An Extension of the Linnik Phenomenon

YOICHI MOTOHASHI

Abstract: We shall prove an extension of Yu.V. Linnik’s phenomenon concerning C.L. Siegel’s exceptional zeros of Dirichlet $L$-functions. Also we shall prove a new version of the Brun–Titchmarsh theorem. These two subjects are closely related to each other. The basic tool is our theory of A. Selberg’s $\Lambda^2$-sieve, which is developed in our old Tata lecture notes [12]. Constants which are involved either explicitly or implicitly in our discussion are all effectively computable, save for the one in the statement of a theorem of Siegel which has no relation with our main assertions. We add that the discussion in Sections 1–4 is a substantially revised and augmented version of Section 9.3 of our recent monograph [15, Vol. I] and that of Section 5 an extraction from our article [16] posted in arXiv.

Keywords: Siegel’s zeros; $\Lambda^2$-sieve; Brun–Titchmarsh theorem

1. Our motivation in a historical perspective. We shall first make explicit our notion of exceptional zeros: Let $\chi$ denote a genetic Dirichlet character, with which the $L$-function $L(s, \chi)$ is associated. Let $T > 0$ be sufficiently large, and let $Z_T = \{\rho = \beta + i\gamma\}$ be the set of all non-trivial zeros $\rho$ in the region $|\text{Im} \ s| \leq T$ of the function $\prod_{q \leq T} \prod_{* \chi \mod q} L(s, \chi)$; here and in what follows the asterisk means that relevant characters are primitive. Then we have that there exists an effectively computable absolute constant $a_0 > 0$ such that

$$\max_{\rho \in Z_T} \beta < 1 - \frac{a_0}{\log T}, \text{ save for a possibly existing } \beta_T \in Z_T. \quad (1.1)$$

If $\beta_T$ exists, it should be real and simple, and we designate both itself and the relevant unique primitive character $\chi_T$, with $L(\beta_T, \chi_T) = 0$, as $T$-exceptional. It is known that $\chi_T$ is real. As a matter of fact, we have more precisely

$$1 - \frac{a_0}{\log T} \leq \beta_T \leq 1 - \frac{1}{\sqrt{T}(\log T)^3}. \quad (1.2)$$

The upper bound is a consequence of the Dirichlet class number formula for quadratic number fields. However, the formula can be dispensed with. See [7, Section 2, Chapter IX] as well as [15, Section 4.5, Vol. I]; in both monographs are comprehensible accounts of the theory of the zeta and Dirichlet $L$-functions and the theory of the distribution of primes.

One of the most tantalising problems in number theory is the elimination of the possibility of the existence of exceptional zeros. It is generally believed that they do not exist at all. A way to confirm this is to improve the Brun–Titchmarsh theorem in the manner to be made explicit at a later part of Section 5, which remains, however, to be one of the most difficult problems in number theory. Also one may have a hope in employing our Theorem 1 for this purpose; we shall try to be precise in one of the remarks there.
Exceptional zeros, if exist, would mar the quality of the prime number theorem for arithmetic progressions, especially when the uniformity is taken into account with respect to varying moduli. Whence is acclaimed Siegel’s assertion (1936) that for any fixed $\varepsilon > 0$ there exists a $c(\varepsilon) > 0$ such that

$$\beta_T < 1 - \frac{c(\varepsilon)}{T^\varepsilon},$$

since this yields a fine uniformity as far as moduli are relatively small yet meaningful as demonstrated in a variety of fundamental works such as I.M. Vinogradov’s resolution (1937) of the ternary Goldbach conjecture and the mean prime number theorem of E. Bombieri and A.I. Vinogradov (1965, independently).

However, $c(\varepsilon)$ is ineffective; that is, all known proofs of Siegel’s theorem assert only the existence of $c(\varepsilon)$ for each $\varepsilon$ and do not provide any means to evaluate its actual values. This causes severe difficulties in various basic problems. The most outstanding among them is the estimation of the size of the least prime that appears in a given arithmetic progression. To resolve this particular difficulty, Linnik (1944) greatly refined Siegel’s theorem (1.3) by providing a quantitative version of the Deuring–Heilbronn phenomenon or the repelling effect of an exceptional zero towards other zeros that had been observed in the efforts to solve C.F. Gauss’ conjecture concerning the class numbers of imaginary quadratic number fields. Linnik’s theorem or rather his phenomenon asserts that the inequality

$$\beta \leq 1 - \frac{a_0}{\log T} \log \frac{a_0 e}{(1 - \beta_T) \log T}$$

holds for all $\rho \neq \beta_T$ in $\mathbb{Z}_T$. With this and a zero density result of a special type which is another important contribution by him, Linnik could reach his famed Least Prime Number Theorem for arithmetic progressions.

Linnik’s argument is, however, exceptionally involved, especially in his intriguing use of an analytic convexity argument. After K.A. Rodosskii’s simplification, P. Turán and S. Knapowski (1961-62) developed a relatively accessible alternative approach to Linnik’s two assertions via Turán’s power sum method which superseded the convexity argument. Their theory was later improved considerably by E. Fogels (1965).

Then, a totally different approach was devised by A. Selberg (1973-74), which was in essence yet another replacement of the convexity argument by an infusion of his quasi-characters into the theory of $L$-functions and the theory of the large sieve of Linnik; see Théorème 7 of [1]. We observed later that Selberg’s quasi-characters come from the optimal $\lambda$-weights for the simplest situation of his $\Lambda^2$-sieve, that is, sifting integers in an interval with residue classes to be discarded being equal to 0. With this, we were able to extend the notion of quasi-characters by considering the $\Lambda^2$-sieve applied to sequences of values of a class of multiplicative functions. Combined the result with the large sieve, we obtained a fairly simplified proof [11] of Linnik’s least prime number theorem or more precisely its extended version due to P.X. Gallagher (1970) who had used the power sum method following Fogels. In passing, we mention that M. Jutila [6] developed an argument which is similar to but less general than ours.
In retrospect, the rôle played by the sieve method in our argument is in one way an enhancement of a sieve aspect in Linnik’s method as well as Turán–Knapowksi’s, Fogels’ and Gallagher’s. Namely those people applied the Brun–Titchmarsh theorem at respective crucial stages in their arguments.

Now, the aim of the present work is to extend the Linnik phenomenon (1.4) to zeros of \( L \)-functions which may belong to a family of functions much wider than hitherto considered, by employing this sieve argument of ours. Thus, if any in the family has a zero that violates (1.4), then \( \beta_T \) should not exist.

Numerous extensions of the Deuring–Heilbronn phenomenon and Siegel’s theorem have been considered, but the present work of ours appears to be the first concerning Linnik’s. Among the former is D.M. Goldfeld’s fundamental work [2]. He related exceptional zeros with the vanishing of the central values of \( L \)-functions attached to elliptic curves and thus the Birch–Swinnerton-Dyer conjecture came into the scene. An appropriate case of the conjecture was confirmed later by B. Gross and D. Zagier (1986), which thus settled Gauss’ conjecture mentioned above in an effective fashion. However, despite its amazing depth, their work does not yield any significant consequence in the theory of the distribution of primes, since it does not eliminate the possibility of the existence of exceptional zeros and the implied zero-free region of Dirichlet \( L \)-functions is unfortunately too weak to be applied.

2. Basic tools. Hereafter until the end of Section 4, we shall assume that the \( T \)-exceptional character exists. We put

\[
\beta_T = 1 - \delta, \tag{2.1}
\]

and introduce the multiplicative function

\[
f(n) = \sum_{d \mid n} \chi_T(d)d^{-\delta}, \tag{2.2}
\]

which is always positive. We consider the \( \Lambda^2 \)-sieve situation

\[
\sum_{n < N} f(n) \left( \sum_{d \mid n} \lambda_d \right)^2, \tag{2.3}
\]

where \( \lambda_1 = 1 \) and \( \lambda_d = 0 \) for \( d \geq z \), with a large parameter \( z \). The optimal \( \{ \lambda_d \} \) are given by

\[
\lambda_d = \mu(d) F_d \frac{G_d(z/d)}{G_1(z)}, \tag{2.4}
\]

where \( \mu \) is the Möbius function,

\[
F_d = \prod_{p \mid d} F_p, \quad F_p = \left( 1 - \frac{1}{p} \right)^{-1} \left( 1 - \frac{\chi_T(p)}{p^{1+\delta}} \right)^{-1}, \tag{2.5}
\]
and
\[ G_d(x) = \sum_{d < x} \frac{\mu^2(r)}{g(r)}, \quad g(r) = \prod_{p \mid r} (F_p - 1)^{-1}, \quad (2.6) \]
with \( p \) denoting a generic prime. The \( \{\lambda_d\} \) defined by (2.4) lead us to the multiplicative function
\[ \Phi_r(n) = \mu((r, n))g((r, n)), \quad (2.7) \]
where \((r, n)\) denotes the greatest common divisor of \( r \) and \( n \); in fact, we have
\[ \sum_{d \mid n} \lambda_d = \frac{1}{G_1(z)} \sum_{r < z} \frac{\mu^2(r)}{g(r)} \Phi_r(n). \quad (2.8) \]
The functions \( \{\Phi_r\} \) have a quasi-orthogonality like Selberg’s quasi-characters as is exhibited in

**Lemma 1.** We have, for any complex numbers \( \{a_n\} \) and for arbitrary \( M, N, z \) with \( N \ll M, z \geq T \),
\[ \sum_{r < z} \frac{\mu^2(r)}{g(r)} \left| \sum_{M \leq n < M + N} \Phi_r(n)f(n)a_n \right|^2 \leq \left( FN + O(z^5M^{2/3}) \right) \sum_{M \leq n < M + N} f(n)|a_n|^2, \quad (2.9) \]
where \( F = L(1 + \delta, \chi_T) \).

**Proof.** See [12, p. 33 and p. 187]; in fact a stronger result is given there, but this simplified version suffices for our present purpose. The argument is, in essence, an application of the duality principle \( \|D\| = \|D^*\| \) concerning an arbitrary bounded linear operator \( D \) and its adjoint \( D^* \) over a Hilbert space.

**Lemma 2.** We have, for any \( z \geq T \),
\[ \delta^{-1}F \ll G_1(z), \quad (2.10) \]

**Proof.** See [12, p. 187]. The quantity \( 1/G_1(z) \) represents the sieve effect of the optimal \( \{\lambda_d\} \).

The following assertion is essential for our purpose but is independent of the above:

**Lemma 3.** Let \( v \) be a large positive parameter and \( \vartheta > 0 \) an arbitrary constant. We put, with an integer \( l \geq 1 \),
\[ \Xi_d^{(l)} = \frac{1}{l!} (\vartheta \log v)^{-l} \sum_{j=0}^{l} (-1)^{l-j} \binom{l}{j} \xi_d^{(j, l)} \quad (2.11) \]
where
\[
\xi_d^{(j,l)} = \begin{cases} 
\mu(d)(\log \frac{v^{1+j\vartheta}}{d})^l & d < v^{1+j\vartheta}, \\
0 & d \geq v^{1+j\vartheta}.
\end{cases}
\] (2.12)

Then we have
\[
\Xi_d^{(l)} = \mu(d), \quad d < v
\] (2.13)
as well as
\[
\sum_{n=1}^{\infty} d_l(n) \left( \sum_{d|n} \xi_d^{(l)} \right)^2 n^{-\omega} \ll 1,
\] (2.14)
whenever \(\omega \geq 1 + 1/\log v\); here \(d_l(n)\) is as usual the number of ways of representing \(n\) as a product of \(l\) integral factors.

**Proof.** This is Theorem 4 on [12, p. 24].

### 3. The hypothesis.
We then deal with a multiplicative function \(h\) that is to be convolved with \(f\). We shall impose conditions upon \(h\). For this sake we need to introduce the following notations:

\[
F(s, h) = \prod_p F_p(s, h), \quad F_p(s, h) = \sum_{l=0}^{\infty} f(p^l)h(p^l)p^{-ls},
\]
\[
H(s) = \prod_p H_p(s), \quad H_p(s) = \sum_{l=0}^{\infty} h(p^l)p^{-ls},
\] (3.1)
\[
H(s, \chi T) = \prod_p H_p(s, \chi T), \quad H_p(s, \chi T) = \sum_{l=0}^{\infty} \chi_T(p^l)h(p^l)p^{-ls},
\]

where \(\text{Re } s\) is assumed to be large initially. Then the conditions are:

(A) There is an \(L \in \mathbb{N}\) such that \(h(n) \ll d_L(n)\) for all \(n \in \mathbb{N}\);

(B) There exist constants \(\frac{1}{2} < \eta < 1\) and \(\gamma > 0\) such that the functions \(H(s)\) and \(H(s, \chi T)\) are regular and \(\ll (T|s|)^{\gamma}\) for \(\text{Re } s > \eta\);

(C) For any prime \(p\), the functions \(H_p(s)\), \(H_p(s, \chi T)\), and \(F_p(s, h)\) do not vanish for \(\text{Re } s > \eta\);

Since (C) is fulfilled by all sufficiently large primes, one may dispense with it.

**Lemma 4.** Provided (A), (C), \(\mu^2(r) = 1\), and \(\xi_d \ll \mu^2(d)\), we have, for \(\text{Re } s > 1\),
\[
\sum_{n=1}^{\infty} f(n)h(n)\Phi_r(n) \left( \sum_{d|n} \xi_d \right)n^{-s} = F(s, h)M_r(s, h; \xi),
\] (3.2)
where

\[ M_r(s, h; \xi) = g(r) \sum_{d=1}^{\infty} \xi_d \mu((r, d)) \prod_{p|d} (1 - F_p(s, h)^{-1}) \prod_{p|d} (F_p(s, h)^{-1} F_p - 1). \quad (3.3) \]

**Proof.** This is a simple modification of Lemma 5 on [12, p. 35], and the proof is analogous.

**Lemma 5.** If \( h \) in (3.2) is replaced by the constant function 1, then we have

\[ \sum_{r < v^{1+\vartheta}} \mu^2(r) \frac{\mu(r)}{g(r)} \{ M_r(1, 1; \Xi^{(l)}) \}_{2} \ll 1/(F \log v), \quad (3.4) \]

provided \( \log v \gg \log T \).

**Proof.** This is Lemma 6 on [12, p. 36].

4. **Main theorem.** We proceed to the proof of our main assertion which is given at the end of this section. We shall assume (A)–(C).

First we note that for \( \text{Re} \ s > \eta \)

\[ F(s, h) = H(s)H(s + \delta, \chi_T)W_0(s)W_1(s), \quad (4.1) \]

where

\[ W_0(s) = \prod_{p < P_0} \frac{F_p(s, h)}{H_p(s)H_p(s + \delta, \chi_T)}, \quad W_1(s) = \prod_{p \geq P_0} \frac{F_p(s, h)}{H_p(s)H_p(s + \delta, \chi_T)}, \quad (4.2) \]

with any sufficiently large constant \( P_0 \). The factor \( W_1(s) \) is obviously regular and bounded for \( \text{Re} \ s > \frac{1}{2} \), and \( W_0(s) \) is dealt with (C). Thus by (B) the function \( F(s, h) \) is regular for \( \text{Re} \ s > \eta \). We set, in the previous sections,

\[ z = T, \ v = T^A, \ \vartheta = 1, \ \xi_d = \Xi^{(l)}_d, \quad (4.3) \]

with a sufficiently large \( A \); actually \( A = 20 \) suffices. The constant \( l \) is to be fixed later. Then, as it follows from (3.3), the (C) implies that \( M_r(s, h; \Xi^{(l)}) \) is regular and \( \ll T^B \) for \( \text{Re} \ s > \eta \) with an effective constant \( B \).

With this, let us consider the expression

\[ X_r = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} F(\rho + w, h)M_r(\rho + w, h; \Xi^{(l)}) \Gamma(w)V^w dw. \quad (4.4) \]

Here \( H(\rho) = 0 \), \( \rho = \nu + i\tau \), with

\[ \frac{1}{2}(\eta + 1) \leq \nu \leq 1, \ |\tau| < T, \quad (4.5) \]
and \( V = T^C \) with a sufficiently large \( C \). Note that we have

\[
F(\rho, h) = 0, \tag{4.6}
\]

because of (4.1). We have, by (2.13) and (3.2),

\[
X_r = e^{-1/V} + \sum_{n \geq v} \Phi_r(n) f(n) h(n) \left( \sum_{d \mid n} \Xi^{(l)}_d \right) n^{-\rho} e^{-n/V}. \tag{4.7}
\]

We shift the line of integration in (4.4) to \( \text{Re } w = -\frac{2}{3}(\nu - \eta) \) and by (4.6) get immediately

\[
X_r \ll V^{-(\nu-\eta)/2}, \tag{4.8}
\]

which implies that

\[
\frac{1}{4} \leq \left| \sum_{v \leq n \leq V^2} \Phi_r(n) f(n) h(n) \left( \sum_{d \mid n} \Xi^{(l)}_d \right) n^{-\rho} e^{-n/V} \right|^2. \tag{4.9}
\]

We multiply both sides by the factor \( \mu^2(r)/g(r) \) and sum over \( r < z \), getting

\[
\frac{G_1(z)}{\log T} \ll \sum_{N=2^m}^{N \leq N < V^2} \sum_{r < z} \frac{\mu^2(r)}{g(r)} \times \left| \sum_{N \leq n < 2N} \Phi_r(n) f(n) h(n) \left( \sum_{d \mid n} \Xi^{(l)}_d \right) n^{-\rho} e^{-n/V} \right|^2, \tag{4.10}
\]

where \( m \in \mathbb{N} \). By virtue of Lemma 1, we have

\[
\frac{G_1(z)}{\log T} \ll \mathcal{F} V^{2(1-\nu)} \sum_{n=1}^{\infty} f(n)|h(n)|^2 \left( \sum_{d \mid n} \Xi^{(l)}_d \right)^2 n^{-\omega_0}, \tag{4.11}
\]

with \( \omega_0 = 1 + (\log T)^{-1} \). Further, by (2.14) with \( l = 2L^2 \) we find that

\[
\frac{G_1(z)}{\log T} \ll \mathcal{F} V^{2(1-\nu)}. \tag{4.12}
\]

In view of Lemma 2, we have proved

**Theorem 1.** We assume the existence of the \( T \)-exceptional zero in the sense (1.1). Then, under the hypothesis (A)–(C), all the zeros of the function \( H(s) \) in the region (4.5) satisfy (1.4), provided the constant \( a_0 \) is adjusted appropriately.

A few remarks are in order:

(1) We are able to include the situation where \( H(s) \) has a simple pole at \( s = 1 \), which occurs, for instance, if we consider \( h \equiv 1 \), i.e., \( H(s) = \zeta(s) \). This violates (B). Nevertheless, the above argument works well, since we have Lemma 5, although we skip details. Thus,
our theorem is applicable to Dedekind zeta and Hecke $L$-functions of algebraic number fields as well. The case where $H(s)$ has multiple poles at $s = 1$ can also be included by an appropriate modification of Lemma 5.

(2) Hecke $L$-functions associated with holomorphic cusp forms on the hyperbolic upper half plane or more general $L$-functions of similar nature can be included in Theorem 1, provided they admit the twist by the real character $\chi_T$. On the other hand the situation with Maass forms is an open question, as the condition (A) is then hard to confirm. We surmise that there should be an appropriate modification of Lemma 4.

(3) It should be worth remarking that according to A. Ogg [18] there are Dirichlet series with Euler products which satisfy both (A) and (C) and vanish at $s = 1$. Thus, if we are ever able to apply Theorem 1 to any of his functions, then the existence of the exceptional zeros will be eliminated once and for all. However, what really matters is the confirmation of the regularity condition (B), which does not seem feasible. We wonder nevertheless if it is absurd to try to find a function which satisfies (A)–(C) and vanishes at $s = 1$.

(4) It is possible to extend the notion of exceptional zeros to those zeros lying close to the line $\text{Re } s = 1$ of any $L$-function which is not necessarily in Dirichlet’s family. In fact, one may extend Lemmas 1–4 for this purpose. This line of consideration yielded even a novel way [12, Section 4.1] to discuss zero-free regions of $\zeta(s)$; in fact, it gave later the assertion [14] that appears to be beyond the reach of the convexity argument of E. Borel and C. Carathéodory.

5. An extension of the Brun-Titchmarsh Theorem. In the rest we shall give a version of large sieve extensions of the Brun–Titchmarsh theorem. This is in fact a rework of our old file left unpublished since early 1980’s which we originally intended to include into our lecture notes [12]. We publish it here, as it might have now some fresh interest in the light of J. Maynard’s recent work [8] as well as what we have developed above.

Let

$$\pi(x; k, \ell) = \sum_{\substack{p \leq x \\ p \equiv \ell \mod k}} 1, \quad (k, \ell) = 1. \quad (5.1)$$

Then Theorem 13 on [12, p. 140] asserts, among other things, that we have, uniformly for $kQ^2 \leq x^{9/20-\varepsilon},$

$$\sum_{q \leq Q} \sum_{\chi \mod q} \left| \sum_{\substack{p \leq x \\ p \equiv \ell \mod k}} \chi(p) \right|^2 \leq \frac{(2 + o(1))x}{\varphi(k) \log(x/(kQ^2)^{3/8})} \pi(x; k, \ell), \quad (5.2)$$

provided $x$ is larger than a constant which is effectively computable for each sufficiently small $\varepsilon > 0$, where $\varphi$ is the Euler totient function. In particular we have

$$\pi(x; k, \ell) \leq \frac{(2 + o(1))x}{\varphi(k) \log(x/k^{3/8})}, \quad k \leq x^{9/20-\varepsilon}, \quad (5.3)$$
which surpasses partly the famed bound
\[
\pi(x + y; k, \ell) - \pi(x; k, \ell) \leq \frac{2y}{\varphi(k) \log(y/k)}, \quad k < y,
\]  \hspace{1cm} (5.4)
due to H.L. Montgomery and R.C. Vaughan [9]. In contrast to this, Maynard [8] asserts in essence that
\[
\pi(x; k, \ell) \leq \frac{2x}{\varphi(k) \log x}, \quad k \leq x^{1/8},
\]  \hspace{1cm} (5.5)
provided \(x\) is larger than an effectively computable constant. He gives also a lower bound, though we skip it in order to make our presentation simple; for the same reason, we also skip mentioning former improvements upon (5.4) other than (5.3). The bound (5.5) has been known as a kind of folklore among specialists, but with a much less precision about the range of moduli.

We shall refine (5.2) by

**Theorem 2.** There exists an effectively computable constant \(\Omega\) such that we have, uniformly for \(kQ \leq x^\Omega\),
\[
\sum_{q \leq Q} \sum_{\substack{\chi \mod q \, \star \, |}} \left| \sum_{p \leq x, p \equiv \ell \mod k} \chi(p) \right|^2 \leq \frac{2x}{\varphi(k) \log x} \pi(x; k, \ell).
\]  \hspace{1cm} (5.6)

**Proof.** This assertion is in fact a simple consequence of our version on [12, p. 185] of the Linnik–Fogels–Gallagher prime number theorem mentioned above. We put
\[
\psi(x, \chi) = \sum_{n \leq x} \chi(n) \Lambda(n),
\]  \hspace{1cm} (5.7)
with the von Mangold function \(\Lambda\); and let
\[
\tilde{\psi}(x, \chi) = \begin{cases} 
\psi(x, \chi) - x & \text{if } \chi \text{ is principal}, \\
\psi(x, \chi) + x^{\beta_T}/\beta_T & \text{if } \chi = \chi_T, \\
\psi(x, \chi) & \text{otherwise}.
\end{cases}
\]  \hspace{1cm} (5.8)
Then, Theorem 17, loc.cit., asserts that there exist effectively computable absolute constants \(a_1, a_2, a_3 > 0\) such that provided \(T^{a_3} \leq x \leq \exp((\log T)^2)\)
\[
\sum_{q \leq T} \sum_{\chi \mod q \, \star} |\tilde{\psi}(x, \chi)| \leq a_1 x \Delta_T \exp \left( - a_2 \log x / \log T \right),
\]  \hspace{1cm} (5.9)
where
\[
\Delta_T = \begin{cases} 
1 & \text{if } \chi_T \text{ does not exist}, \\
(1 - \beta_T) \log T & \text{if } \chi_T \text{ exists}.
\end{cases}
\]  \hspace{1cm} (5.10)
Then we note that
\[
\sum_{\substack{p \leq x \\
p \equiv \ell \mod k}} \chi(n)\Lambda(n) = \frac{1}{\varphi(k)} \sum_{\xi \mod k} \xi(\ell)\psi(x, \xi) \\
= \frac{1}{\varphi(k)} \sum_{\xi \mod k} \xi(\ell)\psi(x, \xi^\ell \chi) + O(\nu(k) \log x),
\]  
(5.11)
where \(\xi^\ell\) is the primitive character inducing the Dirichlet character \(\xi\), and \(\nu(k)\) the number of distinct prime factors of \(k\). Here \(\xi^\ell \chi\) stands for a unique primitive character whose conductor is not larger than \(k\). We have thus
\[
\sum_{\substack{q \leq Q \\
(q, k) = 1}} \sum_{\chi \mod q}^* \left| \sum_{\substack{p \leq x \\
p \equiv \ell \mod k}} \chi(n)\Lambda(n) \right| \\
\leq \frac{1}{\varphi(k)} \left\{ x + x^{\beta_0} / \beta_0 + E(x, kQ) \right\} + O(\nu(k)Q^2 \log x),
\]  
(5.12)
where \(\beta_0\) is the \(kQ\)-exceptional zero if exists; and \(E(x, kQ)\) is the left side of (5.9) for \(T = kQ\). If \(\beta_0\) exists, then (5.9) implies that
\[
x^{-1} \left( x^{\beta_0} / \beta_0 + E(x, kQ) \right) \leq \exp \left( - \Delta_T \log x / \log T \right) / (1 - \Delta_T / \log T) \\
+ a_1 \Delta_T \exp \left( - a_2 \log x / \log T \right),
\]  
(5.13)
provided \(T^{a_3} \leq x \leq ((\log T)^2)\). The right side is
\[
\leq \exp \left( - a_3 \Delta_T \right) / (1 - \Delta_T / \log T) + a_1 \Delta_T \exp(-a_2a_3) \\
\leq \exp(-a_3\Delta_T) + \Delta_T (1/(2 \log T) + a_1 \exp(-a_2a_3)) \\
< 1 - \frac{1}{2} a_3 \Delta_T,
\]  
(5.14)
since we may assume that \(a_3 \Delta_T\) is small while \(a_3\) is large. Hence we have proved that if \(\beta_0\) exists, then
\[
\sum_{\substack{q \leq Q \\
(q, k) = 1}} \sum_{\chi \mod q}^* \left| \sum_{\substack{p \leq x \\
p \equiv \ell \mod k}} \chi(n)\Lambda(n) \right| \leq 2 \frac{x}{\varphi(k)} \left( 1 - \frac{1}{5} a_3 \Delta_T \right),
\]  
(5.15)
provided \(a_3 \Delta_T\) is small and \(x \geq T^c = (kQ)^c\) with an effective absolute constant \(c > 0\). The case where the \(kQ\)-exceptional zero does not exist is analogous; in fact, simpler. The rest of the proof may be skipped, as it is a routine application of integration by parts.

Obviously (5.6) contains (5.5) but for \(k \leq x^\Omega\). It remains thus to find a good lower bound for \(\Omega\). We are certain that Maynard’s argument will extend to the direction indicated by (5.6) and yield (5.5) as a particular instance, since the basic structure of his
argument is essentially the same as ours that is developed in [11][12], although the intricate part of [8] corresponding to the numerical precision should be overhauled accordingly. Further, we add that it is possible to prove a short interval version of (5.6).

If the coefficient 2 in (5.2)–(5.6) is replace by any smaller effective constant, then the exceptional zeros should not exist; a proof can be found in [12, Section 4.3]. This means that the coefficient 2 will then be essentially halved immediately, which shows well a tantalising nature of the problem of improving the Brun–Titchmarsh theorem.

Here some more comments are in order: The bounds (5.2)–(5.4) are sieve results; that is, they are proved using mainly sieve arguments, without the zero-density theory or the Linnik phenomenon. The proof in [12] of the assertion (5.2) depends on Iwaniec’s work [4] on the bilinear structure in the error term of the combinatorial linear sieve; an alternative approach to his result itself can be found in [12] (see also [3]). Prior to [4], a bilinear structure in the error term of the $\Lambda^2$ sieve was observed in [10] and the first improvement upon (5.4) was achieved; see [5] as well. Later the development [13] made it possible to prove (5.2) via the $\Lambda^2$ sieve; see [17] for a further development. On the other hand, the bound (5.9) depends on our large sieve extension of the $\Lambda^2$ sieve that is devised via the duality principle and the quasi-character derived from optimal $\Lambda^2$-weights, as is already mentioned in the first section. Finally we remark that Selberg [19, Section 22] improved (5.4) by replacing the denominator by $\varphi(k)(\log(y/k) + 2.8)$. This is, however, definitely weaker than (5.3) in the range of variables indicated there. Apparently he did not notice the fact that (5.3) was attainable with his $\Lambda^2$-sieve.

References

[1] E. Bombieri. *Le Grand Crible dans la Théorie Analytique des Nombres.* (seconde édition revue et augmentée). Astérisque 18, Soc. Math. France, Paris 1987/1974.
[2] D.M. Goldfeld. The class number of quadratic fields and the conjectures of Birch and Swinnerton-Dyer. Annali della Scuola Norm. Sup. di Pisa Cl. Sci., (4) 3, 623–663 (1976).
[3] G. Greaves. *Sieves in Number Theory.* Springer-Verlag, Berlin 2001.
[4] H. Iwaniec. A new form of the error term in the linear sieve. Acta Arith., 37 (1980), 307–320.
[5] —. Sieve methods. In: *Intern. Congress of Math. Proc., Helsinki 1978*, Acad. Sci. Fennica, Helsinki 1980, pp. 357–364.
[6] M. Jutila. On Linnik’s constant. Math. Scand., 41 (1977), 45–62.
[7] A.A. Karatsuba. *Elements of Analytic Number Theory.* Nauka, Moscow 1975. (Russian)
[8] J. Maynard. On the Brun–Titchmarsh theorem. arXiv:1201.1777 (math.NT).
[9] H.L. Montgomery and R.C. Vaughan. The large sieve. Mathematika, 20 (1973), 119–134.
[10] Y. Motohashi. On some improvements of the Brun–Titchmarsh theorem. J. Math. Soc. Japan, 26 (1974), 306–323.
[11] —. Primes in arithmetic progressions. Invent. math., 44 (1978), 163–178.
[12] —. Sieve Methods and Prime Number Theory. Lect. Notes in Math. Phys., 72, Tata IFR and Springer-Verlag, Bombay 1983.

[13] —. On the error term in the Selberg sieve. In: Number Theory in Progress: A. Schinzel Festschrift, Walter de Gruyter, Berlin 1999, pp. 1053–1064.

[14] —. An observation on the zero-free region of the Riemann zeta-function. A. Sárközy Festschrift, Periodica Math. Hungarica, 42 (2001), 117–122.

[15] —. Analytic Number Theory. I. Distribution of Prime Numbers. Asakura Books, Tokyo 2009; II. Zeta Analysis. ibid, 2011. (Japanese)

[16] —. On some improvements of the Brun–Titchmarsh theorem. IV. arXiv:1201.3134v1 [math.NT].

[17] Y. Motohashi and J. Pintz. A smoothed GPY sieve. Bull. London Math. Soc., 40 (2008), pp. 298–310.

[18] A. Ogg. On a convolution of L-series. Invent. math., 7 (1969), 297–312.

[19] A. Selberg. Lectures on sieves. In: Collected Papers, II, Springer-Verlag, Berlin 1991, pp. 65–247.

Department of Mathematics,
Nihon University,
Surugadai, Tokyo 101-8308, JAPAN