Small Gaps between Primes Exist

By D.A. Goldston, Y. Motohashi, J. Pintz, and C.Y. Yıldırım

Abstract. In the recent preprint [3], Goldston, Pintz, and Yıldırım established, among other things,

$$\liminf_{n\to\infty} \frac{p_{n+1}-p_n}{\log p_n} = 0, \tag{0}$$

with p_n the nth prime. In the present article, which is essentially self-contained, we shall develop a simplified account of the method used in [3]. While [3] also includes quantitative versions of (0), we are concerned here solely with proving the qualitative (0), which still exhibits all the essentials of the method. We also show here that an improvement of the Bombieri–Vinogradov prime number theorem would give rise infinitely often to bounded differences between consecutive primes. We include a short expository last section. Detailed discussions of quantitative results and a historical review will appear in the publication version of [3] and its continuations.

1. Basic Lemma

In this section we shall prove an asymptotic formula relevant to Selberg's sieve, which is to be modified so as to involve primes in the next section. The two asymptotic formulas thus obtained will be combined in a simple weighted sieve setting, and give rise to (0) in the third section.

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

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

and

$$\text{integers } k, \ell > 0 \text{ are arbitrary but bounded.} \tag{1.2}$$

To prove a quantitative assertion superseding (0), we need to regard k, ℓ as functions of N; but for our present purpose the above is sufficient. 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.

Let

$$\mathcal{H} = \{h_1, h_2, \ldots, h_k\} \subseteq [1, H] \cap \mathbb{Z}, \tag{1.3}$$

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

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

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

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

1991 Mathematics Subject Classification: Primary 11N05; Secondary 11P32

Key Words and phrases: prime number

The first author was supported by NSF grant DMS-0300563, the NSF Focused Research Group grant 0244660, and the American Institute of Mathematics; the second author by KAKENHI 15540047 and Nihon University Research grant; the third author by OTKA grants No. T38396, T43623, T49693 and Balaton program; the fourth author by TÜBITAK

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

$$d|P(n;\mathcal{H}), \quad P(n;\mathcal{H}) = (n+h_1)(n+h_2)\cdots(n+h_k). \tag{1.6}$$

Also, we put, with μ the Möbius function,

$$\lambda_R(d;a) = \begin{cases} 0 & \text{if } d > R, \\ \dfrac{1}{a!}\mu(d)(\log R/d)^a & \text{if } d \le R, \end{cases} \tag{1.7}$$

and

$$\Lambda_R(n;\mathcal{H},a) = \sum_{n\in\Omega(d)} \lambda_R(d;a) = \frac{1}{a!}\sum_{\substack{d|P(n;\mathcal{H})\\ d\le R}} \mu(d)(\log R/d)^a. \tag{1.8}$$

With this, we shall consider the evaluation of

$$\sum_{N<n\le 2N} \Lambda_R(n;\mathcal{H},k+\ell)^2. \tag{1.9}$$

Expanding out the square, we have

$$\sum_{d_1,d_2} \lambda_R(d_1;k+\ell)\lambda_R(d_2;k+\ell) \sum_{\substack{N<n\le 2N\\ n\in\Omega(d_1), n\in\Omega(d_2)}} 1. \tag{1.10}$$

The condition on n is equivalent to $n \in \Omega([d_1,d_2])$, with $[d_1,d_2]$ the least common multiple of the two integers; and

$$\sum_{N<n\le 2N} \Lambda_R(n;\mathcal{H},k+\ell)^2 = N\mathcal{T} + O\left(\left(\sum_d |\Omega(d)||\lambda_R(d;k+\ell)|\right)^2\right), \tag{1.11}$$

in which

$$\mathcal{T} = \sum_{d_1,d_2} \frac{|\Omega([d_1,d_2])|}{[d_1,d_2]}\lambda_R(d_1;k+\ell)\lambda_R(d_2;k+\ell). \tag{1.12}$$

We have $|\Omega(d)| \le \tau_k(d)$ with the divisor function τ_k. Thus

$$\sum_{N<n\le 2N} \Lambda_R(n;\mathcal{H},k+\ell)^2 = N\mathcal{T} + O\left(R^2(\log R)^c\right). \tag{1.13}$$

On noting that for $a \ge 1$

$$\lambda_R(d;a) = \frac{\mu(d)}{2\pi i}\int_{(1)} \left(\frac{R}{d}\right)^s \frac{ds}{s^{a+1}}, \tag{1.14}$$

with (α) the vertical line in the complex plane passing through α, we have

$$\mathcal{T} = \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)} F(s_1,s_2;\Omega)\frac{R^{s_1+s_2}}{(s_1 s_2)^{k+\ell+1}} ds_1 ds_2, \tag{1.15}$$

where

$$F(s_1,s_2;\Omega)=\sum_{d_1,d_2}\mu(d_1)\mu(d_2)\frac{|\Omega([d_1,d_2])|}{[d_1,d_2]d_1^{s_1}d_2^{s_2}} \tag{1.16}$$
$$=\prod_p\left(1-\frac{|\Omega(p)|}{p}\left(\frac{1}{p^{s_1}}+\frac{1}{p^{s_2}}-\frac{1}{p^{s_1+s_2}}\right)\right)$$

in the region of absolute convergence.

Since $|\Omega(p)|=k$ for $p>H$, we put

$$G(s_1,s_2;\Omega)=F(s_1,s_2;\Omega)\left(\frac{\zeta(s_1+1)\zeta(s_2+1)}{\zeta(s_1+s_2+1)}\right)^k, \tag{1.17}$$

with ζ the Riemann zeta-function. This is regular and bounded for $\operatorname{Re}s_1,\operatorname{Re}s_2>-c$. In particular, we have the singular series

$$\mathfrak{S}(\mathcal{H})=G(0,0;\Omega)=\prod_p\left(1-\frac{|\Omega(p)|}{p}\right)\left(1-\frac{1}{p}\right)^{-k}, \tag{1.18}$$

which does not vanish because of (1.5). We have the bound

$$G(s_1,s_2;\Omega)\ll\exp(c(\log N)^{-2\sigma}\log\log\log N), \tag{1.19}$$

with $\min(\operatorname{Re}s_1,\operatorname{Re}s_2,0)=\sigma\geq -c$, as can be seen via the Euler product expansion of the right side of (1.17). In fact, the part corresponding to those $p>H$ is uniformly bounded in the indicated region since $|\Omega(p)|=k$. For $k^2<p\leq H$, the logarithm of each p-factor is estimated to be $\ll H^{-2\sigma}\sum_{p\leq H}p^{-1}$; and the treatment of those $p\leq k^2$ is trivial. Note that the restrictions (1.1) and (1.2) are relevant here.

Now, we have in place of (1.15)

$$\mathcal{T}=\frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)}G(s_1,s_2;\Omega)\left(\frac{\zeta(s_1+s_2+1)}{\zeta(s_1+1)\zeta(s_2+1)}\right)^k\frac{R^{s_1+s_2}}{(s_1s_2)^{k+\ell+1}}ds_1ds_2. \tag{1.20}$$

Let us put $U=\exp(\sqrt{\log N})$, and shift the s_1 and s_2-contours to the vertical lines $(\log U)^{-1}+it$ and to $(2\log U))^{-1}+it$, $t\in\mathbb{R}$, respectively. We truncate them to $|t|\leq U$ and $|t|\leq U/2$, and denote the results by L_1 and L_2, respectively. On noting (1.1) and (1.19), we have readily that

$$\mathcal{T}=\frac{1}{(2\pi i)^2}\int_{L_2}\int_{L_1}G(s_1,s_2;\Omega)\left(\frac{\zeta(s_1+s_2+1)}{\zeta(s_1+1)\zeta(s_2+1)}\right)^k\frac{R^{s_1+s_2}}{(s_1s_2)^{k+\ell+1}}ds_1ds_2 \tag{1.21}$$
$$+O\left(\exp(-c\sqrt{\log N})\right).$$

We then shift the s_1-contour to $L_3:-(\log U)^{-1}+it$, $|t|\leq U$; necessary facts about the functions ζ and $1/\zeta$ can be found in [4, p. 53] (or p. 60 in the Second Edition). We encounter singularities at $s_1=0$ and $s_1=-s_2$, which are poles of orders $\ell+1$ and k, respectively. We have

$$\mathcal{T}=\frac{1}{2\pi i}\int_{L_2}\left\{\operatorname*{Res}_{s_1=0}+\operatorname*{Res}_{s_1=-s_2}\right\}ds_2+O\left(\exp(-c\sqrt{\log N})\right), \tag{1.22}$$

in which we have used (1.19).

We rewrite the residue, and have

$$\operatorname*{Res}_{s_1=-s_2}=\frac{1}{2\pi i}\int_{C(s_2)}G(s_1,s_2;\Omega)\left(\frac{\zeta(s_1+s_2+1)}{\zeta(s_1+1)\zeta(s_2+1)}\right)^k\frac{R^{s_1+s_2}}{(s_1s_2)^{k+\ell+1}}ds_1, \tag{1.23}$$

with the circle $C(s_2): |s_1+s_2| = (\log N)^{-1}$. By (1.19), we have $G(s_1, s_2; \Omega) \ll (\log\log N)^c$; and $R^{s_1+s_2} \ll 1$, $\zeta(s_1+s_2+1) \ll \log N$. Also, since $|s_2| \ll |s_1| \ll |s_2|$, we have $(s_1\zeta(s_1+1))^{-1} \ll (|s_2|+1)^{-1}\log(|s_2|+2)$, *loc.cit*. Thus

$$\operatorname*{Res}_{s_1=-s_2} \ll (\log N)^{k-1}(\log\log N)^c\left(\frac{\log(|s_2|+2)}{|s_2|+1}\right)^{2k}|s_2|^{-2\ell-2}. \tag{1.24}$$

Inserting this into (1.22), we get

$$\mathcal{T} = \frac{1}{2\pi i}\int_{L_2}\left\{\operatorname*{Res}_{s_1=0}\right\}ds_2 + O\left((\log N)^{k+\ell}\right). \tag{1.25}$$

To evaluate the last integral, we put

$$Z(s_1, s_2) = G(s_1, s_2; \Omega)\left(\frac{(s_1+s_2)\zeta(s_1+s_2+1)}{s_1\zeta(s_1+1)s_2\zeta(s_2+1)}\right)^k, \tag{1.26}$$

which is regular in a neighborhood of the point $(0,0)$. Then we have

$$\operatorname*{Res}_{s_1=0} = \frac{R^{s_2}}{\ell! s_2^{\ell+1}}\left(\frac{\partial}{\partial s_1}\right)^{\ell}_{s_1=0}\left\{\frac{Z(s_1,s_2)}{(s_1+s_2)^k}R^{s_1}\right\}. \tag{1.27}$$

We insert this into (1.25) and shift the contour to $L_4: -(2\log U)^{-1} + it$, $|t| \le U/2$. We see the new integral is $O(\exp(-c\sqrt{\log N}))$; the necessary bound for the integrand could be obtained in much the same way as (1.24). Thus

$$\begin{aligned}\mathcal{T} &= \operatorname*{Res}_{s_2=0}\operatorname*{Res}_{s_1=0} + O((\log N)^{k+\ell}) \\ &= \frac{1}{(2\pi i)^2}\int_{C_2}\int_{C_1}\frac{Z(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{\ell+1}}ds_1ds_2 + O((\log N)^{k+\ell}),\end{aligned} \tag{1.28}$$

where C_1, C_2 are the circles $|s_1| = \rho$, $|s_2| = 2\rho$, with a small $\rho > 0$. We write $s_1 = s$, $s_2 = s\xi$. Then the double integral is equal to

$$\frac{1}{(2\pi i)^2}\int_{C_3}\int_{C_1}\frac{Z(s, s\xi)R^{s(\xi+1)}}{(\xi+1)^k\xi^{\ell+1}s^{k+2\ell+1}}ds d\xi, \tag{1.29}$$

where C_3 is the circle $|\xi| = 2$. This is equal to

$$\frac{Z(0,0)}{2\pi i(k+2\ell)!}(\log R)^{k+2\ell}\int_{C_3}\frac{(\xi+1)^{2\ell}}{\xi^{\ell+1}}d\xi + O((\log N)^{k+2\ell-1}(\log\log N)^c), \tag{1.30}$$

where we have used (1.19).

Hence, we have obtained our basic implement:

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*

$$\sum_{N<n\le 2N}\Lambda_R(n; \mathcal{H}, k+\ell)^2 = \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell)!}\binom{2\ell}{\ell}N(\log R)^{k+2\ell} + O(N(\log N)^{k+2\ell-1}(\log\log N)^c). \tag{1.31}$$

2. Twist with Primes

Next, let us put

$$\varpi(n) = \begin{cases} \log n & \text{if } n \text{ is a prime,} \\ 0 & \text{otherwise;} \end{cases} \tag{2.1}$$

and consider the evaluation of the sum

$$\sum_{N<n\le 2N} \varpi(n+h)\Lambda_R(n;\mathcal{H},k+\ell)^2, \tag{2.2}$$

with an arbitrary positive integer $h \le H$. We observe that by (1.6) this is equal to

$$\sum_{N<n\le 2N} \varpi(n+h)\Lambda_R(n;\mathcal{H}\backslash\{h\},k+\ell)^2, \tag{2.3}$$

provided $R < N$; in fact, if $\varpi(n+h) \ne 0$ and $h \in \mathcal{H}$, then the factor $n+h$ of $P(n;\mathcal{H})$ is irrelevant in computing $\Lambda_R(n;\mathcal{H};k+\ell)$.

We shall work on the assumption: There exists an absolute constant $0 < \theta < 1$ such that we have, for any fixed $A > 0$,

$$\sum_{q\le x^\theta} \max_{y\le x} \max_{\substack{a\\(a,q)=1}} |\vartheta^*(y;a,q) - y/\varphi(q)| \ll x/(\log x)^A, \quad \vartheta^*(y;a,q) = \sum_{\substack{y<n\le 2y\\ n\equiv a \bmod q}} \varpi(n), \tag{2.4}$$

with the implicit constant depending only on A. We assume that

$$R \le N^{\theta/2}/(\log N)^A. \tag{2.5}$$

In particular, (2.3) implies that we may assume also that $h \notin \mathcal{H}$.

With this, expanding out the square in (2.2), we see that the sum is equal to

$$\sum_{d_1,d_2} \lambda_R(d_1;k+\ell)\lambda_R(d_2;k+\ell) \sum_{b\in\Omega([d_1,d_2])} \delta((b+h,[d_1,d_2]))\vartheta^*(N;b+h,[d_1,d_2]), \tag{2.6}$$

where $\delta(x)$ is the unit measure placed at $x = 1$, because $\vartheta^*(N,b+h,[d_1,d_2]) = 0$ if $b+h$ and $[d_1,d_2]$ are not coprime. Then, by (2.4), this is equal to

$$N\mathcal{T}^* + O(N/(\log N)^{A/3}), \tag{2.7}$$

with

$$\mathcal{T}^* = \sum_{d_1,d_2} \frac{\lambda_R(d_1;k+\ell)\lambda_R(d_2;k+\ell)}{\varphi([d_1,d_2])} \sum_{b\in\Omega([d_1,d_2])} \delta((b+h,[d_1,d_2])). \tag{2.8}$$

The error term in (2.7) might require an explanation: We divide the sum in (2.6) into two parts according as $|\Omega([d_1,d_2])| = \tau_k([d_1,d_2]) \le (\log N)^{A/2}$ and otherwise. To the first part we apply (2.4), while the second part is

$$\ll N(\log R)^{2(k+\ell)} \log N \sum_{d_1,d_2\le R} \frac{\tau_k([d_1,d_2])}{(\log N)^{A/2}} \frac{|\Omega([d_1,d_2])|}{[d_1,d_2]} \ll N/(\log N)^{A/3}, \tag{2.9}$$

provided A is sufficiently large.

It remains for us to evaluate $\mathcal{T}^*$. The inner sum of (2.8) is equal to

$$\prod_{p|[d_1,d_2]}\left(\sum_{b\in\Omega(p)}\delta((b+h,p))\right)=\prod_{p|[d_1,d_2]}(|\Omega^+(p)|-1). \tag{2.10}$$

Here Ω^+ corresponds to the set $\mathcal{H}^+=\mathcal{H}\cup\{h\}$. In fact, $\delta((b+h,p))$ vanishes if and only if $-h\in\Omega(p)$; and the latter is equivalent to $\Omega(p)=\Omega^+(p)$. Note that the analogue of (1.5) for Ω^+ could be violated. Nevertheless, we have, as before,

$$\mathcal{T}^*=\frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)}\prod_p\left(1-\frac{|\Omega^+(p)|-1}{p-1}\left(\frac{1}{p^{s_1}}+\frac{1}{p^{s_2}}-\frac{1}{p^{s_1+s_2}}\right)\right)\frac{R^{s_1+s_2}}{(s_1s_2)^{k+\ell+1}}ds_1ds_2. \tag{2.11}$$

For $p>H$, we have $|\Omega^+(p)|=k+1$, since $h\notin\mathcal{H}$. Thus, we consider the function

$$\prod_p(\cdots)\left(\frac{\zeta(s_1+1)\zeta(s_2+1)}{\zeta(s_1+s_2+1)}\right)^k \tag{2.12}$$

as in (1.17). If $\mathcal{H}^+$ is admissible, the singular series is $\mathfrak{S}(\mathcal{H}^+)$ and the argument and computation of residues is analogous to above. Thus we find that provided $h\notin\mathcal{H}$

$$\mathcal{T}^*=\frac{\mathfrak{S}(\mathcal{H}^+)}{(k+2\ell)!}\binom{2\ell}{\ell}(\log R)^{k+2\ell}+O((\log N)^{k+2\ell-1}(\log\log N)^c). \tag{2.13}$$

On the other hand, if $\mathcal{H}^+$ is not admissible or $\mathfrak{S}(\mathcal{H}^+)=0$, then the Euler product in (2.11) vanishes at either $s_1=0$ or $s_2=0$ to the order equal to the number of primes such that $|\Omega^+(p)|=p$. However, since we have then $p\le k+1$, the necessary change to the above reasoning results only in the lack of the main term in (2.13) and the error term remains to be the same or actually smaller.

Finally, if $h\in\mathcal{H}$, the above evaluation applies with the translation $k\mapsto k-1$, $\ell\mapsto\ell+1$ to (2.13) because of (2.3).

From this, we obtain

Lemma 2. *Provided* (1.1), (1.2), *and* (2.4) *hold, we have, for $R\le N^{\theta/2}/(\log N)^C$ with a sufficiently large $C>0$ depending only on k and ℓ,*

$$\begin{aligned}&\sum_{N<n\le 2N}\varpi(n+h)\Lambda_R(n;\mathcal{H},k+\ell)^2\\ &=\begin{cases}\dfrac{\mathfrak{S}(\mathcal{H}\cup\{h\})}{(k+2\ell)!}\dbinom{2\ell}{\ell}N(\log R)^{k+2\ell}+O(N(\log N)^{k+2\ell-1}(\log\log N)^c) & \textit{if } h\notin\mathcal{H},\\ \dfrac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!}\dbinom{2(\ell+1)}{\ell+1}N(\log R)^{k+2\ell+1}+O(N(\log N)^{k+2\ell}(\log\log N)^c) & \textit{if } h\in\mathcal{H}.\end{cases}\end{aligned} \tag{2.14}$$

3. Proof of (0)

We are now ready to prove (0). We shall evaluate the expression

$$\sum_{\substack{\mathcal{H}\subseteq[1,H]\\|\mathcal{H}|=k}}\sum_{N<n\le 2N}\left(\sum_{h\le H}\varpi(n+h)-\log 3N\right)\Lambda_R(n;\mathcal{H},k+\ell)^2. \tag{3.1}$$

If this turns out to be positive, then there exists an integer $n \in (N, 2N]$ such that

$$\sum_{h \le H} \varpi(n+h) - \log 3N > 0. \tag{3.2}$$

That is, there exists a subinterval of length H in $(N, 2N+H]$ which contains *two* primes; hence

$$\min_{N < p_r \le 2N+H} (p_{r+1} - p_r) \le H. \tag{3.3}$$

Here, we need to quote, from [2],

$$\sum_{\substack{\mathcal{H} \subseteq [1,H] \\ |\mathcal{H}|=k}} \mathfrak{S}(\mathcal{H}) = (1+o(1))H^k, \tag{3.4}$$

as H tends to infinity. With this and Lemma 1, we see that (3.1) is asymptotically equal to

$$\begin{aligned} \sum_{\substack{\mathcal{H} \subseteq [1,H] \\ |\mathcal{H}|=k}} \sum_{N<n\le 2N} \left\{ \sum_{\substack{h \le H \\ h \notin \mathcal{H}}} + \sum_{\substack{h \le H \\ h \in \mathcal{H}}} \right\} \varpi(n+h) \Lambda_R(n; \mathcal{H}, k+\ell)^2 \\ - \frac{1}{(k+2\ell)!} \binom{2\ell}{\ell} N H^k (\log N)(\log R)^{k+2\ell}, \end{aligned} \tag{3.5}$$

with an admissible error of the size of $o\left(NH^k (\log N)^{k+2\ell+1}\right)$. By Lemma 2 and (3.4) with an appropriate replacement of $\mathcal{H}$, this is asymptotically equal, in the same sense, to

$$\begin{aligned} &\frac{1}{(k+2\ell)!}\binom{2\ell}{\ell} NH^{k+1}(\log R)^{k+2\ell} + \frac{k}{(k+2\ell+1)!}\binom{2(\ell+1)}{\ell+1} NH^k (\log R)^{k+2\ell+1} \\ &\qquad - \frac{1}{(k+2\ell)!}\binom{2\ell}{\ell} NH^k(\log N)(\log R)^{k+2\ell} \\ &= \left\{ H + \frac{k}{k+2\ell+1} \cdot \frac{2(2\ell+1)}{\ell+1} \cdot \log R - \log N \right\} \frac{1}{(k+2\ell)!}\binom{2\ell}{\ell} NH^k (\log R)^{k+2\ell}. \end{aligned} \tag{3.6}$$

Hence (3.1) is positive, provided

$$H \ge \left(1 + \varepsilon - \frac{k}{k+2\ell+1} \cdot \frac{2(2\ell+1)}{\ell+1} \cdot \frac{\theta}{2} \right) \log N, \tag{3.7}$$

with any fixed $\varepsilon > 0$. Therefore, with $\ell = [\sqrt{k}]$, say, we obtain

$$\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} \le \max\{0, 1 - 2\theta\}. \tag{3.8}$$

In particular, the Bombieri–Vinogradov Prime Number Theorem [1, Théorème 17] gives rise to the assertion (0). This ends the proof.

Finally, we shall exhibit two conjectural assertions:

1) If we have $\theta > \frac{1}{2}$, then there will be infinitely many n such that $p_{n+1} - p_n \le c(\theta)$ with an absolute constant $c(\theta)$. In fact, we would be able to suppose $H > c(\theta)$ in the above as far as (3.6), and the assertion follows immediately.

2) If we have $\theta > \frac{20}{21}$, then we will be able to assert the simultaneous appearance of primes in admissible 7-tuples. For instance, at least *two* of the seven integers $\{n, n+2, n+6, n+8, n+12, n+18, n+20\}$ will

be primes for infinitely many n. In particular, $p_{n+1} - p_n \le 20$ infinitely often. To prove this, let $\mathcal{H}$ be such a tuple, and consider, in place of (3.1),

$$\sum_{N<n\le 2N} \left(\sum_{h\in\mathcal{H}} \varpi(n+h) - \log 3N \right) \Lambda_R(n;\mathcal{H},8)^2. \tag{3.9}$$

Lemmas 1 and 2, with $k = 7$, $\ell = 1$, imply that under the present assumption on θ this is asymptotically equal to

$$\frac{2}{9!}\mathfrak{S}(\mathcal{H}) \left(\frac{21}{10} \log R - \log N \right) N(\log R)^9 > 0, \tag{3.10}$$

provided $R = N^{10/21+\xi}$ with sufficiently large N and a small $\xi > 0$. Hence the assertion follows.

4. Expository

The principal idea in [3] is the amazing effect induced by the introduction of the parameter ℓ in (1.9). The sieve weight $\mu(d)(\log m/d)^{k+\ell}$, $d|m$, applied to the polynomial $m = P(n;\mathcal{H})$ detects n with which $P(n;\mathcal{H})$ has $k+\ell$ distinct prime factors at most, implying that the integers $n + h_j$, $j \le k$, are mostly primes, provided k is large compared with ℓ. By a standard method in this field, we approximate these sieve weights by $\lambda_R(d; k+\ell)$, and consider the Selberg sieve situation (1.9), with the parameters ℓ and R at our disposal. An asymptotic formula for the sum (1.9) is given in (1.31). Then, to detect at least two primes among $n + h_j$, $j \le k$, a usual weighted sieve situation is considered at (3.1); for this the other asymptotic formula (2.14) is required. The upshot is condensed in (3.6) and (3.7). The proof of (0) requires that both k and ℓ can be taken appropriately and the Bombieri–Vinogradov prime number theorem is available.

Rendering the above more technically, the reason for success lies not only in the introduction of the parameter ℓ but also in the trivial fact (2.3), which brings forth the translation $\ell \mapsto \ell + 1$ as remarked in the proof of Lemma 2. This introduces the factor $\binom{2(\ell+1)}{\ell+1}$ on the right of (2.14). One should note that $\binom{2(\ell+1)}{\ell+1}/\binom{2\ell}{\ell} = \frac{2(2\ell+1)}{(\ell+1)}$, which tends to 4 as $\ell \to \infty$. This is extremely critical when appealing to the Bombieri–Vinogradov prime number theorem. On the other hand, the translation $k \mapsto k-1$ does not cause any effect as long as k is much larger than ℓ.

References

[1] E. Bombieri. Le Grand Crible dans la Théorie Analytique des Nombres (Seconde Édition). Astérisque **18**, Paris 1987.

[2] P.X. Gallagher. On the distribution of primes in short intervals. Mathematika, **23** (1976), 4–9. Corrigendum, ibid **28** (1981), 86.

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

[4] E.C. Titchmarsh. The Theory of the Riemann Zeta-Function. Oxford at the Clarendon Press, First Edition, 1951; Second Edition, 1986.

Department of Mathematics, San Jose State University, San Jose, CA 95192, USA
E-mail address: goldston@math.sjsu.edu

Department of Mathematics, College of Science and Technology, Nihon University, Surugadai, Tokyo 101-8308, JAPAN
E-mail address: ymoto@math.cst.nihon-u.ac.jp

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

Department of Mathematics, Boğaziçi University, Bebek, Istanbul 34342
& Feza Gürsey Enstitüsü, Çengelköy, Istanbul, P.K. 6, 81220 TURKEY
E-mail address: yalciny@boun.edu.tr