TWIN PRIMES VIA EXCEPTIONAL CHARACTERS

JOHN B. FRIEDLANDER

Department of Mathematics, University of Toronto
Toronto, Ontario M5S 2E4, Canada
E-mail: frdlndr@math.toronto.edu

HENRYK IWANIEC

Department of Mathematics, Rutgers University
Piscataway, NJ 08903, USA
E-mail: iwaniec@comcast.net

Dedicated to Jerzy Kaczorowski in celebration
of his sixtieth birthday

Abstract. We give an account of the arguments that lead from the assumption of the existence of exceptional characters to the asymptotics in related ranges for the counting function of twin primes.

1. Introduction. The twin prime conjecture asserts that there are infinitely many primes p such that $p + 2$ is also prime. More generally, given a positive even number h, we expect the asymptotic formula

$$S_h(x) = \sum_{n \leqslant x} \Lambda(n)\Lambda(n + h) \sim BC(h)x \tag{1.1}$$

to hold as $x \to \infty$, where B is the absolute constant

$$B = 2\prod_{p>2}\left(1 - \frac{1}{p-1}\right)\left(1 - \frac{1}{p}\right)^{-1} \tag{1.2}$$

[95]

and $C(h)$ depends mildly on h, namely

$$C(h) = \prod_{\substack{p|h \\ p>2}} \left(1 - \frac{1}{p-1}\right)^{-1}. \tag{1.3}$$

This is a conjecture of Hardy and Littlewood. Actually, it is easy to predict more general formulas, such as (3.7) of [FI1] by exploiting the assumption of randomness of the Möbius function in conjunction with sieve ideas.

In these notes we present elementary arguments of sieve type which yield the asymptotic formula (1.1) with an error term estimated by means of the series

$$L(1, \chi) = \sum_{n=1}^{\infty} \chi(n) n^{-1} \tag{1.4}$$

with a primitive, real character $\chi \pmod{D}$.

THEOREM 1.1. *Let $x \geqslant D^{336000}$. For any even positive number h we have*

$$S_h(x) = BC(h)x + O\big(L(1,\chi)x \log x + x/\log x\big) \tag{1.5}$$

where the implied constant depends only on h.

The result is unconditional but is meaningful only if $L(1,\chi)$ is sufficiently small. We put

$$\eta(D) = L(1,\chi) \log D. \tag{1.6}$$

If $\eta(D)$ is small we call the character χ exceptional. For such characters the error term in (1.5), say $xE(x)$, is also relatively small, namely

$$E(x) \ll \eta(D)A + 1/\log D, \qquad \text{if } D^{336000} \leqslant x \leqslant D^A. \tag{1.7}$$

Hence, we conclude the following result of Heath-Brown [H-B].

COROLLARY 1.2. *If there are infinitely many exceptional characters then there are infinitely many twin prime numbers.*

2. A partition of $S_h(x)$. First, for notational simplicity, we write

$$S_h(x) = \sum_{m-n=h}\sum \Lambda(m)\Lambda(n) + O\big(h(\log x)^2\big)$$

where here and in the following we understand, but do not display, the conditions $0 < m, n \leqslant x$, $(mn, h) = 1$. Next, into $\Lambda(m)$, $\Lambda(n)$ we introduce upper-bound sieve factors $\theta(m)$, $\theta(n)$ which are almost redundant. Specifically, let (ξ_q) be sieve weights of level y and range $P(z)$ which is the product of all primes $p < z$, p not dividing h. This means we have real numbers ξ_q with $\xi_1 = 1$, $|\xi_q| \leqslant 1$, and $\xi_q = 0$ unless $q \mid P(z)$, $q < y$,

and such that for any positive integer m

$$\theta(m) = \sum_{q|m} \xi_q \geqslant 0. \tag{2.1}$$

Note that $\theta(m)$ is bounded by the divisor function $\tau(m)$. We get

$$S_h(x) = \sum_{m-n=h}\sum \theta(m)\theta(n)\Lambda(m)\Lambda(n) + O\big((h+z)(\log x)^3\big). \tag{2.2}$$

The level y and range z of the sieve are chosen as small powers of x, specifically

$$z^{700} = y = x^{1/120}.$$

We assume that $D^4 \leqslant z$. All implied constants in Sections 2, 3, 4, 5 are absolute.

Now we are ready to decompose Λ in terms of the following Dirichlet convolutions:

$$\lambda = \chi * 1, \qquad \lambda' = \chi * \log = \lambda * \Lambda, \qquad \nu = \mu\chi * \mu. \tag{2.3}$$

Note that $0 \leqslant \Lambda \leqslant \lambda' \leqslant \tau \log$ and $|\nu| \leqslant \lambda \leqslant \tau$. By Möbius inversion

$$\Lambda = \lambda' * \nu. \tag{2.4}$$

Having in mind that ν can be lacunary, we split $\Lambda = \Lambda^* + \Lambda_*$ with

$$\Lambda^*(n) = \sum_{\substack{ab=n \\ b<y}} \lambda'(a)\nu(b), \qquad \Lambda_*(n) = \sum_{\substack{ab=n \\ b\geqslant y}} \lambda'(a)\nu(b). \tag{2.5}$$

Note that, for economy of notation, we have chosen the splitting parameter y to be the same as that for the sieve level. The splitting parameter will be relatively small so the variable b in $\Lambda^*(n)$ is quite short and $\Lambda^*(n)$ looks like the divisor-type function λ' which can be analyzed by various means (such as by Dirichlet's switching-divisors technique, the δ-method, the circle method, the spectral decomposition of Poincaré series). On the other hand, the variable b in $\Lambda_*(n)$ is sufficiently long that the lacunarity of the factor $\nu(b)$ will kick in to yield crude but admissible estimates.

Writing $\Lambda\Lambda = \Lambda^*\Lambda^* + \frac{1}{2}(\Lambda + \Lambda^*)\Lambda_* + \frac{1}{2}\Lambda_*(\Lambda + \Lambda^*)$, we get

$$S_h(x) = S_h^*(x) + \tfrac{1}{2}T_h(x) + \tfrac{1}{2}T_{-h}(x) + O\big((h+z)(\log x)^3\big), \tag{2.6}$$

where

$$S_h^*(x) = \sum_{m-n=h}\sum \theta(m)\theta(n)\Lambda^*(m)\Lambda^*(n), \tag{2.7}$$

$$T_h(x) = \sum_{m-n=h}\sum \theta(m)\theta(n)\big(\Lambda(m) + \Lambda^*(m)\big)\Lambda_*(n). \tag{2.8}$$

3. Estimates for divisor-like functions.

In $T_h(x)$ we estimate $|\Lambda^*(m)| \leqslant b(m)\log x$ with

$$b(m) = \sum_{\substack{ab=m \\ b<y}} \tau(a)\tau(b).$$

Here we have (see Proposition 22.10 of [FI1]),

$$\tau(a) \leqslant \sum_{\substack{c|a \\ c\leqslant a^{1/4}}} \big(2\tau(c)\big)^4 \leqslant 16 \sum_{\substack{c|a \\ c\leqslant a^{1/4}}} \tau_{16}(c),$$

because

$$\tau_r(p^\alpha)\tau(p^\alpha)^{-4} = \binom{\alpha+r-1}{r-1}(\alpha+1)^{-4} \geqslant \binom{r}{r-1}2^{-4} = 1$$

if $r = 16$. Hence

$$b(m) \leqslant 16 \sum_{\substack{d|m \\ d<m^{1/4}y^{3/4}}} \tau_{18}(d).$$

For $m \leqslant x$ and $y \leqslant x^{1/9}$ we have divisors $d < x^{1/3}$. Moreover, we have

$$\tau_{18}(d) \leqslant 8\tau_{19}(d)\varphi(d)/d.$$

Finally, we estimate $\tau_{19}(d)\varphi(d)/d$ by

$$\gamma(d) = \frac{\varphi(d)}{d} \sum_{d_1\dots d_r=d} d_1^{\varepsilon_1}\dots d_r^{\varepsilon_r}, \tag{3.1}$$

where $\varepsilon_1,\dots,\varepsilon_r$ are distinct small positive numbers with $r = 19$. We obtain $b(m) \leqslant 2^7 c(m)$ with

$$c(m) = \sum_{\substack{d|m \\ d<x^{1/3}}} \gamma(d). \tag{3.2}$$

The reason for making a slight deformation of the divisor function $\tau_{18}(d)$ in the above argument is so as to reach a simple generating Dirichlet series, specifically

$$D(s) = \sum_d \frac{\gamma(d)}{\varphi(d)} d^{-s} = \zeta(s+1-\varepsilon_1)\dots\zeta(s+1-\varepsilon_r). \tag{3.3}$$

We take $\varepsilon_i = i/\log z$ for $1 \leqslant i \leqslant r = 19$. We also introduce $\varepsilon_0 = 0$.

Next we estimate $\Lambda_*(n)$ as follows:

$$|\Lambda_*(n)| \leqslant \sum_{ab=n,b\geqslant y} \lambda'(a)|\nu(b)| = \sum_{abc=n,b\geqslant y} \lambda(a)\Lambda(c)|\nu(b)|.$$

Note that $\nu(b)$ is multiplicative with $\nu(p) = -\lambda(p)$, $\nu(p^2) = \chi(p)$ and $\nu(p^\alpha) = 0$ for $\alpha \geqslant 3$. Moreover, we have $\lambda(p^\alpha) = 1 + \chi(p) + \dots + \chi(p)^\alpha$ if $\alpha \geqslant 1$, so it is easy to check that

$$\sum_{ab=k} \lambda(a)|\nu(b)| \leqslant 4^{\omega(k)}\lambda(k),$$

where $\omega(k)$ denotes the number of distinct prime divisors of k. Hence, we obtain $|\Lambda_*(n)| \leqslant a(n)$ with

$$a(n) = \sum_{bc=n,b\geqslant y} 4^{\omega(b)}\lambda(b)\Lambda(c). \tag{3.4}$$

From the above estimates we conclude that

$$|T_h(x)| \leqslant 2^8 V_h(x)\log x, \tag{3.5}$$

with

$$V_h(x) = \sum\sum_{m-n=h} \theta(m)\theta(n)c(m)a(n). \tag{3.6}$$

Recall that we are assuming $1 \leqslant m, n \leqslant x$, $(mn, h) = 1$.

4. Estimation of $V_h(x)$. By (2.1), (3.3) and (3.4) we arrange $V_h(x)$ into sums of congruence sums

$$V_h(x) = \sum_{\substack{q<y \\ (q,h)=1}} \xi_q \sum_{\substack{d<x^{1/3} \\ (d,h)=1}} \gamma(d) \sum_{\substack{n\equiv -h(\mathrm{mod}\,[d,q]) \\ (n,h)=1}} \theta(n)a(n), \tag{4.1}$$

where n runs over the segment $\max(0,-h) < n \leqslant \min(x, x-h)$. We extend this segment to $0 < n \leqslant x$ up to an error term of size $O(|h|(\log x)^{42} + \sqrt{x})$. Note that the sequence $\mathcal{A} = (\theta(n)a(n))$ with $(n,h)=1$, $0 < n \leqslant x$, admits a level of distribution $x^{1/2-\varepsilon}$. This can be achieved by the large sieve method. Therefore, our congruence sums are equidistributed over reduced residue classes apart from an error term $O(x^{1-\delta})$ provided that $Dy^2 \leqslant x^{1/9}$. We have

$$V_h(x) = M_h(x) + O\big(|h|(\log x)^{42} + x^{1-\delta}\big) \tag{4.2}$$

with

$$M_h(x) = \sum_{\substack{q<y \\ (q,h)=1}} \xi_q \sum_{\substack{d<x^{1/3} \\ (d,h)=1}} \frac{\gamma(d)}{\varphi([d,q])} \sum_{\substack{0<n\leqslant x \\ (n,hdq)=1}} \theta(n)a(n), \tag{4.3}$$

where $\delta > 0$ and the implied constants are absolute.

We are going to execute the summation over d first, which for given n and q, is equal to

$$L(x) = \sum_{\substack{d<x^{1/3} \\ (d,hn)=1}} \frac{\gamma(d)}{\varphi([d,q])} = \frac{1}{\varphi(q)} \sum_{c|q^\infty} \gamma(c) \frac{(c,q)}{c} \sum_{\substack{d<x^{1/3}/c \\ (d,hnq)=1}} \frac{\gamma(d)}{\varphi(d)}. \tag{4.4}$$

Note that the divisors c of q^∞ with $c > x^{1/9}$ contribute a negligible amount. For smaller c we evaluate the sum over d by contour integration of the series

$$D_v(s) = \sum_{(d,v)=1} \frac{\gamma(d)}{\varphi(d)} d^{-s} = P_v(s)D(s),$$

where $D(s) = D_1(s)$ is given by the product of zeta-functions (3.3) and $P_v(s)$ removes the local factors of $D(s)$ at primes $p|v$;

$$P_v(s) = \prod_{p|v} \big(1 - p^{\varepsilon_1-s-1}\big) \dots \big(1 - p^{\varepsilon_r-s-1}\big). \tag{4.5}$$

Hence $D_v(s)/s$ has simple poles at $s = \varepsilon_0 = 0$ and at $s = \varepsilon_1, \dots, \varepsilon_r$ with residue

$$R_v(i) = R(i)P_v(\varepsilon_i)$$

where $R(i)$ is the residue of $D(s)/s$ at $s = \varepsilon_i$, that is

$$R(i) = \kappa(i) \prod_{\substack{j=0 \\ j\neq i}}^{r} \zeta(1 + \varepsilon_i - \varepsilon_j)$$

with $\kappa(0) = 1$ and $\kappa(i) = 1/\varepsilon_i$ if $1 \leqslant i \leqslant r$. Note that

$$R(i) \asymp (\log z)^r. \tag{4.6}$$

By complex integration, the inner sum over $(d, hnq) = 1$ in (4.4) is equal to

$$\sum_{i=0}^{r} R(i) P_{hnq}(\varepsilon_i) \left(x^{1/3}/c \right)^{\varepsilon_i} \tag{4.7}$$

up to a small, negligible error term. Next we compute the resulting multiplicative functions (the sieve density)

$$g_\varepsilon(q) = \frac{P_q(\varepsilon)}{\varphi(q)} \sum_{c \mid q^\infty} \gamma(c)(c, q) c^{-\varepsilon - 1} \tag{4.8}$$

for every $\varepsilon = \varepsilon_i$, $0 \leqslant i \leqslant r$. At primes we have

$$g_\varepsilon(p) = \frac{P_p(\varepsilon)}{p-1} \left(1 + p \sum_{\alpha=1}^{\infty} \gamma(p^\alpha) p^{-\alpha(\varepsilon+1)} \right) = 1 - \frac{p-2}{p-1} P_p(\varepsilon).$$

Note that $0 < P_p(\varepsilon) \leqslant 1$, so $0 < g_\varepsilon(p) < 1$ if $p \neq 2$, which is the case since $p \mid q$, $(q, h) = 1$, h even. More precisely, we have

$$P_p(\varepsilon) = 1 - \sum_{j=1}^{r} p^{\varepsilon_j - \varepsilon - 1} + O\left(p^{-2}\right),$$

so

$$g_\varepsilon(p) = p^{-1} + \sum_{j=1}^{r} p^{\varepsilon_j - \varepsilon - 1} + O\left(p^{-2}\right), \tag{4.9}$$

for every $\varepsilon = \varepsilon_i$, $0 \leqslant i \leqslant r$. Since the $\varepsilon_i = i/\log z$ are small we have essentially a sieve problem of dimension $r + 1$.

By the above computations, we get

$$L(x) = \sum_{i=0}^{r} R(i) P_{hn}(\varepsilon_i) g_{\varepsilon_i}(q) x^{\varepsilon_i/3},$$

up to a small, negligible error term. Hence, (4.3) becomes

$$M_h(x) = \sum_{i=0}^{r} R(i) P_h(\varepsilon_i) x^{\varepsilon_i/3} \sum_{\substack{n \leqslant x \\ (n,h)=1}} \theta(n) a(n) P_n(\varepsilon_i) G_{hn}(\varepsilon_i) \tag{4.10}$$

where

$$G_v(\varepsilon) = \sum_{\substack{q < y \\ (q,v)=1}} \xi_q g_\varepsilon(q) \tag{4.11}$$

up to a small, negligible error term. By sieve methods,

$$G_v(\varepsilon) \asymp \prod_{\substack{p < z \\ p \nmid v}} \left(1 - g_\varepsilon(p)\right) \asymp \prod_{\substack{p < z \\ p \nmid v}} \left(1 - p^{-1}\right)\left(1 - p^{\varepsilon_1 - \varepsilon - 1}\right) \dots \left(1 - p^{\varepsilon_r - \varepsilon - 1}\right),$$

provided $z^{4r} \leqslant y$. Actually, this condition is not necessary if we claimed only the upper bound for $G_v(\varepsilon)$. For $s = 1 + O(1/\log z)$ we have

$$\prod_{p < z} \left(1 - p^{-s}\right) \asymp (\log z)^{-1}.$$

Put $\psi(v) = v/\varphi(v)$. Thus

$$G_v(\varepsilon) \asymp \psi(v) P_v(\varepsilon)^{-1} (\log z)^{-1}. \tag{4.12}$$

Inserting this into (4.10) we obtain an upper bound for $M_h(x)$, specifically

$$M_h(x) \ll \psi(h) A(x) (\log z)^{-1}. \tag{4.13}$$

where

$$A(x) = \sum_{n \leqslant x} \psi(n) \theta(n) a(n). \tag{4.14}$$

Recall that $\theta(n)$ is given by (2.1), $a(n)$ is given by (3.4) and $(n, h) = 1$.

First we examine the sieve factor $\theta = 1 * \xi$. We need a crude but very transparent bound for $\theta(n)$. Note that $\theta(n) = \theta\big((n, P(z))\big)$ because the sieve weights ξ_q are supported on divisors of $P(z)$. Assume (ξ_q) is the combinatorial upper-bound sieve with truncation parameters $y_r = (y/p_1 \ldots p_r)^{1/\beta}$, see (6.54) of [FI1], where $\beta \geqslant 2$ will be chosen later. Assume $z^{\beta+1} \leqslant y$. Then, for $m | P(z), m \neq 1$, we have

$$\theta(m) \leqslant \sum_{r \text{ odd}} \Big(\sum_{\substack{ap_1 \ldots p_r = m \\ z_r \leqslant p_r < \ldots < p_1 < z \\ p | a \Rightarrow p_r < p < z}} 1 \Big)$$

where

$$z_r = z^{((\beta-1)/(\beta+1))^{r/2}},$$

see (6.41) and (6.64) of [FI1]. Let $s = \omega(m) \geqslant 1$ be the number of prime factors of m. Then, we can choose r prime factors $p_r < \ldots < p_1$ in $\binom{s}{r}$ ways, $1 \leqslant r \leqslant s$. We have $z_s \leqslant z_r$ and

$$2 \sum_{\substack{1 \leqslant r \leqslant s \\ r \text{ odd}}} \binom{s}{r} = 2^s = \tau(m).$$

Hence $2\theta(m) \leqslant \tau(m)$ if m has all its prime factors in the segment $z_s \leqslant p < z$ and $\theta(m) = 0$ otherwise.

Now we are ready to estimate the sum (4.14). Our goal is

$$A(x) \ll L(1, \chi) x \log x. \tag{4.15}$$

Applying (3.4) for $a(n)$ and the combinatorial estimation for $\theta(m)$ described above, we estimate as follows:

$$A(x) \leqslant \sum_{\substack{bc \leqslant x, \, b \geqslant y \\ (bc, h) = 1}} 4^{\omega(b)} \lambda(b) \Lambda(c) \psi(bc) \theta(bc)$$

$$\leqslant \sum_{\substack{b'c'bc \leqslant x, \, b'b \geqslant y \\ p | bc \Rightarrow p | P(z) \\ p | b'c' \Rightarrow p \geqslant z}} 4^{\omega(b'b)} \lambda(b'b) \Lambda(c'c) \psi(b'bc'c) \theta(bc)$$

$$\leqslant \sum_{\substack{b'bc \leqslant x, \, b'b \geqslant y \\ p | b \Rightarrow z_{r+1} \leqslant p < z \\ p | b' \Rightarrow p \geqslant z}} 2^{2\omega(b')+3r} \lambda(b'b) \Lambda(c) \psi(b'bc)$$

where $r = \omega(b)$ is the number of prime factors of b.

Let d be the product of the first (that is smallest) $s = [r/2]$ distinct prime factors of b so that d is squarefree and $d \leqslant \sqrt{b}$. Note that $2s \leqslant r \leqslant 2s + 1$ and $z_{r+1} \geqslant z_{2s+2}$. Moreover, we have $\omega(b') \ll 1$ and $\psi(b'bc) \ll \psi(b)$. Hence, we obtain

$$A(x) \ll \sum_{\substack{dbc \leqslant x,\, b \geqslant \sqrt{y} \\ p|d \Rightarrow z_{2s+2} \leqslant p < z}} 2^{6s} \lambda(dc) \Lambda(c) \psi(b)$$

$$\ll x \sum_{\substack{d \leqslant x/\sqrt{y} \\ p|d \Rightarrow z_{2s+2} \leqslant p < z}} \tau(d)^6 d^{-1} \sum_{\sqrt{y} \leqslant b \leqslant x} \lambda(db) \psi(b) b^{-1}.$$

The sum over b is $\ll \tau(d)L(1, \chi) \log x$ using elementary arguments, so

$$A(x) \ll L(1, \chi) x (\log x) V$$

where

$$V = \sum_{p|d \Rightarrow z_{2s+2} \leqslant p < z} \tau(d)^7 d^{-1}.$$

Recall that $s = \omega(d)$ is the number of prime factors of d and that d runs over squarefree integers. This inequality will imply the bound (4.15) if we show that V is absolutely bounded. To see this we proceed as in (6.64)–(6.70) of [FI1], obtaining

$$V \leqslant 1 + \sum_{s \geqslant 1} \frac{1}{s!} \left(\sum_{z_{2s+2} \leqslant p < z} \frac{128}{p} \right)^s \leqslant \sum_{s \geqslant 1} \frac{1}{s!} \left(128s \log \tfrac{\beta+1}{\beta-1} + O(1) \right)^s$$

$$\leqslant \sum_{s \geqslant 1} \left(128e \log \tfrac{\beta+1}{\beta-1} + O(s^{-1}) \right)^s \ll 1,$$

(4.16)

provided that $128e \log \tfrac{\beta+1}{\beta-1} < 1$, which does indeed hold for $\beta = 696$.

Inserting (4.15) into (4.13) we get

$$M_h(x) \ll \psi(h) L(1, \chi) x.$$

Inserting this in turn into (4.2) and then into (3.5), we conclude by (2.6) the following nice approximation of $S_h(x)$ by $S_h^*(x)$.

PROPOSITION 4.1. *Let* $z^{700} = y = x^{1/120}$ *and* $D^4 \leqslant z$. *Let* $S_h(x)$ *and* $S_h^*(x)$ *be given by* (1.1) *and* (2.7) *respectively. We have*

$$S_h(x) = S_h^*(x) + x E_h(x)$$

(4.17)

with the error term satisfying

$$E_h(x) \ll \psi(h) L(1, \chi) \log x + (\log x)^{-2}$$

(4.18)

for every even h, $0 < h \leqslant H = x(\log x)^{-48}$, *where the implied constant is absolute.*

5. Variation of $S_h(x)$ in the shift. Opening the sieve functions $\theta(m)$, $\theta(n)$ in (2.7) and inserting (2.5) for $\Lambda^*(m)$, $\Lambda^*(n)$ we get

$$S_h^*(x) = \sum_{q_1, q_2 < y} \xi_{q_1} \xi_{q_2} \sum_{b_1, b_2 < y} \nu(b_1) \nu(b_2) A_h\big(x; q_1, b_1, q_2, b_2\big),$$

with

$$A_h(x; q_1, b_1, q_2, b_2) = \sum_{\substack{a_1 b_1 - a_2 b_2 = h \\ m = a_1 b_1 \equiv 0 (\mathrm{mod}\, q_1) \\ n = a_2 b_2 \equiv 0 (\mathrm{mod}\, q_2)}} \lambda'(a_1) \lambda'(a_2). \tag{5.1}$$

Recall that m, n run over $0 < m, n \leqslant x$ with $(mn, h) = 1$. Note that the last condition implies automatically that $(q_1 q_2 b_1 b_2, h) = 1$. Now we need an asymptotic formula for the "congruence sums" (5.1) which holds uniformly for $0 < h \leqslant H$ and $q_1 b_1 q_2 b_2 \leqslant y^4$. This is a problem of shifted convolution type for the divisor-like function $\lambda' = \chi * \log$. There is a vast literature on related subjects (see for example [DFI1, Theorem 1], [DFI2, Proposition 15.1] or [KMV]) but no result is stated which would exactly cover our sum (5.1). The closest seems to be Theorem 4.4 of [CI], which we adopt here without repeating the involved arguments (Kloosterman circle method with Weil's bound for Kloosterman sums). Fortunately, we do not need to use the results in an explicit form. In our current situation these arguments yield

$$\begin{aligned} A_h(x; q_1, b_1, q_2, b_2) = &\left\{ \mathfrak{S}(h) B(x) + O\big(\tau(h) (\log x)^{-20} \big) \right\} x / [q_1, b_1][q_2, b_2] \\ &+ O\big(\tau(h) (q_1 b_1 q_2 b_2)^6 x^{3/4} (\log x)^4 \big), \end{aligned} \tag{5.2}$$

where $\mathfrak{S}(h) = \mathfrak{S}(h; q_1, b_1, q_2, b_2)$ is a singular series and

$$B(x) = B(x; q_1, b_1, q_2, b_2) \ll x (\log x)^2$$

by trivial estimations. The error terms are negligible. The main term depends on h only via the singular series $\mathfrak{S}(h)$. The key feature of $\mathfrak{S}(h)$ is that it varies only slightly with respect to large prime divisors of h. It is given by an absolutely convergent series of Ramanujan sums. For our application the dependence on q_1, b_1, q_2, b_2 does not matter. Specifically, we can write $\mathfrak{S} = 1 * \delta$ with

$$\delta(d) \ll \tau(d) d^{-1}. \tag{5.3}$$

Hence, for $(k, h) = 1$ we derive

$$\mathfrak{S}(hk) - \mathfrak{S}(h) = \sum_{d | hk} \delta(d) - \sum_{d | h} \delta(d) = \sum_{a | h} \sum_{\substack{c | k \\ c > 1}} \delta(ac)$$

$$\ll \left(\sum_{a | h} \frac{\tau(a)}{a} \right) \left(\sum_{\substack{c | k \\ c > 1}} \frac{\tau(c)}{c} \right) \ll \psi(h)^2 \psi(k)(\psi(k) - 1),$$

where we recall that $\psi(k) = k / \varphi(k)$. Note that for k free of small prime divisors $\psi(k) - 1$ is small.

By the above observations we conclude the following

PROPOSITION 5.1. *Assume the conditions as in Proposition* 4.1. *Then, for every k with $(h, k) = 1$, $1 \leqslant k \leqslant H/h$, we have*

$$S_h^*(x) = S_{hk}^*(x) + O\big((\psi(k) - 1) x (\log x)^9 + x (\log x)^{-2} \big) \tag{5.4}$$

where the implied constant is absolute.

Combining (5.4) with (4.17) we obtain (under the above conditions)

$$S_h^*(x) = S_{hk}^*(x) + x E_{h,k}(x) \tag{5.5}$$

where the error term satisfies

$$E_{h,k}(x) \ll \psi(hk)L(1,\chi)\log x + (\psi(k)-1)x(\log x)^9 + x(\log x)^{-2}. \qquad (5.6)$$

6. Proof of Theorem 1.1. Fix a positive even number h. We shall average the relation (5.5) with respect to k over the set

$$\mathcal{K} = \{k : 0 < k \leqslant K, \ (k, hP) = 1\}$$

where $K = H/h$ and $P = P(w)$ is the product of all primes $p \nmid h$, $p < w$ with $w = (\log x)^{2016}$. Note that (by elementary sieve methods)

$$|\mathcal{K}| = K \prod_{p|hP} \left(1 - \frac{1}{p}\right)\{1 + O(1/\log x)\}, \qquad (6.1)$$

$$\sum_{k \in \mathcal{K}} \psi(k) \ll |\mathcal{K}|, \qquad (6.2)$$

and

$$\sum_{k \in \mathcal{K}} \bigl(\psi(k)-1\bigr) \ll \sum_{k \in \mathcal{K}} \Bigl(\sum_{\substack{c|k \\ c>1}} c^{-1}\Bigr) \leqslant K \sum_{\substack{(c,hP)=1 \\ c>1}} c^{-2} \ll K/w.$$

Hence, (5.5) yields

$$S_h(x) = \frac{1}{|\mathcal{K}|} \sum_{k \in \mathcal{K}} S_{hk}(x) + O\bigl(L(1,\chi)x\log x + x(\log x)^{-2}\bigr). \qquad (6.3)$$

Here we have

$$\sum_{k \in \mathcal{K}} S_{hk}(x) = \sum_{\substack{n \leqslant x \\ (n,hP)=1}} \Lambda(n) \sum_{k \in \mathcal{K}} \Lambda(n+hk) + O(xw)$$

where the error term $O(xw)$ takes care of the condition $(n, hP) = 1$ which is introduced here for technical reasons.

Next, we are going to execute the summation of $\Lambda(n + hk)$ over $k \in \mathcal{K}$. To this end we relax the condition $(k, hP) = 1$ by means of upper-bound and lower-bound sieves of level $\Delta = x^{1/2016}$. First we get an upper bound as follows:

$$\sum_{k \in \mathcal{K}} \Lambda(n+hk) \leqslant \sum_{\substack{d<\Delta \\ d|hP}} \xi_d \sum_{\substack{n<\ell\leqslant n+hK \\ \ell\equiv n\,(\mathrm{mod}\,dh)}} \Lambda(\ell)$$

$$= hK \sum_{\substack{d<\Delta \\ d|hP}} \xi_d/\varphi(dh) + O\bigl(x(\log x)^{-A}\bigr)$$

by the Bombieri–Vinogradov theorem. Here $g(d) = \varphi(h)/\varphi(dh)$ is the relevant multiplicative sieve density function for which we get

$$\sum_{\substack{d<\Delta \\ d|hP}} \xi_d g(d) = \bigl\{1 + O(1/\log x)\bigr\} \prod_{p|hP} \bigl(1 - g(p)\bigr).$$

Similarly we deal with the lower bound. Combining the two we obtain

$$\sum_{k\in\mathcal{K}} \Lambda(n+hk) = \big\{1 + O(1/\log x)\big\} \frac{Kh}{\varphi(h)} \prod_{p|hP} \big(1 - g(p)\big)$$

$$= BC(h)|\mathcal{K}|\{1 + O(1/\log x)\}$$

where B and $C(h)$ are given by (1.2), (1.3). Summing this over n we find

$$\frac{1}{|\mathcal{K}|} \sum_{k\in\mathcal{K}} S_{hk}(x) = BC(h)x\big\{1 + O(1/\log x)\big\}. \tag{6.4}$$

This, together with (6.3), completes the proof of Theorem 1.1.

References

[CI]   J. B. Conrey, H. Iwaniec, *Spacing of zeros of Hecke L-functions and the class number problem*, Acta Arith. 103 (2002), 259–312.

[DFI1] W. Duke, J. B. Friedlander, H. Iwaniec, *A quadratic divisor problem*, Invent. Math. 115 (1994), 209–217.

[DFI2] W. Duke, J. B. Friedlander, H. Iwaniec, *The subconvexity problem for Artin L-functions*, Invent. Math. 149 (2002), 489–577.

[FI1]   J. B. Friedlander, H. Iwaniec, *Opera de Cribro*, Amer. Math. Soc. Colloq. Publ. 57, Amer. Math. Soc., Providence, 2010.

[FI2]   J. B. Friedlander, H. Iwaniec, *Exceptional discriminants are the sum of a square and a prime*, Q. J. Math. 64 (2013), 1099–1107.

[H-B]   D. R. Heath-Brown, *Prime twins and Siegel zeros*, Proc. London Math. Soc. 47 (1983), 193–224.

[KMV]   E. Kowalski, P. Michel, J. VanderKam, *Rankin–Selberg L-functions in the level aspect*, Duke J. Math. 114 (2002), 123–191.