)^2}{\zeta(4)} = 5.$$

We take in turn $X = 1, 2, 4, 8, \dots, 2^t$, where $t \in \mathbb{Z}$ is such that

$$2^t < N^{3/4} \quad \text{and} \quad 2^{t+1} \ge N^{3/4}.$$

Summing over these various values of X, we find for sufficiently large N that

$$\sum_{2\le l\le N^{3/4}} \frac{16}{N\phi(l)} \le \frac{80(t+1)}{N} \le \frac{90\log N}{N} < 0.001. \tag{6}$$

The third sum in (5) clearly depends on L. Observe that

$$\sum_{d|l} \frac{\mu^2(d)}{\phi(d)} = \prod_{p|l} \Big(1+\frac{1}{\phi(p)}\Big) = \prod_{p|l} \Big(1+\frac{1}{p-1}\Big) = \frac{1}{\prod_{p|l}(1-1/p)} = \frac{l}{\phi(l)}.$$

Hence, for $X > 0$,

$$\sum_{X<l\le 2X} \frac{1}{l\phi(l)} = \sum_{X<l\le 2X} \frac{l}{l^2\phi(l)} \le \frac{1}{X^2} \sum_{X<l\le 2X} \frac{l}{\phi(l)}$$

$$= \frac{1}{X^2} \sum_{X<l\le 2X} \sum_{d|l} \frac{\mu^2(d)}{\phi(d)} \le \frac{1}{X^2} \sum_{d\le 2X} \sum_{\substack{l\le 2X \\ d|l}} \frac{\mu^2(d)}{\phi(d)}$$

$$\le \frac{1}{X^2} \sum_{d\le 2X} \frac{2X}{d} \cdot \frac{\mu^2(d)}{\phi(d)} \le \frac{2}{X} \sum_{d=1}^{\infty} \frac{\mu^2(d)}{d\phi(d)}$$

$$= \frac{2}{X} \prod_{p} \Big(1+\frac{1}{p(p-1)}\Big) = \frac{2}{X} \cdot \frac{\zeta(2)\zeta(3)}{\zeta(6)},$$

where the last equation follows by using the Euler product formulation of $\zeta(s)$. Taking $X = L, 2L, 4L, \dots$ and summing over X gives

$$\sum_{l>L} \frac{1}{l\phi(l)} \le \frac{2\zeta(2)\zeta(3)}{\zeta(6)} \Big(\frac{1}{L} + \frac{1}{2L} + \frac{1}{4L} + \cdots\Big) = \frac{4\zeta(2)\zeta(3)}{\zeta(6)L}.$$

We deduce that

$$\sum_{L<l\le N^{3/4}} \frac{16}{l\phi(l)} < \frac{64\zeta(2)\zeta(3)}{\zeta(6)L} < \frac{125}{L}. \tag{7}$$

We are now ready to attend to revising (5), in particular to adjust the first two sums there to give us the desired result. Recall that we are considering pairs (m,n) for which $m-n$ is a prime $\ge M-2N \ge 2N$ and, hence, $m-n$ is odd.

LEMMA 5. *Let $l \ge 2$ be an integer and p a prime satisfying $p \ge 3$ if $l \in \{2,4\}$, and $p \ge 5$ otherwise. Suppose $\Phi_l(x)$ divides $G_{m,n}(x)$ for some positive integers m and n with $m-n$ odd. If $\Phi_{pl}(x)$ divides $G_{m',n'}(x)$ for some positive integers m' and n' with $m'-n'$ odd, then $\Phi_l(x)$ also divides $G_{m',n'}(x)$.*

Proof. Since $\Phi_l(x)$ divides $x^m - x^n + g(x)$ and $\Phi_{pl}(x)$ divides $x^{m'} - x^{n'} + g(x)$, we deduce that

$$x^{m'} - x^{n'} - x^m + x^n = \Phi_{pl}(x)u(x) - \Phi_l(x)v(x)$$

for some $u(x)$ and $v(x)$ in $\mathbb{Z}[x]$. Setting $x = \zeta_l$, we obtain

$$\zeta_l^{m'} - \zeta_l^{n'} - \zeta_l^m + \zeta_l^n = \Phi_{pl}(\zeta_l)u(\zeta_l). \tag{8}$$

Observe that if the right-hand side is 0, then ζ_l is a root of $G_{m',n'}(x)$. Hence, in this case, $\Phi_l(x)$ also divides $G_{m',n'}(x)$, which is what we want to show. So suppose the right-hand side of (8) is non-zero. The second author in [5, Lemma 3] showed that if a, b and t are positive integers and p is a prime for which $a = p^t b$, then $p \,|\, \Phi_a(\zeta_b)$ in $\mathbb{Z}[\zeta_b]$. Taking $a = pl$ and $b = l$ we deduce that $p \,|\, \Phi_{pl}(\zeta_l)$ in $\mathbb{Z}[\zeta_l]$. Taking norms above, we deduce that $p^{\phi(l)}$ divides the rational integer

$$N_{\mathbb{Q}(\zeta_l)/\mathbb{Q}}(\zeta_l^{m'} - \zeta_l^{n'} - \zeta_l^m + \zeta_l^n).$$

We deduce that

$$p^{\phi(l)} \le \prod_{\substack{1\le j\le l-1\\ \gcd(j,l)=1}} |\zeta_l^{jm'} - \zeta_l^{jn'} - \zeta_l^{jm} + \zeta_l^{jn}| \le 4^{\phi(l)},$$

where we have used the fact that the product is non-zero since the right-hand side of (8) is non-zero. This finishes the proof except for the two cases that $l = 2$ and $l = 4$, both with $p = 3$.

For these remaining cases, we recall that we are considering $m-n$ and $m'-n'$ odd. For $l = 2$, the condition $\zeta_l^{m'} - \zeta_l^{n'} - \zeta_l^m + \zeta_l^n \ne 0$ implies that

$$\zeta_l^{m'} - \zeta_l^{n'} - \zeta_l^m + \zeta_l^n = \pm 4,$$

which is clearly not divisible by 3 in $\mathbb{Z}[\zeta_l] = \mathbb{Z}$. For $l = 4$, one similarly

deduces that

$$\zeta_l^{m'} - \zeta_l^{n'} - \zeta_l^m + \zeta_l^n \in \{\pm 2, \pm 2i, \pm 2(1 \pm i)\},$$

and we easily deduce here (for example, by taking norms) that $\zeta_l^{m'} - \zeta_l^{n'} - \zeta_l^m + \zeta_l^n$ is not a multiple of 3 in $\mathbb{Z}[\zeta_l] = \mathbb{Z}[i]$. Hence, the proof is complete. ∎

We note that it is possible to extend the last part of our arguments to strengthen the lemma, but this will not be needed to obtain Theorem 2. Set

$$\mathcal{T} = \{l \in \mathbb{Z}^+ : \exists m, n \in \mathbb{Z}^+ \text{ such that } \gcd(m-n, 6) = 1 \\ \text{and } l \text{ is the minimal positive integer such that } \Phi_l(x) \mid G_{m,n}(x)\}.$$

Thus, $\mathcal{T}$ consists of those positive integers l for which there are positive integers m and n with $\gcd(m-n, 6) = 1$ and satisfying $\Phi_l(x) \mid G_{m,n}(x)$ and $\Phi_{l'}(x) \nmid G_{m,n}(x)$ for every positive integer $l' < l$. The condition $\gcd(m-n, 6) = 1$ will be used later in establishing Lemma 7, but we note again here that we are interested in the case where $(m, n) \in (M, 2M] \times (N, 2N]$ and $m - n \geq M - 2N \geq 2N$ is prime. As observed already, we only need to consider N large. Hence, for all such pairs (m, n) the condition $\gcd(m-n, 6) = 1$ follows.

The definition of $\mathcal{T}$ implies that we can restrict our attention to $l \in \mathcal{T}$ in (5) so that all we need to show is the inequality

$$\sum_{\substack{2 \leq l \leq L \\ l \in \mathcal{T}}} \frac{1}{l\phi(l)} + \sum_{\substack{2 \leq l \leq L/2 \\ l \in \mathcal{T}}} \frac{1}{2l\phi(2l)} + \sum_{L < l \leq N^{3/4}} \frac{16}{l\phi(l)} + \sum_{2 \leq l \leq N^{3/4}} \frac{16}{N\phi(l)} < 0.99. \tag{9}$$

From Lemma 5, we find that if $l \geq 2$ and $l \in \mathcal{T}$, then $pl \notin \mathcal{T}$ for all primes $p \geq 5$. Further, if $2 \in \mathcal{T}$, then $6 \notin \mathcal{T}$; and if $4 \in \mathcal{T}$, then $12 \notin \mathcal{T}$.

Next, we give another improvement that allows us to further reduce the size of the second sum in (9). Recall that the second sum is present since for even $l > 2$, Lemma 4 only allows us to conclude that there are at most two different pairs (a, b) with $0 \leq a, b \leq l - 1$ such that $\Phi_l(x)$ divides $G_{a,b}(x)$. We show that some of these pairs are already accounted for in the case that $\Phi_2(x)$ divides some $G_{m,n}(x)$. More precisely, we prove the following:

LEMMA 6. *Suppose $2 \in \mathcal{T}$. Let $l = 4k$ where k is a positive integer. Suppose a and b are integers in $[0, l-1]$ for which $\Phi_l(x)$ divides $G_{a,b}(x)$. Then either $\Phi_2(x)$ divides $G_{m,n}(x)$ for any positive integers m and n with $m \equiv a \pmod{l}$, $n \equiv b \pmod{l}$ and $m - n$ odd, or $\Phi_2(x)$ divides $G_{m,n}(x)$ for any positive integers m and n with $m \equiv b + l/2 \pmod{l}$, $n \equiv a + l/2 \pmod{l}$ and $m - n$ odd.*

Proof. Since $\Phi_2(x) = x + 1$ divides some $G_{m',n'}(x)$ with $m' - n'$ odd and $G_{m',n'}(-1) = \pm 2 + g(-1)$, we deduce that $g(-1) = 2$ or -2. Note that l even implies that any m and n as in the lemma satisfy $m - n \equiv a - b \pmod{2}$.

Since also $m-n$ is odd, we may and do suppose that $a-b$ is odd. We deduce that either

$$(-1)^a-(-1)^b=2 \quad \text{and} \quad (-1)^{b+l/2}-(-1)^{a+l/2}=-2$$

or

$$(-1)^a-(-1)^b=-2 \quad \text{and} \quad (-1)^{b+l/2}-(-1)^{a+l/2}=2.$$

As $g(-1)=\pm 2$, at least one of $(m,n)=(a,b)$ and $(m,n)=(b+l/2,a+l/2)$ satisfies $G_{m,n}(-1)=0$. The lemma easily follows. ∎

The above lemma implies that if $\Phi_2(x)$ divides some $G_{m,n}(x)$, then one of the two solutions given by Lemma 4 when l is a multiple of 4 corresponds to pairs (m,n) for which $G_{m,n}(x)$ is divisible by $\Phi_2(x)$. We need not consider the contribution of such pairs (m,n) twice. In other words, we can replace (9) with showing both

$$\sum_{\substack{2\le l\le L\\ l\in\mathcal{T}}}\frac{1}{l\phi(l)}+\sum_{\substack{2\le l\le L/2\\ l\in\mathcal{T},\,2\nmid l}}\frac{1}{2l\phi(2l)}+\sum_{L<l\le N^{3/4}}\frac{16}{l\phi(l)}+\sum_{2\le l\le N^{3/4}}\frac{16}{N\phi(l)}<0.99 \tag{10}$$

and

$$\sum_{\substack{3\le l\le L\\ l\in\mathcal{T}}}\frac{1}{l\phi(l)}+\sum_{\substack{2\le l\le L/2\\ l\in\mathcal{T}}}\frac{1}{2l\phi(2l)}+\sum_{L<l\le N^{3/4}}\frac{16}{l\phi(l)}+\sum_{2\le l\le N^{3/4}}\frac{16}{N\phi(l)}<0.99, \tag{11}$$

where the first of these corresponds to the case that $\Phi_2(x)$ divides some $G_{m,n}(x)$ and the second to the case that $\Phi_2(x)$ divides no $G_{m,n}(x)$.

Our next lemma will allow us to reduce the size of the left-hand sides of (10) and (11) a little further. This lemma is the closest we get to executing our original sieve approach (or perhaps more appropriately an inclusion-exclusion approach). It basically takes into account the fact that if l and l' are relatively prime elements of $\mathcal{T}$, then there are many pairs $(m,n)\in\mathcal{S}$ for which $G_{m,n}(x)$ is divisible by both $\Phi_l(x)$ and $\Phi_{l'}(x)$. The lemma clarifies, however, that we only need to take this into account for the pairs $(l,l')\in\{(2,3),(3,4)\}$. Note that a more serious sieve approach would need to take into account further commonality of cyclotomic divisors of $G_{m,n}(x)$; more precisely, we would also want information about $(m,n)\in\mathcal{S}$ for which $G_{m,n}(x)$ is divisible by $\Phi_{l_1}(x),\dots,\Phi_{l_k}(x)$ for arbitrary positive integers $l_1,\dots,l_k$. We prove simply

LEMMA 7. *Suppose both* 2 *and* 3 *are in* $\mathcal{T}$. *Then there exist asymptotically* $MN/(12\log M)$ *distinct pairs* $(m,n)\in\mathcal{S}$ *such that* $G_{m,n}(x)$ *is divisible by both* $\Phi_2(x)$ *and* $\Phi_3(x)$. *Similarly, if both* 3 *and* 4 *are in* $\mathcal{T}$, *then there exist asymptotically either* $MN/(48\log M)$ *or* $MN/(24\log M)$ *distinct pairs* $(m,n)\in\mathcal{S}$ *such that* $G_{m,n}(x)$ *is divisible by both* $\Phi_3(x)$ *and* $\Phi_4(x)$ *depending*

on whether there are one or two pairs (a,b), with $0 \le a,b \le 3$, such that $\Phi_4(x)$ divides $G_{a,b}(x)$, respectively.

Proof. Consider the case that 2 and 3 are in $\mathcal{T}$. By Lemma 4, there are positive integer pairs (a_2,b_2) and (a_3,b_3) such that $\Phi_2(x)$ divides $G_{m,n}(x)$ if and only if $m \equiv a_2 \pmod 2$ and $n \equiv b_2 \pmod 2$, and $\Phi_3(x)$ divides $G_{m,n}(x)$ if and only if $m \equiv a_3 \pmod 3$ and $n \equiv b_3 \pmod 3$. Observe that 2 and 3 in $\mathcal{T}$ imply $\gcd(a_2-b_2,2)=1$ and $\gcd(a_3-b_3,3)=1$. By the Chinese Remainder Theorem, there is a unique integer pair (a,b) with $0 \le a,b \le 5$ such that $a \equiv a_2 \pmod 2$, $a \equiv a_3 \pmod 3$, $b \equiv b_2 \pmod 2$ and $b \equiv b_3 \pmod 3$. We deduce that both $\Phi_2(x)$ and $\Phi_3(x)$ divide $G_{m,n}(x)$ if and only if $m \equiv a \pmod 6$ and $n \equiv b \pmod 6$. Since $\gcd(a_2-b_2,2)=1$ and $\gcd(a_3-b_3,3)=1$, we deduce

$$\gcd(a-b,6)=1.$$

We are interested now in estimating the number of pairs $(m,n) \in (M,2M] \times (N,2N]$ with $m \equiv a \pmod 6$, $n \equiv b \pmod 6$ and $m-n$ prime. Fixing $n \in (N,2N]$ with $n \equiv b \pmod 6$ in one of $\lfloor N/6 \rfloor + O(1)$ ways, we count the number of primes $p \in (M-n,2M-n]$ with $p \equiv a-b \pmod 6$. Since we have $\gcd(a-b,6)=1$, the Prime Number Theorem implies that there are asymptotically $M/(2\log M)$ such primes. The first part of the lemma easily follows. The second part follows along similar lines. ∎

We are now ready to complete the proof of Theorem 2. We begin here by finishing our justification that there is an integer pair $(m,n) \in (M,2M] \times (N,2N]$ such that $G_{m,n}(x)$ is free of reciprocal factors by considering several cases. To simplify notation, we clarify that whenever p appears below, it represents a prime. We make use of (10) and (11) with $L=20000$ throughout.

Case 1: Each of 2, 3 and 4 is in $\mathcal{T}$. By Lemma 5, numbers of the form $2p$ and $4p$ for primes $p \ge 3$ are not in $\mathcal{T}$. We are interested in verifying inequality (10) except that Lemma 7 allows us to subtract $1/12$ from the left-hand side due to double counting of cases where both $\Phi_2(x)$ and $\Phi_3(x)$ divide $G_{m,n}(x)$. Recalling $L=20000$ and using (6) and (7), we see that the left-hand side of (10) minus $1/12$ is

$$\le \sum_{\substack{2 \le l \le 20000 \\ l \notin \{2p,4p\,:\,p \ge 3\}}} \frac{1}{l\phi(l)} + \sum_{\substack{2 \le l \le 10000 \\ 2 \nmid l \\ l \notin \{p\,:\,p \ge 3\}}} \frac{1}{2l\phi(2l)} + \frac{125}{20000} + 0.001 - \frac{1}{12} < 0.99.$$

Hence, we are done in this case.

Case 2: Each of 2 and 3 is in $\mathcal{T}$ but $4 \notin \mathcal{T}$. Lemma 5 implies here that the numbers of the form $2p$ for primes $p \ge 3$ are not in $\mathcal{T}$ either. Lemma 7

allows us again to subtract $1/12$ from the left-hand side of (10). Since

$$\sum_{\substack{2\le l\le 20000\\ l\notin\{2p\,:\,p\ge 2\}}}\frac{1}{l\phi(l)}+\sum_{\substack{2\le l\le 10000\\ 2\nmid l\\ l\notin\{p\,:\,p\ge 3\}}}\frac{1}{2l\phi(2l)}+\frac{125}{20000}+0.001-\frac{1}{12}<0.9,$$

this case is also settled.

Case 3: Each of 2 and 4 is in $\mathcal{T}$ but $3\notin\mathcal{T}$. Lemma 5 implies here that numbers of the form $2p$ and $4p$ for primes $p\ge 3$ are not in $\mathcal{T}$, but we cannot apply Lemma 7. On the other hand, since $3\notin\mathcal{T}$, we do not need to consider $1/(3\phi(3))=1/6$ in the first sum in (10). Hence, we can use here the same estimate as in Case 1 but with $-1/12$ replaced by $-1/6$. The estimate in Case 1 therefore completes this case.

Case 4: The number 2 is in $\mathcal{T}$ but 3 and 4 are not in $\mathcal{T}$. This case is the same as Case 2, but analogously to Case 3, we can replace $-1/12$ in the estimate in Case 2 with $-1/6$. Hence, we are done in this case.

Case 5: The numbers 2 and 4 are not in $\mathcal{T}$. In this case, we verify (11) using (6) and (7). Since

$$\sum_{\substack{3\le l\le 20000\\ l\ne 4}}\frac{1}{l\phi(l)}+\sum_{3\le l\le 10000}\frac{1}{2l\phi(2l)}+\frac{125}{20000}+0.001<0.85,$$

this case is complete.

Case 6: The number 2 is not in $\mathcal{T}$ but 3 and 4 are in $\mathcal{T}$. Lemma 5 implies numbers of the form $3p$ for primes $p\ge 5$ and numbers of the form $4p$ for primes $p\ge 3$ are not in $\mathcal{T}$. Lemma 7 implies that we can subtract at least $1/48$ from the left-hand side of (11) due to over-counting $G_{m,n}(x)$ that are divisible by both $\Phi_3(x)$ and $\Phi_4(x)$. Hence, the estimate

$$\sum_{\substack{3\le l\le 20000\\ l\notin\{3p\,:\,p\ge 5\}\\ l\notin\{4p\,:\,p\ge 3\}}}\frac{1}{l\phi(l)}+\sum_{\substack{2\le l\le 10000\\ l\notin\{2p\,:\,p\ge 3\}}}\frac{1}{2l\phi(2l)}+\frac{125}{20000}+0.001-\frac{1}{48}<0.99$$

resolves this case. (Although it is not needed, we note that this estimate is easily improved. Lemma 7 actually allows us either to subtract $1/24$ from the left-hand side of (11) or to remove the term $1/(4\phi(4))=1/8$ from the second sum.)

Case 7: The numbers 2 and 3 are not in $\mathcal{T}$ but 4 is in $\mathcal{T}$. This case is settled by the inequality

$$\sum_{\substack{4\le l\le 20000\\ l\notin\{4p\,:\,p\ge 3\}}}\frac{1}{l\phi(l)}+\sum_{\substack{2\le l\le 10000\\ l\notin\{2p\,:\,p\ge 3\}}}\frac{1}{2l\phi(2l)}+\frac{125}{20000}+0.001<0.86.$$

To finish Theorem 2, we simply apply Lemma 1. For N and M as in (1), we have justified that there exist $n \in (N, 2N]$ and $m \in (M, 2M]$ for which the non-reciprocal part of $G_{m,n}(x)$ is $G_{m,m}(x)$ itself. As noted after the statement of Lemma 1, (i) and (ii) of that lemma do not hold. Observe that

$$m \geq M = 4N \exp((\log 5)(8\|f\|^2 + 9)) \geq 2n \exp((\log 5)(8\|f\|^2 + 9)).$$

Hence, $G_{m,n}(x)$ is irreducible. By taking $w(x) = G_{m,n}(x) - f(x)$, the first part of Theorem 2 follows. Since $m \leq 2M = 8N \exp((\log 5)(8\|f\|^2 + 9))$, the second part of Theorem 2 follows by taking $N = \max\{r + 3, n_0\}$.

References

[1] A. Bérczes and L. Hajdu, *Computational experiences on the distances of polynomials to irreducible polynomials*, Math. Comp. 66 (1997), 391–398.

[2] —, —, *On a problem of P. Turán concerning irreducible polynomials*, in: Number Theory (Eger, 1996), de Gruyter, Berlin, 1998, 95–100.

[3] J. H. Conway and A. J. Jones, *Trigonometric Diophantine equations* (*On vanishing sums of roots of unity*), Acta Arith. 30 (1976), 229–240.

[4] E. Dobrowolski, *On a question of Lehmer and the number of irreducible factors of a polynomial*, ibid. 34 (1979), 391–401.

[5] M. Filaseta, *Coverings of the integers associated with an irreducibility theorem of A. Schinzel*, in: Number Theory for the Millennium, II (Urbana, IL, 2000), A K Peters, Natick, MA, 2002, 1–24.

[6] M. Filaseta, K. Ford, and S. Konyagin, *On an irreducibility theorem of A. Schinzel associated with coverings of the integers*, Illinois J. Math. 44 (2000), 633–643.

[7] M. Filaseta and A. Schinzel, *On testing the divisibility of lacunary polynomials by cyclotomic polynomials*, Math. Comp. 73 (2004), 957–965.

[8] L. Kronecker, *Zwei Sätze über Gleichungen mit ganzzahligen Koeffizienten*, J. Reine Angew. Math. 53 (1857), 173–175.

[9] H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika 20 (1973), 119–134.

[10] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. 6 (1962), 64–94.

[11] A. Schinzel, *Reducibility of polynomials and covering systems of congruences*, Acta Arith. 13 (1967), 91–101.

[12] —, *Reducibility of lacunary polynomials I*, ibid. 16 (1969), 123–159.

[13] —, *Reducibility of lacunary polynomials II*, ibid. 16 (1970), 371–392.

[14] Ł. Włodarski, *On the equation* $\cos\alpha_1 + \cos\alpha_2 + \cos\alpha_3 + \cos\alpha_4 = 0$, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. 12 (1969), 147–155

Pradipto Banerjee, Michael Filaseta
Department of Mathematics
University of South Carolina
Columbia, SC 29208, U.S.A.
E-mail: filaseta@mailbox.sc.edu

Received on 2.9.2008
and in revised form on 13.3.2009 (5789)