TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 353, Number 7, Pages 2705–2724
S 0002-9947(01)02692-7
Article electronically published on March 12, 2001

PRIMES IN SHORT ARITHMETIC PROGRESSIONS WITH RAPIDLY INCREASING DIFFERENCES

P. D. T. A. ELLIOTT

Abstract. Primes are, on average, well distributed in short segments of arithmetic progressions, even if the associated moduli grow rapidly.

1. Statement of results

In this paper I establish two results concerning the distribution of prime numbers in short segments of residue classes to widely separated moduli.

Let $\Lambda(n)$ denote von Mangoldt's function, $\log p$ if n is a power of a prime p, zero otherwise. For integers $D > 0$ and r, let $\psi(x, D, r)$ denote the sum of the $\Lambda(n)$ over the positive integers not exceeding x which lie in the residue class $r \,(\mathrm{mod}\, D)$.

Let f be a polynomial of deg(ree) at least one, with integer coefficients, leading coefficient positive.

Theorem 1. *If $A > 0$, $4\beta \deg f < 1$, then*

$$\sum_{D\le x^\beta} \frac{\phi(f(D))}{D} \max_{(r,f(D))=1} \max_{y\le x} \left| \psi(y, f(D), r) - \frac{y}{\phi(f(D))} \right| \ll x(\log x)^{-A},$$

the summation confined to integers D for which the modulus $f(D)$ is positive.

In a recent paper, [5], investigating the solution of polynomial equations in primes, I employ a lower bound for a doublesum

$$\sum_{\substack{p+1=mf(q),x/2<m\le x\\ x^\alpha/2<q\le x^\alpha}}\sum r(m)$$

the sum taken over representations $m = (p + 1)f(q)^{-1}$ of integers m by primes p, q. My appeal in that paper to elaborate results of Linnik, Fogels, Gallagher and others may be replaced by the application of Theorem 1. In the notation of that paper, the lower bound improves to an asymptotic estimate

$$\left(1 + O\left(\frac{1}{(\log x)}\right)\right) \frac{\phi(\Delta)Px^{1+\alpha}}{4\alpha(1 + \alpha \deg f)\Delta\phi(P)(\log x)^2},$$

valid for each positive α not exceeding the β of Theorem 1.

Received by the editors January 7, 1999 and, in revised form, February 26, 2000.
2000 *Mathematics Subject Classification.* Primary 11N13; Secondary 11B25.
Partially supported by NSF contract DMS-9530690.

When f is a linear polynomial, the Bombieri–Vinogradov theorem guarantees the estimate of Theorem 1 for any $\beta < 1/2$ [3]. Under the tighter restriction $4\beta \deg f < 1$ a bound

$$\sum_{\frac{1}{2}x^\beta < D \le x^\beta} \max_{(r,f(D))=1} \max_{y\le x} \left| \psi(y, f(D), r) - \frac{y}{\phi(f(D))} \right| \ll x(\log x)^{-A}$$

in the case of a polynomial of degree 2 or more is immediate, but it is considerably weaker than the estimate $\ll x^{1+\beta-\beta \deg f}$ that follows from the well-known Brun–Titchmarsh theorem. Theorem 1 delivers the more appropriate $\ll x^{1+\beta-\beta\deg f}(\log x)^{-A}$.

The following result will illustrate the extent to which I employ the polynomial nature of the moduli $f(D)$.

Theorem 2. *Let $\delta > 0$, $A > 0$. Then*

$$\sum_{j=1}^{J} \max_{(r,sD_j)=1} \max_{y \le x} \left| \psi(y, sD_j, r) - \frac{\psi(y,s,r)}{\phi(D_j)} \right|$$
$$\ll \frac{x}{\phi(s)(\log x)^A} \sum_{j=1}^{J} \frac{1}{\phi(D_j)} + \frac{x}{s} \left(\sum_{\substack{u=1 \\ (D_u, D_v)>1}}^{J} \sum_{v=1}^{J} \frac{1}{D_u D_v} \right)^{1/2} (\log x)^{4200}$$

for all $x \ge 2$ and positive integers s, D_j satisfying $sD_j \le x^{\frac{1}{4}-\delta}$, $1 \le s \le x^\delta$, and $(s, D_j) = 1$, $j = 1, \dots, J$.

The general argument is a largely self-contained careful application of Linnik's Dispersion Method, without appeal to Fourier Analysis on $\mathbb{R}/\mathbb{Z}$.

I conclude the paper with a discussion of interesting problems that arise in the pursuit of wider uniformities in Theorems 1 and 2.

2. Introduction of multilinear forms

For a positive integer k and complex number s define

$$M(s) = \sum_{n \le x^{1/k}} \mu(n) n^{-s}.$$

The following identity is clear (Heath–Brown [9], cf. Linnik [11], Introduction, pp. 21–22).

Lemma 1.

$$\frac{\zeta'(s)}{\zeta(s)} = \sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} \zeta(s)^{j-1} \zeta'(s) M(s)^j + \frac{\zeta'(s)}{\zeta(s)} (1 - \zeta(s)M(s))^k$$

whenever $Re\,(s) > 1$.

Since the Dirichlet series $(1 - \zeta(s)M(s))^k$ has no terms $a_n n^{-s}$ with $n \le x$, for any arithmetic function g whatsoever,

$$\sum_{\substack{n \le x \\ n \equiv b (\mathrm{mod} D)}} g(n)\Lambda(n) = \sum_{r=1}^{k} (-1)^{r-1} \binom{k}{r} S_r,$$

where

$$S_r = \sum_{\substack{n_1\cdots n_r m_1\cdots m_r \le x \\ n_1\cdots n_r m_1\cdots m_r \equiv b (\mathrm{mod} D)}} \cdots \sum g(n_1\cdots n_r m_1\cdots m_r)\log n_1 \mu(m_1)\cdots\mu(m_r).$$

The treatment of a typical multilinear form differs according to the size of (the product) $n_1 \cdots n_j$. We first assume it large.

3. Divisor functions on arithmetic progressions

Although the approach is classical, pioneered by Linnik [10], I tailor the results to the situation at hand.

For the duration of this section $L(s,\chi)$ will denote the series $\sum_{n=1}^{\infty} \chi(n)n^{-s}$ formed with the Dirichlet character χ defined on the multiplicative group of reduced residue classes $(\mathrm{mod}\, D)$, $D \ge 1$.

Lemma 2.

$$\sum_{\chi (\mathrm{mod}\, D)} \int_{-T}^{T} \left|\sum a_n \chi(n) n^{it}\right|^2 dt \ll \sum |a_n|^2 (n + T\phi(D))$$

uniformly for $D \ge 1$, $T > 0$ and complex a_n for which the series $\sum |a_n|^2 n$ converges.

Proof. It will suffice to establish the inequality for an arbitrary finite collection of complex numbers a_n. With $\tau = e^{1/T}$,

$$\int_{-T}^{T} \left|\sum a_n n^{it}\right|^2 dt \ll T^2 \int_0^\infty \left|\sum_{y<n\le y\tau} a_n\right|^2 \frac{dy}{y},$$

Gallagher [8].

From the orthogonality of Dirichlet characters $(\mathrm{mod}\, D)$,

$$\sum_{\chi (\mathrm{mod}\, D)} \left|\sum_{u<n\le v} a_n \chi(n)\right|^2 \ll (v - u + \phi(D)) \sum_{u<n\le v} |a_n|^2,$$

treating the cases $v - u < 1$, $v - u \ge 1$ separately. The sum to be estimated in the lemma is then

$$\ll T^2 \int_0^\infty \sum_{y<n\le y\tau} |a_n|^2((\tau-1)y + \phi(D))\frac{dy}{y}$$
$$\ll T^2 \sum |a_n|^2 \int_{n/\tau}^{n} ((\tau-1)y + \phi(D))\frac{dy}{y} \ll \sum |a_n|^2(n + T\phi(D)).$$

□

Lemma 3.

$$\sum_{\chi\ne\chi_0 (\mathrm{mod} D)} \int_{-T}^{T} |L(\sigma+it,\chi)|^4 dt \ll \phi(D)(T+2)(\log D(T+2))^6$$

uniformly for $|\sigma - 1/2| \le 2(\log D(T+2))^{-1}$.

The power of the logarithm may be reduced and the principal character χ_0 included in the summation, but we shall not need these refinements.

Proof. Since

$$L(s,\chi)^2 = \sum_{n=1}^{\infty} \tau(n)\chi(n)n^{-s}, \qquad \mathrm{Re}\,(s) > 1,$$

we may interchange integration and summation to obtain

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} L(s+z,\chi)^2\Gamma(z)Y^z dz = \sum_{n=1}^{\infty} \tau(n)\chi(n)n^{-s}e^{-n/Y}$$

for real $Y > 0$, $c > 0$. For non-principal characters both sides are analytic and so equal in the whole complex s-plane.

Moving the integral to the line $\mathrm{Re}\,(z) = c$ with $-1 < c < 0$, we pass over the simple pole of $\Gamma(z)$ at $z = 0$ which gives to the integrand a residue $L(s,\chi)^2$.

Let $R = D(2+T)$. In terms of the functional equation

$$L(s,\chi) = \psi(s,\chi)L(1-s,\bar{\chi})$$

we may decompose the (new) integral as

$$\begin{aligned}&\frac{1}{2\pi i}\int_{d-i\infty}^{d+i\infty} \psi(s+z,\chi)^2 \sum_{n\le R} \tau(n)\bar{\chi}(n)n^{-(1-s-z)}\Gamma(z)Y^z dz\\ &\quad + \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \psi(x+z,\chi)^2 \sum_{n>R} \tau(n)\bar{\chi}(n)n^{-(1-s-z)}\Gamma(z)Y^z dz\end{aligned}$$

with d, initially the same as c, permitted any value in the interval $(-1,0)$.

Set $c = -1/2 - 3(\log R)^{-1}$, $d = -(\log R)^{-1}$, $Y = R$.

If χ is induced by the primitive character $\chi_1 (\mathrm{mod}\, D_1)$, $D_1 \mid D$, then

$$L(s,\chi) = L(s,\chi_1)\prod_{p|D}\left(1 - \frac{\chi_1(p)}{p^s}\right),$$

so that $\psi(s,\chi)$ becomes

$$\psi(s,\chi_1)\prod_{p|D}\left(1-\frac{\chi_1(p)}{p^s}\right)\left(1-\frac{\overline{\chi_1(p)}}{p^{1-s}}\right)^{-1}.$$

For $3/8 \le \mathrm{Re}\,(s) \le 5/8$ the product over the divisors of D has a representation

$$\exp\left(\sum_{p|D}\left\{-\frac{\chi_1(p)}{p^s} + \frac{\bar{\chi}_1(p)}{p^{1-s}} + \frac{\chi_1(p)^2}{2p^{2s}} - \frac{\bar{\chi}_1(p)^2}{2p^{2(1-s)}} + O(1)\right\}\right).$$

In the band $|\mathrm{Re}\,(s) - 1/2| \le 2(\log R)^{-1}$, $p^{-(1-s)} - p^{-\bar{s}} \ll p^{-1/2}\left(\exp\left(\frac{4\log p}{\log D}\right) - 1\right)$. The real part of the sum over the prime divisors of D in the exponential is

$$\ll \sum_{p|D}\frac{\log p}{p^{1/2}\log D} \ll 1.$$

On $\mathrm{Re}\,(z) = d$ we obtain $\psi(s+z,\chi) \ll (2+|z|)^{1/\log R}$, cf. Prachar [13], VII, Satz 1.1, p. 207 and Anhang Satz 6.2, p. 395, hence $\psi(x+z,\chi) \ll (2+|z|)^2$ for s in the narrow strip about $\mathrm{Re}\,(s) = 1/2$.

On $\text{Re}\,(z) = c$ we obtain $\psi(s+z,\chi_1) \ll D_1^{1/2}(T+2+|z|)^{1/2+5/\log R}$. The modifying product does not exceed $c_1^{\nu(D/D_1)}$ for a certain positive c_1, and so is $\ll (D/D_1)^{1/2}$. Hence $\psi(s+z,\chi) \ll R^{1/2}(2+|z|)^{11}$ for s in the narrow strip.

Thus

$$\begin{aligned} L(s,\chi)^2 \ll & \left|\sum_{n=1}^{\infty} \tau(n)\chi(n)n^{-s}e^{-n/R}\right| \\ & + \log R \int_{d-i\infty}^{d+i\infty} \left|\sum_{n\le R} \tau(n)\bar{\chi}(n)n^{-(1-s-z)}\right| |\Gamma(z+1)|(2+|z|)^4 d\text{Im}\,(z) \\ & + R^{1/2}\int_{c-i\infty}^{c+i\infty} \left|\sum_{n>R} \tau(n)\bar{\chi}(n)n^{-(1-s-z)}\right| |\Gamma(z+1)|(2+|z|)^{22} d\text{Im}\,(z) \end{aligned}$$

uniformly for $|\text{Re}\,(s) - 1/2| \le 2(\log R)^{-1}$.

Since

$$\int_{h-i\infty}^{h+i\infty} |\Gamma(z+1)|(2+|z|)^{22} d\text{Im}\,(z)$$

is bounded uniformly for $-1 < h < 0$, applications of the Cauchy–Schwarz inequality and the previous lemma complete the proof. For example, on $\text{Re}\,(z) = d$ with $s = \sigma + it$

$$\begin{aligned} & \sum_{\chi\ne\chi_0 (\text{mod} D)} \int_{-T}^{T} \left|\sum_{n\le R} \tau(n)\chi(n)n^{-(1-z-\sigma)+it}\right|^2 dt \\ & \ll \sum_{n\le R} \tau(n)^2 n^{-1}(n+T\phi(D)) \ll \phi(D)(T+2)(\log R)^4. \end{aligned}$$

The factor $\log R$ in the corresponding term bounding $L(s,\chi)^2$ is wastefully large. □

Since $|L(s,\chi)|^r \le 1 + |L(s,\chi)|^4$ for $0 < r < 4$, the upper bound of Lemma 3 remains valid with any positive power of $|L(\sigma+it,\chi)|$ up to the fourth.

Lemma 4.

$$\sum_{\chi\ne\chi_0 (\text{mod}\, D)} \int_{-T}^{T} |L(1/2+it,\chi)^r L(1/2+it,\chi)'| dt \ll \phi(D)(T+2)(\log D(T+2))^7$$

for $r = 1, 2$ *and* 3.

Proof. With R again $D(T+2)$, Cauchy's integral formula gives

$$L(s,\chi)^r L(s,\chi)' = \frac{1}{r+1}(L(s,\chi)^{r+1})' = \frac{1}{2\pi i(r+1)}\int \frac{L(w,\chi)^{r+1}}{(w-s)^2}dw,$$

the integral taken over the disc $|w-s| = (\log R)^{-1}$. Taking absolute values we integrate over t and change the order of the two integrations. We see that the expression to be estimated is

$$\ll \log R \int_0^{2\pi} \sum_{\chi} \int_{-T}^{T} |L(\sigma+it+e^{i\theta}(\log R)^{-1},\chi)|^r dt d\theta$$

where $|\sigma + (\log R)^{-1}\cos\theta - 1/2| \le 2(\log R)^{-1}$ so that we may apply the previous lemma. □

Versions of Lemmas 3 and 4 may be derived from Lemma 10.5 of Montgomery [12]. I have adapted (with minor corrections) the more nearly self-contained argument, due to Ramachandra, that is presented in §10 of Bombieri [3].

Lemma 5. *let $0 < \alpha < 1/2$, k a positive integer. There is a number β, depending at most upon k, so that*

$$\sum_{x_1 < Dm+\ell \le x} \tau(Dm+\ell)^k \ll D^{-1}(x - x_1)(\log x)^\beta$$

uniformly for $(\ell, D) = 1$, $D \le x^{1-\alpha}$, $x_1 < x$, $x - x_1 \ge x^{1-\alpha/2}$.

For $x_1 = 0$ the upper bound may be replaced by $xD^{-1}(D^{-1}\phi(D)\log x)^{2^k-1}$.

Proof. See Linnik [11], Lemma 1.1.5 and Lemma 1.1.4. Variant proofs will be indicated following Lemma 14 of the present paper. □

Lemma 6. *Let $A > 0$, $\delta > 0$. Then*

$$\sum_{\substack{n_1\cdots n_r \le x \\ n_1\cdots n_r \equiv b (\operatorname{mod} D)}} \log n_1 - \frac{1}{\phi(D)} \sum_{\substack{n_1\cdots n_r \le x \\ (n_1\cdots n_r, D)=1}} \log n_1 \ll \frac{x(\log x)^{-A}}{\phi(D)} + x^{1/2}(\log x)^8$$

uniformly for $(b, D) = 1$, $1 \le D \le x^{1-\delta}$, $x \ge 2$, $r = 1, 2, 3$ or 4.

Proof. Without loss of generality we may assume x not to be an integer. The expression, $S(x)$, to be estimated has a representation

$$\frac{1}{\phi(D)} \sum_{\chi \ne \chi_0 (\operatorname{mod} D)} \frac{\bar{\chi}(b)}{2\pi i} \int_{c-i\infty}^{c+i\infty} L(s,\chi)^{r-1} L(s,\chi)' \frac{x^s}{s} ds,$$

where the integral may be taken over any line $\operatorname{Re}(s) = c > 1$.

Let $B > 0$. Replace x by y and average with respect to y over the interval $(x-(x\log x)^{-B}, x]$. Since $\tau_r(n) \le \tau(n)^{r-1}$, the previous lemma shows that typically

$$S(x) - S(y) \ll \frac{x}{D}(\log x)^{c_0 - B},$$

the constant c_0 absolute, the implied constant depending upon B and δ. Within a similar error we may replace $x^s s^{-1}$ in the integral by the kernel

$$K(s) = \frac{(\log x)^B}{sx} \int_{x - x(\log x)^{-B}}^{x} y^s dy.$$

Of its properties we need only that it is $\ll \min(|s|^{-1}x^\sigma, |s(s+1)|^{-1}x^\sigma(\log x)^B)$ uniformly for $\sigma = \operatorname{Re}(s) > 0$.

We move the contour $\operatorname{Re}(s) = c$ to $\operatorname{Re}(s) = 1/2$. This may be justified by replacing $L(s,\chi)$ with its approximant

$$\sum_{n \le Z} \chi(n) n^{-s} + O(|s| Z^{-\sigma} D^{1/2} \log 2D),$$

valid in $\sigma = \operatorname{Re}(s) \ge 1/4$, likewise replacing $L(s,\chi)'$, and letting Z become large.

Appeal to Lemma 4 shows that the range $|t| \leq x$ contributes

$$\ll \frac{1}{\phi(D)} \sum_{\chi \neq \chi_0 (\operatorname{mod} D)} \int_{-1}^{1} |L(s,\chi)^{r-1} L(s,\chi)'| \frac{x^{1/2}}{|s|} dt$$
$$+ \sum_{0 \leq k \leq \frac{\log 2x}{\log 2}} \frac{1}{\phi(D)} \sum_{\chi \neq \chi_0 (\operatorname{mod} D)} 2^{-k} \int_{-2^k}^{2^k} |L(s,\chi)^{r-1} L(s,\chi)'| x^{1/2} dt$$

which is $\ll x^{1/2}(\log x)^8$. For the range $|t| > x$ we argue similarly, employing the second of the bounds on $K(s)$, and obtain a much smaller contribution.

Choosing $B = A + c_0$ completes the proof. □

The error term in Lemma 6 is $\ll x(\phi(D)(\log x)^A)^{-1}$ provided that $D \leq x^{1/2}(\log x)^{-A-8}$, and this will be satisfactory.

A result of the same form as Lemma 6 but with a weaker error term may be obtained for any fixed integer r, even by purely elementary means. However, if side conditions are to be placed upon the individual variables n_j, $1 \leq j \leq r$, particularly by introducing factors $n_j^{-s_j}$ with new complex variables s_j, then the foregoing argument is altogether more flexible.

When the product $n_1 \cdots n_r$ in a multilinear form of §2 is small, we may expect the complementary product $m_1 \cdots m_r$ to be large. Since each variable m_j is constrained to be small, we may then partition variables and reduce ourselves from a multilinear to a bilinear form.

4. Bilinear forms on residue classes

Let s be a positive integer, and D_j, $1 \leq j \leq J$, further positive integers, coprime to s.

Let

$$\Delta = \sum_j \left| \sum_{mn \equiv u_j (\operatorname{mod} D_j s)} a_m b_n - \frac{1}{D_j} \sum_{\substack{mn \equiv u_j (\operatorname{mod} s) \\ (n, D_j) = 1}} a_m b_n \right|$$

where $(u_j, D_j s) = 1$ for each j, the a_m, b_n are real numbers supported on $(M/2, M]$, $(N/2, N]$ respectively. Note that the second doublesum between the solidi has summation condition $(n, D_j) = 1$ rather than $(mn, D_j) = 1$, and weight D_j^{-1} rather than $\phi(D_j)^{-1}$. I shall assume that $M \geq 2, N \geq 2$

Lemma 7. *Let $0 < \delta < 1$, $1 \leq s \leq N^{1-\delta}$, and assume that $|b_n| \leq (\tau(n)\log 2n)^c$ for a non-negative integer c and all positive n. Then*

$$\Delta^2 \ll \|\mathbf{a}\|^2 (\log N)^\gamma \left(\frac{MN^2}{s^2} \sum\sum_{(D_u, D_v) > 1} \frac{1}{D_u D_v} + \frac{MN^{1+\delta}}{s} \sum\sum_{s(D_u, D_v) > N^{1-\delta}} \frac{1}{[D_u, D_v]} + \frac{J^2 N^2}{s} \right)$$

with $\|\mathbf{a}\|^2 = \sum |a_n|^2$, $\gamma = 2^{c+1} + 2c - 1$.

Proof. I employ Linnik's dispersion method [11]. We write

$$\Delta = \sum_j c_j \left(\sum_{mn \equiv u_j (\mathrm{mod} D_j s)} a_m b_n - \cdots \right)$$

with $|c_j| \le 1$ and interchange the summations to obtain a representation

$$\Delta = \sum_m a_m \left(\sum_{\substack{j \\ mn \equiv u_j (\mathrm{mod} D_j s)}} \sum_n c_j b_n - \sum_j \frac{c_j}{D_j} \sum_{\substack{mn \equiv u_j (\mathrm{mod} s) \\ (n, D_j) = 1}} b_n \right).$$

Then $|\Delta|^2 \le \|\mathbf{a}\|^2 Z$ where $Z = A - 2B + C$ with

$$A = \sum_{j_1} \sum_{j_2} c_{j_1} c_{j_2} \sum_{n_1} \sum_{n_2} b_{n_1} b_{n_2} \sum_{\substack{mn_1 \equiv u_{j_1} (\mathrm{mod} D_{j_1} s) \\ mn_2 \equiv u_{j_2} (\mathrm{mod} D_{j_2} s)}} 1,$$

$$B = \sum_{j_1} \sum_{j_2} \frac{c_{j_1} c_{j_2}}{D_{j_2}} \sum_{n_1} \sum_{n_2} b_{n_1} b_{n_2} \sum_{\substack{mn_1 \equiv u_{j_1} (\mathrm{mod} D_{j_1} s) \\ mn_2 \equiv u_{j_2} (\mathrm{mod} s) \\ (n_2, D_{j_2}) = 1}} 1,$$

$$C = \sum_{j_1} \sum_{j_2} \frac{c_{j_1}}{D_{j_1}} \frac{c_{j_2}}{D_{j_2}} \sum_{n_1} \sum_{n_2} b_{n_1} b_{n_2} \sum_{\substack{mn_1 \equiv u_{j_1} (\mathrm{mod} s) \\ mn_2 \equiv u_{j_2} (\mathrm{mod} s) \\ (n_1, D_{j_1}) = 1, (n_2, D_{j_2}) = 1}} 1.$$

Consider A. The congruence condition on m cannot be fulfilled unless $(n_1, D_{j_1} s) = 1 = (n_2, D_{j_2} s)$ and $n_2 u_{j_1} \equiv n_1 u_{j_2} (\mathrm{mod} s(D_{j_1}, D_{j_2}))$. The innermost sum is then $M(2s[D_{j_1}, D_{j_2}])^{-1} + \theta$ with $|\theta| \le 1$. Therefore,

$$\begin{aligned}
&\left| A - \frac{M}{2s} \sum_{(D_{j_1}, D_{j_2}) = 1} \sum \frac{c_{j_1} c_{j_2}}{D_{j_1} D_{j_2}} \sum_{\substack{u_{j_1} n_2 \equiv u_{j_2} n_1 (\mathrm{mod} s) \\ (n_1, D_{j_1} s) = 1 = (n_2, D_{j_2} s)}} \sum b_{n_1} b_{n_2} \right| \\
&\le \frac{M}{2s} \sum_{(D_{j_1}, D_{j_2}) > 1} \frac{1}{[D_{j_1}, D_{j_2}]} \sum_{\substack{u_{j_1} n_2 \equiv u_{j_2} n_1 (\mathrm{mod} s (D_{j_1}, D_{j_2})) \\ (n_1, D_{j_1} s) = 1 = (n_2, D_{j_2} s)}} \sum |b_{n_1} b_{n_2}| \\
&\quad + \sum_{j_1} \sum_{j_2} \sum_{\substack{u_{j_1} n_2 \equiv u_{j_2} n_1 (\mathrm{mod} s) \\ (n_1, D_{j_1} s) = 1 = (n_2, D_{j_1} s)}} \sum |b_{n_1} b_{n_2}|.
\end{aligned}$$

Similar (in fact stronger) estimates hold for B and C with the leading majorant replaced by

$$\frac{M}{2s} \sum_{(D_{j_1}, D_{j_2}) > 1} \sum \frac{1}{D_{j_1} D_{j_2}} \sum_{\substack{u_{j_1} n_2 \equiv u_{j_2} n_1 (\mathrm{mod} s) \\ (n_1, D_{j_1} s) = 1 = (n_2, D_{j_2} s)}} \sum |b_{n_1} b_{n_2}|.$$

The main terms in these estimates for A, B and C coincide and Z does not exceed

$$\frac{M}{2s}\sum_{(D_{j_1},D_{j_2})>1}\sum \frac{1}{[D_{j_1},D_{j_2}]}\sum_{\substack{u_{j_1}n_2\equiv u_{j_2}n_1(\mathrm{mod} s(D_{j_1},D_{j_2}))\\(n_1,D_{j_1}s)=1=(n_2,D_{j_2}s)}}\sum |b_{n_1}b_{n_2}|$$
$$+\frac{3M}{2s}\sum_{(D_{j_1},D_{j_2})>1}\sum \frac{1}{D_{j_1}D_{j_2}}\sum_{\substack{u_{j_1}n_2\equiv u_{j_2}n_1(\mathrm{mod} s)\\(n_1,D_{j_1}s)=1=(n_2,D_{j_2}s)}}\sum |b_{n_1}b_{n_2}|$$
$$+4J^2\sum_{\substack{u_{j_1}n_2\equiv u_{j_2}n_1(\mathrm{mod} s)\\(n_1n_2,s)=1}}\sum |b_{n_1}b_{n_2}|.$$

The innermost sum of the first of these three bounding terms is

$$\leq \sum_{\substack{n_2\leq N\\ n_2\equiv n_0(\mathrm{mod} t)}}|b_{n_2}|\leq (\log 2N)^c\sum_{\substack{n\leq N\\ n\equiv n_0(\mathrm{mod}\, t)}}\tau(n)^c,$$

where $t=s(D_{j_1},D_{j_2})$ and $u_{j_1}n_0\equiv u_{j_2}n_1(\mathrm{mod} t)$; recall that $(u_{j_1},D_{j_1}s)=1$ so that $(u_j,t)=1$. If $t>N^{1-\delta/2}$, then this last sum is

$$\ll N^{\delta/2}\sum_{n\leq N,n\equiv n_0(\mathrm{mod} t)}1\ll N^{\delta/2}(Nt^{-1}+1)\ll N^{\delta}.$$

Otherwise, Lemma 5 shows it to be $\ll Nt^{-1}(\log N)^{2^c+c}$. Hence,

$$\sum_{\substack{u_{j_2}n_1\equiv u_{j_1}n_2(\mathrm{mod} s(D_{j_1},D_{j_2}))\\(n_1,D_{j_1}s)=1=(n_2,D_{j_2}s)}}\sum |b_{n_1}b_{n_2}|\ll N(\log N)^{\gamma}(Nt^{-1}+N^{\delta})$$

with $\gamma=2^{c+1}+2c-1$ and where the term N^{δ} may be omitted if $t\leq N^{1-\delta/2}$. The first term in the bound on Z is thus

$$\ll \frac{MN^2(\log N)^{\gamma}}{s^2}\sum_{(D_u,D_v)>1}\sum \frac{1}{D_uD_v}+\frac{MN^{1+\delta}(\log N)^{\gamma}}{s}\sum_{s(D_u,D_v)>N^{1-\delta}}\sum \frac{1}{[D_u,D_v]}.$$

This estimate also serves for the second sum since in that case the rôle of t is played by s, and by hypothesis $s\leq N^{1-\delta}$. The third term in the bound on Z is $\ll s^{-1}J^2N^2(\log N)^{\gamma}$.

The proof of Lemma 7 is complete. □

Lemma 8. *The inequality of Lemma 7 remains valid if we replace Δ by*

$$\sum_j\left|\sum_{mn\equiv u_j(\mathrm{mod} D_js)}a_mb_n-\frac{1}{\phi(D_j)}\sum_{\substack{mn\equiv u_j(\mathrm{mod} s)\\(mn,D_j)=1}}a_mb_n\right|.$$

Proof. We follow the argument for Lemma 7. There is no change in the treatment of A, and to treat B and C we apply the estimate

$$\left| \sum_{\substack{u<n\le v\\ n\equiv r(\mathrm{mod}\, a)\\ (n,b)=1}} 1 - \left(\frac{v-u}{ab}\right)\sum_{d|b}\mu(d)\right| \le \sum_{d|b}|\mu(d)| = 2^{\nu(b)},$$

valid whenever $u \le v$ and the positive integers a and b are mutually prime. Since $2^{\nu(D)}\phi(D)^{-1} \le 2$ for every positive integer D, it will suffice to raise the coefficient $4J^2$ to $8J^2$. The only remaining (and otherwise substantial) change is the introduction into the second term bounding Z of a factor $\Pi(1-p^{-1})^{-1}$ taken over the prime divisors of (D_{j_1}, D_{j_1}).

If $s(D_{j_1}, D_{j_2}) \le N^{1-\delta/2}$, $t = s(D_{j_1}, D_{j_2})$, then again by Lemma 5,

$$\sum_{\substack{n_2\le N, n_2\equiv \bar{u}_{j_1}u_{j_2}n_1(\mathrm{mod} s)\\ n\equiv w(\mathrm{mod}(D_{j_1},D_{j_2}))}} |b_{n_2}|$$

with $u_{j_1}\bar{u}_{j_1} \equiv 1(\mathrm{mod} s)$, is $\ll Nt^{-1}(\log N)^{2^c+c}$ uniformly for each w coprime to (D_{j_1}, D_{j_2}). Allowing w to run over a complete set of reduced residue class representatives $\mathrm{mod}(D_{j_1}, D_{j_2})$ introduces a scaling factor $\phi((D_{j_1}, D_{j_2}))(D_{j_1}, D_{j_2})^{-1}$ that cancels the above extra product. The corresponding contribution towards the modified Δ falls within the first majorant of Lemma 7.

If $s(D_{j_1}, D_{j_1}) > N^{1-\delta/2}$, then we omit the condition involving w. The corresponding contribution to Z is

$$\ll \frac{MN^2(\log N)^\gamma}{s^2} \sum\sum_{s(D_u,D_v)>N^{1-\delta/2}} \frac{1}{D_uD_v}\prod_{p|(D_u,D_v)}\left(1-\frac{1}{p}\right)^{-1}.$$

Since $\phi(m) \gg m(\log m)^{-1}$ uniformly for $m \ge 2$, the reciprocal of a typical summand is

$$[D_u, D_v]\phi((D_u, D_v)) \gg [D_u, D_v]\phi(s)^{-1}N^{1-\delta/2}(\log N)^{-1},$$

and the sum is

$$\ll \frac{MN^{1+\delta/2}(\log N)^{\gamma+1}}{s} \sum\sum_{s(D_u,D_v)>N^{1-\delta/2}} \frac{1}{[D_u, D_v]},$$

leading to a contribution towards the modified Δ that falls within the second majorant of Lemma 7.

This completes the proof of Lemma 8. □

We replace the L^2 norm on **a** by (essentially) an L^∞ norm. This will allow the interchange of the rôles of M and N.

Lemma 9. *Let $0 < \delta < 1$, $1 \le s \le N^{1-\delta}$ and assume that for some positive integer c, $|a_m| \le (\tau(m)\log 2m)^c$, $|b_n| \le (\tau(n)\log 2n)^c$ on their respective intervals*

$(M/2, M]$, $(N/2, N]$. *Then*

$$\sum_j \left| \sum\sum_{mn\equiv u_j (\text{mod} D_j s)} a_m b_n - \frac{1}{\phi(D_j)} \sum\sum_{\substack{mn\equiv u_j (\text{mod} s) \\ (mn, D_j)=1}} a_m b_n \right| \ll \lambda^{1/2} s^{-1} MN (\log MN)^\gamma$$

with $\gamma = 2^{2c} + 2c - 1$ *and*

$$\lambda = \sum\sum_{(D_u, D_v)>1} \frac{1}{D_u D_v} + \frac{s}{N^{1-\delta}} \sum\sum_{s(D_u,D_v)>N^{1-\delta}} \frac{1}{[D_u, D_v]} + \frac{J^2 s}{M},$$

the implied constant depending at most upon δ *and* c.

Proof. Estimate $\|\mathbf{a}\|^2$ by Lemma 5.

Lemma 9 will suffice for my present purposes. The following result will help implement it.

Lemma 10. *For complex numbers* $a(m, j)$, $1 \le j \le J$, $1 \le m \le x$, *where* $x \ge 2$,

$$\sum_{j=1}^{J} \max_{h\le x} \left| \sum_{m\le h} a(m, j) \right| \ll \max_{Re(z)=(\log x)^{-1}} \sum_{j=1}^{J} \left| \sum_{m\le x} a(m, j) m^{-z} \right|,$$

the implied constant absolute.

Proof. Without loss of generality we may assume each inner maximum of the first sum taken at h half an odd positive integer. For any $\theta > 0$,

$$\sum_{m\le h} a(m, j) = \frac{1}{2\pi i} \int_{\text{Re}(z)=\theta} \frac{y^z}{z} \sum_{m\le x} a(m, j) m^{-z} dz$$

certainly holds if $h - 1/4 \le y \le h$. Averaging over this interval we see that

$$\sum_{m\le h} a(m, j) = \frac{1}{2\pi i} \int_{\text{Re}(z)=\theta} K(h, z) \sum_{m\le x} a(m, j) m^{-z} dz$$

with

$$K(h, z) = \frac{4}{z} \int_{h-1/4}^{h} y^z dy = \frac{4(h^{z+1} - (h - 1/4)^{z+1})}{z(z + 1)}.$$

In particular, $|K(h, z)| \le \min(h^{\text{Re} z}|z|^{-1},\ h^{\text{Re} z+1}(|z(z + 1)|)^{-1})$ uniformly in the half-plane $\text{Re}(z) > 0$.

The lemma follows from the fact that with $\theta = (\log x)^{-1}$,

$$\int_{\text{Re}(z)=\theta} \max_{h\le x} |K(h, z)| d\text{Im}(z) \ll \int_{|t|\le x} \frac{dt}{\sqrt{\theta^2 + t^2}} + \int_{|t|>x} \frac{x dt}{t^2} \ll \log x.$$

We smooth by averaging over a short interval here, but over a long interval in Lemma 6. □

5. Proof of Theorem 2

A term which falls within the bound of Theorem 2 will be called an admissible error. Let $0 < \delta < 1$. The Brun–Titchmarsh estimate $\psi(y, D, r) \ll y\phi(D)^{-1}$, valid uniformly for $(r, D) = 1$, $1 \leqq D \leq y^{1-\delta}$, shows that within an admissible error we may replace the maximum over $y \leq x$ for each modulus sD_j by the choice of half an odd integer $y = y_j$ in the interval $(x(\log x)^{-A}, x]$. Moreover, in $\psi(y_j, sD_j, r)$ and $\psi(y_j, s, r)$ we need only count those prime powers exceeding $x(\log x)^{-A}$ and coprime to D_j.

We apply Lemma 1 with $k = 4$. In the notation of the remark following that lemma, with g identically 1, a typical sum over the products $n_1 \cdots n_r m_1 \cdots m_r$ may be expressed in the form

$$\sum_{m_1 \cdots m_r \leq x} \mu(m_1) \cdots \mu(m_r) \sum_{\substack{n_1 \cdots n_r \leq x(m_1 \cdots m_r)^{-1} \\ n_1 \cdots n_r \equiv b\bar{m}_1 \cdots \bar{m}_r (\mathrm{mod} D)}} \log n_1$$

where $v\bar{v} \equiv 1 (\mathrm{mod} D)$. Lemma 6 allows us to estimate the terms for which $m_1 \cdots m_r \leq x^{1/4}$ within an error of

$$\begin{aligned} &\ll \sum_{m_1 \cdots m_r \leq x^{1/4}} x(m_1 \cdots m_r)^{-1}(\phi(D)(\log x)^{A+4})^{-1} \\ &\quad + \sum_{m_1 \cdots m_r \leq x^{1/4}} x^{1/2}(m_1 \cdots m_r)^{-1/2}(\log x)^8 \\ &\ll x\phi(D)^{-1}(\log x)^{-A} + x^{5/8}(\log x)^{11}. \end{aligned}$$

It will suffice to establish that for each r, $1 \leq r \leq 4$,

$$\sum_{j=1}^{J} \left| \sum_{m_1 \cdots m_r n_1 \cdots n_r \equiv u_j (\mathrm{mod} sD_j)} \mu(m_1) \cdots \mu(m_r) \log n_1 \right.$$
$$\left. - \frac{1}{\phi(D_j)} \sum_{\substack{m_1 \cdots m_r n_1 \cdots n_r \equiv u_j (\mathrm{mod} s) \\ (m_1 \cdots m_r n_1 \cdots n_r, D_j) = 1}} \mu(m_1) \cdots \mu(m_r) \log n_1 \right|$$

is an admissible error, where each $(u_j, sD_j) = 1$, every product $m_1 \ldots m_r$ exceeds $x^{1/2}$ and $x(\log x)^{-A} < m_1 \cdots m_r n_1 \cdots n_r \leq y_j$ for the sums involving D_j.

We may dominate this sum over j by $O((\log x)^{2r})$ similar sums with the variables m_u, n_v restricted by $\frac{1}{2}M_u < m_u \leq M_u$, $\frac{1}{2}N_v < n_v \leq N_v$ respectively, $u = 1, \ldots, r$, $v = 1, \ldots, r$. We may assume that every M_u does not exceed $x^{1/4}$, $M_1 \cdots M_r > x^{1/4}$ and that $x(\log x)^{-A} < M_1 \cdots M_r N_1 \cdot \cdots \cdot N_r \leq x$. According to Lemma 10, at the expense of an ultimate factor of $O((\log x)^2)$ and the introduction to the innermost summands of factors $(m_1 \cdots m_r)^{-z_1 - z_2}(n_1 \cdots n_r)^{-z_1}$, we may ignore the three conditions on the size of the products involving the m_u, n_v. They are not lost since they have been transferred to the M_u, N_v.

We can find a subset of the M_u, $1 \leq u \leq r$, whose product lies in the interval $[x^{1/4}, x^{1/2})$. Let M denote that product and collect the corresponding variables m_u together. Collect the remaining variables amongst the m_u, n_v together. We are reduced to estimating

$$\sum_{j=1}^{J}\left|\sum_{mn\equiv u_j(\bmod sD_j)} a_m b_n - \frac{1}{\phi(D_j)}\sum_{\substack{mn\equiv u_j(\bmod s)\\ (mn,D_j)=1}} a_m b_n\right|$$

where $|a_m| \le \tau_r(m)$, $|b_n| \le \tau_{2r-1}(n)\log n$, and the respective sequences are supported on $(2^{-r}M, M]$, $(2^{-(2r-1)}N, N]$, with $x^{1/4} \le M < x^{1/2}$, $x(\log x)^{-A} < MN \le x$. We apply Lemma 9 with the rôles of N, M there, played by $M, M_1\cdots M_r N_1\cdots N_r M^{-1}$ here. We may take $c = 6$, so that $\gamma = 2^{12} + 11$. Since $s(D_u, D_v) \le \max sD_j \le x^{1/4-\delta}$ is satisfied by hypothesis, the second of the sums defining λ is empty. Moreover, in the notation of that lemma we presently have

$$\max_j \phi(sD_j)\left(\frac{s}{M}\right)^{1/2} \ll x^{1/4-\delta}\left(\frac{s(\log x)^A}{x^{1/2}}\right)^{1/2} \ll x^{-\delta/4},$$

so that the third term in the definition of λ gives rise to an amount

$$\ll x^{1-\delta/8}\sum_{j=1}^{J}\phi(sD_j)^{-1}$$

in the final upper bound.

This completes the proof of Theorem 2. □

6. Preparation for Theorem 1

Let $\omega = \exp((\log\log x)^2)$. Let $\exp(2(\log\log x)^6) < J \le x^{\beta}$. Express the moduli $f(j)$, $J/2 < j \le J$, in the form $s_j D_j$, where s_j is made up of primes not exceeding ω, D_j of primes greater than ω. We begin by showing that we may assume the s_j to be comparatively small.

Lemma 11. *Let X_j, $j = 1, \dots, h$, be independent random variables that satisfy $|X_j| \le c$ almost surely. Define*

$$S_h = X_1 + \cdots + X_h, \qquad m = \sum_{j=1}^{h} \textit{Expect}\ (|X_j|).$$

Then the inequality

$$P(S_h \ge u) \le \exp\left(-\frac{u}{c}\left\{\log\frac{u}{m} - \log\log\frac{u}{m} - 1\right\}\right)$$

holds uniformly for $u \ge me$.

Proof. This is Lemma 3.3 of [4]. □

Lemma 12. *The number of positive integers n not exceeding z for which*

$$\prod_{p|f(n),p\le r} p > y$$

is

$$\ll z\exp\left(-\frac{\log y}{50\log r}\log\left(\frac{\log y}{\log r}\right)\right) + z^{3/4}$$

uniformly for $2 \le r \le y \le z$.

Proof. Let $\rho(D)$ denote the number of solutions to $f(n) \equiv 0(\text{mod} D)$. Of its properties we need that it is multiplicative, $\rho(p^\alpha)$ in uniformly bounded, and $\rho(p) \leqq \deg f < p$ for $p > p_0$. If ℓ is any arithmetic function and the independent random variables W_p, $p_0 < p \leq r$, are distributed according to

$$W_p = \begin{cases} \ell(p) & \text{with probability} \quad \dfrac{\rho(p)}{p}, \\ \quad 0 & \text{with probability} \quad 1 - \dfrac{\rho(p)}{p}, \end{cases}$$

then the frequency of the integers n not exceeding z for which $\sum \ell(p)$, taken over the prime divisors of $f(n)$ that lie in the interval $(p_0, r]$, belongs to $a(ny)$ set F has a representation

$$P\left(\sum_{p_0<p\leq r} W_p \in F\right) + O\left(\exp\left(-\frac{\log z}{50\log r}\log\left(\frac{\log z}{\log r}\right)\right)\right) + O(z^{-1/4}),$$

[4], Lemma 3.6. An earlier result with a weaker error is implicit in Uzdavinis, [14]. Our estimate is developed from ideas of Erdös, Kac and Kubilius in Probabilistic Number Theory.

We set $\ell(p) = \log p$ on the primes, F the set of reals exceeding $\frac{1}{2}\log y$. Then

$$\sum_{p_0<p\leq r} \text{Expect }(|W_p|) = \sum_{p_0<p\leq r} \frac{\rho(p)\log p}{p} \ll \log r,$$

and the above probability is

$$\ll \exp\left(-\frac{\log y}{3\log r}\log\left(\frac{\log y}{\log r}\right)\right),$$

by Lemma 11 with $c = \log r$, $u = \frac{1}{2}\log y$ if $\log y/\log r$ is suitably large, and trivially otherwise.

Lemma 12 follows rapidly. □

Lemma 13. *There are* $\ll J\exp(-(\log\log x)^2)$ *moduli for which*

$$s_j > \exp((\log\log x)^6).$$

Proof. According to Lemma 12 the number of those j for which the product of all primes in s_j exceeds $\exp(\frac{1}{2}(\log\log x)^4)$ falls well within the asserted bound. Moreover, if $a = 2[(\log\log x)^2]$, then of the remaining s_j those divisible by the ath-power of a prime are

$$\ll \sum_{p\leq\omega} \sum_{\substack{j\leq J \\ f(j)\equiv 0(\text{mod} p^a)}} 1 \ll J\sum_{p\leq\omega} p^{-a} \ll J2^{-a}$$

in number.

Lemma 13 is clear. □

Consider those j for which s_j has a particular value, s, the value not exceeding $\exp((\log\log x)^6)$. An application of Theorem 2 reduces to the consideration of the sum

$$\sum_{t>\omega}\left(\sum_{\substack{J/2<j\le J\\ f(j)\equiv 0(\mathrm{mod}\, st)}} 1\right)^2.$$

Ignoring the square we obtain a doublesum that does not exceed

$$\sum_{\substack{j\le J\\ f(j)\equiv 0(\mathrm{mod}\, s)}} \tau(s^{-1}f(j)).$$

For $s=1$ there is a long history of the treatment of such sums, beginning with van der Corput [16], who obtained a bound $\ll x(\log x)^c$ for a certain $c>0$. Assuming f irreducible over the rationals, Erdös [6] improved c to the best possible value $c=1$. His argument was modified and extended by Barban [1], Wolke [16], including the sum $\sum \tau(a_n)$ for a rather general sequence a_n. The treatment of Wolke would modify to give a bound $\ll s^{-1}J\log J$ for the above sum with irreducible f. For my present purposes it is enough to have a bound of van der Corput quality with the condition $f(j)\equiv 0(\mathrm{mod} s)$ taken into account. Wolke gave a new proof of van der Corput's result, [18]. The argument which follows contains another (and very short) proof of van der Corput's result. The polynomial need not be irreducible.

Lemma 14. *Let* $0<\delta<1$. *Then*

$$\sum_{\substack{k\le y\\ f(k)\equiv 0(\mathrm{mod} s)}} \tau(s^{-1}f(k)) \ll \frac{\rho(s)}{s} y(\log y)^\gamma$$

with $(\frac{1}{2}\gamma)^\delta = 4^{\deg f}(\deg f)^\delta$, *uniformly for* $1\le s\le y^{1-\delta}$, $y\ge 2$.

Proof. Let $0<\theta<1$. Order the prime divisors of each positive integer n according to size. We may marshall those not exceeding $n^{\theta/2}$ as a product $m_1\cdots m_t$ with at most one of the m_i not in the interval $(n^{\theta/2}, n^\theta]$. Hence, $t\le 1+2\theta^{-1}$. There are no more than $2\theta^{-1}$ prime divisors of n exceeding $N^{\theta/2}$. Bearing in mind that $\tau(ab)\le\tau(a)\tau(b)$ whether the integers a and b are mutually prime or not, we see that $\tau(n)\ll\tau(m_i)^t$ for some i, the implied constant depending at most upon θ.

Choosing $\theta=\delta(\deg f)^{-1}$ it follows that the sum which we wish to estimate is

$$\ll \sum_{m\ll y}\tau(m)^t \sum_{\substack{n\le y\\ f(n)\equiv 0(\mathrm{mod} sm)}} 1 \ll \frac{y\rho(s)}{s}\sum_{m\ll y}\frac{\tau(m)^t\rho(m)}{m},$$

with the third sum

$$\ll \prod_{p\ll y}\left(1+\frac{2^t\rho(p)}{p}+O\left(\frac{1}{p^2}\right)\right) \ll (\log y)^{2^t\deg f}.$$

This completes the proof of Lemma 14. ∎

A similar argument yields a proof of the first part of the Linnik–Vinogradov result in Lemma 5. The second part of that lemma may be justified as in the earliest of the references to Wolke.

7. Proof of Theorem 1

For positive integers D, real $x \geq 2$, define

$$E(D) = \max_{(r,D)=1} \max_{y \leq x} \left| \psi(y, D, r) - \frac{y}{\phi(D)} \right|.$$

Application of the Brun–Titchmarsh theorem with Lemma 13 shows that the contribution to $(\Delta(J) =)$

$$\sum_{J/2<j\leq J} \frac{f(j)}{j} E(f(j))$$

from the j with $s_j > \exp((\log\log x)^6)$ is $\ll x\exp(-(\log\log x)^2)$. After Theorem 2, those of the remaining terms with $s_j = s$ contribute

$$\ll (x(\log x)^{-A} + \phi(s)E(s))J^{-1} \sum_{\substack{j\leq J \\ f(j)\equiv 0 (\text{mod} s)}} 1$$

$$+ xJ^{-1} \left(\sum_{w<r\ll J^{\deg f}} \left(\sum_{\substack{j\leq J \\ f(j)\equiv 0 (\text{mod} sr)}} 1 \right)^2 \right)^{1/2} (\log x)^{4200}.$$

Bearing in mind the constraint upon the size of J,

$$\sum_{\substack{j\leq J \\ f(j)\equiv 0 (\text{mod}\, sr)}} 1 \ll \rho(sr)(J(sr)^{-1} + 1) \ll s^{-1}\rho(s)J\omega^{-1/2}$$

uniformly for the whole range of r. By Lemma 14 the second of the typical bounding terms is $\ll s^{-1}\rho(s)x\omega^{-1/4}(\log x)^{4200+\gamma/2}$. Altogether

$$\Delta(J) \ll x(\log x)^{-A} \sum_s s^{-1}\rho(s) + \sum_s \rho(s)E(s)$$

where the summations are allowed to run over (otherwise unrestricted) positive integers not exceeding $\exp((\log\log x)^6)$.

Results obtained by the theory of Dirichlet L-series (e.g. Prachar [13], IX, Satz 2.2, cf. Satz 2.3) guarantee that $E(s) \ll s^{-1}x\exp(-(\log x)^{1/4})$ uniformly for all such s save possibly for the multiples of an s_0. Moreover, according to the Siegel–Walfisz theorem we may assume s_0 to exceed an appropriate multiple of a chosen power of $\log x$. After another application of the Brun–Titchmarsh theorem

$$\sum_s \rho(x)E(x) \ll x\exp(-(\log x)^{1/4}) \sum_{s\leq x} \rho(s)s^{-1} + x \sum_{t\leq x} \rho(ts_0)(ts_0)^{-1}.$$

Since $\rho(D) \leq b^{\nu(D)}$ for some positive b and all D, the second of these dominating terms is

$$\ll xs_0^{-1}b^{\nu(s_0)} \sum_{t\leq x} b^{\nu(t)}t^{-1} \ll s_0^{-1/2}x(\log x)^b.$$

Summing over $J = 2^{-k}x^\beta$ with $k = 0, 1, \ldots$, we see that the moduli $f(D)$ in Theorem 1 with $D > \exp((\log\log x)^6)$ contribute $\ll x(\log x)^{-A+\deg f+1}$.

We treat the complementary moduli in Theorem 1 as we did the moduli s in $E(s)$, the only change being in the sum over the multiples of s_0, which becomes a sum

$$\ll x \sum_{\substack{D \le \exp((\log\log x)^6) \\ f(D) \equiv 0 (\bmod s_0)}} D^{-1} \ll s \sum_{t=1}^{\rho(s_0)} \sum_{\substack{D \le x \\ D \equiv D_t (\bmod s_0)}} D^{-1}$$

where the D_t comprise a set of least positive solutions to the congruence $f(D) \equiv 0 (\bmod s_0)$. Typically $s_0 \le |f(D_t)| \ll D_t^{\deg f}$ and the corresponding innersum is

$$\le \sum_{0 \le u \le x} (D_t + u s_0)^{-1} \ll D_t^{-1} + s_0^{-1} \log x.$$

The final doublesum is $\ll s_0^{-1/(2\deg f)} x \log x$, which suffices.

The proof of Theorem 1 is complete. □

8. Wider uniformities

It is natural to approach Theorem 1 via the Large Sieve. Introducing Dirichlet characters we may dominate $\Delta(J)$ by

$$\sum_{J/2 < j \le J} j^{-1} \sum_{\chi \ne \chi_0 (\bmod f(j))} \max_{y \le x} \left| \sum_{n \le y} \chi(n) \Lambda(n) \right|.$$

Replacing each character by the primitive character which induces it, reduces us to the consideration of

$$\sum_{w \ll J^{\deg f}} \sideset{}{^*}\sum_{\chi (\bmod w)} \max_{y \le x} \left| \sum_{n \le y} \chi(n) \Lambda(n) \right| \sum_{\substack{J/2 < j \le J \\ f(j) \equiv 0 (\bmod w)}} j^{-1}.$$

The innersum weight is $\ll \rho(w)(w^{-1} + J^{-1})$, a bound that for $w > J$ is poor. The sum of these weights is not more than $2J^{-1} \sum \tau(f(j))$, $j \le J$, and by a remark preceding Lemma 14, $\ll \log J$. Very few moduli w need be considered but the standard versions of the Large Sieve are concerned more with the size than the distribution of the moduli, and naturally approach

$$L(J) = J^{-1} \sum_{w \ll J^{\deg f}} \sideset{}{^*}\sum_{\chi (\bmod w)} \max_{y \le x} \left| \sum_{n \le x} \chi(n) \Lambda(n) \right|$$

where the moduli w are (otherwise) allowed to traverse all positive integers.

With $d = \deg f$, the proof of the Bombieri-Vinogradov theorem given by Vaughan [15] delivers a bound

$$L(J) \ll J^{-1}(x + x^{5/6} J^d + x^{1/2} J^{2d})(\log xJ)^4.$$

For polynomials of degree two or more we gain a version of Theorem 1 under the restriction $6(\deg f - 1)\beta < 1$; a worse result.

Ignoring the difficulties of the reduction to bilinear forms, the Large Sieve gives for the sum considered in Lemma 8 at best a bound of the form

$$((M + \max D_j^2)(N + \max D_j^2) \sum |a_m|^2 \sum |b_n|^2)^{1/2} (\log MN)^{c_1}, \quad c_1 > 0.$$

Applied to estimate $L(J)$ this will yield at best a bound

$$\ll x(\log x)^{c_2}(M^{-1/2}J^{d-1}+N^{-1/2}J^{d-1}+x^{-1/2}J^{2d-1}+J^{-1})$$

with some 'worst case' values of M,N that satisfy $x(\log x)^{-A}<MN\le x$.

We see that the case $d=1$, which arises in the Bombieri–Vinogradov theorem, is particularly well served; M and N need only exceed a suitable power of $\log x$, and J not quite reach $x^{1/2}$. Once $d>1$, M and N must exceed J^{2d-1} and be appreciably large. Moreover, M and N should be approximately equal. Unless we take k in Lemma 1 large, this is not easily arranged. Implicitly, we would need Lemma 6 for large values of k with little degradation of the error term, a result that is at present beyond reach.

Disregarding all other difficulties, the best version of Theorem 1 to be reached by the application(s) of the Large Sieve indicated so far would apparently require that $(4\deg f-2)\beta<1$. I show that for polynomials of degree above 8 the argument of the present paper will already achieve better.

In the notation of §5 we set $k=5$. For each non-principal Dirichlet character $(\bmod D)$, integration by parts using the Pólya–Vinogradov inequality gives

$$L(\sigma+it,\chi)\ll((2+|t|)\log((2+|t|)D))^{1/2}D^{1/4}$$

in the strip $|\sigma-1/2|\le 2(\log(D(2+|t|)))^{-1}$. We can obtain a version of Lemma 4 with $r=4$ if we inflate the bound by a factor $D^{1/4}(T+2)^{1/2}(\log D(T+2))^{1/2}$. This legitimizes a version of Lemma 6 with $r=5$ and the second error term replaced by $x^{1/2}D^{1/4}(\log x)^{B+9}$; we employ the second of the bounds on the kernel $K(s)$ over the whole range.

In the proof proper of Theorem 2 the contribution of the terms with $m_1\cdots m_4\le x^{4/15}$ leads to an admissible error provided $\max sD_j\le x^{22/75}(\log x)^{-c}$ for a certain $c=c(A)>0$.

Bearing in mind that each M_i is at most $x^{1/5}$, whenever $M_1\cdots M_r>x^{4/15}$ there is a subset of the M_u whose product lies in $[x^{4/15},x^{7/15})$. We can apply Lemma 9 with (in the notation of that lemma) $M>x^{8/15}(\log x)^{-A}$, $N\ge x^{4/15}$.

Since $4/15<22/75$, Theorem 2 is valid for $\max sD_j\le x^{4/15-\delta}$, $s\le x^{\delta}$, and Theorem 1 for $\beta\deg f<4/15$. This last condition is more generous than $(4\deg f-2)\beta<1$ as soon as $\deg f>8$. With even a weak improvement in Lemma 6, the Dispersion method pulls away from the Large Sieve.

The study of the divisor function $\tau_r(n)$ on arithmetic progressions made by Friedlander and Iwaniec [7], using trilinear forms and Burgess' estimate for character sums, readily modifies to give a version of Lemma 6 valid for $r=6$, $D\le x^{5/12-\varepsilon}$, and $\varepsilon>0$. This allows the choice $k=6$ in the proof of Theorem 2 and the looser restraint $\max sD_j\le x^{5/18-\delta}$. The corresponding bound $\beta\deg f<5/18$ in Theorem 1 improves upon the best possible Large Sieve bound as soon as $\deg f>5$.

For cube-free moduli Friedlander and Iwaniec sharpen their result. We may correspondingly choose $k=9$ in the proof of Theorem 2, allowing $\max sD_j\le x^{\theta-\varepsilon}$ with $\theta=95/324$. The applicability of Theorem 2 towards a proof of Theorem 1 then depends upon the nature of the polynomial f, which may have a polynomial divisor that is already a cube.

With the simple version of the Dispersion Method employed in the present paper the natural limit of the exponents 1/4, 4/15, 5/18,…,95/324 arising in the various versions of Theorem 2 appears to be 1/3.

We may further refine the Dispersion Method and the Large Sieve, too, by the introduction of Fourier Analysis on $\mathbb{R}/\mathbb{Z}$. For the Dispersion Method this devolves to the estimation of sums

$$\sum_{J/2<j_1\le J}\left|\sum_{J/2<j_2\le J}\exp(2\pi ib\overline{f(j_2)}/f(j_1))\right|,$$

possibly without the absolute values. Here $f(j_2)\overline{f(j_2)}\equiv 1(\bmod f(j_1))$, b is an integer parameter and I have slightly simplified matters.

With obvious exceptions the results of Deligne yield an estimate

$$\sum_{j\le J}\exp(2\pi ir(j)/p)\ll p^{1/2}\log p$$

for any rational function $r(x)=f(x)/g(x)$, f,g in $(\mathbb{Z}/p\mathbb{Z})[x]$, provided $1\le J\le p$ and summation is restricted to classes $j(\bmod p)$ for which $g(j)$ does not vanish. We identify $1/g(j)$ with $\overline{g(j)}(\bmod p)$.

An elementary argument, again with some constraints upon the algebraic nature of $r(x)$, yields a similar result with the modulus p replaced by a prime power p^n and the bound by $\ll p^{n/2}\log p^n$.

I have worked out the details of both the arguments. Still, we gain for the inner exponential sum over j_2 a bound that differs by only a few effectively small factors from $f(j_1)^{1/2}$. Since $f(j_1)\gg J^{\deg f}$, for polynomials of degree at least two the bound is no better than the trivial, for cubics and up much worse. A direct application of these estimates appears insufficient.

There are further possibilities but I do not consider them here. The circle of problems attached to the study of primes in short arithmetic progressions with widely spaced differences appears challenging and interesting.

References

1. Barban, M.B. Multiplicative functions on $\sum_R$-equidistributed sequences, *Izv. Akad. Nauk Uz. SSR* (l964), 6, 13–19. MR **31:**1239
2. Barban, M.B. The "large sieve" method and its applications to number theory, *Uspekhi Mat. Nauk*. **21** (1966), 51–102 = *Russian Math. Surveys* **21** (1966), 49–104. MR **33:**7320
3. Bombieri, E. On the large sieve, *Mathematika* **12** (1965), 201–225. MR **33:**5590
4. Elliott, P.D.T.A. *Probabilistic Number Theory I: Mean Value Theorems*, Grund. der Math. Wiss. **239**, Springer Verlag, New York, Heidelberg, Berlin, 1979. MR **82h:**10002a
5. Elliott, P.D.T.A. Primes, products and polynomials, preprint, to appear.
6. Erdös, P. On the sum $\sum_{k=1}^{x} d(f(k))$, *J. London Math. Soc.* **27** (1952), 7–15. MR **13:**438f
7. Friedlander, J.R., Iwaniec, H. The divisor problem for arithmetic progressions, *Acta Arithmetica* **45** (1985), 273–277. MR **87b:**11087
8. Gallagher, P.X. A large sieve density estimate near $\sigma=1$, *Invent. Math.* **11** (1970), 329–339. MR **43:**4775
9. Heath–Brown, R. Sieve identities and gaps between primes, *Journées Arithmétiques*, Metz, 1981, *Astérisque* **94**, Soc. Math. de France (1982), 61–65.
10. Linnik, U.V. All large numbers are sums of a prime and two squares (A problem of Hardy and Littlewood) I, *Mat. Sb.* (N.S.), **52** (94) (1960), 661–700. MR **22:**10963
11. Linnik, U.V. *The Dispersion Method in Binary Additive Problems*, Leningrad, 1961, = AMS Transl. of Math. Monographs **4**, Providence, Rhode Island, 1963. MR **25:**3920; MR **29:**5804
12. Montgomery, H.L., *Topics in Multiplicative Number Theory*, Lecture Notes in Math. **227**, Springer-Verlag, Berlin, Heidelberg, New York, 1971. MR **49:**2616
13. Prachar, K. *Primzahlverteilung*, Grund. der Math. Wiss. **91**, Springer Verlag, Berlin, Göttingen, Heidelberg, 1957. MR **19:**393b

14. Uzdavinis, R.V. On the joint distribution of values of additive arithmetic functions of integral polynomials, *Liet. TSR Mokslu Akad. darbai*, Ser. B2 (18), 1959 = *Trud. Akad. Nauk. Lit. SSR*, Ser. B2 (18) (1960), 9–29. MR **26:**100
15. Vaughan, R.C. An elementary method in prime number theory, *Acta Arithmetica* **37** (1980), 111–115. MR **82c:**10055
16. van der Corput, J.G. "Une inégalité relative aux nombres des diviseurs", *Koninklijke Nederlandsche Akademie von Wetenschapen, Proceedings*, **42** (1939), 547–553. MR **1:**41b
17. Wolke, D. Multiplikative Funktionen auf schnell wachsenden Folgen, *J. Reine Angew. Math.* **251** (1971), 54–67. MR **44:**6629
18. Wolke, D. A new proof of a theorem of van der Corput, *J. London Math. Soc.* (2), **5** (1972), 609–612. MR **47:**3336

Department of Mathematics, University of Colorado Boulder, Boulder, Colorado 80309–0395

E-mail address: pdtae@euclid.colorado.edu