Small gaps between primes exist

By Daniel Alan GOLDSTON,*) Yoichi MOTOHASHI,**) János PINTZ,***) and Cem Yalçın YILDIRIM****)

(Communicated by Shigefumi MORI, M. J. A., April 12, 2006)

Abstract: In the 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]. We include a short expository last section.

Key word: Prime number.

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

(1.2) integers $k, \ell > 0$ are arbitrary but bounded.

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 (this aspect is to be discussed in the publication version of [3] and its continuations). 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

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

with $h_i \ne h_j$ for $i \ne j$. For a prime p, let $\Omega(p)$ be the set of different residue classes among $-h \pmod p$, $h \in \mathcal{H}$, and write $n \in \Omega(p)$ instead of $n \pmod p \in \Omega(p)$. We call $\mathcal{H}$ admissible if

$$|\Omega(p)| < p \quad \text{for all} \quad p, \tag{1.3}$$

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 $d|P(n;\mathcal{H})$ with $P(n;\mathcal{H}) = (n+h_1)(n+h_2)\cdots(n+h_k)$. Also, we put, with μ the Möbius function,

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

and

$$\begin{aligned}\Lambda_R(n;\mathcal{H},a) &= \sum_{n\in\Omega(d)} \lambda_R(d;a) \\ &= \frac{1}{a!}\sum_{\substack{d|P(n;\mathcal{H})\\ d\le R}} \mu(d)\left(\log\frac{R}{d}\right)^a.\end{aligned}$$

With this, we shall consider the evaluation of

$$\sum_{N<n\le 2N} \Lambda_R(n;\mathcal{H},k+\ell)^2. \tag{1.4}$$

Expanding out the square, we have

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

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

*) Department of Mathematics, San Jose State University, San Jose, CA 95192, USA.

**) Department of Mathematics, College of Science and Technology, Nihon University, Surugadai, Tokyo 101-8308, Japan.

***) Rényi Mathematical Institute of the Hungarian Academy of Sciences, H-1364 Budapest, P.O.B. 127, Hungary.

****) Department of Mathematics, Boğaziçi University, Bebek, Istanbul 34342 & Feza Gürsey Enstitüsü, Çengelköy, Istanbul, P.K. 6, 81220, Turkey.

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 (1.4) is equal to

$$N\mathcal{T} + O\left(\left(\sum_d |\Omega(d)|\,|\lambda_R(d; k+\ell)|\right)^2\right),$$

in which

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

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

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

On noting that for $a \ge 1$

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

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

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

where

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

in the region of absolute convergence.

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

$$\begin{aligned} &G(s_1, s_2; \Omega) \\ &\quad = F(s_1, s_2; \Omega)\left(\frac{\zeta(s_1+1)\zeta(s_2+1)}{\zeta(s_1+s_2+1)}\right)^k, \end{aligned} \tag{1.5}$$

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

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

which does not vanish, because of (1.3). We have the bound

$$\begin{aligned} &G(s_1, s_2; \Omega) \\ &\quad \ll \exp\left(c(\log N)^{-2\sigma} \log\log\log N\right), \end{aligned} \tag{1.6}$$

with $\min(\operatorname{Re} s_1, \operatorname{Re} s_2, 0) = \sigma \ge -c$, as can be seen via the Euler product expansion of the right side of (1.5). In fact, the part corresponding to those $p > H$ is uniformly bounded in the indicated region since $|\Omega(p)| = k$; and the logarithm of the remaining part is estimated to be $\ll H^{-2\sigma}\sum_{p\le H} p^{-1}$. Note that the restrictions (1.1) and (1.2) are relevant here.

Now, we have

$$\begin{aligned} \mathcal{T} = \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)} G(s_1,s_2;\Omega)&\left(\frac{\zeta(s_1+s_2+1)}{\zeta(s_1+1)\zeta(s_2+1)}\right)^k \\ &\times \frac{R^{s_1+s_2}}{(s_1s_2)^{k+\ell+1}}\, ds_1 ds_2. \end{aligned}$$

Let us put $U = \exp(\sqrt{\log N})$, and shift the s_1 and s_2-contours to the vertical lines $c_0(\log U)^{-1} + it$ and to $c_0(2\log U)^{-1} + it$, $t \in \mathbf{R}$, respectively, with a sufficiently small constant $c_0 > 0$; necessary facts about the functions ζ and $1/\zeta$ can be found in [4, Section 3.11], from which the choice of c_0 transpires. We truncate the contours to $|t| \le U$ and $|t| \le U/2$, and denote the results by L_1 and L_2, respectively. On noting (1.1) and (1.6), we have readily that

$$\begin{aligned} \mathcal{T} = &\frac{1}{(2\pi i)^2}\int_{L_2}\int_{L_1} G(s_1,s_2;\Omega) \\ &\times \left(\frac{\zeta(s_1+s_2+1)}{\zeta(s_1+1)\zeta(s_2+1)}\right)^k \frac{R^{s_1+s_2}}{(s_1s_2)^{k+\ell+1}} ds_1 ds_2 \\ &+ O\left(\exp\left(-c\sqrt{\log N}\right)\right). \end{aligned}$$

We then shift the s_1-contour to L_3: $-c_0(\log U)^{-1} + it$, $|t| \le U$. We encounter singularities at $s_1 = 0$ and $s_1 = -s_2$, which are poles of orders $\ell + 1$ and k, respectively. We have

$$\begin{aligned} \mathcal{T} = &\frac{1}{2\pi i}\int_{L_2} \left\{\operatorname*{Res}_{s_1=0} + \operatorname*{Res}_{s_1=-s_2}\right\} ds_2 \\ &+ O\left(\exp\left(-c\sqrt{\log N}\right)\right), \end{aligned} \tag{1.7}$$

in which we have used (1.6). We rewrite the residue, and have

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

with the circle $C(s_2)$: $|s_1+s_2| = (\log N)^{-1}$. By (1.6), we have $G(s_1, s_2; \Omega) \ll (\log\log N)^c$; $\zeta(s_1+s_2+1) \ll \log N$; $R^{s_1+s_2} \ll 1$. 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

$$\operatorname*{Res}_{s_1=-s_2} \ll (\log N)^{k-1}(\log\log N)^c \times \left(\frac{\log(|s_2|+2)}{|s_2|+1}\right)^{2k} |s_2|^{-2\ell-2}. \tag{1.8}$$

Inserting this into (1.7), we get

$$\mathcal{T} = \frac{1}{2\pi i}\int_{L_2}\left\{\operatorname*{Res}_{s_1=0}\right\} ds_2 + O\left((\log N)^{k+\ell-1/2}(\log\log N)^c\right). \tag{1.9}$$

To evaluate the last integral, we put

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

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

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

We insert this into (1.9) and shift the contour to L_4: $-c_0(\log U)^{-1}+it$, $|t|\le U/2$. We see the new integral is $O\big(\exp\big(-c\sqrt{\log N}\big)\big)$; the necessary bound for the integrand can be obtained in much the same way as in (1.8). Thus

$$\begin{aligned}\mathcal{T} &= \operatorname*{Res}_{s_2=0}\operatorname*{Res}_{s_1=0} + O\big((\log N)^{k+\ell}\big)\\ &= \frac{1}{(2\pi i)^2}\int_{C_2}\int_{C_1}\frac{Z(s_1,s_2)R^{s_1+s_2}}{(s_1+s_2)^k(s_1s_2)^{\ell+1}}\,ds_1\,ds_2 + O\big((\log N)^{k+\ell}\big),\end{aligned}$$

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

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

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

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

where we have used (1.6); note that $Z(0,0)=\mathfrak{S}(\mathcal{H})$.

Hence, we have obtained our basic implement:

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

$$\sum_{N<n\le 2N}\Lambda_R(n;\mathcal{H},k+\ell)^2 = \frac{\mathfrak{S}(\mathcal{H})}{(k+2\ell)!}\binom{2\ell}{\ell}N(\log R)^{k+2\ell} + O\big(N(\log N)^{k+2\ell-1}(\log\log N)^c\big). \tag{1.10}$$

2. Twist with primes. Next, let $\varpi(n)$ be equal to $\log n$ if n is a prime, and to 0 otherwise; and let us consider the evaluation of the sum

$$\sum_{N<n\le 2N}\varpi(n+h)\Lambda_R(n;\mathcal{H},k+\ell)^2, \tag{2.1}$$

with an arbitrary positive integer $h\le H$. We first observe that this is equal to

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

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 $A>0$,

$$\sum_{q\le x^\theta}\max_{y\le x}\max_{\substack{a\\(a,q)=1}}\left|\vartheta^*(y;a,q)-\frac{y}{\varphi(q)}\right| \ll \frac{x}{(\log x)^A}, \tag{2.3}$$

with

$$\vartheta^*(y;a,q) = \sum_{\substack{y<n\le 2y\\ n\equiv a \bmod q}}\varpi(n),$$

where φ is the Euler totient function, and the implicit constant depends only on A.

We assume that

$$R\le N^{\theta/2}.$$

In particular, we may assume also that $h\notin\mathcal{H}$ in (2.1).

With this, expanding out the square in (2.1), we see that the sum is equal to

$$\sum_{d_1,d_2}\lambda_R(d_1;k+\ell)\lambda_R(d_2;k+\ell) \times \sum_{b\in\Omega([d_1,d_2])}\delta((b+h,[d_1,d_2]))\vartheta^*(N;b+h,[d_1,d_2]) + O\big(R^2(\log N)^c\big), \tag{2.4}$$

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 (2.3), this is equal to

$$NT^* + O\left(\frac{N}{(\log N)^{A/3}}\right), \tag{2.5}$$

with

$$\mathcal{T}^* = \sum_{d_1,d_2} \frac{\lambda_R(d_1; k+\ell)\lambda_R(d_2; k+\ell)}{\varphi([d_1,d_2])} \times \sum_{b\in\Omega([d_1,d_2])} \delta((b+h,[d_1,d_2])). \tag{2.6}$$

The error term in (2.5) might require an explanation: We consider first the part of (2.4) in which $|\Omega([d_1,d_2])| \le \tau_k([d_1,d_2]) < (\log N)^{A/2}$. Note that then $|\{d_1,d_2\colon [d_1,d_2]=d\}| = \tau_3(d)$ is less than $(\log N)^{(A\log 3)/(2\log k)}$. With this, we appeal to (2.3). On the other hand the remaining part is

$$\begin{aligned}&\ll N(\log R)^{2(k+\ell)}\log N \\ &\quad\times \sum_{d_1,d_2\le R} \frac{\tau_k([d_1,d_2])}{(\log N)^{A/2}} \frac{|\Omega([d_1,d_2])|}{[d_1,d_2]} \\ &\ll \frac{N}{(\log N)^{A/3}},\end{aligned}$$

provided A is sufficiently large.

It remains for us to evaluate $\mathcal{T}^*$. The inner sum of (2.6) is equal to

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

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 (1.3) for Ω^+ can be violated. Nevertheless, we have, as before,

(2.7)
$$\begin{aligned}\mathcal{T}^* = \frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)}\prod_p\biggl(1-\frac{|\Omega^+(p)|-1}{p-1} \\ \cdot\left(\frac{1}{p^{s_1}}+\frac{1}{p^{s_2}}-\frac{1}{p^{s_1+s_2}}\right)\biggr)\frac{R^{s_1+s_2}}{(s_1s_2)^{k+\ell+1}}\,ds_1\,ds_2.\end{aligned}$$

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

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

as in (1.5). If $\mathcal{H}^+$ is admissible, the singular series is $\mathfrak{S}(\mathcal{H}^+)$ and the argument and computation of residues is analogous to the above. We find that provided $h\notin\mathcal{H}$

$$\begin{aligned}\mathcal{T}^* &= \frac{\mathfrak{S}(\mathcal{H}^+)}{(k+2\ell)!}\binom{2\ell}{\ell}(\log R)^{k+2\ell} \\ &\quad+ O\left((\log N)^{k+2\ell-1}(\log\log N)^c\right).\end{aligned} \tag{2.8}$$

On the other hand, if $\mathcal{H}^+$ is not admissible, i.e., $\mathfrak{S}(\mathcal{H}^+)=0$, then the Euler product in (2.7) 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 (2.8) and the error term remains to be the same or actually smaller.

Finally, if $h\in\mathcal{H}$, then the above evaluation applies with the translation $k\mapsto k-1$, $\ell\mapsto\ell+1$ to (2.8) because of (2.2).

From this, we obtain

Lemma 2. *If* (1.1), (1.2), *and* (2.3) *hold, then we have, for* $R\le N^{\theta/2}$,

(2.9)
$$\begin{aligned}&\sum_{N<n\le 2N}\varpi(n+h)\Lambda_R(n;\mathcal{H},k+\ell)^2 \\ &= \begin{cases} \dfrac{\mathfrak{S}(\mathcal{H}\cup\{h\})}{(k+2\ell)!}\dbinom{2\ell}{\ell}N(\log R)^{k+2\ell} \\ \quad + O(N(\log N)^{k+2\ell-1}(\log\log N)^c) & \text{if } h\notin\mathcal{H}, \\ \dfrac{\mathfrak{S}(\mathcal{H})}{(k+2\ell+1)!}\dbinom{2(\ell+1)}{\ell+1}N(\log R)^{k+2\ell+1} \\ \quad + O(N(\log N)^{k+2\ell}(\log\log N)^c) & \text{if } h\in\mathcal{H}.\end{cases}\end{aligned}$$

3. Proof of (0).

We are now ready to prove the assertion (0). To this end, we shall evaluate the expression

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

where we set $R=N^{\theta/2}$ so that both (1.10) and (2.9) hold. If (3.1) turns out to be positive, then there exists an integer $n\in(N,2N]$ such that

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

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

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

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

$$\sum_{\substack{\mathcal{H}\subseteq[1,H]\\|\mathcal{H}|=k}}\mathfrak{S}(\mathcal{H}) = (1+o(1))H^k, \tag{3.2}$$

as H tends to infinity; note that the permutations of the elements $\{h_1, \dots, h_k\}$ are counted on the left side. With this and Lemma 1, we see that (3.1) is asymptotically equal to

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

with an error of the size of $o\big(NH^k(\log N)^{k+2\ell+1}\big)$. By Lemma 2 and (3.2) with an appropriate replacement of $\mathcal{H}$, this is asymptotically equal, in the same sense, to

(3.3)

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

Hence (3.1) is positive, provided

$$\text{(3.4)} \qquad \frac{H}{\log N} \ge 1 + \varepsilon - \frac{k}{k+2\ell+1}\cdot\frac{2(2\ell+1)}{\ell+1}\cdot\frac{\theta}{2},$$

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

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

In particular, the Bombieri-Vinogradov prime number theorem [1, Théorème 17] gives rise to the assertion (0), since we may choose θ to be any number less than 1/2.

Finally, we shall exhibit a conditional assertion: If (2.3) holds with an absolute constant $\theta > 1/2$, 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 (3.3), and the assertion follows immediately.

4. Exposition. The principal idea in [3] is the amazing effect induced by the introduction of the parameter ℓ in (1.4). The sieve with 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, and $P(n;\mathcal{H})$ is squarefree; note that those n such that $P(n;\mathcal{H})$ is not squarefree are easily excluded. By a standard method in this field, we approximate these weights by $\lambda_R(d;k+\ell)$, and consider the Selberg sieve situation (1.4), with the parameters ℓ and R at our disposal. An asymptotic formula for the sum (1.4) is given in (1.10). 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 asymptotic formula (2.9) is needed. The upshot is condensed in (3.3) and (3.4). The proof of (0) requires that both k and ℓ can be taken appropriately and the Bombieri-Vinogradov prime number theorem be 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 (2.2), 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 (2.9). One should note that $\binom{2(\ell+1)}{\ell+1}\big/\binom{2\ell}{\ell} = 2(2\ell+1)/(\ell+1)$, 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.

Acknowledgements. 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 Grants-in-Aid for Scientific Research (C) 15540047; the third author by OTKA grants No. T38396, T43623, T49693 and Balaton program; the fourth author by Tübitak.

References

[1] E. Bombieri, *Le grand crible dans la théorie analytique des nombres*, second édition revue et augmentée, Astérisque, 18, Soc. Math. France, Paris, 1987.

[2] P. X. Gallagher, On the distribution of primes in short intervals, Mathematika **23** (1976), no. 1, 4–9; Corrigendum, ibid, **28** (1981), 86.

[3] D. A. Goldston, J. Pintz, and C. Y. Yıldırım, Small gaps between primes II (Preliminary). (February 8, 2005). See also [2005-19 of `http://aimath.org/preprints.html`].

[4] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, Clarendon Press, Oxford, 1951.