A Smoothed GPY Sieve

by

Yoichi Motohashi & János Pintz

Abstract: Combining the arguments developed in [2] and [7], we introduce a smoothing device to the sieve procedure [3] of D.A. Goldston, J. Pintz, and C.Y. Yıldırım (see [4] for its simplified version). Our assertions embodied in Lemmas 3 and 4 imply that an improvement of the prime number theorem of E. Bombieri, J.B. Friedlander and H. Iwaniec [1] should give rise infinitely often to bounded differences between primes.

To this end, a rework of the main part of [7] is developed in Sections 2–3; thus the present article is essentially self-contained, except for the first section which is an excerpt from [4].

1. Let N be a parameter increasing monotonically to infinity. There are four other basic parameters H, R, k, ℓ in our discussion; the last two are integers. We impose the following conditions to them:

$$H \ll \log N \ll \log R \le \log N, \tag{1.1}$$

and

$$1 \le \ell \le k \ll 1. \tag{1.2}$$

All implicit constants in the sequel are possibly dependent on k, ℓ at most; and besides, the symbol c stands for a positive constant with the same dependency, whose value may differ at each occurrence. It suffices to have (1.2), since our eventual aim is to look into the possibility to detect the bounded differences between primes with a certain modification of the GPY sieve. We surmise that such a modification might be obtained by introducing a smoothing device. The present article is, however, only to indicate that the GPY sieve admits indeed a smoothing; it is yet to be seen if this particular smoothing contributes to our eventual aim.

Let

$$\mathcal{H} = \{h_1, h_2, \ldots, h_k\} \subseteq [-H, H] \cap \mathbb{Z}, \tag{1.3}$$

with $h_i \ne h_j$ for $i \ne j$. Let us put, for a prime p,

$$\Omega(p) = \{\text{different residue classes among } -h (\operatorname{mod} p),\ h \in \mathcal{H}\} \tag{1.4}$$

2000 Mathematics Subject Classification: Primary 11N05; Secondary 11P32

The first author was supported by KAKENHI 15540047, and the second author by OTKA grants No. T38396, T43623, T49693 and Balaton program.

and write $n \in \Omega(p)$ instead of $n \pmod p \in \Omega(p)$. We call $\mathcal{H}$ admissible if

$$|\Omega(p)| < p \quad \text{for all } p, \tag{1.5}$$

and assume this unless otherwise stated. We extend Ω multiplicatively, so that $n \in \Omega(d)$ with square-free d if and only if $n \in \Omega(p)$ for all $p|d$, which is equivalent to

$$(n+h_1)(n+h_2)\cdots(n+h_k) \equiv 0 \pmod d. \tag{1.6}$$

We put, with μ the Möbius function,

$$\lambda_R(d;\ell) = \begin{cases} 0 & \text{if } d > R, \\ \dfrac{\mu(d)}{(k+\ell)!}\left(\log \dfrac{R}{d}\right)^{k+\ell} & \text{if } d \le R, \end{cases} \tag{1.7}$$

and

$$\Lambda_R(n;\mathcal{H},\ell) = \sum_{n\in\Omega(d)} \lambda_R(d;\ell). \tag{1.8}$$

Also, let

$$E^*(y;a,q) = \vartheta^*(y;a,q) - \frac{y}{\varphi(q)}, \quad \vartheta^*(y;a,q) = \sum_{\substack{y<n\le 2y \\ n\equiv a \ (\mathrm{mod}\, q)}} \varpi(n), \tag{1.9}$$

where φ is the Euler totient function; and $\varpi(n) = \log n$ if n is a prime, and $= 0$ otherwise. In all of the existing accounts [2]–[4] of the GPY sieve, it is assumed that

$$\sum_{q\le x^\theta} \max_{(a,q)=1} \max_{y\le x} |E^*(y;a,q)| \ll \frac{x}{(\log x)^{C_0}}, \tag{1.10}$$

with a certain absolute constant $\theta \in (0,1)$ and an arbitrary fixed $C_0 > 0$; the implied constant depending only on C_0.

The following asymptotic formulas (1.12) and (1.14) are the implements with which Goldston, Pintz and Yıldırım established

$$\liminf_{n\to\infty} \frac{\mathrm{p}_{n+1} - \mathrm{p}_n}{\log \mathrm{p}_n} = 0, \tag{1.11}$$

where p_n is the nth prime.

Lemma 1. *Provided* (1.1), (1.2), *and* $R \le N^{1/2}/(\log N)^C$ *hold with a sufficiently large* $C > 0$ *depending only on* k *and* ℓ, *we have*

$$\begin{aligned} &\sum_{N<n\le 2N} \Lambda_R(n;\mathcal{H},\ell)^2 \\ &\quad = \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell)!}\binom{2\ell}{\ell} N(\log R)^{k+2\ell} + O(N(\log N)^{k+2\ell-1}(\log\log N)^c), \end{aligned} \tag{1.12}$$

where

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{|\Omega(p)|}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}. \tag{1.13}$$

Lemma 2. *Provided* (1.1), (1.2), (1.10), *and* $R \le N^{\theta/2}/(\log N)^C$ *hold with a sufficiently large* $C > 0$ *depending only on* k *and* ℓ,

$$\begin{aligned}\sum_{N<n\le 2N} &\varpi(n+h)\Lambda_R(n;\mathcal{H},\ell)^2 \\ &= \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!}\binom{2(\ell+1)}{\ell+1} N(\log R)^{k+2\ell+1} + O(N(\log N)^{k+2\ell}(\log\log N)^c),\end{aligned} \tag{1.14}$$

whenever $h \in \mathcal{H}$.

A short self-contained treatment of the assertions (1.11)–(1.14) can be found in [4].

Note that the case $h \notin \mathcal{H}$ in the last lemma, which is included in [2]–[4], is irrelevant for our present purpose. In fact, a combination of (1.10), (1.12), and (1.14) gives, for $R = N^{\theta/2}/(\log N)^{C_0}$,

$$\begin{aligned}\sum_{N<n\le 2N} &\left\{\sum_{h\in\mathcal{H}} \varpi(n+h) - \log 3N\right\}\Lambda_R(n;\mathcal{H},\ell)^2 \\ =&(1+o(1))\frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell)!}\binom{2\ell}{\ell} N(\log R)^{k+2\ell}(\log N)\left(\frac{k}{k+2\ell+1}\cdot\frac{2(2\ell+1)}{\ell+1}\cdot\frac{\theta}{2} - 1\right).\end{aligned} \tag{1.15}$$

Thus, the k-tuple $(n+h_1,\ldots,n+h_k)$ with any fixed admissible $\mathcal{H}$ should contain two primes for infinitely many n, if the last factor in (1.15) is positive. Namely, with an appropriate choice of k, ℓ depending on θ we would be able to conclude that

$$\liminf_{n\to\infty} (\mathrm{p}_{n+1} - \mathrm{p}_n) < \infty, \tag{1.16}$$

provided $\theta > \frac{1}{2}$.

The aim of the present work is to prove a smoothed version of (1.12) and (1.14) in order to look into the possibility of replacing (1.10) with a $\theta > \frac{1}{2}$ by any less stringent hypothesis.

In passing, we note that the historical aspect of the Selberg sieve and the bilinear structure of its error term can be found in [8], including that of smoothed sieves which came later, and are naturally relevant to our present work.

Convention. All symbols and conditions introduced above are retained. We assume additionally that

$$H = H(k,\ell) \text{ is bounded,} \tag{1.17}$$

which should not cause any loss of generality under the present circumstance. Implicit constants may depend on k at most, but they can be regarded to be absolute once the least possible value of k is fixed. Thus the dependency on k of estimations will not be mentioned repeatedly, excepting at (4.15), (4.16), (5.15), and (6.1).

2. We shall first rework the main part of [7] in the present and the next sections (cf. [6, Sections 2.3 and 3.4]).

Thus let us put

$$R_0 = \exp\left(\frac{\log R}{(\log\log R)^{1/5}}\right), \quad R_1 = \exp\left(\frac{\log R}{(\log\log R)^{9/10}}\right). \tag{2.1}$$

We divide the half-line (R_0, ∞) into intervals $(R_0 R_1^{j-1}, R_0 R_1^j]$, $j = 1, 2, \ldots$, denoting them by P, with or without suffix. We let $|P|$ be the right end point of P.

Let

$$R_0 R_1 \le z \le R. \tag{2.2}$$

We consider the commutative semi-group $\mathcal{Y}(z)$ generated by all P such that $|P| \le z$. Let $D = P_1 P_2 \cdots P_r$ be an element of $\mathcal{Y}(z)$. Then the notation $d \in D$ indicates that d has the prime decomposition $d = p_1 p_2 \cdots p_r$ with $p_j \in P_j$ $(1 \le j \le r)$. We use the convention $1 \in D$ if and only if D is the empty product. Also, $|D|$ stands for $|P_1| \cdots |P_r|$. Naturally, $|D| = 1$ if D is empty.

Let ξ be a real valued function over $\mathcal{Y}(z)$, which satisfies the following conditions:

$$\xi(D) = \begin{cases} 0, & \text{if } |D| > R, \\ 0, & \text{if } D \text{ is not square-free}, \\ \text{arbitrary}, & \text{otherwise}, \end{cases} \tag{2.3}$$

with an obvious abuse of terminology. We are concerned with the quadratic form

$$\mathcal{J} = \sum_{D_1, D_2} \xi(D_1)\xi(D_2) \sum_{d_1 \in D_1,\, d_2 \in D_2} \frac{|\Omega([d_1, d_2])|}{[d_1, d_2]}, \tag{2.4}$$

where $[d_1, d_2]$ is the least common multiple of d_1, d_2.

In the inner double sum of (2.4), D_1 and D_2 can be supposed to be square-free, and by multiplicativity the sum is equal to

$$\begin{aligned} &\prod_{P_1|D_1,\, P_2|D_2} \left(\sum_{p_1 \in P_1,\, p_2 \in P_2} \frac{|\Omega([p_1, p_2])|}{[p_1, p_2]} \right) \\ &= \prod_{P_1|D_1} \left(\sum_{p_1 \in P_1} \frac{|\Omega(p_1)|}{p_1} \right) \prod_{P_2|D_2} \left(\sum_{p_2 \in P_2} \frac{|\Omega(p_2)|}{p_2} \right) \prod_{\substack{P|D_1 \\ P|D_2}} \frac{\left(\sum_{p, p' \in P} \frac{|\Omega([p, p'])|}{[p, p']} \right)}{\left(\sum_{p \in P} \frac{|\Omega(p)|}{p} \right)^2}, \end{aligned} \tag{2.5}$$

with primes p_1, p_2, p, p'. We then introduce

$$\Delta(D) = \prod_{P|D} \left(\sum_{p \in P} \frac{|\Omega(p)|}{p} \right), \tag{2.6}$$

$$\Phi(D) = \frac{1}{\Delta(D)^2} \prod_{P|D} \left(\sum_{p \in P} \frac{|\Omega(p)|}{p} \left(1 - \frac{|\Omega(p)|}{p} \right) \right). \tag{2.7}$$

Obviously Φ does not vanish; actually we have here $|\Omega(p)| = k$ but we retain the notation because of a future purpose. We have, for any square-free D,

$$\sum_{K|D} \Phi(K) = \frac{1}{\Delta(D)^2} \prod_{P|D} \left(\sum_{p,p' \in P} \frac{|\Omega([p,p'])|}{[p,p']} \right), \tag{2.8}$$

which is to be compared with the last factor in (2.5).

From (2.5)–(2.8), we get the diagonalisation

$$\mathcal{J} = \sum_{K} \Phi(K) \Xi(K)^2, \tag{2.9}$$

with

$$\Xi(K) = \sum_{K|D} \Delta(D) \xi(D). \tag{2.10}$$

Note that (2.3) induces the restriction that K be square-free and $|K| \le R$ in (2.9). Reversing (2.10), we have, with an obvious generalisation of the Möbius function,

$$\xi(D) = \frac{1}{\Delta(D)} \sum_{K} \mu(K) \Xi(KD); \tag{2.11}$$

and the case $D = \emptyset$ the empty product implies that

$$\mathcal{J} = \sum_{K} \Phi(K) \left(\Xi(K) - \frac{\xi(\emptyset)}{G(R,z)} \frac{\mu(K)}{\Phi(K)} \right)^2 + \frac{\xi(\emptyset)^2}{G(R,z)}, \tag{2.12}$$

where

$$G(y,z) = \sum_{|K| \le y} \frac{\mu(K)^2}{\Phi(K)}. \tag{2.13}$$

Note that the appearance of z here indicates that $K \in \mathcal{Y}(z)$.

We now set

$$\Xi(K) = \xi(\emptyset) \frac{\mu(K)}{G(R,z)\Phi(K)} \tag{2.14}$$

or by (2.11)

$$\xi(D) = \frac{\xi(\emptyset)}{G(R,z)} \frac{\mu(D)}{\Delta(D)\Phi(D)} \sum_{\substack{|K| \le R/|D| \\ (K,D)=1}} \frac{\mu(K)^2}{\Phi(K)}. \tag{2.15}$$

Then we have

$$\mathcal{J} = \frac{\xi(\emptyset)^2}{G(R,z)}. \tag{2.16}$$

Hereafter we shall work with (2.15), as (2.3) is obviously satisfied. It should be noted that we have now

$$|\xi(D)| \le |\xi(\emptyset)|, \tag{2.17}$$

since

$$G(R,z) \ge \sum_{L|D} \frac{\mu(L)^2}{\Phi(L)} \sum_{\substack{|K| \le R/|D| \\ (K,D)=1}} \frac{\mu(K)^2}{\Phi(K)} \tag{2.18}$$

and by (2.8)

$$\begin{aligned} \sum_{L|D} \frac{\mu(L)^2}{\Phi(L)} &= \frac{1}{\Phi(D)} \sum_{L|D} \mu^2(L)\Phi(L) \\ &= \frac{1}{\Delta(D)\Phi(D)} \cdot \frac{1}{\Delta(D)} \prod_{P|D} \left(\sum_{p,p' \in P} \frac{|\Omega([p,p'])}{[p,p']} \right) \\ &\ge \frac{1}{\Delta(D)\Phi(D)}. \end{aligned} \tag{2.19}$$

3. In this section we shall evaluate $G(z) = G(z,z)$ asymptotically; we are still working with $\mathcal{Y}(z)$. In fact, we shall treat more generally $G(z;Q)$ with $Q \in \mathcal{Y}(z)$, $\log|Q| \ll \log R$, which is the result of imposing the restriction $y = z$ and $(K,Q) = 1$ to the sum (2.13).

We define $G(y,z;Q)$ analogously, and introduce

$$T(y,z;Q) = \int_1^y G(t,z;Q)\frac{dt}{t}, \quad T_1(y,z;Q) = \sum_{\substack{|K| \le y \\ (K,Q)=1}} \frac{\mu(K)^2}{\Phi(K)} \log|K|, \tag{3.1}$$

so that

$$G(y,z;Q)\log y = T(y,z;Q) + T_1(y,z;Q). \tag{3.2}$$

Observe that for $1 \le y < R_0 R_1$

$$G(y,z;Q) = 1, \quad T(y,z;Q) = \log y, \quad T_1(y,z;Q) = 0. \tag{3.3}$$

Since $\log|K| = \sum_{P|K} \log|P|$ for square-free K, we have

$$T_1(y,z;Q) = \sum_{\substack{|P|\le z \\ P\nmid Q}} \frac{\log|P|}{\Phi(P)} G(y/|P|, z; PQ). \tag{3.4}$$

On the other hand we see readily that for any $P \nmid Q$, $|P| \le z$,

$$G(y,z;Q) = G(y,z;PQ) + \frac{1}{\Phi(P)} G(y/|P|, z; PQ). \tag{3.5}$$

Let

$$\Psi(P) = (1 + \Phi(P))^{-1} \ \text{ or } \ \Phi(P)\Psi(P) = 1 - \Psi(P), \tag{3.6}$$

and rewrite (3.5). In the result we replace y by $y/|P|$, and get

$$\begin{aligned} G(y/|P|, z; PQ) &= \Psi(P)\Phi(P)G(y/|P|, z; Q) \\ &\quad + \Psi(P)\left\{G(y/|P|, z; PQ) - G(y/|P|^2, z; PQ)\right\}. \end{aligned} \tag{3.7}$$

Inserting this into (3.4), we have that

$$\begin{aligned} T_1(y,z;Q) = &\sum_{\substack{|K|\le y \\ (K,Q)=1}} \frac{\mu(K)^2}{\Phi(K)} \sum_{\substack{|P|\le y/|K| \\ P\nmid Q}} \Psi(P)\log|P| \\ &+ \sum_{\substack{y/z^2<|K|\le y \\ (K,Q)=1}} \frac{\mu(K)^2}{\Phi(K)} \sum_{\substack{\sqrt{y/|K|}<|P|\le y/|K| \\ P\nmid KQ}} \frac{\Psi(P)}{\Phi(P)} \log|P|, \end{aligned} \tag{3.8}$$

where the additional condition $R_0 R_1 \le |P| \le z$ is implicit.

Now, to evaluate the first sum over P on the right side of (3.8), we observe that since by (2.1) we have

$$\Delta(P) \ll \frac{\log R_1}{\log|P|}, \tag{3.9}$$

it holds that

$$\Psi(P) = \Delta(P)\left(1 + O\left(\Delta(P)\right)\right). \tag{3.10}$$

This implies that for $\log(R_0R_1) \le \log x \ll \log R$

$$\sum_{\substack{|P|\le x\\ P\nmid Q}} \Psi(P)\log|P| = k\log x + O\left(\log R_0\right), \tag{3.11}$$

where the implied constant is independent of Q. In fact, the left side is equal to

$$\begin{aligned}
&\sum_{|P|\le x} \Delta(P)\log|P| + O\left(\sum_{P|Q}\log R_1 + \sum_{|P|\le x}\frac{(\log R_1)^2}{\log|P|}\right) \\
&= \sum_{|P|\le x}\sum_{p\in P}\frac{|\Omega(p)|}{p}\left(\log p + \log(|P|/p)\right) \\
&\qquad + O\left(\frac{\log|Q|}{\log R_0}\log R_1 + \sum_{\substack{j\\ R_0\le R_0R_1^j\ll x}}\frac{(\log R_1)^2}{\log(R_0R_1^j)}\right) \\
&= k\log\frac{x}{R_0} + O\left(\frac{\log R}{\log R_0}\log R_1\right).
\end{aligned} \tag{3.12}$$

Also, by (2.7) and (3.9)–(3.10),

$$\sum_{\substack{\sqrt{y/|K|}<|P|\le y/|K|\\ P\nmid KQ}} \frac{\Psi(P)}{\Phi(P)}\log|P| \ll \sum_{\sqrt{y/|K|}<|P|\le y/|K|}\frac{(\log R_1)^2}{\log|P|} \ll \log R_1. \tag{3.13}$$

We insert (3.11) and (3.13) into (3.8), on noting the implicit condition mentioned there. We see that

$$T_1(y,z;Q) = k\sum_{\substack{|K|\le y\\ (K,Q)=1}}\frac{\mu(K)^2}{\Phi(K)}\log\frac{y}{|K|} - k\sum_{\substack{|K|\le y/z\\ (K,Q)=1}}\frac{\mu(K)^2}{\Phi(K)}\log\frac{y/z}{|K|} + U(y,z;Q), \tag{3.14}$$

with

$$U(y,z;Q) \ll G(y,z;Q)\log R_0, \tag{3.15}$$

provided $\log(R_0R_1) \le \log y \ll \log R$ and $\log|Q| \ll \log R$; the implied constant is independent of Q.

We set $y = z$ in (3.2) and (3.14), and get

$$G(z;Q)\log z = (k+1)T(z,z;Q) + U(z,z;Q). \tag{3.16}$$

We are then led to the assertion that uniformly in $Q \in \mathcal{Y}(z)$, $\log|Q| \ll \log R$,

$$G(z;Q) = \frac{W(R_0)}{k!\mathfrak{S}(\mathcal{H})}(\log z)^k\left(1 + O\left(\frac{\log R_0}{\log z}\right)\right), \tag{3.17}$$

where

$$W(R_0) = \prod_{p \le R_0} \left(1 - \frac{|\Omega(p)|}{p}\right). \tag{3.18}$$

The deduction of (3.17) from (3.16) is standard; cf. [5, Section 2.2.2]. We should remark in this context that

$$\frac{W(R_0)}{\mathfrak{S}(\mathcal{H})} = \left(1 + O\left(\frac{\log R_1}{\log R_0}\right)\right) \lim_{s \to 0^+} \zeta(s+1)^{-k} \prod_{P \nmid Q} \left(1 + \frac{1}{|P|^s \Phi(P)}\right); \tag{3.19}$$

see [7, pp. 1060–1601] together with a minor correction. Here ζ is the Riemann zeta-function. Note that the left side of (3.19) is independent of Q.

4. With this, we are now ready to start smoothing the assertions of Lemmas 1 and 2. Hereafter we shall work with $\mathcal{Y}(w)$ in place of $\mathcal{Y}(z)$, where

$$w = R^{\omega}. \tag{4.1}$$

The constant ω is to satisfy

$$3 \log k \le k\omega \le \tfrac{1}{2}k, \tag{4.2}$$

while k is assumed to be sufficiently large, though bounded.

We put

$$\tilde{\lambda}_R(D;\ell) = \frac{\mathfrak{S}(\mathcal{H})}{\ell! W(R_0)} \frac{\mu(D)}{\Phi(D)\Delta(D)} \sum_{\substack{|K| \le R/|D| \\ (K,D)=1}} \frac{\mu(K)^2}{\Phi(K)} \left(\log \frac{R/|D|}{|K|}\right)^{\ell}, \tag{4.3}$$

where $D, K \in \mathcal{Y}(w)$. This is to be compared with (2.15) specialised by $z = w$ and

$$\xi(\emptyset) = \mathfrak{S}(\mathcal{H}) \frac{G(R,w)}{\ell! W(R_0)}. \tag{4.4}$$

The side condition (2.3) is obviously satisfied; also, by (2.17) and (4.4),

$$|\tilde{\lambda}_R(D;\ell)| \le \mathfrak{S}(\mathcal{H}) \frac{G(R,w)}{\ell! W(R_0)} (\log R)^{\ell} \ll (\log R)^{k+\ell}, \tag{4.5}$$

where (3.17) is used via $G(R,w) \le G(R)$. Our counterpart of (1.8) is now defined to be

$$\tilde{\Lambda}_R(n;\mathcal{H},\ell) = \sum_D \tilde{\lambda}_R(D;\ell) \sum_{\substack{d \in D \\ n \in \Omega(d)}} 1. \tag{4.6}$$

As to the interval $[1, R_0]$, which is excluded in the above, we appeal to the Fundamental Lemma in the sieve method (see [5, p. 92]). Thus, there exists a function ϱ, supported on the set of square-free integers, such that $\varrho(d) = 0$ or ± 1 for any $d \geq 1$, and $\varrho(d) = 0$ either if $d \geq R_0^\tau$ with τ to be fixed later or if d has a prime factor greater than or equal to R_0, and that for any $n \geq 1$

$$\gamma(n;\mathcal{H}) = \sum_{n\in\Omega(d)} \varrho(d) \geq 0 \tag{4.7}$$

as well as

$$\sum_d \frac{\varrho(d)}{d}|\Omega(d)| = W(R_0)\left(1 + O(e^{-\tau})\right). \tag{4.8}$$

We set

$$\tau = (\log\log R)^{1/10}. \tag{4.9}$$

Now our task is to evaluate asymptotically the sum

$$\sum_{N<n\leq 2N} \gamma(n;\mathcal{H})\tilde{\Lambda}_R(n;\mathcal{H},\ell)^2, \tag{4.10}$$

which is to replace the left side of (1.12). By (4.5) and (4.8), this is equal to

$$NW(R_0)\tilde{\mathcal{T}}\left(1 + O(e^{-\tau})\right) + O\left(R_0^\tau R^2(\log N)^c\right), \tag{4.11}$$

where $\tilde{\mathcal{T}}$ is defined analogously to (2.4). We have

$$\begin{aligned}\tilde{\mathcal{T}} &= \sum_{|D|\leq R} \Phi(D)\left(\sum_{D|K} \Delta(K)\tilde{\lambda}_R(K;\ell)\right)^2 \\ &= \left(\frac{\mathfrak{S}(\mathcal{H})}{\ell! W(R_0)}\right)^2 \sum_{|K|\leq R} \frac{\mu(K)^2}{\Phi(K)}\left(\log\frac{R}{|K|}\right)^{2\ell};\end{aligned} \tag{4.12}$$

the second line is due to the relation similar to that among (2.10), (2.14) and (2.15).

The last sum is

$$\begin{aligned}&\leq \sum_{\substack{|K|\leq R \\ P|K\Rightarrow|P|\leq R}} \frac{\mu(K)^2}{\Phi(K)}\left(\log\frac{R}{|K|}\right)^{2\ell} \\ &= \int_{R_0R_1}^{R} (\log R/t)^{2\ell} dG(t) + (\log R)^{2\ell} \\ &= \frac{(2\ell)!}{(k+2\ell)!}\frac{W(R_0)}{\mathfrak{S}(\mathcal{H})}(\log R)^{k+2\ell}\left(1 + O\left((\log\log R)^{-1/5}\right)\right).\end{aligned} \tag{4.13}$$

In the first line we have moved to the semi-group $\mathcal{Y}(R)$; the second line depends on $G(t,R) = G(t)$ for $t \le R$, and the last on (3.17) with $Q = \emptyset$. On the other hand, the Buchstab identity implies that the sum in question is equal to

$$\sum_{\substack{|K|\le R\\ P|K\Rightarrow |P|\le R}} \frac{\mu(K)^2}{\Phi(K)}\left(\log\frac{R}{|K|}\right)^{2\ell} - \sum_{w<|P|\le R}\frac{1}{\Phi(P)} \sum_{\substack{|K|\le R/|P|\\ P'|K\Rightarrow |P'|<|P|}} \frac{\mu(K)^2}{\Phi(K)}\left(\log\frac{R/|P|}{|K|}\right)^{2\ell}. \tag{4.14}$$

The last double sum is

$$\begin{aligned} &\le \sum_{w<|P|\le R}\frac{1}{\Phi(P)} \sum_{\substack{|K|\le R/w\\ P'|K\Rightarrow |P'|<R/w}} \frac{\mu(K)^2}{\Phi(K)}\left(\log\frac{R/w}{|K|}\right)^{2\ell}\\ &\ll k|\log\omega|\frac{(2\ell)!}{(k+2\ell)!}\frac{W(R_0)}{\mathfrak{S}(\mathcal{H})}(\log R/w)^{k+2\ell}\\ &\ll e^{-k\omega/3}\frac{(2\ell)!}{(k+2\ell)!}\frac{W(R_0)}{\mathfrak{S}(\mathcal{H})}(\log R)^{k+2\ell}, \end{aligned} \tag{4.15}$$

where (4.2) has been invoked, and the implied constants are absolute.

Hence collecting (4.11)–(4.15) we obtain the following smoothed version of Lemma 1:

Lemma 3. *With* (1.17), (2.1), (4.1), (4.2), (4.3) (4.6), (4.7), (4.9) *and the same assumption as in Lemma 1, we have, as $N \to \infty$,*

$$\begin{aligned} &\sum_{N<n\le 2N}\gamma(n;\mathcal{H})\tilde{\Lambda}_R(n;\mathcal{H},\ell)^2\\ &\quad = \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell)!}\binom{2\ell}{\ell}N(\log R)^{k+2\ell}\left(1+O(e^{-k\omega/3})\right), \end{aligned} \tag{4.16}$$

where the implied constant is absolute.

5. Next, we shall consider a twist of (4.16) with primes:

$$\begin{aligned} &\sum_{N<n\le 2N}\varpi(n+h)\gamma(n;\mathcal{H})\tilde{\Lambda}_R(n;\mathcal{H},\ell)^2\\ &\quad = \sum_{N<n\le 2N}\varpi(n+h)\gamma(n;\mathcal{H}\backslash\{h\})\tilde{\Lambda}_R(n;\mathcal{H}\backslash\{h\},\ell)^2, \end{aligned} \tag{5.1}$$

as it is assumed that $h \in \mathcal{H}$, $R < N$. Note that we are working with $\mathcal{Y}(w)$. Expanding out the square, we see that this is equal to

$$\begin{aligned} &\sum_{D_1,D_2}\tilde{\lambda}_R(D_1;\ell)\tilde{\lambda}_R(D_2;\ell)\sum_d \varrho(d)\\ &\times \sum_{d_1\in D_1, d_2\in D_2}\ \sum_{\substack{a\in\Omega^-(d[d_1,d_2])\\ (a+h,d[d_1,d_2])=1}} \vartheta^*(N;a+h,d[d_1,d_2]) + O(R_0^\tau R^2(\log N)^c), \end{aligned} \tag{5.2}$$

where Ω^- corresponds to $\mathcal{H}\backslash\{h\}$, and (4.5) has been applied. The condition in the inner-most sum induces the introduction of

$$\Omega^*(p) = \Omega^-(p)\backslash\{-h \bmod p\} = \Omega(p)\backslash\{-h \bmod p\}. \tag{5.3}$$

Note that $|\Omega^*(p)| = |\Omega(p)|-1$, which we may assume does not vanish, provided p is sufficiently large.

The sum in (5.2) is equal to

$$N\mathcal{T}^* \sum_d \frac{\varrho(d)}{\varphi(d)}|\Omega^*(d)| + \mathcal{E}, \tag{5.4}$$

where

$$\mathcal{T}^* = \sum_{D_1,D_2} \tilde{\lambda}_R(D_1;\ell)\tilde{\lambda}_R(D_2;\ell) \sum_{d_1\in D_1,\, d_2\in D_2} \frac{|\Omega^*([d_1,d_2])|}{\varphi([d_1,d_2])} \tag{5.5}$$

and

$$\mathcal{E} = \sum_{D_1,D_2} \tilde{\lambda}_R(D_1;\ell)\tilde{\lambda}_R(D_2;\ell) \sum_d \varrho(d) \sum_{d_1\in D_1, d_2\in D_2} \sum_{a\in\Omega^*(d[d_1,d_2])} E^*(N;a,d[d_1,d_2]). \tag{5.6}$$

Corresponding to (4.8), we have

$$\sum_d \frac{\varrho(d)}{\varphi(d)}|\Omega^*(d)| = \frac{W(R_0)}{V(R_0)}\left(1+O(e^{-\tau})\right), \quad V(R_0) = \prod_{p\le R_0}\left(1-\frac{1}{p}\right), \tag{5.7}$$

via the same reasoning. Also we have

$$\mathcal{T}^* = \sum_{|D|\le R} \Phi^*(D)\left(\sum_{D|K} \Delta^*(K)\tilde{\lambda}_R(K;\ell)\right)^2, \tag{5.8}$$

where

$$\Delta^*(D) = \prod_{P|D}\left(\sum_{p\in P}\frac{|\Omega^*(p)|}{p-1}\right), \tag{5.9}$$

and

$$\Phi^*(D) = \frac{1}{\Delta^*(D)^2}\prod_{P|D}\left(\sum_{p\in P}\frac{|\Omega^*(p)|}{p-1}\left(1-\frac{|\Omega^*(p)|}{p-1}\right)\right). \tag{5.10}$$

Here we have actually $|\Omega^*(p)| = k-1$.

We are about to show an effective lower bound of $\mathcal{T}^*$. We first note the trivial inequality

$$\mathcal{T}^* \ge \mathcal{T}^{**}, \tag{5.11}$$

where the right side is the restriction of that of (5.8) to $R/w \le |D| \le R$. Inserting (4.3) into (5.8), we get

$$\begin{aligned}\mathcal{T}^{**} = &\left(\frac{\mathfrak{S}(\mathcal{H})}{\ell! W(R_0)}\right)^2 \sum_{R/w\le|D|\le R} \mu(D)^2 \left(\frac{\Delta^*(D)}{\Delta(D)}\right)^2 \frac{\Phi^*(D)}{\Phi(D)^2}\\ &\times \left(\sum_{\substack{|K|\le R/|D|\\(K,D)=1}} \frac{\mu^2(K)}{\Phi(K)} \prod_{P|K}\left(1-\frac{\Delta^*(P)}{\Delta(P)}\right)\left(\log\frac{R/|D|}{|K|}\right)^\ell\right)^2.\end{aligned} \tag{5.12}$$

This sum over K can be handled with a simple modification of the argument leading to (3.17) besides employing (3.19) with an obvious change. In fact, we may drop the condition $K \in \mathcal{Y}(w)$, since $R/|D| \le w$. We have, for $\log R_0R_1 \le \log y \ll \log R$,

$$\sum_{\substack{|K|\le y\\ P|K\Rightarrow P\nmid D,\,|P|\le y}} \frac{\mu^2(K)}{\Phi(K)} \prod_{P|K}\left(1-\frac{\Delta^*(P)}{\Delta(P)}\right) = V(R_0)\log y\left(1+O\left(\frac{\log R_0}{\log y}\right)\right), \tag{5.13}$$

uniformly in D. The sum in question is then computed by integration by parts, and the result is inserted into (5.12) to give that

$$\begin{aligned}\mathcal{T}^{**} = &\left(\frac{\mathfrak{S}(\mathcal{H})}{(\ell+1)!}\cdot\frac{V(R_0)}{W(R_0)}\right)^2\left(1+O((\log\log R)^{-1/5})\right)\\ &\times\left\{\sum_{\substack{|D|\le R\\ P|D\Rightarrow|P|\le w}} - \sum_{\substack{|D|\le R/w\\ P|D\Rightarrow|P|\le w}}\right\}\mu(D)^2\left(\frac{\Delta^*(D)}{\Delta(D)}\right)^2\frac{\Phi^*(D)}{\Phi(D)^2}\left(\log\frac{R}{|D|}\right)^{2(\ell+1)}\end{aligned} \tag{5.14}$$

To estimate the part over $|D| \le R$, we proceed exactly as in (4.12)–(4.15); and the part over $|D| \le R/w$ as in (4.13) or rather (4.15), appealing to (3.17) and (3.19) with an obvious change. In this way we find that

$$\mathcal{T}^{**} = \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!}\binom{2(\ell+1)}{\ell+1}\frac{V(R_0)}{W(R_0)}(\log R)^{k+2\ell+1}\left(1+O(e^{-k\omega/3})\right), \tag{5.15}$$

which ends our treatment of the main term of (5.4).

6. We still need to consider the structure of $\mathcal{E}$, and it is embodied in the assertion (6.2) below.

Lemma 4. *Under* (1.1), (1.2), (1.17), (2.1), (4.1), (4.2), (4.3) (4.6), (4.7), (4.9), *it holds for any* $h \in \mathcal{H}$ *that*

$$\begin{aligned}&\sum_{N<n\le 2N} \varpi(n+h)\gamma(n;\mathcal{H})\tilde\Lambda_R(n;\mathcal{H},\ell)^2\\ &\ge \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!}\binom{2(\ell+1)}{\ell+1}N(\log R)^{k+2\ell+1}\left(1+O(e^{-k\omega/3})\right) - \mathcal{E}_h(N;\mathcal{H}),\end{aligned} \tag{6.1}$$

as $N \to \infty$. Here we have, for any $A, B \ge 1$ such that $AB = R_0^{2\tau} R^{2+\omega}$,

$$\mathcal{E}_h(N;\mathcal{H}) \le (\log N)^{2(k+\ell)+1} \sup_{\alpha,\,\beta} \left| \sum_{a\le A,\, b\le B} \alpha_a\beta_b \sum_{r\in\Omega^*(ab)} E^*(N;r,ab) \right|, \tag{6.2}$$

with $\Omega^*(p) = \Omega(p) \setminus \{-h \bmod p\}$ and $\Omega^*(p^v) = \emptyset$ $(v \ge 2)$, where α, β run over vectors such that $|\alpha_a| \le 1$, $|\beta_b| \le 1$.

Remark. The above convention on $\Omega^*(ab)$ for non-square-free ab can in fact be replaced appropriately in practice. This is due to the fact that in our construction below the situation $p^2|ab$ is possible only with $p \ge R_0$, and the elimination of the contribution of those moduli is immediate. It should also be stressed that we have in fact $\alpha_a = 0$ or 1, and $\beta_b = 0$ or $\varrho(d)$, $d\|b$, with d being as in (4.7).

The first term on the right of (6.1) follows from (5.4), (5.7), (5.11), and (5.15). As to (6.2) we argue as follows: Returning to (5.6), we consider a generic pair D_1, D_2. Let F be an arbitrary divisor of (D_1, D_2), the greatest common divisor of the pair. We restrict ourselves to the situation in (5.6) where $d_1 \in D_1$, $d_2 \in D_2$ and $(d_1, d_2) \in F$. Let $D_1D_2/F = P_1P_2\cdots P_s$ with $|P_j| \le |P_{j+1}|$. Note that there can be some j such that $P_j = P_{j+1}$; in fact this is the case where P_j divides $(D_1, D_2)/F$. We define u to be such that $|P_1|\cdots|P_u| \le A$ but $|P_1|\cdots|P_{u+1}| > A$. It is possible that there does not exist such u; then we are done. Otherwise, let $a \in P_1\cdots P_u$ and $a' \in P_{u+1}\cdots P_s$. Obviously we have $aa' \le R^2$. On the other hand, we have $a \ge |P_1|\cdots|P_u|R_1^{-u} > A|P_{u+1}|^{-1}R_1^{-u}$, because of the definition of the intervals given after (2.1). Thus $a' < R^{2+\omega}R_1^u/A$, as $|P_{u+1}| \le R^\omega$. Let d be as in (5.6), and put $b = a'd$ we have $b < R^{2+\omega}R_0^\tau R_1^u/A < B$, since $u \ll (\log R)/\log R_0 \ll (\log\log R)^{1/5}$. We are about to designate these a, b as to be the same as in (6.2); note that $d[d_1, d_2]$ in (5.6) are among the set of ab. Then we need to exclude those ab which are not square-free, for only those moduli are superfluous. One way to employ here is to introduce a convention about $\Omega^*(p^v)$ $(v \ge 2)$ as is done above. Finally, on noting (4.5) as well as that the number of triples D_1, D_2, F is not larger than $\exp((\log\log R)^{9/10}\log 3)$, we end the proof of (6.2).

In the possible application to the problem about the gaps between primes, we may assume that k is large, and ω can be so small as $3(\log k)/k$. Hence the size of AB is essentially $R^{2+\varepsilon}$ with an arbitrarily small constant $\varepsilon > 0$. With this, we see that a combination of Lemmas 3 and 4 implies that if there exists a $C_1 \ge 2(k+\ell+1)$ such that uniformly for h in any admissible $\mathcal{H}$

$$\mathcal{E}_h(N;\mathcal{H}) \ll \frac{N}{(\log N)^{C_1}}, \quad R = N^{\theta/2} \text{ with an absolute constant } \theta > \tfrac{1}{2}, \tag{6.3}$$

then (1.16) should follow. This hypothesis is certainly less stringent than (1.10) with $\theta > \frac{1}{2}$. What is interesting is that (6.3) is true if the condition $r \in \Omega^*(ab)$ is replaced by $r \equiv r_0 \pmod{ab}$ with a fixed integer r_0, as is proved in [1]. It is, however, unclear how to extend the argument of [1] to the situation with many residue classes as we require.

Concluding Remark. The argument of our paper can be employed in a more general setting: With a *large* two-sided sifting density κ (see, e.g., [5, p. 29]), the remainder term in the Selberg sieve admits a flexible bilinear form similar to the one proved by H. Iwaniec for Rosser's linear sieve, although the level condition $MN \le D$, in the now common notation, has to be replaced by the slightly weaker $MN \le D^{1+\delta}$ with $\delta \ll (\log\kappa)/\kappa$, which is to be compared with (6.2). In fact, this assertion was obtained by the first author in early 1980's; however, any possible application of it was not in his view then and even later when the relevant article [7] was written. He realised recently that his old method could be applied to smoothing both Lemmas 1 and 2, and reached an earlier version of Lemmas 3 and 4. Simultaneously and independently, the second author obtained the same.

Acknowledgements. The present work is an outcome of the workshop '*Gaps between Primes*' (November 28–December 3, 2005) at the American Institute of Mathematics. The authors are indebted to the institute for the financial aid that made it possible for them to attend the workshop.

References

[1] E. Bombieri, J.B. Friedlander and H. Iwaniec. Primes in arithmetic progressions to large moduli. Acta Math., **156** (1986), 203–251.

[2] G.A. Goldston, S.W. Graham, J. Pintz, and C.Y. Yıldırım. Small gaps between primes or almost primes. Preprint. arXiv: math.NT/506067.

[3] D.A. Goldston, J. Pintz, and C.Y. Yıldırım. Small gaps between primes II (Preliminary). Preprint, February 8, 2005.

[4] G.A. Goldston, Y. Motohashi, J. Pintz, and C.Y. Yıldırım. Small gaps between primes exist. Submitted.

[5] G. Greaves. Sieves in Number Theory. Ergebnisse Math., **43**, Springer Verlag, Berlin etc. 2001.

[6] Y. Motohashi. Sieve Methods and Prime Number Theory. Tata IFR LN, vol. **72**, Springer Verlag, Berlin etc. 1983.

[7] Y. Motohashi. On the remainder term in the Selberg sieve. Number Theory in Progress - A. Schinzel Festschrift, Walter de Gruyter, Berlin · New York 1999, pp. 1053–1064.

[8] Y. Motohashi. An overview of the sieve method and its history. Preprint (to appear in Sugaku Expositions, AMS). arXiv: math.NT/0505521.

Yoichi Motohashi
Department of Mathematics, Nihon University,
Surugadai, Tokyo 101-8308, JAPAN.
E-mail: ymoto@math.cst.nihon-u.ac.jp

János Pintz
Rényi Mathematical Institute of the Hungarian Academy of Sciences,
H-1364 Budapest, P.O.B. 127, HUNGARY
E-mail: pintz@renyi.hu

Addendum

We are indebted to Professor Terrence Tao for kindly pointing out an inadequacy in the explanation following the formula (5.14) of this paper of ours, designated as MP. The aim of the present addendum is to provide a rectification to our argument. This entails a replacement of the condition (4.2) of MP on the basic parameters k, ℓ, ω by a more restrictive one; see $(**)$ below. It should be stressed, however, that being minor our correction does not make any difference to the main assertion of MP that the never-ending appearances of bounded gaps between primes could be established by assuming the possibility of going past level $\frac{1}{2}$ in a mean prime number theorem of the Bombieri–Vinogradov type in which the moduli are restricted to smooth numbers, i.e., those q whose prime factors are all less than q^{ϖ} with an arbitrary small but fixed $\varpi > 0$. We should repeat here the statement in Lemma 4 of MP that our smoothing of the GPY sieve induces a flexibility in the formulation of this conjectural mean prime number theorem, which allows for an appeal to the method due to Bombieri, Fouvry, Friedlander and Iwaniec concerning the distribution of primes in arithmetic progressions to large moduli.

As is widely known, Yitang Zhang proved recently this conjectural part of MP, which may now be called Zhang's mean prime number theorem; his argument depends on the works of the four people above as we envisaged. Therefore the bounded gaps between primes do occur infinitely often. It should be remarked that in Zhang's work a smoothing of the GPY sieve is given as an essential ingredient of his argument; but, including its implication for the structure mentioned above of the mean prime number theorem to have been established, his relevant reasoning is largely identical to the one we developed in MP several years earlier. It appears to us that without the smoothing of the GPY sieve the assertion on bounded gaps between primes would not be deducible from Zhang's mean prime number theorem.

Now, turning to MP, immediately after (5.14) it is stated that (i) to estimate the part over $|D| \le R$, we proceed exactly as in (4.12)–(4.15) and (ii) the part over $|D| \le R/w$ as in (4.13) or rather (4.15), appealing to (3.17) and (3.19) with an obvious change.

Neither (i) nor (ii) are quite adequate. In order to rectify these, we note first the simple fact that $\Delta^*(P)\Phi^*(P) > \Delta(P)\Phi(P)$ for any P, because of (2.1). Namely, we may consider, instead of (5.14), the expression

$$\begin{aligned}\mathcal{T}^{**} &\ge \left(\frac{\mathfrak{S}(\mathcal{H})}{(\ell+1)!}\cdot\frac{V(R_0)}{W(R_0)}\right)^2 \Big(1 + O((\log\log R)^{-1/5})\Big) \\ &\quad \times \left\{ \sum_{\substack{|D|\le R \\ P|D \Rightarrow |P|\le w}} - \sum_{\substack{|D|\le R/w \\ P|D \Rightarrow |P|\le w}} \right\} \frac{\mu(D)^2}{\Phi^*(D)} \left(\log\frac{R}{|D|}\right)^{2(\ell+1)}.\end{aligned}$$

With this, the assertion (i) is quite correct. As to (ii), we replace the factor $(\log R/|D|)^{2(\ell+1)}$

by

$$\sum_{f=0}^{2(\ell+1)} \binom{2(\ell+1)}{f} (\log w)^f \left(\log \frac{R/w}{|D|}\right)^{2(\ell+1)-f},$$

and consider the sum

$$\sum_{|D| \le R/w} \frac{\mu(D)^2}{\Phi^*(D)} \left(\log \frac{R/w}{|D|}\right)^{2(\ell+1)-f}.$$

Note that the condition $P|D \Rightarrow |P| \le w$ has been dropped. This is equal to

$$\frac{(2(\ell+1)-f)!}{(k+2\ell+1-f)!} \frac{W(R_0)}{\mathfrak{S}(\mathcal{H})V(R_0)} \big(\log R/w\big)^{k+2\ell+1-f} \left(1 + O((\log\log R)^{-1/5})\right).$$

Hence we have

$$\begin{aligned}
&\sum_{\substack{|D| \le R/w \\ P|D \Rightarrow |P| \le w}} \frac{\mu(D)^2}{\Phi^*(D)} \left(\log \frac{R}{|D|}\right)^{2(\ell+1)} \\
&\le Y(k,\ell;\omega) \cdot \frac{(2(\ell+1))!}{(k+2\ell+1)!} \frac{W(R_0)}{\mathfrak{S}(\mathcal{H})V(R_0)} (\log R)^{k+2\ell+1} \left(1 + O((\log\log R)^{-1/5})\right),
\end{aligned}$$

with

$$Y(k,\ell;\omega) = \sum_{f=0}^{2(\ell+1)} \binom{k+2\ell+1}{f} \omega^f (1-\omega)^{k+2\ell+1-f}; \quad w = R^\omega.$$

We have thus

$$\mathcal{T}^{**} \ge \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!} \binom{2(\ell+1)}{\ell+1} \frac{V(R_0)}{W(R_0)} (\log R)^{k+2\ell+1} \left(1 - Y(k,\ell;\omega) + O(e^{-k\omega/3})\right).$$

Denoting the fth term of $Y(k,\ell;\omega)$ by $A(f)$, we have

$$A(f+1)/A(f) = \frac{k+2\ell+1-f}{f+1} \frac{\omega}{1-\omega}.$$

This is decreasing as f increases; and assuming

$$(*) \qquad \frac{k}{2(\ell+1)} \frac{\omega}{1-\omega} \ge 1,$$

we have

$$Y(k,\ell;\omega) \le (2\ell+3) \binom{k+2\ell+1}{2(\ell+1)} \omega^{2(\ell+1)} (1-\omega)^{k-1}.$$

Then, we make a drastic but practical specialisation: We put, with an arbitrary constant $\alpha > 0$,

$$(**) \qquad \omega = 8\frac{\alpha}{\sqrt{k}}, \quad \ell = \alpha\sqrt{k},$$

which amply satisfies $(*)$, provided k is sufficiently large. We find, by means of Stirling's formula, that

$$Y(k,\ell;\varpi) < e^{-3k\omega/8}.$$

Hence, we have obtained (6.1) of MP on $(**)$.

Ending our discussion, we remark that $(**)$ induces an obvious change in the paragraphs of MP following the proof of (6.2); however, this does not affect our overall conclusion in the paper.