POLYNOMIAL DIFFERENCES IN THE PRIMES

NEIL LYALL ALEX RICE

Abstract. We establish, utilizing the Hardy-Littlewood Circle Method, an asymptotic formula for the number of pairs of primes whose differences lie in the image of a fixed polynomial. We also include a generalization of this result where differences are replaced with any integer linear combination of two primes.

1. Introduction

Given a natural number N, how many pairs of primes less than or equal to N differ by a perfect square?

More generally, for an arbitrary polynomial $f \in \mathbb{Z}[x]$, we define

$$r_f(N) = \#\{(p_1, p_2) \in \mathcal{P}_N^2 \,:\, p_1 - p_2 \in f(\mathbb{N})\}, \tag{1}$$

where $\mathcal{P}$ denotes the primes and $\mathcal{P}_N = \mathcal{P} \cap \{1, \dots, N\}$. In this article we will first provide a heuristic, then argue rigorously with the Hardy-Littlewood Circle Method, for the following result.

Theorem 1. *If $f(x) = c_k x^k + \cdots + c_1 x + c_0 \in \mathbb{Z}[x]$, $c_k > 0$, and $k \geq 1$, then*

$$r_f(N) = \prod_{p \in \mathcal{P}} \left(1 + \frac{z_f(p) - 1}{(p-1)^2}\right) \frac{1}{c_k^{1/k}} \frac{k}{k+1} \frac{N^{(k+1)/k}}{\log^2 N} + O\left(\frac{N^{(k+1)/k} \log\log N}{\log^3 N}\right), \tag{2}$$

where $z_f(p)$ denotes the number of roots of f modulo p, and the implied constant depends only on f.

We note that due to the sufficient decay of $(z_f(p)-1)/(p-1)^2$, the product in the leading term is always finite, and is only 0 if one of the individual terms is 0. This only occurs if $z_f(2) = 0$, yielding a meaningful result as long as f has a root modulo 2. Also, if $f(x) = cx^k$, then the product collapses to $\prod_{p|c} \frac{p}{p-1}$, including its complete disappearance if $f(x) = x^k$.

Remark. The condition that f has a root modulo 2 is one piece of the more stringent condition that f is an "intersective polynomial", i.e. f has a root modulo n for all $n \in \mathbb{N}$. This additional assumption on f is necessary and sufficient to conclude the same order of growth for the count analogous to $r_f(N)$ for any subset of the primes of positive relative upper density, shown in a recent result of Thái Hoàng Lé [4].

While there is an extensive literature devoted to questions of this type going back some 60 or 70 years, namely Goldbach type problems for polynomials (see for example [3], [10], [7] and [1]), it appears that the precise result stated in Theorem 1 above has not been considered previously.

2. Heuristics

We begin by addressing the original motivating question, temporarily letting $r(N) = r_{x^2}(N)$. If we consider the discrete derivative $r'(N) = r(N) - r(N-1)$, we see that $r'(N) = 0$ unless $N \in \mathcal{P}$, which by the Prime Number Theorem occurs with probability $1/\log N$. If N does happen to be prime, then it induces $|\mathcal{P}_{N-1}| \sim N/\log N$ new positive differences of primes, which are each perfect squares with probability $1/\sqrt{N}$.

Therefore, we may well expect $r'(N)$ to grow on average like $\sqrt{N}/\log^2 N$, which in turns leads, via a simple summation by parts argument, to the prediction that

$$r(N) = \sum_{k=1}^{N} r'(k) = \frac{2}{3} \frac{N^{3/2}}{\log^2 N} + \text{error}. \tag{3}$$

2000 *Mathematics Subject Classification.* 11P05, 11P32, 11P55.

This heuristic immediately generalizes to predict

$$r_{x^k}(N) = \frac{k}{k+1}\frac{N^{(k+1)/k}}{\log^2 N} + \text{error}, \tag{4}$$

but it fails to account for potential congruence biases.

In the spirit of generalization, we fix $f(x) = c_k x^k + \cdots + c_1 x + c_0 \in \mathbb{Z}[x]$. Note that $r_f(N) = r_{-f}(N)$, so we can assume without loss of generality that $c_k > 0$. We must consider how $r_f(N)$ should compare to the prediction (4) above. One difference is that the leading coefficient c_k causes the number of points in the image of f that are at most N to be reduced asymptotically by a factor of $c_k^{-1/k}$. We must also consider local congruence biases of both the primes and the polynomial. For example, if we took $f(x) = 2x^k + 1$, then the second prime in each pair would have to be 2, $r_f(N)$ would clearly then be bounded by $|\mathcal{P}_N| \sim N/\log N$, and we couldn't possibly obtain the same order of growth as predicted in (4). The reason for this collapse is of course that almost all primes are odd, so almost no differences of primes are odd. More generally, we can use knowledge about the distribution of primes to gain knowledge about the distribution of differences of primes, and use these facts to investigate which polynomials these differences favor and avoid.

In order to exhaust all possible congruence biases, we must consider the distribution of differences of primes and the image of our polynomial modulo n for each $n \in \mathbb{N}$. However, if two moduli n and m are coprime, then congruence modulo n and congruence modulo m are independent events (by the Chinese Remainder Theorem). Therefore, we can determine all bias by investigating modulo arbitrarily large powers of each prime. Toward this end, we consider an arbitrary prime p, and we proceed probabilistically. We note that if two sets A and B are uniformly distributed modulo p^n, and we randomly select $a \in A$ and $b \in B$, then the probability that $a \equiv b$ modulo p^n, which we denote by $\mathbf{P}(a \equiv b \bmod p^n)$, is $1/p^n$. We compare the analogous probability for our specific sets with this expectation by defining a "bias factor"

$$b_f(p^n) = \frac{\mathbf{P}(p_1 - p_2 \equiv f(d) \bmod p^n)}{1/p^n} \tag{5}$$

and further noting that

$$\begin{aligned} b_f(p) &= p^n \sum_{a=0}^{p^n-1} \mathbf{P}(p_1 - p_2 \equiv a \bmod p^n)\mathbf{P}(f(d) \equiv a \bmod p^n) \\ &= \sum_{a=0}^{p^n-1} \mathbf{P}(p_1 - p_2 \equiv a \bmod p^n)(\#\text{ solutions to } f(x) = a \text{ in } \mathbb{Z}/p^n\mathbb{Z}). \end{aligned}$$

Once we determine these biases, we can make a prediction of the form

$$r_f(N) = \prod_{p\in\mathcal{P}} \limsup_{n\to\infty} b_f(p^n)\,\frac{1}{c_k^{1/k}}\frac{k}{k+1}\frac{N^{(k+1)/k}}{\log^2 N} + \text{error}. \tag{6}$$

This formulation turns out to be unnecessarily frightening. The Prime Number Theorem for Arithmetic Progressions due to Siegel and Walfisz states that for any modulus m, the primes are evenly distributed among the congruence classes coprime to m. Therefore, the only biases of the primes, and hence the only biases of differences of primes, are related the coprimality. However, we know that $(a, p) = 1$ if and only if $(a, p^n) = 1$ for all n, which means that $\mathbf{P}(p_1 - p_2 \equiv a \bmod p^n) = \mathbf{P}(p_1 - p_2 \equiv b \bmod p^n)$ if $a \equiv b \bmod p$. One can easily show from the definition above that this implies $b_f(p^n) = b_f(p)$ for all n, greatly simplifying prediction (6) to

$$r_f(N) = \prod_{p\in\mathcal{P}} b_f(p)\frac{1}{c_k^{1/k}}\frac{k}{k+1}\frac{N^{(k+1)/k}}{\log^2 N} + \text{error}. \tag{7}$$

Now, for each fixed prime p, it follows from the Siegel-Walfisz Theorem that primes are congruent to $1, 2, \ldots, p-1$ modulo p each with probability $1/(p-1)$ and consequently that the difference of two randomly selected primes is congruent to 0 modulo p with probability $1/(p-1)$, and congruent to $1, 2, \ldots, p-1$

modulo p each with probability $(p-2)/(p-1)^2$. From this observation it follows that the bias factor $b_f(p)$ is completely determined by the number of roots of f modulo p, a quantity we shall denote by $z_f(p)$. In fact,

$$b_f(p) = \frac{1}{p-1} z_f(p) + \frac{p-2}{(p-1)^2}\left(p - z_f(p)\right) = 1 + \frac{z_f(p)-1}{(p-1)^2}. \tag{8}$$

This completes the heuristic for the result in Theorem 1, and indicates that the only "fatal" obstruction toward the expected order of growth for $r_f(N)$ is the modulo 2 consideration noted above.

3. Rigorous Treatment via the Hardy-Littlewood Circle Method

In this section we give a proof of Theorem 1 using the circle method developed by Hardy and Littlewood. As is standard, we begin by weighting the characteristic function of the primes with a logarithm to obtain a more uniform distribution. This yields a weighted count intimately related to $r_f(N)$, which we define below using a truncated von Mangoldt function.

Definition 1. *We define $\Lambda_N : \mathbb{Z} \to [0, \infty)$ by*

$$\Lambda_N(n) = \begin{cases} \log p & \text{if } n = p^k \leq N, p \in \mathcal{P}, k \in \mathbb{N} \\ 0 & \text{else} \end{cases}, \tag{9}$$

and for $f \in \mathbb{Z}[x]$ as in Theorem 1, we define

$$R_f(N) = \sum_{d=1}^{M} \sum_{n \in \mathbb{Z}} \Lambda_N(n)\Lambda_N(n - f(d)), \tag{10}$$

where $M = (N/c_k)^{1/k}$.

The main result of this section, and indeed the whole article, is the following.

Theorem 2. *For $f \in \mathbb{Z}[x]$ as in Theorem 1 and any $A > 0$, we have*

$$R_f(N) = \mathfrak{S}(f) \frac{1}{c_k^{1/k}} \frac{k}{k+1} N^{(k+1)/k} + O\left(\frac{N^{(k+1)/k}}{\log^A N}\right), \tag{11}$$

where the implied constant depends only on f and A, and

$$\mathfrak{S}(f) = \sum_{q=1}^{\infty} \frac{\mu(q)^2}{q\phi(q)^2} \sum_{\substack{0 \leq a < q \\ (a,q)=1}} \sum_{r=0}^{q-1} e^{2\pi i f(r) a/q}, \tag{12}$$

where μ is the Möbius function, ϕ is the Euler totient function.

It is easy to see that Theorem 1 will follow from Theorem 2 and the following two lemmas.

Lemma 1. *If $f \in \mathbb{Z}[x]$ is as in Theorem 1 and $z_f(p)$ denotes the number of roots of f modulo p, then*

$$\sum_{q=1}^{\infty} \frac{\mu(q)^2}{q\phi(q)^2} \sum_{\substack{0 \leq a < q \\ (a,q)=1}} \sum_{r=0}^{q-1} e^{2\pi i f(r) a/q} = \prod_{p \in \mathcal{P}} \left(1 + \frac{z_f(p) - 1}{(p-1)^2}\right). \tag{13}$$

Lemma 2. *If $f \in \mathbb{Z}[x]$ is as in Theorem 1, then*

$$r_f(N) = \frac{R_f(N)}{\log^2 N} + O\left(\frac{N^{(k+1)/k} \log\log N}{\log^3 N}\right). \tag{14}$$

The proof of both Lemma 1 and Lemma 2 are standard exercises, for completeness however we include their proofs in Section 4 below. In the remainder of this section we present the proof of Theorem 2.

3.1. **Proof of Theorem 2.** We fix f as in Theorem 1 and $A > 0$.

In order to apply the circle method, we need to reformulate the weighted count $R_f(N)$ on the transform side as an integral over the circle. Specifically, it follows from the usual orthogonality relation

$$\int_0^1 e^{2\pi i n\alpha}\, d\alpha = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n \in \mathbb{Z} \setminus \{0\} \end{cases} \tag{15}$$

that

$$R_f(N) = \int_0^1 |\widehat{\Lambda_N}(\alpha)|^2 S_M(\alpha)\, d\alpha, \tag{16}$$

where

$$\widehat{\Lambda_N}(\alpha) = \sum_{n\in\mathbb{Z}} \Lambda_N(n) e^{-2\pi i n\alpha} \tag{17}$$

and

$$S_M(\alpha) = \sum_{d=1}^{M} e^{2\pi i f(d)\alpha}. \tag{18}$$

As usual, we wish to partition the circle into two pieces, one the collection of points near rationals with small denominator, from which the primary contribution to the integral (16) will stem, and the other simply the complement. In order to appropriately make this partition official, we need a parameter yielded from the following estimate on high moments of the Weyl sum S_M.

Lemma 3. *There exists $s_0(k) = O(k^2 \log k)$ such that if $s \geq s_0(k)$, then*

$$\int_0^1 |S_M(\alpha)|^s\, d\alpha = O(M^{s-k}). \tag{19}$$

For a proof of this estimate see Proposition 3.3 in [5]. We remark that it follows from recent work of Trevor Wooley [9], that one can in fact take $s_0(k) = 2k(k+1)$.

Definition 2. *Fixing $s = \max\{s_0(k), 4k\}$ we let $B = s(A+1)+8$ and define $\mathfrak{M}$, the **major arcs**, by*

$$\mathfrak{M} = \bigcup_{q=1}^{\log^B N} \bigcup_{\substack{0\leq a<q \\ (a,q)=1}} \mathbf{M}_{a/q}, \tag{20}$$

where

$$\mathbf{M}_{a/q} = \left\{ \alpha \in [0,1] \,:\, \left|\alpha - \frac{a}{q}\right| < \frac{\log^B N}{N} \right\}, \tag{21}$$

*and $\mathfrak{m}$, the **minor arcs**, by $\mathfrak{m} = [0,1] \setminus \mathfrak{M}$.*

3.1.1. *Estimates on the minor arcs.* We first focus our attention on the minor arcs, in an attempt to absorb their contribution to the integral into the error term in Theorem 2. To this end, we invoke the minor arc estimate on $\widehat{\Lambda_N}$ that was Vinogradov's main achievement in solving the ternary Goldbach problem unconditionally.

Lemma 4 (Vinogradov, Vaughan)**.** *If $\alpha \in [0,1]$, $|\alpha - a/q| < 1/q^2$ and $(a,q) = 1$, then*

$$\widehat{\Lambda_N}(\alpha) = O\left(\log^4 N \left(N/q^{1/2} + N^{4/5} + N^{1/2} q^{1/2}\right)\right). \tag{22}$$

The argument for this estimate was greatly simplified by Vaughan, and an exposition of it can be found in [8], for example. It is now a relatively straightforward matter to see that Lemmas 3 and 4 combine to give the following desired estimate for our integral over the minor arcs.

Corollary 1 (Minor arc estimate)**.**

$$\int_{\mathfrak{m}} |\widehat{\Lambda_N}(\alpha)|^2 S_M(\alpha)\, d\alpha = O\left(\frac{N^{(k+1)/k}}{\log^A N}\right). \tag{23}$$

Proof. First we fix an arbitrary $\alpha \in \mathfrak{m}$. By the pigeonhole principle, there must exist $1 \leq q \leq N/\log^B N$ such that

$$\left|\alpha - \frac{a}{q}\right| < \frac{\log^B N}{qN} \leq \frac{1}{q^2}.$$

However, by the definition of $\mathfrak{m}$, it must be the case that $q \geq \log^B N$. Combining these bounds on q with Lemma 4 gives

$$\widehat{\Lambda_N}(\alpha) = O\left(\frac{N}{\log^{(B-8)/2} N}\right). \tag{24}$$

Now, we invoke Hölder's Inequality in order to utilize our control on the higher moments of S_M. Indeed, applying Hölder's inequality, Lemma 3 and (24) we obtain

$$\begin{aligned}\left|\int_{\mathfrak{m}} |\widehat{\Lambda_N}(\alpha)|^2 S_M(\alpha)\, d\alpha\right| &\leq \left(\int_{\mathfrak{m}} |\widehat{\Lambda_N}(\alpha)|^{2s/(s-1)}\, d\alpha\right)^{(s-1)/s} \left(\int_0^1 |S_M(\alpha)|^s\, d\alpha\right)^{1/s} \\ &= O\left(\left(\frac{N}{\log^{(B-8)/2} N}\right)^{2/s} \left(\int_0^1 |\widehat{\Lambda_N}(\alpha)|^2\, d\alpha\right)^{(s-1)/s} M^{\frac{s-k}{s}}\right) \\ &= O\left(\frac{N^{(k+1)/k}}{\log^A N}\right).\end{aligned}$$

In the last line above we have used the fact that $M = O(N^{1/k})$ and $B = s(A+1)+8$, together with the deeper fact that

$$\int_0^1 |\widehat{\Lambda_N}(\alpha)|^2\, d\alpha = O(N \log N),$$

which is a standard consequence of the Prime Number Theorem (and Plancherel's Identity). □

3.1.2. *Estimates on the major arcs.* For the major arcs, we invoke the most classical of estimates.

Lemma 5 (Major arc estimates)**.** *If $\alpha = a/q + \beta$ with $1 \leq q \leq \log^B N$, $(a,q) = 1$ and $|\beta| < \log^B N/N$, then*

$$S_M(\alpha) = q^{-1} S_{a/q} I_M(\beta) + O(\log^{2B} N) \tag{25}$$

and

$$\widehat{\Lambda_N}(\alpha) = \frac{\mu(q)}{\phi(q)} \nu_N(\beta) + O(N e^{-c\sqrt{\log N}}) \tag{26}$$

for some $c = c(B) > 0$, where μ is the Möbius function, ϕ is the Euler totient function,

$$S_{a/q} = \sum_{r=0}^{q-1} e^{2\pi i f(r) a/q}, \quad I_M(\beta) = \int_0^M e^{2\pi i f(x)\beta}\, dx, \quad \text{and} \quad \nu_N(\beta) = \sum_{n=1}^{N} e^{-2\pi i n\beta}.$$

Estimate (25) can be found in any discussion of Waring's Problem (Lemma 4.2 of [2] for example), while estimate (26) follows from the Siegel-Walfisz Theorem on primes in arithmetic progressions, and is included in any discussion of Vinogradov's Three Primes Theorem (see for example Lemma 3.1 in [8]). We further note that

$$\nu_N(\beta) = \int_0^N e^{-2\pi i x\beta} dx + O(\log^B N), \tag{27}$$

thus we maintain the same asymptotic by replacing the sum with the integral.

We are now ready to combine all of our estimates and establish Theorem 2.

3.1.3. *Proof of Theorem 2.* From (16) and Corollary 1, and noting that for large N the individual major arcs are pairwise disjoint, we have

$$R_f(N) = \sum_{q=1}^{\log^B N} \sum_{\substack{0\le a<q\\(a,q)=1}} \int_{\mathbf{M}_{a/q}} |\widehat{\Lambda_N}(\alpha)|^2 S_M(\alpha)\, d\alpha + O\left(\frac{N^{(k+1)/k}}{\log^A N}\right). \tag{28}$$

From (25), (26), and (27), we then obtain

$$R_f(N) = \underbrace{\sum_{q=1}^{\log^B N} \frac{\mu(q)^2}{q\phi(q)^2} \sum_{\substack{0\le a<q\\(a,q)=1}} S_{a/q}}_{(\star)} \underbrace{\int_{|\beta|<(\log^B N)/N} \left|\int_0^N e^{-2\pi i y\beta}\, dy\right|^2 I_M(\beta)\, d\beta}_{(\star\star)} + O\left(\frac{N^{(k+1)/k}}{\log^A N}\right). \tag{29}$$

The reader can now easily see that Theorem 2 will be an immediate consequence of the following estimates for $(\star)$ and $(\star\star)$ respectively:

$$\sum_{q=1}^{\log^B N} \frac{\mu(q)^2}{q\phi(q)^2} \sum_{\substack{0\le a<q\\(a,q)=1}} S_{a/q} = \sum_{q=1}^{\infty} \frac{\mu(q)^2}{q\phi(q)^2} \sum_{\substack{0\le a<q\\(a,q)=1}} S_{a/q} + O\left(\frac{1}{\log^A N}\right) \tag{30}$$

and

$$\int_{|\beta|<(\log^B N)/N} \left|\int_0^N e^{-2\pi i y\beta}\, dy\right|^2 I_M(\beta)\, d\beta = \frac{k}{k+1} MN + O\left(\frac{MN}{\log^B N}\right). \tag{31}$$

In order to establish (30) we need an estimate on the magnitude of the Gauss sum $S_{a/q}$ that beats the trivial bound of q. Theorem 7.1 in [8] tells us that

$$S_{a/q} = O_\epsilon(q^{(1-1/k)+\epsilon}) \text{ for all } \epsilon > 0, \tag{32}$$

so in particular, we know that $S_{a/q} = O(q^{1-1/2k})$, and since $\phi(q) \ge Cq^{1-\frac{1}{4k}}$ (trivially), we can deduce from this that

$$\begin{aligned}\Big|\sum_{q=\log^B N}^{\infty} \frac{\mu(q)^2}{q\phi(q)^2} \sum_{\substack{0\le a<q\\(a,q)=1}} S_{a/q}\Big| &= O\left(\sum_{q=\log^B N}^{\infty} \frac{1}{q^{1+1/4k}}\right)\\ &= O\left(\frac{1}{\log^{B/4k} N}\right).\end{aligned}$$

Noting that $B/4k > A$, we see that (30) immediately follows.

In order to establish (31) we initially note that

$$I_M(\beta) - \int_0^M e^{2\pi i c_k x^k \beta}\, dx = O\left(\int_0^M x^{k-1}\beta\, dx\right) = O(\log^B N) \tag{33}$$

and that substituting this into the left hand side of (31) gives

$$\int_{|\beta|<(\log^B N)/N} \left|\int_0^N e^{-2\pi i y\beta}\, dy\right|^2 I_M(\beta)\, d\beta = \int_{|\beta|<(\log^B N)/N} \left|\int_0^N e^{-2\pi i y\beta}\, dy\right|^2 \left(\int_0^M e^{2\pi i c_k x^k \beta} dx\right) d\beta + O(N\log^{2B} N).$$

We note further that after three changes of variables (namely $x := x/M$, $\beta := N\beta$, $y := y/N$), the integral on the right hand side of the above identity can be seen to satisfy

$$\int_{|\beta|<\log^B N} \left|\int_0^N e^{-2\pi i y\beta}\, dy\right|^2 \left(\int_0^M e^{2\pi i c_k x^k \beta} dx\right) d\beta = MN \int_{|\beta|<\log^B N} |\widehat{1_{[0,1]}}(\beta)|^2 \left(\int_0^1 e^{2\pi i x^k \beta} dx\right) d\beta, \tag{34}$$

where $\widehat{1_{[0,1]}}(\beta) = \int_0^1 e^{-2\pi i y\beta}\,dy$ is the usual Euclidean Fourier transform of the characteristic function of the unit interval. Utilizing the standard (and easily verified) fact that $|\widehat{1_{[0,1]}}(\beta)|^2 = O(1+|\beta|^2)^{-1}$, it then follows that

$$\int\limits_{|\beta|\geq \log^B N} |\widehat{1_{[0,1]}}(\beta)|^2 \left(\int_0^1 e^{2\pi i x^k \beta} dx\right) d\beta = O\left(\frac{1}{\log^B N}\right) \tag{35}$$

and hence that

$$\int\limits_{|\beta|< \log^B N} |\widehat{1_{[0,1]}}(\beta)|^2 \left(\int_0^1 e^{2\pi i x^k \beta} dx\right) d\beta = \int_{\mathbb{R}} |\widehat{1_{[0,1]}}(\beta)|^2 \left(\int_0^1 e^{2\pi i x^k \beta} dx\right) d\beta + O\left(\frac{1}{\log^B N}\right). \tag{36}$$

Thus, in order to establish (31) it finally suffices to show

$$\int_{\mathbb{R}} |\widehat{1_{[0,1]}}(\beta)|^2 \left(\int_0^1 e^{2\pi i x^k \beta} dx\right) d\beta = \frac{k}{k+1}. \tag{37}$$

Changing variables again ($x := x^k$), letting $g(x) = x^{(1-k)/k}/k\, 1_{[0,1]}(x)$, and applying Parseval's Identity, it follows that

$$\begin{aligned}
\int_{\mathbb{R}} |\widehat{1_{[0,1]}}(\beta)|^2 \left(\int_0^1 e^{2\pi i x^k \beta} dx\right) d\beta &= \int_{\mathbb{R}} |\widehat{1_{[0,1]}}(\beta)|^2\, \overline{\widehat{g}(\beta)}\, d\beta \\
&= \int_{\mathbb{R}} 1_{[0,1]}(x) \int_{\mathbb{R}} g(y) 1_{[0,1]}(x-y)\, dy\, dx \\
&= \int_0^1 \int_0^x \frac{y^{(1-k)/k}}{k}\, dy\, dx \\
&= \frac{k}{k+1}
\end{aligned}$$

as required. □

4. Proof of Lemmas 1 and 2

4.1. **Proof of Lemma 1.** First we write

$$\sum_{q=1}^{\infty} \frac{\mu(q)^2}{q\phi(q)^2} \sum_{\substack{0\leq a<q \\ (a,q)=1}} \sum_{r=0}^{q-1} e^{2\pi i f(r) a/q} = \sum_{q=1}^{\infty} F(q), \tag{38}$$

where

$$F(q) = \frac{\mu(q)^2}{q\phi(q)^2} \sum_{\substack{0\leq a<q \\ (a,q)=1}} S_{a/q} \quad \text{and} \quad S_{a/q} = \sum_{r=0}^{q-1} e^{2\pi i f(r) a/q}. \tag{39}$$

Since q, $\mu(q)$, and $\phi(q)$ are all multiplicative functions, and it is a standard exercise to verify that

$$\sum_{\substack{0\leq a<q \\ (a,q)=1}} S_{a/q} \tag{40}$$

is also multiplicative (see for example Lemma 5.1 of [2] for a proof), it follows that $F(q)$ is a multiplicative function. Furthermore, during the proof of Theorem 2, we used estimate (32) to establish that $F(q) = O(q^{-1-1/4k})$, so in particular we know that $\sum_{q=1}^{\infty} F(q)$ is an absolutely convergent sum.

These two facts combine to yield the usual Euler product formula, namely

$$\sum_{q=1}^{\infty} F(q) = \prod_{p\in\mathcal{P}} \sum_{n=0}^{\infty} F(p^n) = \prod_{p\in\mathcal{P}} (1 + F(p)), \tag{41}$$

with the great simplification occurring due the presence of the Möbius function, which forces $F(p^n) = 0$ for every prime p, provided $n \geq 2$.

Noting that

$$F(p) = \frac{1}{p(p-1)^2} \sum_{a=1}^{p-1} \sum_{r=o}^{p-1} e^{2\pi i f(r)a/p}, \tag{42}$$

we see that it now suffices to simply verify that

$$\sum_{r=0}^{p-1} \sum_{a=1}^{p-1} e^{2\pi i f(r)a/p} = p(z_f(p) - 1) \tag{43}$$

for all primes p. This is straightforward, since if $f(r) \equiv 0$ modulo p, then the exponential term is identically 1, and

$$\sum_{a=1}^{p-1} e^{2\pi i f(r)a/p} = p - 1.$$

While if $f(r) \not\equiv 0$ modulo p, then the inner sum goes over all p-th roots of unity except 1, hence

$$\sum_{a=1}^{p-1} e^{2\pi i f(r)a/p} = -1.$$

Therefore, since $z_f(p)$ denotes the number of roots of f mod p, it follows that

$$\sum_{r=0}^{p-1} \sum_{a=1}^{p-1} e^{2\pi i f(r)a/p} = z_f(p)(p-1) - (p - z_f(p)) = p(z_f(p) - 1) \tag{44}$$

as required. □

4.2. **Proof of Lemma 2.** First, we let

$$M' = \min\{n \in \mathbb{N} \mid f(d) > N \text{ for all } d > n\}, \tag{45}$$

so

$$\sum_{d=1}^{M'} \sum_{n \in \mathbb{Z}} 1_{\mathcal{P}_N}(n) 1_{\mathcal{P}_N}(n - f(d)) = \#\{(p, d) \in \mathcal{P}_N \times \mathbb{N} : p - f(d) \in \mathcal{P}_N\} = r_f(N) + O\left(\frac{N}{\log N}\right), \tag{46}$$

where the error term accounts for the finitely many instances of "double-counting" in the polynomial. This convenience is the primary reason for restricting to the image of $\mathbb{N}$ as opposed to $\mathbb{Z}$, making all nonconstant polynomials "eventually injective". We also note that $M' = M + O_\epsilon(N^\epsilon)$ for every $\epsilon > 0$, and in particular,

$$r_f(N) = \sum_{d=1}^{M} \sum_{n \in \mathbb{Z}} 1_{\mathcal{P}_N}(n) 1_{\mathcal{P}_N}(n - f(d)) + O\left(N^{1+1/4k}\right). \tag{47}$$

We now make the usual observation that the contribution to the von Mangoldt function from proper prime powers is negligible. Namely,

$$R_f(N) = \sum_{d=1}^{M} \sum_{n \in \mathbb{Z}} \log(n) \log(n - f(d)) 1_{\mathcal{P}_N}(n) 1_{\mathcal{P}_N}(n - f(d)) + O\left(N^{(k+2)/2k} \log^2 N\right). \tag{48}$$

By (46), (47), and (48), noting the trivial upper bound

$$\sum_{d=1}^{M} \sum_{n \in \mathbb{Z}} 1_{\mathcal{P}_N}(n) 1_{\mathcal{P}_N}(n - f(d)) \geq \frac{1}{\log^2 N} \sum_{d=1}^{M} \sum_{n \in \mathbb{Z}} \log(n) \log(n - f(d)) 1_{\mathcal{P}_N}(n) 1_{\mathcal{P}_N}(n - f(d)), \tag{49}$$

and changing variables ($n := n - f(d)$), it now suffices to show

$$\sum_{d=1}^{M}\sum_{n\in\mathbb{Z}} 1_{\mathcal{P}_N}(n)1_{\mathcal{P}_N}(n+f(d)) \leq \frac{1}{\log^2 N}\sum_{d=1}^{M}\sum_{n\in\mathbb{Z}} \log(n)\log(n+f(d))1_{\mathcal{P}_N}(n)1_{\mathcal{P}_N}(n+f(d)) + O\left(\frac{N^{(k+1)/k}\log\log N}{\log^3 N}\right). \tag{50}$$

Toward this end, we see that for any $\delta > 0$,

$$\begin{aligned}\sum_{d=1}^{M}\sum_{n\in\mathbb{Z}} 1_{\mathcal{P}_N}(n)1_{\mathcal{P}_N}(n+f(d)) &= \sum_{d=1}^{M}\sum_{n=N^{1-\delta}}^{N} 1_{\mathcal{P}_N}(n)1_{\mathcal{P}_N}(n+f(d)) + \sum_{d=1}^{M}\sum_{n=1}^{N^{1-\delta}} 1_{\mathcal{P}_N}(n)1_{\mathcal{P}_N}(n+f(d)) \\ &\leq \frac{1}{(1-\delta)^2\log^2 N}\sum_{d=1}^{M}\sum_{n\in\mathbb{Z}} \log(n)\log(n+f(d))1_{\mathcal{P}_N}(n)1_{\mathcal{P}_N}(n+f(d)) \\ &\quad + O\left(\frac{N^{(k+1)/k}}{N^{\delta}\log N}\right).\end{aligned}$$

Setting

$$\delta = \frac{2\log\log N}{\log N},$$

and noting from Theorem 2 and (48) that

$$\sum_{d=1}^{M}\sum_{n\in\mathbb{Z}} \log(n)\log(n+f(d))1_{\mathcal{P}_N}(n)1_{\mathcal{P}_N}(n+f(d)) = O\left(N^{(k+1)/k}\right), \tag{51}$$

the lemma follows. □

5. Further Generalization

A natural question: How would our results change if, instead of differences of primes, we counted sums of primes that lie in the image of a fixed polynomial? Even more generally, for $f \in \mathbb{Z}[x]$, $N \in \mathbb{N}$, and arbitrary integers $a_1, a_2 \neq 0$, we define

$$r_{f,a_1,a_2}(N) = \#\left\{(p_1,p_2) \in \mathcal{P}_N^2 \,:\, a_1p_1 + a_2p_2 \in f(\mathbb{N})\right\}. \tag{52}$$

Again we note that $r_{f,a_1,a_2}(N) = r_{-f,-a_1,-a_2}(N)$, so we can assume that f has positive leading coefficient. Under this assumption, if a_1 and a_2 were both negative, one can see that r_{f,a_1,a_2} would be uniformly bounded in N, so we assume without loss of generality that $a_1 > 0$.

Careful adaptation of our initial heuristic with the discrete derivative yields an additional factor of

$$C_k(a_1,a_2) = \begin{cases} \dfrac{(a_1+a_2)^{1/k}-a_1^{1/k}}{a_2} + \dfrac{(a_1+a_2)^{1/k}-a_2^{1/k}}{a_1} & \text{if } a_2 > 0 \\ \dfrac{(a_1+a_2)^{1/k}-a_1^{1/k}}{a_2} + \dfrac{(a_1+a_2)^{1/k}}{a_1} & \text{if } 0 > a_2 \geq -a_1 \\ -\dfrac{a_1^{1/k}}{a_2} & \text{if } a_2 \leq -a_1 \end{cases}. \tag{53}$$

Note for example that $C_k(1,-1) = 1$ for all k, as it should to agree with our previous discussion, and $C_k(1,1) = 2(2^{1/k}-1)$.

To adapt our heuristic for local congruence biases, it is useful to partition the primes based on how many of the coefficients a_1, a_2 they divide. Namely, we define

$$\begin{aligned}\mathcal{P}_0 &= \{p \in \mathcal{P} : p \nmid a_1a_2\} \\ \mathcal{P}_1 &= \{p \in \mathcal{P} : p \mid a_1a_2,\ (p \nmid a_1 \text{ or } p \nmid a_2)\} \\ \mathcal{P}_2 &= \{p \in \mathcal{P} : p \mid a_1,\ p \mid a_2\}.\end{aligned} \tag{54}$$

Hopefully context will prevent any confusion between this notation and the notation $\mathcal{P}_N = \mathcal{P} \cap \{1, \dots, N\}$.

Adaptation of our heuristic with the Siegel-Walfisz Theorem yields local bias factors

$$b_{f,a_1,a_2}(p) = \begin{cases} 1 + \dfrac{z_f(p) - 1}{(p-1)^2} & \text{if } p \in \mathcal{P}_0 \\ \dfrac{p - z_f(p)}{p-1} & \text{if } p \in \mathcal{P}_1 \\ z_f(p) & \text{if } p \in \mathcal{P}_2 \end{cases}. \tag{55}$$

Note that we still have $b_{x^k,a_1,a_2}(p) = 1$ for all $p \in \mathcal{P}$, $k \in \mathbb{N}$.

Indeed, one can establish the following generalization of Theorem 1.

Theorem 3. *If $a_1, a_2 \in \mathbb{Z}$, $a_1 > 0$, $a_2 \neq 0$, $f(x) = c_k x^k + \cdots + c_1 x + c_0 \in \mathbb{Z}[x]$, $c_k > 0$, and $k \geq 1$, then*

$$\begin{aligned} r_{f,a_1,a_2}(N) = &\prod_{p \in \mathcal{P}_0} \left(1 + \frac{z_f(p) - 1}{(p-1)^2}\right) \prod_{p \in \mathcal{P}_1} \frac{p - z_f(p)}{p-1} \prod_{p \in \mathcal{P}_2} z_f(p) \frac{C_k(a_1, a_2)}{c_k^{1/k}} \frac{k}{k+1} \frac{N^{(k+1)/k}}{\log^2 N} \\ &+ O\left(\frac{N^{(k+1)/k} \log\log N}{\log^3 N}\right), \end{aligned} \tag{56}$$

where the implied constant depends on f, a_1, and a_2.

The vast majority of the argument for Theorem 3 is immediately analogous to the proof of Theorem 1, so rather than starting from scratch, we proceed by highlighting the key differences. Much like (16), we start by giving a weighted, transform side formulation of our count, defining

$$R_{f,a_1,a_2} = \int_0^1 \widehat{\Lambda_N}(a_1\alpha)\widehat{\Lambda_N}(a_2\alpha) S_M(\alpha) d\alpha, \tag{57}$$

where now

$$M = \left(\frac{\max\{a_1, a_1 + a_2\} N}{c_k}\right)^{1/k}. \tag{58}$$

The procedure for absorbing the minor arcs into the error term goes through just as before, but some subtlety arises in the major arcs. For example, if α is in some major arc $\mathbf{M}_{a/q}$, then (26) still gives an estimate for $\widehat{\Lambda_N}(a_1\alpha)$, just potentially with a different denominator, based on the factors shared by a_1 and the original denominator q. More officially, we decompose each squarefree number based on our previous partition of the primes, i.e. we define

$$\mathbb{N}_i = \{p_1 p_2 \cdots p_\ell : p_j \in \mathcal{P}_i \text{ distinct}\} \tag{59}$$

for $i = 0, 1, 2$, where each set includes 1, so every squarefree number can be written uniquely as qm_1m_2 with $q \in \mathbb{N}_0$, $m_1 \in \mathbb{N}_1$, and $m_2 \in \mathbb{N}_2$. (Recall that due to the presence of the Möbius function, only major arcs with squarefree denominators contribute to the main term.) Now, one can check that by (26), if $\alpha = a/qm_1m_2 + \beta$, $qm_1m_2 \leq \log^B N$, $(a, qm_1m_2) = 1$ and $|\beta| < \log^B N/N$, then

$$\widehat{\Lambda_N}(a_1\alpha)\widehat{\Lambda_N}(a_2\alpha) = \frac{\mu(q)^2\mu(m_1)}{\phi(q)^2\phi(m_1)} \nu_N(a_1\beta)\nu_N(a_2\beta) + O(N^2 e^{-c\sqrt{\log n}}) \tag{60}$$

for some $c = c(B, a_1, a_2) > 0$.

In place of the integral observed in (34), we now have

$$\int_{|\beta| < \log^B N/N} \left(\int_0^N e^{-2\pi i a_1 y\beta}\, dy\right)\left(\int_0^N e^{-2\pi i a_2 z\beta}\, dz\right)\left(\int_0^M e^{2\pi i c_k x^k \beta}\, dx\right) d\beta, \tag{61}$$

which with similar tricks of changing variables and applying Parseval's Identity can be seen to equal

$$C_k(a_1, a_2) c_k^{-1/k} \frac{k}{k+1} N^{(k+1)/k} + O\left(\frac{N^{(k+1)/k}}{\log^B N}\right). \tag{62}$$

In place of the singular series $\mathfrak{S}(f)$, we now have

$$\mathfrak{S}(f,a_1,a_2) = \sum_{\substack{q\in\mathbb{N}_0\\ m_1\in\mathbb{N}_1\\ m_2\in\mathbb{N}_2}} \frac{\mu(q)^2\mu(m_1)}{\phi(q)^2\phi(m_1)}\frac{1}{qm_1m_2} \sum_{\substack{0\le a<qm_1m_2\\ (a,qm_1m_2)=1}} S_{a/qm_1m_2}, \tag{63}$$

which is still an absolutely convergent sum of a multiplicative function that is zero for non-squarefree numbers, hence we still have the Euler product formula

$$\mathfrak{S}(f,a_1,a_2) = \prod_{p\in\mathcal{P}}(1+F(p)),$$

where now

$$F(p) = \begin{cases} \dfrac{1}{p(p-1)^2}\displaystyle\sum_{a=0}^{p-1} S_{a/p} & \text{if } p\in\mathcal{P}_0 \\ -\dfrac{1}{p(p-1)}\displaystyle\sum_{a=0}^{p-1} S_{a/p} & \text{if } p\in\mathcal{P}_1 \\ \dfrac{1}{p}\displaystyle\sum_{a=0}^{p-1} S_{a/p} & \text{if } p\in\mathcal{P}_2 \end{cases}. \tag{64}$$

Recalling our previous calculation that

$$\sum_{a=0}^{p-1} S_{a/p} = p(z_f(p)-1),$$

the result follows. □

Remark. Generalization of this result to linear combinations of more than two primes is definitely possible, but not particularly useful, as the circle method machinery is already capable of tackling the strictly harder task of counting solutions to $a_1p_1+\cdots+a_\ell p_\ell = N$ for fixed N and $\ell\ge 3$. A useful pursuit, however, may be to extend the result to count solutions of the form $a_1p_1^{k_1}+a_2p_2^{k_2} = f(d)$, where k_1 and k_2 are arbitrary fixed natural numbers.

References

[1] J. Brüdern, K. Kawada, T. D. Wooley, *Additive representation in thin sequences. II. The binary Goldbach problem*, Mathematika 47 (2000), no. 1-2, 117125 (2002).

[2] H. Davenport, *Analytic Methods for Diophantine Equations and Diophantine Inequalities*, Cambridge Univ. Press, Second Edition, 2005.

[3] H. Halberstam, *On the representation of large numbers as sums of squares, higher powers, and primes*, Proc. LMS 53 (1951), 363-380.

[4] Thái Hoàng Lé, *Intersective Polynomials and the Primes*, J. Number Theory 130 (2010), no. 8, 1705-1717.

[5] N. Lyall, Á. Magyar, *Simultaneous Polynomial Recurrence*, to appear in the Bulletin of the LMS, arXiv: 1010.345.

[6] M. Nathanson, *Additive Number Theory: The Classical Bases*, Springer GTM #164, 1996.

[7] A. Perelli, *Goldbach numbers represented by polynomials*, Rev. Mat. Iberoamericana 12 (1996), no. 2, 477490.

[8] R. C. Vaughan, *The Hardy-Littlewood Method*, Cambridge Univ. Press, Second Edition, 1997.

[9] T. D. Wooley, *Vinogradov's mean value theorem via efficient congruencing*, to appear in the Annals of Math.

[10] A. Zulauf, *On sums and differences of primes and squares, Compositio Math*, 13 (1958), 103–112

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

E-mail address: lyall@math.uga.edu

Department of Mathematics, The University of Georgia, Athens, GA 30602, USA

E-mail address: arice@math.uga.edu