A DISTRIBUTION PROBLEM FOR POWERFREE VALUES OF IRREDUCIBLE POLYNOMIALS

Brian Beasley*and Michael Filaseta**

dedicated to András Sárközy on the occasion of his 60^{th} *birthday*

Abstract: To better understand the distribution of gaps between k-free numbers, Erdős posed the problem of establishing an asymptotic formula for the sum of the powers of the lengths of the gaps between k-free numbers. This paper generalizes the problem of Erdős by considering moments of gaps between positive integers m for which $f(m)$ is k-free. Here, $f(x)$ denotes an irreducible polynomial with integer coefficients with some necessary conditions imposed on it. Some results in this general setting are obtained that are analogous to those that have been obtained for the original problem of Erdős.

AMS Mathematics Subject Classification: 11N25, 11N37, 11C08, 11R09

1. Introduction

Let k be an integer ≥ 2, and let $f(x)$ be an irreducible polynomial in $\mathbb{Z}[x]$ satisfying $\gcd_{m\in\mathbb{Z}}(f(m))$ is k-free (i.e., $f(m)$ has no fixed divisors of the form p^k with p prime). Under these conditions, we expect that there are infinitely many integers m for which $f(m)$ is k-free, but this is far from being established. Set $g = \deg f$. Then Erdős [2] (improving on work of Nagel [11]) showed that there are infinitely many integers m for which $f(m)$ is k-free provided $k \geq g-1$, and later Hooley [6] provided a proof that such m have the expected asymptotic density among the set of positive integers. That infinitely many such m exist for $k \geq g-1$ is the best lower bound on k known when g is small. A better result was obtained by Nair [12] when g is large. He showed that for every $g \geq 2$ and $k \geq (\sqrt{2}-1/2)g$, there are infinitely many m for which $f(m)$ is k-free. His result

*The research for this paper constituted part of the first author's requirement for a doctoral degree at the University of South Carolina.

**The second author was supported by grants from the National Science Foundation and the National Security Agency.

can be made slightly stronger by introducing smaller degree terms in g, but the constant $\sqrt{2}-1/2$ is the best constant obtainable from his methods (cf. [12], [13], and [9]).

Let s_n denote the nth positive integer m (if it exists) for which $f(m)$ is k-free. In this paper we investigate the problem of determining for what γ one can establish

$$\sum_{s_{n+1}\le X} \big(s_{n+1}-s_n\big)^{\gamma} \sim B(\gamma,f,k)X \tag{1}$$

where $B(\gamma,f,k)$ is a constant depending on γ, f, and k. Erdős [1] introduced this problem in the case $f(x)=x$, and subsequent work has been done by Hooley [7], Filaseta [3], Graham [5], and Huxley [8]. The best result, due to Huxley [8], is that if $f(x)=x$, then (1) holds provided $0<\gamma<2k-1+2/(k+1)$. The more general problem in the setting of $f(x)$ described above does not appear yet to have been investigated.

Our main result is the following:

Theorem 1. *Let $g\ge 2$, and let $k\ge(\sqrt{2}-1/2)g$. Let*

$$\phi_1=\frac{(2s+g)(k-s)-g(g-1)}{(2s+g)(k-s)+g(2s+1)},$$

where

$$s=\begin{cases}1 & \text{if } 2\le g\le 4\\ \big[(\sqrt{2}-1)g/2\big] & \text{if } g\ge 5.\end{cases}$$

Let

$$\phi_2=\begin{cases}\dfrac{8g(g-1)}{(2k+g)^2-4} & \text{if } (\sqrt{2}-1/2)g\le k\le g\\ \dfrac{g}{2k-g+r} & \text{if } k\ge g+1,\end{cases}$$

where r is the largest positive integer such that $r(r-1)<2g$. Then $\phi_1>0$, $\phi_2>0$, and if

$$0\le\gamma<\min\left\{\frac{1}{\phi_2},1+\frac{\phi_1}{\phi_2},k\right\},$$

then

$$\sum_{s_{n+1}\le X}(s_{n+1}-s_n)^{\gamma}\sim B(\gamma,f,k)X,$$

for some constant $B(\gamma,f,k)$ depending only on γ, $f(x)$, and k.

The value of ϕ_2 above is based on gap results for powerfree values that already appear in the literature; improvements on these gap results would lead to a corresponding improvement in the theorem above. The gap problem is to find the minimum $h=h(f,X,k)$ such that for X sufficiently large, the interval $(X,X+h]$ contains an integer m with $f(m)$ k-free. Writing $h=cX^{\theta}$, we list several results for the case when $k>g$: $\theta=g/(2k-g+1)$ by Nair [13] in 1979; $\theta=g/(2k-g+2)$ for $g\ge 2$ by Huxley and Nair [9] in 1980; $\theta=g/(2k-g+r)$, where r is the greatest integer such that $r(r-1)<2g$, by Filaseta [4] in 1993. In the case when $(\sqrt{2}-1/2)g \le k \le g$, Huxley and Nair [9] obtained $\theta = 8g(g-1)/((2k+g)^2-4)$ in 1980; small improvements on this value of θ are possible, as noted in [4]. There is an obvious connection between this gap problem and the result obtained in Theorem 1. If one could take $\gamma > 1/\phi_2$ in Theorem 1 and end with the same asymptotic estimates, then a stronger gap result would hold than those currently available. In addition, if $k \ge g+1$ and $g \ge 2$, then the bound $\gamma < k$ can be viewed as essential. Indeed, if one could obtain Theorem 1 in this case with $\gamma > k$, then the result would be quite remarkable.

The comments in the previous paragraph suggest a strong connection between the size of θ in the gap problem for k-free values of irreducible polynomials and the size of γ permissible in (1). It is simple to see that if γ can be taken arbitrarily large in (1), then θ can be taken to be arbitrarily small in the gap problem. The converse is also true: if θ can be taken arbitrarily small in the gap problem, then γ can be taken to be arbitrarily large in (1) (this is a consequence of Theorem 2 below). This generalizes an observation made by Filaseta [3] in the case $f(x) = x$.

Additionally, we generalize a result of Filaseta's in [3] and show that γ may be allowed to be arbitrarily large provided that we restrict ourselves to *small* gaps. More precisely, we establish:

Theorem 2. *Let $k \ge (\sqrt{2}-1/2)g$. Given any $\gamma > 0$, there exists a $\delta = \delta(\gamma) > 0$ such that*

$$\sum_{\substack{s_{n+1} \le X \\ s_{n+1}-s_n \le X^{\delta}}} (s_{n+1}-s_n)^{\gamma} \sim B(\gamma, f, k)X,$$

for some constant $B(\gamma, f, k)$ depending only on γ, $f(x)$, and k.

To clarify, the constants $B(\gamma, f, k)$'s in the above results are the same.

To obtain the above results, we will establish and make use of the following partial generalization of a theorem of Mirsky [10].

Theorem 3. *Let $k \ge (\sqrt{2}-1/2)g$. For a fixed positive integer d, set*

$$\begin{aligned} N_d(X) = |\{m \in \mathbb{Z}^+ : m \le X-d, f(m) \text{ and } f(m+d) \text{ are } k\text{-free}, \\ f(m+1), f(m+2), \ldots, f(m+d-1) \text{ are not } k\text{-free}\}|. \end{aligned}$$

Suppose that for some positive integer j, $s_{j+1} - s_j = d$. Then there is a constant $c_d > 0$, depending on d, for which

$$N_d(X) \sim c_d X.$$

2. A Preliminary Remark

We will make use of notation and arguments similar to that used in the work of Graham [5]. We begin with

Definition 1. Given $f(x) \in \mathbb{Z}[x]$, let $s_n = s_n(f)$ be the nth positive integer m such that $f(m)$ is k-free. Let

$$L(h) = L(h, X) = |\{n \in \mathbb{Z}^+ : h < s_{n+1} - s_n \le 2h, X/2 < s_{n+1} \le X\}|.$$

Throughout, we will suppose that $f(x)$ is irreducible and that $\gcd_{m\in\mathbb{Z}}(f(m))$ is k-free. Our goal will be to show that for h sufficiently large, we have

$$L(2h) \ll \frac{X}{h^{\gamma+\varepsilon}}, \tag{2}$$

where γ is as in Theorem 1 or Theorem 2 and where $\varepsilon > 0$ is sufficiently small. Implied constants here and elsewhere may depend on γ, k, $f(x)$, and ε. We return to establishing (2) in Sections 4 and 5. In this section, we show that Theorem 3 together with (2) imply Theorem 1 and Theorem 2. We note here that for Theorem 2 we will only require (2) holding for sufficiently large $h \le X^{\delta'}$ where $\delta' > 0$ is some number (any number will do) depending on γ and ε. We return to this matter later in this section.

We begin by explaining how Theorem 1 follows from (2) for sufficiently large h and Theorem 3. Let r denote the positive integer satisfying $2^{r-1} < X \le 2^r$. By (2),

$$\sum_{h<d\le 2h} N_d(X) d^\gamma \le \sum_{j=0}^{r} L(h, X/2^j)(2h)^\gamma \ll \frac{X}{h^{\gamma+\varepsilon}}(2h)^\gamma \ll \frac{X}{h^\varepsilon}. \tag{3}$$

The above holds provided that h is sufficiently large. Let D_0 be such that (3) is satisfied for $h > D_0$. We write

$$\sum_{s_{n+1}\le X} (s_{n+1} - s_n)^\gamma = \sum_{d=1}^{\infty} N_d(X) d^\gamma = \sum_{d\le D_0} N_d(X) d^\gamma + \sum_{d>D_0} N_d(X) d^\gamma.$$

We deduce from (3) that

$$\sum_{d>D_0} N_d(X) d^\gamma = O\left(\frac{X}{D_0^\varepsilon}\right).$$

Theorem 3 implies

$$\sum_{d=1}^{\infty} N_d(X) d^\gamma = X \sum_{d\le D_0} c_d d^\gamma + o(X) + O\left(\frac{X}{D_0^\varepsilon}\right).$$

Hence,

$$\frac{1}{X}\sum_{d=1}^{\infty} N_d(X) d^\gamma = \sum_{d\le D_0} c_d d^\gamma + O\left(\frac{1}{D_0^\varepsilon}\right) + o(1).$$

Thus,

$$\limsup_{X\to\infty} \frac{1}{X}\sum_{d=1}^{\infty} N_d(X) d^\gamma = \sum_{d\le D_0} c_d d^\gamma + O\left(\frac{1}{D_0^\varepsilon}\right)$$

and

$$\liminf_{X\to\infty} \frac{1}{X}\sum_{d=1}^{\infty} N_d(X)d^\gamma = \sum_{d\le D_0} c_d d^\gamma + O\left(\frac{1}{D_0^\varepsilon}\right).$$

By letting D_0 tend to infinity, we deduce that $\lim_{X\to\infty}(1/X)\sum_{d=1}^\infty N_d(X)d^\gamma$ exists and $\sum_{d=1}^\infty c_d d^\gamma$ converges. We obtain

$$\sum_{s_{n+1}\le X} (s_{n+1}-s_n)^\gamma \sim B(\gamma,f,k)X$$

where

$$B(\gamma,f,k) = \sum_{d=1}^\infty c_d d^\gamma.$$

The argument for Theorem 2 is similar. We make use of Theorem 3 and the assumption that (2) holds for sufficiently large $h \le X^{\delta'}$ where $0 < \delta' < 1/2$. We choose $\delta = \min\{\delta'/(2\gamma+2\varepsilon), \delta'/2\}$. We consider r as before and an integer s satisfying $2^s \le X^{\delta'} < 2^{s+1}$. If $0 \le j \le s$, then we deduce

$$\left(\frac{X}{2^j}\right)^{\delta'} \ge \left(\frac{X}{2^s}\right)^{\delta'} \ge X^{(1-\delta')\delta'} \ge X^{\delta'/2} \ge X^\delta.$$

Thus, if $h \le X^\delta$, then $h \le (X/2^j)^{\delta'}$ for $0 \le j \le s$. Hence, for h sufficiently large satisfying $h \le X^\delta$, we deduce from (2) that

$$L(h, X/2^j) \ll \frac{X}{2^j h^{\gamma+\varepsilon}} \qquad \text{for } 0 \le j \le s.$$

Observe that the definition of $L(h,X)$ easily implies $L(h,X) \ll X/h$. Hence, for $s+1 \le j \le r$, we obtain

$$L(h, X/2^j) \ll \frac{X}{2^j h} \ll \frac{X}{2^s h} \ll \frac{X^{1-\delta'}}{h}.$$

If $h \le X^\delta$, then $h^{\gamma+\varepsilon} \le X^{\delta(\gamma+\varepsilon)} \le X^{\delta'/2}$. Thus, if $h \le X^\delta$, then

$$L(h, X/2^j) \ll \frac{X^{1-(\delta'/2)}}{hX^{\delta'/2}} \ll \frac{X^{1-(\delta'/2)}}{h^{\gamma+\varepsilon}} \qquad \text{for } s+1 \le j \le r.$$

Note that $r \ll \log X$. One easily deduces that (3) follows for sufficiently large $h \le X^\delta$ (indeed, something stronger holds). The rest of the argument for Theorem 2 follows along the lines of the argument given above for Theorem 1. We omit the details.

3. Estimates for $L(2h)$

As before, we consider $f(x)$ to be irreducible in $\mathbb{Z}[x]$ with $\gcd_{m\in\mathbb{Z}}(f(m))$ being k-free and with $g = \deg f$. In particular, it follows that the resultant of f and f', denoted $D = R(f, f')$, is non-zero. We fix $z = 4g$ and consider the finite set

$$\mathcal{A} = \{p : p|D \text{ or } p \le z\}.$$

We define

$$Q = \prod_{p\in\mathcal{A}} p^k.$$

We take

$$H = \frac{h\log h}{8gQ}$$

with h a sufficiently large integer. For a positive integer q, we define

$$\rho(q) = \rho(f, q) = |\{a \in \mathbb{Z}_q : f(a) \equiv 0 \pmod q\}|.$$

Observe that the condition $\gcd_{m\in\mathbb{Z}}(f(m))$ is k-free implies that $\rho(p^k) < p^k$ for all primes p.

Lemma 1. *There is an integer a such that if $\mathcal{B} = \{y \in (n, n+h] : y \equiv a \pmod Q\}$, then*

$$\sum_{p\le H}\sum_{\substack{m\in\mathcal{B}\\ p^k|f(m)}} 1 \le \sum_{\substack{z<p\le H\\ p\nmid D}} \left(\frac{h}{Qp^k}+1\right)\rho(p^k).$$

Proof. For each $p \in \mathcal{A}$, there exists an integer a_p such that

$$f(a_p) \not\equiv 0 \pmod{p^k}.$$

By the Chinese Remainder Theorem, there exists an integer a satisfying $a \equiv a_p \pmod{p^k}$ for every $p \in \mathcal{A}$. Hence,

$$f(a) \not\equiv 0 \pmod{p^k} \quad \text{for every } p \in \mathcal{A}.$$

Furthermore, if $y \equiv a \pmod Q$, then $f(y) \not\equiv 0 \pmod{p^k}$ for each $p \in \mathcal{A}$. With this choice of a, we deduce that

$$\sum_{p\le H}\sum_{\substack{m\in\mathcal{B}\\ p^k|f(m)}} 1 = \sum_{\substack{z<p\le H\\ p\nmid D}}\sum_{\substack{m\in\mathcal{B}\\ p^k|f(m)}} 1,$$

since $m \in \mathcal{B}$ with $p^k|f(m)$ implies $p \notin \mathcal{A}$. With $z < p \le H$ and p not dividing D, we have $\gcd(Q, p^k) = 1$. It follows that a given set of p^k consecutive elements in $\mathcal{B}$ has its members distinct modulo p^k. Thus

$$\sum_{\substack{z<p\le H\\ p\nmid D}}\sum_{\substack{m\in\mathcal{B}\\ p^k|f(m)}} 1 \le \sum_{\substack{z<p\le H\\ p\nmid D}} \left(\frac{h}{Qp^k}+1\right)\rho(p^k). \quad \blacksquare$$

Lemma 2. *If $p \nmid D$, then $\rho(p^k) \leq g$ for every positive integer k.*

The above lemma is well-known, and we omit its proof. We note however that it is not difficult to show, under the same hypotheses, that $\rho(p^k) = \rho(p^{k-1})$ for each $k \geq 2$.

Lemma 3. *Given the notation above, $\sum_{z<p\leq H, p\nmid D} \rho(p^k)/p^k \leq 1/4$.*

Proof. By Lemma 2,

$$\begin{aligned}\sum_{\substack{z<p\leq H\\ p\nmid D}} \frac{\rho(p^k)}{p^k} &\leq g \sum_{\substack{z<p\leq H\\ p\nmid D}} \frac{1}{p^k} \leq g \sum_{p>z} \frac{1}{p^k}\\ &< g \sum_{n>z} \frac{1}{n^2} < g \int_z^\infty \frac{1}{x^2}\,dx = \frac{g}{z} = \frac{1}{4}. \qquad \blacksquare\end{aligned}$$

Lemma 4. *Let*

$$F(n) = F(n,h,f) = \sum_{n<m\leq n+h} \sum_{\substack{p>H\\ p^k \mid f(m)}} 1.$$

If there are no integers m in $(n, n+h]$ such that $f(m)$ is k-free, then

$$F(n) \geq \frac{h}{4Q}.$$

Proof. Given an integer $m > 1$, define

$$\chi_f(m) = \begin{cases} 1 & \text{if } f(m) \text{ is } k-\text{free} \\ 0 & \text{otherwise.} \end{cases}$$

Using the notation of Lemma 1, we have

$$0 = \sum_{m\in\mathcal{B}} \chi_f(m) \geq (h/Q - 1) - \sum_{m\in\mathcal{B}} \sum_{p^k \mid f(m)} 1,$$

so

$$\sum_{m\in\mathcal{B}} \sum_{\substack{p>H\\ p^k\mid f(m)}} 1 + \sum_{m\in\mathcal{B}} \sum_{\substack{p\leq H\\ p^k\mid f(m)}} 1 \geq \frac{h}{Q} - 1.$$

By Lemma 1,

$$\sum_{m\in\mathcal{B}} \sum_{\substack{p\leq H\\ p^k\mid f(m)}} 1 = \sum_{p\leq H} \sum_{\substack{m\in\mathcal{B}\\ p^k\mid f(m)}} 1 \leq \sum_{\substack{z<p\leq H\\ p\nmid D}} \left(\frac{h}{Qp^k} + 1\right) \rho(p^k).$$

Using Lemmas 2 and 3 and the Prime Number Theorem, this latter sum is in turn

$$\leq \frac{h}{Q}\frac{1}{4} + \pi(H)g \quad \leq \frac{h}{4Q} + \frac{2Hg}{\log H} = \frac{h}{4Q} + \frac{h\log h}{4Q\log\left(\dfrac{h\log h}{8gQ}\right)} \quad \leq \frac{h}{2Q}.$$

Thus,

$$F(n) + \frac{h}{2Q} \geq \sum_{m\in\mathcal{B}} \sum_{\substack{p>H \\ p^k|f(m)}} 1 + \sum_{m\in\mathcal{B}} \sum_{\substack{p\leq H \\ p^k|f(m)}} 1 \geq \frac{h}{Q} - 1.$$

Hence, since h is sufficiently large, $F(n) \geq h/(2Q) - 1 \geq h/(4Q)$. ■

We note that we could have used the Prime Ideal Theorem, which implies that

$$\sum_{p\leq H} \rho(p^k) \sim \frac{H}{\log H},$$

to replace $\rho(p^k)$ by 1 (on the average) instead of g.

Recall that we are interested in obtaining upper bounds for $L(2h)$. Accordingly, we consider $F^r(n)$ and establish a relationship between $\sum_{X/2<n\leq X} F^r(n)$ and $L(2h)$.

Lemma 5. *For r a positive integer, define*

$$M_r = M_r(h, X, f) = \sum_{X/2<n\leq X} F^r(n).$$

Then

$$L(2h) \ll_r \frac{M_r}{h^{r+1}} + 1.$$

Proof. If $j \in \mathbb{Z}^+$ is counted by $L(2h)$, then $2h < s_{j+1} - s_j$. Consider $(n, n+h]$ for $n \in \{s_j, s_j+1, \ldots, s_j+h-1\}$. In each such interval, no integer m can have $f(m)$ k-free, as

$$s_{j+1} \geq s_j + 2h > n + h \qquad \text{for each such } n.$$

Also, for each such n,

$$\frac{X}{2} < s_j \leq n < s_{j+1} \leq X$$

except possibly in the case that j is the unique positive integer satisfying $s_j \leq X/2 < s_{j+1}$. By Lemma 4, in each of these h intervals,

$$F(n) = \sum_{n<m\leq n+h} \sum_{\substack{p>H \\ p^k|f(m)}} 1 \geq \frac{h}{4Q}.$$

Then by definition,

$$M_r \geq (L(2h) - 1) \left(\frac{h}{4Q} \right)^r h,$$

so we get

$$L(2h) \leq \frac{(4Q)^r}{h^{r+1}} M_r + 1,$$

from which the lemma follows. ■

We examine $M_r = \sum_{X/2<n\leq X} F^r(n)$ by breaking up $F(n)$ as follows. Observe that if $m \asymp X$, then $f(m) \asymp X^g$. We consider $m \in (X/2, 2X]$. If we further consider primes $p > H$ in intervals of the form $(T, 2T]$, then $p^k | f(m)$ would force $T^k \ll X^g$, so that $T \leq cX^{g/k}$ for some constant c. We define

$$G(n,T) = G(n,T,f) = \sum_{n<m\leq n+h} \sum_{\substack{p^k | f(m) \\ T<p\leq 2T}} 1$$

for $T \in \mathcal{T} = \{2^j H : j = 0, 1, \ldots, J\}$, where $J = \left[\log_2(cX^{g/k}/H)\right]$. Recall that $h \leq X$. We deduce that

$$F(n) = \sum_{T\in\mathcal{T}} G(n,T) \qquad \text{for } X/2 < n \leq X.$$

Next, we note that for each positive integer r,

$$G^r(n,T) \ll G(n,T) + \binom{G(n,T)}{r},$$

where henceforth we allow for implied constants to depend on r. In particular, as noted in [5], if $G(n,T) < r$, then $G^r(n,T) \leq r^{r-1} G(n,T)$, while if $G(n,T) \geq r$, then

$$G^r(n,T) \ll \binom{G(n,T)}{r}. \tag{4}$$

This upper bound on $G^r(n,T)$ motivates our next step.

Definition 2. Let $S_r(T)$ be the number of $2r$-tuples $(p_1, \ldots, p_r, m_1, \ldots, m_r)$ with $T < p_1 < p_2 < \cdots < p_r \leq 2T$ such that, for each $j \in \{1, 2, \ldots, r\}$,

$$|m_1 - m_j| < h, \quad \frac{X}{2} < m_j \leq X + h \leq 2X, \quad \text{and} \quad p_j^k | f(m_j).$$

Lemma 6. *Let r be a positive integer. If $r = 1$, then*

$$\sum_{X/2<n\leq X} \binom{G(n,T)}{r} \leq (2g)^r h S_r(T).$$

If $r > 1$, then

$$\sum_{\substack{X/2<n\le X\\ G(n,T)\ge 2gr}} \binom{G(n,T)}{r} \le (2g)^r h S_r(T).$$

Proof. We define an auxiliary function $\tilde{G}(n,T)$ equal to the number of primes $p \in (T, 2T]$ for which there exists an $m \in (n, n+h]$ such that p^k divides $f(m)$. Since $T > h$, Lemma 2 implies that for any prime $p > T$, there exist at most g integers $m \in (n, n+h]$ such that $p^k | f(m)$. Therefore,

$$G(n,T) = \sum_{n<m\le n+h} \sum_{\substack{p^k|f(m)\\ T<p\le 2T}} 1 = \sum_{T<p\le 2T} \sum_{\substack{p^k|f(m)\\ n<m\le n+h}} 1 \le g\tilde{G}(n,T).$$

Furthermore, we define $\tilde{S}_r(n,T)$ as the number of $2r$-tuples $(p_1, \dots, p_r, m_1, \dots, m_r)$ with $T < p_1 < p_2 < \cdots < p_r \le 2T$ such that $m_j \in (n, n+h]$ and $p_j^k | f(m_j)$ for each $j \in \{1, 2, \dots, r\}$. Observe that the conditions on m_j in the definition of $\tilde{S}_r(n,T)$ imply that as n varies, a given $2r$-tuple $(p_1, \dots, p_r, m_1, \dots, m_r)$ can be counted by at most h different $\tilde{S}_r(n,T)$. Hence,

$$\sum_{X/2<n\le X} \tilde{S}_r(n,T) \le h S_r(T).$$

We view $\binom{\tilde{G}(n,T)}{r}$ as the number of ways of choosing $p_1, \dots, p_r$ such that $T < p_1 < p_2 < \cdots < p_r \le 2T$ and such that there exists an $m_j \in (n, n+h]$ for which $p_j^k | f(m_j)$ for each $j \in \{1, 2, \dots, r\}$. We deduce

$$\binom{\tilde{G}(n,T)}{r} \le \tilde{S}_r(n,T).$$

If $G(n,T) \ge 2gr$, then we deduce that $\tilde{G}(n,T) \ge 2r$. For each $j \in \{0, 1, \dots, r-1\}$, it then follows that $G(n,T) - j \le g\tilde{G}(n,T) - j \le 2g(\tilde{G}(n,T) - j)$. Therefore, if $G(n,T) \ge 2gr$, then

$$\binom{G(n,T)}{r} \le (2g)^r \binom{\tilde{G}(n,T)}{r}.$$

This inequality also holds if $r = 1$ independent of whether or not $G(n,T) \ge 2gr$ is satisfied. We obtain

$$\begin{aligned}\sum_{\substack{X/2<n\le X\\ G(n,T)\ge 2gr}} \binom{G(n,T)}{r} &\le (2g)^r \sum_{\substack{X/2<n\le X\\ G(n,T)\ge 2gr}} \binom{\tilde{G}(n,T)}{r}\\ &\le (2g)^r \sum_{\substack{X/2<n\le X\\ G(n,T)\ge 2gr}} \tilde{S}_r(n,T) \le (2g)^r h S_r(T),\end{aligned}$$

with a similar sequence of inequalities holding in the case $r = 1$ without the condition $G(n,T) \ge 2gr$ in the summations. This completes the proof. ∎

Now that we have obtained the bound in Lemma 6, we seek an upper bound on $S_r(T)$ as well.

Lemma 7. *If $T \geq H$ and $r \in \mathbb{Z}^+$, then*

$$S_r(T) \ll \frac{h^{r-1}X}{T^{(k-1)r}\log^r T} + \frac{h^{r-1}T^r}{\log^r T}.$$

Proof. We count the $2r$-tuples $(p_1, p_2, \dots, p_r, m_1, m_2, \dots, m_r)$ with $T < p_1 < p_2 < \cdots < p_r \leq 2T$ such that, for each $j \in \{1, 2, \dots, r\}$,

$$|m_1 - m_j| < h, \quad \frac{X}{2} < m_j \leq X + h \leq 2X, \quad \text{and} \quad p_j^k | f(m_j).$$

Fix such $p_1, p_2, \dots, p_r$ in $\ll (T/\log T)^r$ ways. Since $f(m_j) \equiv 0 \pmod{p_j^k}$ for each j, we have m_j congruent to one of $\leq g$ incongruent numbers modulo p_j^k for each j. Since $|m_1 - m_j| < h$ for each j, we deduce that m_1 is congruent to one of $\leq g(2h+1)$ different numbers modulo p_j^k for each $j \geq 2$. Then by the Chinese Remainder Theorem, m_1 is congruent to one of

$$\leq g^{r-1}\,(2h+1)^{r-1}$$

different numbers modulo ${p_2}^k \cdots {p_r}^k$. Also, since m_1 is congruent to one of g different numbers modulo ${p_1}^k$, we obtain from the Chinese Remainder Theorem that m_1 is congruent to one of

$$\leq g^r\,(2h+1)^{r-1}$$

different numbers modulo ${p_1}^k{p_2}^k \cdots {p_r}^k$. Since ${p_1}^k{p_2}^k \cdots {p_r}^k \asymp T^{kr}$, we get $\ll_{g,r} (X/T^{kr} + 1)h^{r-1}$ choices for m_1. Hence there are

$$\ll \left(\frac{X}{T^{kr}} + 1\right) h^{r-1} \left(\frac{T}{\log T}\right)^r = \frac{h^{r-1}X}{T^{(k-1)r}\log^r T} + \frac{h^{r-1}T^r}{\log^r T}$$

different choices for $p_1, p_2, \dots, p_r$ and m_1. We use that $|m_1 - m_j| < h$, $p_j^k | f(m_j)$, $p_j > h$, and $\rho(p_j^k) \leq g$ to deduce that for each $j \geq 2$, there are $\leq g$ choices for m_j given the values of $p_1, p_2, \dots, p_r$ and m_1. The lemma now follows. ■

Our basic strategy will be to obtain (2) by combining Lemma 5 with the above estimates and using Hölder's inequality. We will be summing over T in $\mathcal{T}$ to get the desired bound on $L(2h)$, so we mention the following lemma (which has an easy proof that we omit).

Lemma 8. *For $\mathcal{T}$ as previously defined,*

$$\sum_{\substack{T \in \mathcal{T} \\ U < T \leq V}} T^a \ll \begin{cases} U^a & \text{if } a < 0 \\ V^a & \text{if } a > 0. \end{cases}$$

4. A Proof of Theorem 2

In this section, we prove Theorem 2 assuming Theorem 3. More precisely, we establish that (2) holds for $h \le X^{\delta'}$ (by Section 2, this is sufficient).

In the following lemma, we will need a positive real number ε such that

$$\varepsilon < \frac{(k-s)(2s+g) - g(g-1)}{(k-s)(2s+g)},$$

where s is to be chosen from $\{1, 2, \ldots, k-1\}$. To ensure that the numerator of the right-hand side of this inequality is positive, we choose $s \in \{1, 2, \ldots, k-1\}$ to maximize this numerator. We can achieve this by taking $s = (2k - g + c)/4$, where $c \in \{-1, 0, 1, 2\}$. This choice of s is possible since $2k - g < 4(k-1) + 2$ and since $2k - g \ge 2$ for $k \ge (\sqrt{2} - 1/2)g$ and $k \ge 2$. For any such s, the numerator becomes

$$\frac{1}{4}(2k + g - c)\frac{1}{2}(2k + g + c) - g(g-1) = \frac{1}{8}((2k+g)^2 - c^2) - g(g-1).$$

Thus the numerator is positive provided that $(2k + g)^2 + 8g > 8g^2 + c^2$, which follows for any $c \in \{-1, 0, 1, 2\}$ provided $k \ge (\sqrt{2} - 1/2)g$. Furthermore, we note that for our choice of s,

$$\frac{(k-s)(2s+g) - g(g-1)}{(k-s)(2s+g)} = 1 - \frac{8g(g-1)}{(2k+g)^2 - c^2}.$$

It is easy to see that the expression on the right is minimized when $c = 2$. Consequently, if we take

$$\varepsilon < 1 - \frac{8g(g-1)}{(2k+g)^2 - 4},$$

then we know the previous inequality on ε also holds for some choice of $s \in \{1, 2, \ldots, k-1\}$.

Lemma 9. *Let $k \ge (\sqrt{2} - 1/2)g$. Let $\varepsilon \in (0, (k-1)/k]$ such that*

$$\varepsilon < 1 - \frac{8g(g-1)}{(2k+g)^2 - 4}.$$

If $X^{1-\varepsilon} < T \ll X^{g/k}$, then there is a $\xi = \xi(\varepsilon) > 0$ such that

$$S_1(T) \ll X^{1-\xi}.$$

Proof. We suppose as we may that X is sufficiently large. We find an upper bound on the number of pairs (p, m) with $T < p \le 2T$, $X/2 < m \le 2X$, and $p^k | f(m)$. Set

$$H' = c_1 T^{2s(k-s)/(g(2s+1))} X^{1/(2s+1)} \quad \text{and} \quad B = c_2 X^{-1/(2s+1)} T^{(k+s+1)/(g(2s+1))}$$

where c_1 and c_2 are some appropriately chosen positive constants and where s denotes an arbitrary integer in $\{1, 2, \ldots, k-1\}$. We divide the interval $(X/2, 2X]$ into $\ll X/H' + 1$

subintervals of length $\leq H'$. It follows from the work of Huxley and Nair in [9] (also see [4]) that in such a subinterval I of length H', there are $\ll T/B^g + 1$ primes $p \in (T, 2T]$ for which $p^k | f(m)$ for some $m \in I$. Since $T \ll X^{g/k}$, it is straightforward to check that $T \geq B^g$ and $X \geq H'$. The condition $\varepsilon \leq (k-1)/k$ implies $T > X^{1-\varepsilon} \geq X^{1/k}$ from which it follows that $T^k \geq H'$. Hence, for each $p \in (T, 2T]$, there exist at most $\rho(p^k) \leq g$ different integers $m \in I$ such that $p^k | f(m)$. Thus,

$$S_1(T) \ll \left(\frac{X}{H'} + 1 \right) \left(\frac{T}{B^g} + 1 \right)$$

$$\ll \frac{XT}{H'B^g} \ll \frac{X^{(2s+g)/(2s+1)}}{T^{(2s+g)(k-s)/(g(2s+1))}}$$

$$< X^{(g(2s+g)-(1-\varepsilon)(2s+g)(k-s))/(g(2s+1))}.$$

The exponent of X in the last part of this inequality will be < 1 provided that

$$\varepsilon < 1 - \frac{g(g-1)}{(k-s)(2s+g)} = \frac{(k-s)(2s+g) - g(g-1)}{(k-s)(2s+g)}.$$

From the discussion preceding the lemma, we can choose $s \in \{1, 2, \ldots, k-1\}$ so that this inequality on ε holds. Hence the lemma follows. ■

In the next section, we will want a variation of Lemma 9 in which $\xi = \varepsilon$. The above proof would carry through in this case if

$$\frac{g(2s+g) - (1-\varepsilon)(2s+g)(k-s)}{g(2s+1)} \leq 1 - \varepsilon.$$

Equivalently, we can take $\xi = \varepsilon$ in Lemma 9 provided that

$$\varepsilon \leq 1 - \frac{g(2s+g)}{(2s+g)(k-s) + g(2s+1)} = \frac{(2s+g)(k-s) - g(g-1)}{(2s+g)(k-s) + g(2s+1)}$$

for some $s \in \{1, 2, \ldots, k-1\}$. We will need ε positive; choosing $s = \left[(\sqrt{2}-1)g/2\right]$ (or $s = 1$ if $\left[(\sqrt{2}-1)g/2\right] = 0$) will serve our purposes. Writing $s = (\sqrt{2}-1)g/2 + \theta$ and using $k > (\sqrt{2} - 1/2)g$, we note that

$$(2s+g)(k-s) > g^2 - 2\theta^2 \geq g^2 - g$$

provided $2\theta^2 \leq g$. Thus, for $g \geq 2$ (as in Theorem 1), the numerator in the above bound for ε is > 0 for some s (indeed, for s as chosen in Theorem 1) in $\{1, 2, \ldots, k-1\}$. We deduce that we can choose $\xi(\varepsilon) = \varepsilon$ provided $\varepsilon > 0$ is sufficiently small.

In the proof of Theorem 2, we will consider cases based on the size of both T and $G(n, T)$. In the cases where $T > X^{1-\varepsilon}$ for ε as defined in the previous lemma, we will use the bound for $S_1(T)$ obtained in that lemma. Observe that in Lemma 9, the condition $T \ll X^{g/k}$ is not important as $p^k | f(m)$, $p \in (T, 2T]$, and $m \leq 2X$ imply $T \ll X^{g/k}$.

We now prove Theorem 2 (assuming Theorem 3). Fix $\gamma > 0$. Let j be a fixed positive integer with $j \geq \max\{\gamma + \varepsilon', 2\}$, where $\varepsilon' > 0$ is sufficiently small. Let ε be as in Lemma 9. We consider $h \leq X^{\delta'}$, with $\delta' > 0$ to be determined. We will show that $L(2h) \ll X/h^{\gamma+\varepsilon'}$ for $h \leq X^{\delta'}$ so that, as in Section 2, Theorem 2 will follow.

Given $T \in \mathcal{T}$, we consider five cases, based on the size of T and $G(n, T)$:

Let θ_1 represent the case that $T \leq X^{1-\varepsilon}$ and $G(n, T) \leq 2gj$.

Let θ_2 represent the case that $T > X^{1-\varepsilon}$ and $G(n, T) \leq 2gj$.

Let θ_3 represent the case that $T \leq X^{1/(kj)}$ and $G(n, T) > 2gj$.

Let θ_4 represent the case that $X^{1/(kj)} < T \leq X^{1-\varepsilon}$ and $G(n, T) > 2gj$.

Let θ_5 represent the case that $T > X^{1-\varepsilon}$ and $G(n, T) > 2gj$.

For each $i \in \{1, 2, 3, 4, 5\}$, define $F_i(n) = \sum_{\theta_i} G(n, T)$ (so the sum is over $T \in \mathcal{T}$ satisfying the conditions in θ_i). If $\ell \in \mathbb{Z}^+$ is counted by $L(2h)$, then we consider intervals of the form $(n, n+h]$ and allow n to range through $s_\ell, s_\ell + 1, \ldots, s_\ell + h - 1$. As in the proof of Lemma 5, in each such interval, no integer m has $f(m)$ k-free. Then by Lemma 4,

$$F(n) = \sum_{i=1}^{5} F_i(n) \geq \frac{h}{4Q} \qquad \forall n \in \{s_\ell, s_\ell + 1, \ldots, s_\ell + h - 1\}.$$

Thus for each of the h different values of n above, at least one of $F_1(n)$, $F_2(n)$, $F_3(n)$, $F_4(n)$, or $F_5(n)$ is $\geq h/(20Q)$. As we let ℓ vary over the positive integers for which $2h < s_{\ell+1} - s_\ell \leq 4h$ and $X/2 < s_{\ell+1} \leq X$, we deduce that for at least one $i \in \{1, 2, 3, 4, 5\}$, $F_i(n) \geq h/(20Q)$ for at least $h(L(2h) - 1)/5$ different $n \in (X/2, X]$. Hence, for at least one $i \in \{1, 2, 3, 4, 5\}$,

$$\sum_{X/2 < n \leq X} F_i(n) \geq (L(2h) - 1) \left(\frac{h}{20Q} \right) \left(\frac{h}{5} \right).$$

More generally, given any positive integers j_1, j_2, j_3, j_4, and j_5, we have

$$\sum_{X/2 < n \leq X} F_i^{j_i}(n) \geq (L(2h) - 1) \left(\frac{h}{20Q} \right)^{j_i} \left(\frac{h}{5} \right)$$

for some $i \in \{1, 2, 3, 4, 5\}$. Hence,

$$L(2h) \ll 1 + \sum_{i=1}^{5} \frac{1}{h^{j_i+1}} \sum_{X/2 < n \leq X} F_i^{j_i}(n). \tag{5}$$

We take $j_1 = j_2 = j_3 = j$ and $j_4 = j_5 = 1$ and use these choices to show that for each $i \in \{1, 2, 3, 4, 5\}$,

$$\frac{1}{h^{j_i+1}} \sum_{X/2 < n \leq X} F_i^{j_i}(n) \ll \frac{X}{h^{\gamma+\varepsilon'}}.$$

For $i=1$, we write $F_1^j(n) = \left(\sum_{\theta_1} T^{-\varepsilon'/j} T^{\varepsilon'/j} G(n,T)\right)^j$ and apply Hölder's inequality. Taking $P = j/(j-1)$ and $Q = j$ and using Lemma 8, we deduce

$$\begin{aligned}
F_1^j(n) &\le \left\{ \left(\sum_{\theta_1} \left(T^{-\varepsilon'/j} \right)^P \right)^{1/P} \left(\sum_{\theta_1} \left(T^{\varepsilon'/j} G(n,T) \right)^Q \right)^{1/Q} \right\}^j \\
&= \left(\sum_{\theta_1} T^{-\varepsilon'/(j-1)} \right)^{j-1} \sum_{\theta_1} T^{\varepsilon'} G^j(n,T) \\
&\ll H^{-\varepsilon'} \sum_{\theta_1} T^{\varepsilon'} G(n,T),
\end{aligned}$$

since $G(n,T) \le 2gj$. Recalling that $T \le X^{1-\varepsilon}$ in this case, we apply Lemmas 6 (with $r=1$), 7, and 8 to obtain

$$\begin{aligned}
\frac{1}{h^{j+1}} \sum_{X/2<n\le X} F_1^j(n) &\ll \frac{H^{-\varepsilon'}}{h^{j+1}} \sum_{\substack{T\in\mathcal{T} \\ T\le X^{1-\varepsilon}}} T^{\varepsilon'} \sum_{\substack{X/2<n\le X \\ G(n,T)\le 2gj}} G(n,T) \\
&\ll \frac{H^{-\varepsilon'}}{h^{j+1}} \sum_{\substack{T\in\mathcal{T} \\ T\le X^{1-\varepsilon}}} T^{\varepsilon'} h S_1(T) \\
&\ll \frac{H^{-\varepsilon'}}{h^{j}} \sum_{\substack{T\in\mathcal{T} \\ T\le X^{1-\varepsilon}}} T^{\varepsilon'} \left(\frac{X}{T^{k-1}\log T} + \frac{T}{\log T} \right) \\
&\ll \frac{XH^{-\varepsilon'}}{h^j H^{k-1-\varepsilon'}} + \frac{X^{(1+\varepsilon')(1-\varepsilon)} H^{-\varepsilon'}}{h^j} \\
&\ll \frac{X}{h^{j+k-1}} + \frac{X^{(1+\varepsilon')(1-\varepsilon)}}{h^{j+\varepsilon'}},
\end{aligned}$$

which we make $\ll \dfrac{X}{h^{\gamma+\varepsilon'}}$ by taking $\varepsilon' \le \varepsilon/(1-\varepsilon)$ and by noting that $j+k-1 > j \ge \gamma+\varepsilon'$.

For $i=2$, we use the same application of Hölder's inequality to obtain

$$F_2^j(n) \ll H^{-\varepsilon'} \sum_{\theta_2} T^{\varepsilon'} G(n,T).$$

From Lemmas 6, 8, and 9, we obtain for some $\xi = \xi(\varepsilon) > 0$

$$\begin{aligned}
\frac{1}{h^{j+1}} \sum_{X/2<n\le X} F_2^j(n) &\ll \frac{H^{-\varepsilon'}}{h^j} \sum_{\substack{T\in\mathcal{T} \\ T> X^{1-\varepsilon}}} T^{\varepsilon'} \left(X^{1-\xi} \right) \\
&\ll \frac{H^{-\varepsilon'}}{h^j} X^{g\varepsilon'/k+1-\xi} \ll \frac{1}{h^{j+\varepsilon'}} X^{g\varepsilon'/k+1-\xi}.
\end{aligned}$$

Taking ε' such that $\varepsilon' \le k\xi/g$ and recalling $j > \gamma$, we conclude

$$\frac{1}{h^{j+1}} \sum_{X/2<n\le X} F_2^j(n) \ll \frac{X}{h^{\gamma+\varepsilon'}}.$$

For $i = 3$, we again apply Hölder's inequality to obtain

$$F_3^j(n) \ll H^{-\varepsilon'} \sum_{\theta_3} T^{\varepsilon'} G^j(n, T).$$

Since $T \le X^{1/(kj)}$ in this case, Lemma 7 implies that

$$S_j(T) \ll \frac{h^{j-1}X}{T^{(k-1)j}\log^j T}.$$

Combining this result with (4) and with Lemmas 6 and 8 gives

$$\begin{aligned}
\frac{1}{h^{j+1}} \sum_{X/2<n\le X} F_3^j(n) &\ll \frac{H^{-\varepsilon'}}{h^{j+1}} \sum_{\substack{T\in\mathcal{T}\\ T\le X^{1/(kj)}}} T^{\varepsilon'} \sum_{\substack{X/2<n\le X\\ G(n,T)>2gj}} \binom{G(n,T)}{j} \\
&\ll \frac{H^{-\varepsilon'}}{h^{j+1}} \sum_{\substack{T\in\mathcal{T}\\ T\le X^{1/(kj)}}} T^{\varepsilon'} h S_j(T) \\
&\ll \frac{H^{-\varepsilon'}}{h^{j}} \sum_{\substack{T\in\mathcal{T}\\ T\le X^{1/(kj)}}} T^{\varepsilon'} \left(\frac{h^{j-1}X}{T^{(k-1)j}\log^j T}\right) \\
&\ll \frac{XH^{-\varepsilon'}}{hH^{(k-1)j-\varepsilon'}} \ll \frac{X}{h^{(k-1)j+1}} \ll \frac{X}{h^{\gamma+\varepsilon'}},
\end{aligned}$$

since $(k-1)j + 1 > j \ge \gamma + \varepsilon'$.

For $i = 4$, we will require $\delta' > 0$ with $\delta' < (k-1)/(2kj\gamma)$. Then $h \le X^{\delta'} \le X^{(k-1)/(2kj\gamma)}$ and $T > X^{1/(kj)}$ combine to give

$$T > X^{1/(2kj)}X^{1/(2kj)} \ge h^{\gamma/(k-1)}X^{1/(2kj)}.$$

We use that there are $\ll \log X$ different T in $\mathcal{T}$. From Lemmas 6 and 7 (both with $r = 1$), we obtain

$$\begin{aligned}
\frac{1}{h^2}\sum_{X/2<n\le X} F_4(n) &= \frac{1}{h^2}\sum_{\substack{T\in\mathcal{T}\\ X^{1/(kj)}<T\le X^{1-\varepsilon}}}\ \sum_{\substack{X/2<n\le X\\ G(n,T)>2gj}} G(n,T)\\
&\ll \frac{1}{h^2}\sum_{\substack{T\in\mathcal{T}\\ X^{1/(kj)}<T\le X^{1-\varepsilon}}} hS_1(T)\\
&\ll \frac{1}{h}\sum_{\substack{T\in\mathcal{T}\\ X^{1/(kj)}<T\le X^{1-\varepsilon}}} \left(\frac{X}{T^{k-1}\log T}+\frac{T}{\log T}\right)\\
&\ll \frac{X\log X}{h(h^{\gamma/(k-1)}X^{1/(2kj)})^{k-1}\log(X^{1/(kj)})}+\frac{X^{1-\varepsilon}\log X}{h\log(X^{1/(kj)})}\\
&\ll \frac{X}{h^{\gamma+1}}+\frac{X^{1-\varepsilon}}{h}.
\end{aligned}$$

To ensure that $X^{1-\varepsilon}/h \ll X/h^{\gamma+\varepsilon'}$, we require $\delta' \le \varepsilon/(\gamma+\varepsilon'-1)$ if $\gamma \ge 1$ and $\delta' \le \varepsilon/\varepsilon'$ otherwise.

For $i = 5$, we have that $T > X^{1-\varepsilon}$, so we apply Lemmas 6 and 9 to obtain for some $\xi = \xi(\varepsilon) > 0$

$$\begin{aligned}
\frac{1}{h^2}\sum_{X/2<n\le X} F_5(n) &= \frac{1}{h^2}\sum_{\substack{T\in\mathcal{T}\\ T>X^{1-\varepsilon}}}\ \sum_{\substack{X/2<n\le X\\ G(n,T)>2gj}} G(n,T)\\
&\ll \frac{1}{h^2}\sum_{\substack{T\in\mathcal{T}\\ T>X^{1-\varepsilon}}} hS_1(T) \ll \frac{X^{1-\xi}\log X}{h}.
\end{aligned}$$

As in the case for $i = 4$, we ensure that $X^{1-\xi}\log X/h \ll X/h^{\gamma+\varepsilon'}$; we require $\delta' < \xi/(\gamma+\varepsilon'-1)$ if $\gamma \ge 1$ and $\delta' < \xi/\varepsilon'$ otherwise.

To ensure that $1 \ll X/h^{\gamma+\varepsilon'}$, we take $\delta' < 1/(\gamma+\varepsilon')$. Summarizing the above, we obtain from (5) that there exists a $\delta' > 0$ such that $L(2h) \ll X/h^{\gamma+\varepsilon'}$ for $h \le X^{\delta'}$. Thus, we have established (2) for $h \le X^{\delta'}$. Theorem 2 follows.

5. A Proof of Theorem 1

In the following lemma, we fix $\varepsilon > 0$ such that $\varepsilon < 1 - 8g(g-1)/((2k+g)^2-4)$ and use this value to split $\mathcal{T}$ into the sets $\mathcal{T}_1 = \{T \in \mathcal{T} : T < X^{1-\varepsilon}\}$ and $\mathcal{T}_2 = \{T \in \mathcal{T} : T \ge X^{1-\varepsilon}\}$.

Lemma 10. *Fix ε and define $\mathcal{T}_2$ as above. Let $g \ge 2$, and let k be an integer $\ge (\sqrt{2}-1/2)g$. Then for some $\xi = \xi(\varepsilon)$ as in Lemma 9,*

$$\sum_{T\in\mathcal{T}_2}\ \sum_{X/2<n\le X} G(n,T) \ll hX^{1-\xi}\log X.$$

Proof. From Lemma 6 and Lemma 9, we obtain

$$\sum_{T\in\mathcal{T}_2}\sum_{X/2<n\leq X} G(n,T) \ll \sum_{T\in\mathcal{T}_2} hS_1(T) \ll hX^{1-\xi}\sum_{T\in\mathcal{T}_2} 1 \ll hX^{1-\xi}\log X.$$

This establishes the lemma. ■

We are now ready to prove Theorem 1 assuming Theorem 3. One checks that for $2 \leq g \leq 4$, $\phi_1 > 0$. Recall the remarks following the proof of Lemma 9. It follows that $\phi_1 > 0$ for all g, k, and s as in the theorem. Since $k \geq (\sqrt{2}-1/2)g$ implies $2k+g \geq 2\sqrt{2}g$ and $2k \geq g$, we easily deduce $\phi_2 > 0$ as well. Observe that $\gamma < 1+\phi_1/\phi_2$ implies $(\gamma-1)\phi_2 < \phi_1$. We choose ε so that

$$(\gamma-1)\phi_2 < \varepsilon < \phi_1.$$

In particular, we may take $\xi = \varepsilon$ in Lemma 9. Thus, from Lemmas 6, 7, and 10,

$$\begin{aligned}\sum_{X/2<n\leq X} F(n) &= \sum_{T\in\mathcal{T}}\sum_{X/2<n\leq X} G(n,T)\\
&\ll \sum_{T\in\mathcal{T}_2}\sum_{X/2<n\leq X} G(n,T) + \sum_{T\in\mathcal{T}_1}\sum_{X/2<n\leq X} G(n,T)\\
&\ll hX^{1-\varepsilon}\log X + \sum_{T\in\mathcal{T}_1} ghS_1(T)\\
&\ll hX^{1-\varepsilon}\log X + h\sum_{T\in\mathcal{T}_1}\left(\frac{X}{T^{k-1}\log T} + \frac{T}{\log T}\right)\\
&\ll hX^{1-\varepsilon}\log X + \frac{hX}{H^{k-1}\log H} + \frac{hX^{1-\varepsilon}}{\log H}\\
&\ll \frac{X}{h^{k-2}\log^k h} + hX^{1-\varepsilon}\log X.\end{aligned}$$

By Lemma 5 with $r=1$, we obtain

$$L(2h) \ll \frac{1}{h^2}\sum_{X/2<n\leq X} F(n) \; + 1 \ll \frac{X}{h^k\log^k h} + \frac{X^{1-\varepsilon}\log X}{h} + 1.$$

Next, we show that each of these terms is $\ll X/h^{\gamma+\delta}$ for some sufficiently small $\delta > 0$. To establish

$$\frac{X}{h^k\log^k h} \ll \frac{X}{h^{\gamma+\delta}},$$

we simply use that $\gamma < k$. To establish that

$$\frac{X^{1-\varepsilon}\log X}{h} \ll \frac{X}{h^{\gamma+\delta}}, \tag{6}$$

we recall the comments in the introduction that the work of Huxley and Nair [9] for $k \le g$ and the work of Filaseta [4] for $k > g$ imply $h \ll X^{\phi_2}$ (so that (2) holds for $h > cX^{\phi_2}$ for some constant $c > 0$ as then $L(2h) = 0$). Thus,

$$h^{\gamma-1+\delta} \ll X^{(\gamma-1+\delta)\phi_2} \ll \frac{X^{\varepsilon}}{\log X}$$

for some $\delta > 0$. Now, (6) easily follows. Finally, to establish that $1 \ll X/h^{\gamma+\delta}$, we use that $h \ll X^{\phi_2}$ and that $\gamma < 1/\phi_2$. Thus, (2) is established (with δ in place of ε) and Theorem 1 holds.

6. A Proof of Theorem 3

Throughout the argument for Theorem 3, we will consider X to be sufficiently large and suppose that the conditions in the theorem hold so that, in particular, $s_{j+1} - s_j = d$ for some j. Whenever we use p, it will denote a prime. We consider $B = B(d, f)$ to be a sufficiently large number (independent of X). Specifically, we consider B large enough so that for each $i \in \{1, 2, \ldots, d-1\}$ and for some fixed j for which $s_{j+1} - s_j = d$, there is a prime $p \le B$ (the prime depending on i) such that $p^k | f(s_j + i)$.

Let

$$S = \{p : p \le B\} \quad \text{and} \quad M = M(B) = \prod_{p \in S} p^k,$$

and define the sets

$$S' = S'(B) = \{a \in [0, M-1] \cap \mathbb{Z} : p \in S \implies p^k \nmid f(a) \text{ and } p^k \nmid f(a+d)\}$$

and

$$S'' = S''(B) = \{a \in S'(B) : \text{ for each } i \in \{1, 2, \ldots, d-1\},$$
$$\text{there is a } p \in S \text{ such that } p^k | f(a+i)\}.$$

Fix B_0 sufficiently large and $M_0 = M(B_0)$ so that in particular $s_j \in S''(B_0)$ for some fixed j with $s_{j+1} - s_j = d$. Observe that for any integers t and i and any prime $p \le B_0$, $f(s_j + M_0 t + i) \equiv f(s_j + i) \pmod{p^k}$. It easily follows that for each nonnegative integer t, the number $s_j + M_0 t$ belongs to $S''(B)$ provided it is $< M(B)$ and provided that for each prime $p \in (B_0, B]$, $p^k \nmid f(s_j + M_0 t)$ and $p^k \nmid f(s_j + d + M_0 t)$. We deduce from Lemma 2 (for B_0 sufficiently large) that

$$|S''(B)| \ge \frac{M(B)}{M_0} - 2g \sum_{B_0 < p \le B} \frac{M(B)}{p^k M_0}$$
$$\ge \frac{M(B)}{M_0}\left(1 - \sum_{p > B_0} \frac{2g}{p^k}\right) \ge \frac{M(B)}{2M_0}.$$

Thus, as B approaches infinity, we obtain

$$\left(\frac{1}{M(B)}\right)\left(\sum_{a\in S''(B)} 1\right) \geq \frac{1}{\left(2\prod_{p\leq B_0} p^k\right)}.$$

We will show momentarily that the left-hand side above approaches a limit as B tends to infinity; the above will then imply that this limit is positive.

To prove the theorem, we momentarily fix B, $M = M(B)$, and $S'' = S''(B)$ as above. Consider

$$S''(X, M) = |\{n \leq X : n \equiv a \pmod{M} \text{ for some } a \in S''\}|.$$

Observe that if n is counted by $S''(X, M)$, then either n is counted by $N_d(X)$ or one of $f(n)$ or $f(n+d)$ is divisible by p^k for some prime $p > B$. Also, if n is counted by $N_d(X)$, then either n is counted by $S''(X, M)$ or one of the numbers $f(n+1), f(n+2), \ldots, f(n+d-1)$ is divisible by p^k for some prime $p > B$. Therefore, we deduce that

$$N_d(X) = S''(X, M) + O(W(X)),$$

where $W(X)$ denotes the number of positive integers $n \leq X$ for which at least one of the numbers $f(n), f(n+1), \ldots, f(n+d)$ is divisible by p^k for some prime $p > B$. Observe that

$$S''(X, M) = \sum_{a\in S''}\left(\frac{X}{M} + O(1)\right) = \left((1/M)\sum_{a\in S''} 1\right) X + O(M).$$

To estimate $W(X)$, we will make use of Lemma 9. Since B is sufficiently large and $p > B$, we may suppose that p does not divide the discriminant of f. By Lemma 2, there are at most g incongruent roots to the congruence $f(x) \equiv 0 \pmod{p^k}$. Hence, for any $i \in \{0, 1, \ldots, d\}$, we have

$$\begin{aligned}\sum_{n\leq X}\sum_{\substack{B<p\leq X\\ p^k|f(n+i)}} 1 &= \sum_{B<p\leq X}\sum_{\substack{n\leq X\\ p^k|f(n+i)}} 1 \leq \sum_{B<p\leq X}\left(g\,\frac{X}{p^k} + O(g)\right)\\ &= \left(\sum_{B<p\leq\infty}\frac{g}{p^k}\right) X + O(\pi(X)) = O(X/B^{k-1}) + O(X/\log X).\end{aligned}$$

Estimating

$$\sum_{n\leq X}\sum_{\substack{p>X\\ p^k|f(n+i)}} 1$$

is more difficult, but we can appeal to Lemma 9. It follows from that lemma that with $k \geq (\sqrt{2} - 1/2)g$, this last double sum is $o(X)$. Therefore, we deduce

$$N_d(X) = \left((1/M)\sum_{a\in S''} 1\right) X + O(M) + O(dX/B^{k-1}) + o(X).$$

Dividing by X, we obtain

$$\frac{N_d(X)}{X} = \left((1/M) \sum_{a \in S''} 1 \right) + O(M/X) + O(d/B^{k-1}) + o(1).$$

Hence,

$$\limsup_{X \to \infty} \frac{N_d(X)}{X} = \left((1/M) \sum_{a \in S''} 1 \right) + O(d/B^{k-1})$$

and

$$\liminf_{X \to \infty} \frac{N_d(X)}{X} = \left((1/M) \sum_{a \in S''} 1 \right) + O(d/B^{k-1}).$$

Thus the difference between $\limsup_{X\to\infty}(N_d(X)/X)$ and $\liminf_{X\to\infty}(N_d(X)/X)$ can be made arbitrarily small (as B gets large). Hence $\lim_{X\to\infty}(N_d(X)/X)$ exists. Letting B approach infinity, we deduce that $\lim_{B\to\infty}\big((1/M)\sum_{a\in S''} 1\big)$ exists. Setting the value of this limit to be c_d, the theorem follows.

References

1. P. Erdős, *Some problems and results in elementary number theory*, Publ. Math. Debrecen **2** (1951), 103–109.
2. P. Erdős, *Arithmetical properties of polynomials*, J. London Math. Soc. **28** (1953), 416–425.
3. M. Filaseta, *On the distribution of gaps between squarefree numbers*, Mathematika **40** (1993), 88–101.
4. M. Filaseta, *Short interval results for k–free values of irreducible polynomials*, Acta Arith. **64** (1993), 249–270.
5. S. W. Graham, *Moments of gaps between k–free numbers*, Journal of Number Theory **44** (1993), 105–117.
6. C. Hooley, *On the power free values of polynomials*, Mathematika **14** (1967), 21–26.
7. C. Hooley, *On the distribution of square-free numbers*, Canad. J. Math. **25** (1973), 1216–1223.
8. M. N. Huxley, *Moments of differences between square-free numbers*, Sieve Methods, Exponential Sums, and their Applications in Number Theory (eds. G. R. H. Greaves, G. Harman, and M. N. Huxley), 187–204, Cambridge Univ. Press, Cambridge, 1996(7).
9. M. N. Huxley and M. Nair, *Power free values of polynomials III*, Proc. London Math. Soc. **41** (1980), 66–82.
10. L. Mirsky, *Arithmetical pattern problems related to divisibility by rth powers*, Proc. London Math. Soc. **50** (1949), 497–508.
11. T. Nagel, *Zur Arithmetik der Polynome*, Abh. Math. Sem. Hamburg. Univ. **1** (1922), 179–184.
12. M. Nair, *Power free values of polynomials*, Mathematika **23** (1976), 159–183.
13. M. Nair, *Power free values of polynomials II*, Proc. London Math. Soc. (3) **38** (1979), 353–368.

Brian Beasley
Department of Mathematics
Presbyterian College
Clinton, SC 29325
bbeasley@cs1.presby.edu

Michael Filaseta
Mathematics Department
University of South Carolina
Columbia, SC 29208
filaseta@math.sc.edu