18. SHORT GAPS BETWEEN PRIMES.

In this section we shall prove that the gap between two consecutive primes of size x can be much smaller than the average, $\log x$:

Our goal is to show that if $p_1 = 2 < p_2 = 3 < \dots$ is the sequence of prime numbers then

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

Researchers had had little success in showing that there are gaps between consecutive primes that are significantly smaller than the average. However a simple recent method of Goldston, Pintz and Yildirim counts primes in short intervals with simple weighting functions, which themselves can be easily estimated using the Bombieri-Vinogradov theorem, and implies the above. Using deeper techniques the method can be used to prove that there are infinitely many pairs of consecutive primes for which

$$p_{n+1}-p_n \ll (\log p_n)^{1/2+\epsilon}.$$

The Elliott-Halberstam conjecture states that, for any $A > 0$, we have

$$\sum_{q\le Q} \max_{(a,q)=1} \left| \theta(x;q,a) - \frac{x}{\phi(q)} \right| \ll_A \frac{x}{(\log x)^A} \tag{18.1}$$

for any $Q \le x^{1-\epsilon}$. We saw in chapter 13 that this holds for $Q \le x^{1/2}/(\log x)^B$ by the Bombieri-Vinogradov theorem. We warm up for our main result with another striking result of Goldston, Pintz and Yildirim:

Theorem 18.1. *Suppose that* (18.1) *holds for $Q = x^\theta$, for some fixed $\theta > \frac{20}{21}$. Then there are infinitely many pairs of consecutive primes $p_n < p_{n+1}$ such that $p_{n+1} - p_n \le 20$.*

One can improve the 20 here to 16 using other methods. If (18.1) holds for $Q = x^\theta$, for any fixed $\theta > \frac{1}{2}$ then there are infinitely many pairs of consecutive primes $p_n < p_{n+1}$ for which $p_{n+1} - p_n \ll_\theta 1$.

Actually we prove somewhat more than Theorem 18.1: Let $\mathcal{H}$ be a set of linear forms $\{a_i x + b_i : i = 1, 2, \dots, k\}$ where $b_i a_j \ne a_i b_j$ if $i \ne j$; and define $P_{\mathcal{H}}(n) := \prod_{i=i}^k (a_i n + b_i)$. Let $\nu_{\mathcal{H}}(p)$ denote the number of distinct residue classes $n \pmod p$ for which p divides $P_{\mathcal{H}}(n)$, and extend this definition to $\nu_{\mathcal{H}}(d)$ for squarefree integers d by multiplicativity. Define the *singular series*

$$\mathfrak{S}(\mathcal{H}) = \prod_p \left(1 - \frac{1}{p}\right)^{-k} \left(1 - \frac{\nu_{\mathcal{H}}(p)}{p}\right)$$

If $\mathfrak{S}(\mathcal{H}) \ne 0$ then $\mathcal{H}$ is called *admissible*. Thus $\mathcal{H}$ is admissible if and only if $\nu_{\mathcal{H}}(p) < p$ for all p.

Proposition 18.2. *Suppose that* (18.1) *holds for* $Q = x^\theta$, *for some fixed* $\theta > \frac{20}{21}$. *Then for any admissible 7-tuple* $\mathcal{H}$, *there are infinitely many* n *such that at least two members of* $\mathcal{H}\big|_{x=n}$ *are prime.*

To deduce Theorem 18.1 from Proposition 18.2 note that

$$\{x-10, x-8, x-4, x-2, x+2, x+8, x+10\}$$

is an admissible 7-tuple with each $a_i = 1$ (and $|b_7 - b_1|$ as small as possible).

For any non-zero integer b whose prime factors are all ≥ 11, the set $\{210ix + b : 1 \in I\}$ is admissible for any set, I, of seven distinct integers. Taking $I = \{1,2,3,4,5,6,7\}$ we deduce that there exists $r/s > 1$ with $1 \leq s < r \leq 7$ such that there are infinitely many pairs of primes p, q for which $(p-b)/(q-b) = r/s$. Taking $I = \{m^i : 0 \leq i \leq 6\}$ for some integer $m > 1$ we deduce that there exists $j, 1 \leq j \leq 6$ such that there are infinitely many pairs of primes p, q for which $(p-b)/(q-b) = m^j$ (for example $p - 1 = 2^j(q-1)$ for infinitely many pairs of primes p, q where $j = 1,2,3,4,5$ or 6.). There are, presumably, many other such consequences of Proposition 18.2 to be found.

The main result of this chapter is the following:

Theorem 18.3. *For any fixed* $\epsilon > 0$ *there are infinitely many pairs of consecutive primes* $p_n < p_{n+1}$ *for which* $p_{n+1} - p_n < \epsilon \log p_n$. *In other words*

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

18.2. Proof of the main results. Define

$$\Lambda_R(n; \mathcal{H}, \ell) = \frac{1}{(k+\ell)!} \sum_{\substack{d | P_{\mathcal{H}}(n) \\ d \leq R}} \mu(d) \left(\log \frac{R}{d}\right)^{k+\ell}.$$

Our results rest on the following estimates:

Proposition 18.4. *Suppose* $\mathcal{H}$ *is admissible with* $|\mathcal{H}| = k$, *and* $a_0x + b_0 \in \mathcal{H}$; *and let* $h := \max_i\{|a_i|, |b_i|\}$. *For* $R \leq N^{1/2}/(\log N)^{2k}$ *and* $h \leq R^{O(1)}$ *with* $R, N \to \infty$, *and any integer* $\ell \geq 0$, *we have*

$$\sum_{n \leq N} \Lambda_R(n; \mathcal{H}, \ell)^2 = \binom{2\ell}{\ell} \frac{(\log R)^{k+2\ell}}{(k+2\ell)!} (\mathfrak{S}(\mathcal{H}) + o(1))N.$$

If (18.1) *holds with* $Q = R^2$, *and* $h \leq R^\epsilon$ *with* $R, N \to \infty$, *and for any integer* $\ell \geq 0$, *we have*

$$\sum_{n \leq N} \Lambda_R(n; \mathcal{H}, \ell)^2 \vartheta(a_0 n + b_0) = \binom{2\ell+2}{\ell+1} \frac{(\log R)^{k+2\ell+1}}{(k+2\ell+1)!} (\mathfrak{S}(\mathcal{H}) + o(1))N.$$

Note that if a_0n+b_0 is a prime then $\Lambda_R(n;\mathcal{H},\ell)=\Lambda_R(n;\mathcal{H}\cup\{a_0x+b_0\},\ell)$ by definition.

We will indicate the proof of this in section 18.5, after showing how such a result can be applied. We will work out many of the details of the proof (though we will not go into the details of how the theorem holds uniformly in h as claimed).

18.2. Deduction of the main results from the key proposition. By Proposition 18.4 for $\ell \geq 0$ and $R = N^{\vartheta/2-\epsilon/4}$, we have

$$\begin{aligned}
\mathcal{S} :&= \sum_{n=N+1}^{2N} \left(\sum_{i=1}^{k} \vartheta(a_i n+b_i) - \log 3N \right) \Lambda_R(n;\mathcal{H},\ell)^2 \\
&= k\binom{2\ell+2}{\ell+1}\frac{(\log R)^{k+2\ell+1}}{(k+2\ell+1)!}(\mathfrak{S}(\mathcal{H})+o(1))N - \log 3N\binom{2\ell}{\ell}\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}(\mathfrak{S}(\mathcal{H})+o(1))N \\
&= \left(\frac{2k}{k+2\ell+1}\frac{2\ell+1}{\ell+1}\log R - \{1+o(1)\}\log 3N\right)\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H})N \\
&\geq \left(\frac{k}{k+2\ell+1}\frac{\ell+1/2}{\ell+1}\cdot 2\vartheta - 1 - \epsilon\right)\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}\binom{2\ell}{\ell}\mathfrak{S}(\mathcal{H})N\log 3N.
\end{aligned}$$

The term inside the brackets is greater than a positive constant (for ϵ sufficiently small) provided $\frac{k}{k+2\ell+1}\frac{\ell+1/2}{\ell+1}\cdot 2\vartheta > 1$. Evidently one can choose such k and ℓ for any $\vartheta > \frac{1}{2}$; in particular one can take $\ell = 1$ and $k = 7$ when $\vartheta > 20/21$. Thus we deduce Proposition 18.2.

In fact, by the Cauchy-Schwarz inequality, one can show that the number of such $n \leq N$ is $\gg N/(\log N)^C$ for some constant C.

Exercises

18.2a. Suppose that the Elliott-Halberstam conjecture holds for some $\vartheta > 1/2$. Show that there exists an integer k, not much larger than $1/(2\theta-1)$, such that for any admissible k-tuple $\mathcal{H}$, there are infinitely many n for which at least two members of $\mathcal{H}\big|_{x=n}$ are prime.

The Bombieri-Vinogradov theorem allows us to take any $\theta \leq 1/2$ in the above argument so, after exercise 18.2a, we see that we just fail to prove unconditionally that there are bounded gaps between consecutive primes. The question therefore becomes as to how we can "win an ϵ". The idea of Goldston, Pintz and Yildirim is to look for prime values amongst forms a_0x+b_0 that do not belong to $\mathcal{H}$. These each contribute much less to S (by a factor of $c\log N$) than the forms in $\mathcal{H}$ but one wins that needed ϵ by taking enough of them. So now let $\mathcal{F}$ be a larger set of K forms a_ix+b_i, containing $\mathcal{H}$, with $h = \max_{1\leq i\leq K}\{|a_i|,|b_i|\}$.

Using the second part of the key proposition with $\mathcal{H}$ replaced by $\mathcal{H}\cup\{a_ix+b_i\}$, as well as k by $k+1$ and ℓ by $\ell-1$ (as discussed just after the statement) we obtain

$$\begin{aligned}\tilde{\mathcal{S}}_{\mathcal{H}} &:= \sum_{n=N+1}^{2N}\left(\sum_{i=1}^{K}\vartheta(a_in+b_i)-\log 3N\right)\Lambda_R(n;\mathcal{H},\ell)^2,\\ &= \mathcal{S}+\sum_{i=k+1}^{K}\sum_{n=N+1}^{2N}\Lambda_R(n;\mathcal{H},\ell)^2\vartheta(a_in+b_i)\\ &= \mathcal{S}+\sum_{i=k+1}^{K}\binom{2\ell}{\ell}\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}(\mathfrak{S}(\mathcal{H}\cup\{a_ix+b_i\})+o(1))N\end{aligned}$$

Therefore taking $\theta=1/2$ and $\ell=[\sqrt{k}/2]$ with $k>5\epsilon^{-2}$ we obtain that

$$\tilde{\mathcal{S}}_{\mathcal{H}}\Big/\left(\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}\binom{2\ell}{\ell}N\right)\geq\sum_{i=k+1}^{K}\mathfrak{S}(\mathcal{H}\cup\{a_ix+b_i\})-2\epsilon\mathfrak{S}(\mathcal{H})\log 3N$$

Now define $\mathfrak{S}_k(\mathcal{F})$ to be the mean value of $\mathfrak{S}(\mathcal{H})$ taken over all k-term subsets $\mathcal{H}$ of $\mathcal{F}$. Then the average of the last equation over all such $\mathcal{H}$ is

$$(K-k)\mathfrak{S}_{k+1}(\mathcal{F})-2\epsilon\mathfrak{S}_k(\mathcal{F})\log 3N.$$

Now when $a_ix+b_i=x+i$ for each i, Gallagher proved that $\mathfrak{S}_k(\mathcal{F})\sim 1$ as $K\to\infty$, for each fixed k (see Lemma 18.9 below). Therefore the last displayed equation is $>\epsilon\log 3N$ for $K=[4\epsilon\log 3N]$. Thus there exists $n, N<n\leq 2N$ such that there are at least two primes amongst $\{n+i: 1\leq i\leq 4\epsilon\log 3N\}$ (once N is sufficiently large), and the result follows letting $\epsilon\to 0$; that is we have proved Theorem 18.3.

18.3. Lemmas.

Lemma 18.6. (a) *We have*

$$\sum_{\substack{N<n\leq 2N\\ d|P_{\mathcal{H}}(n)}}1=\nu_{\mathcal{H}}(d)\left(\frac{N}{d}+O(1)\right).$$

(b) *Fix $A>0$. We also have, for $d\leq(\log N)^A$,*

$$\sum_{\substack{N<n\leq 2N\\ d|P_{\mathcal{H}}(n)}}\vartheta(n+h_0)=\nu^*_{\mathcal{H}}(d)\frac{N}{\phi(d)}+O\left(\frac{N}{(\log N)^A}\right),$$

where $\nu^*_{\mathcal{H}}(p)=\nu_{\mathcal{H}}(p)-1$ for each prime p, and $\nu^*_{\mathcal{H}}(d)$ is multiplicative in d.

Proof. To see (a) note that for each $m \pmod d$ with $d|P_{\mathcal{H}}(m)$ we get $N/d+O(1)$ values of $n\equiv m \pmod d$ with $N<n\leq 2N$ (and these are the values of n for which $d|P_{\mathcal{H}}(n)$). The number of such m is $\nu_{\mathcal{H}}(d)$.

For (b), if we take this same sum where $1<d<N<n$ with the additional condition that $n+h_0$ is prime with $h_0\in\mathcal{H}$ then d cannot divide $n+h_0$; so for each prime p the number of possibilities for $m \pmod d$ is $\nu^*_{\mathcal{H}}(d)$. The result then follows from the Siegel-Walfisz theorem.

Lemma 18.7. (Variant of Perron's formula) *For any fixed* $c > 0$ *we have*

$$\frac{1}{2\pi i}\int_{(c)} \frac{x^s}{s^{k+1}}\, ds = \begin{cases} 0, & \text{if } \ 0 < x \le 1 \\ \frac{1}{k!}(\log x)^k & \text{if } \ x \ge 1, \end{cases} \tag{18.2}$$

Proof. Exercise 18.3a

Lemma 18.8. *Suppose that* $G(s_1, s_2)$ *is a double Dirichlet series which is absolutely convergent when* $Re(s_1)$, $Re(s_2) > -1/2$, *and suppose also that* $G(0,0) \neq 0$. *Then*

$$\frac{1}{(2\pi i)^2}\int_{(1)}\int_{(1)} G(s_1,s_2)\frac{\zeta(1+s_1+s_2)^k}{\zeta(1+s_1)^k\zeta(1+s_2)^k}\frac{R^{s_1+s_2}}{(s_1s_2)^{k+\ell+1}}ds_1ds_2 \sim \binom{2\ell}{\ell} G(0,0)\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}.$$

Sketch of Proof. Let $\mathcal{L}$ denote the contour given by $s = -\frac{1}{100\log(|t|+2))} + it$: by section 9.6 we know that there are no zeros of $\zeta(1+s)$ on or to the right of $\mathcal{L}$ and so that there are good bounds for $\zeta, 1/\zeta$ and ζ'/ζ in this region. The (difficult) exercise is to bound the contribution of all of the contours to the error terms in this proof.

To compute the above integral write $w = s_1+s_2$ and $s = s_1$ and let $F(s,w) = G(s, w-s)$, so that we have

$$\frac{1}{(2\pi i)^2}\int_{w=(2)}\int_{s=(1)} F(s,w)\frac{\zeta(1+w)^k}{\zeta(1+s)^k\zeta(1+w-s)^k}\frac{R^w}{(s(w-s))^{k+\ell+1}}dsdw.$$

We move the s-contour to the left to $\mathcal{L}$. The only residue is the pole at $s=0$ which contributes

$$\frac{1}{2\pi i}\int_{w=(2)} \frac{1}{\ell!}\frac{\delta^\ell}{\delta s^\ell}\left(\frac{H(s,w)}{(w-s)^{\ell+1}}\right)\Big|_{s=0}\zeta(1+w)^k R^w dw,$$

writing $H(s,w) = F(s,w)/((s\zeta(1+s))((w-s)\zeta(1+w-s)))^k$ which is analytic in this domain, and equals

$$= \frac{1}{2\pi i}\int_{w=(2)} \sum_{i=0}^{\ell}\binom{2\ell-i}{\ell}\frac{H(s,w)^{(i)}}{i!}\Big|_{s=0}\frac{(w\zeta(1+w))^k}{w^{k+2\ell-i+1}}R^w dw.$$

We now move the contour over w to the left to $\mathcal{L}$, so that the main term comes from the unique pole at $w=0$, and contributes

$$\sum_{\substack{i+j+m+n=k+2\ell \\ i\le \ell}} \binom{2\ell-i}{\ell}\frac{H(0,0)^{(i,j)}}{i!j!}\frac{a_m}{m!}\frac{(\log R)^n}{n!}$$

where $H(0,0)^{(i,j)} = \left(\frac{\partial}{\partial s}\right)^i\left(\frac{\partial}{\partial w}\right)^j H(s,w)\big|_{s=0,w=0}$, and we have written $(w\zeta(1+w))^k = \sum_{i\ge 0} a_i w^i$. The largest power of $\log R$ comes from the term with $i=0$ and thus we get a main term of

$$\binom{2\ell}{\ell}H(0,0)\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}(w\zeta(1+w))^k\big|_{w=0} = \binom{2\ell}{\ell}G(0,0)\frac{(\log R)^{k+2\ell}}{(k+2\ell)!}$$

since $\lim_{w\to 0} w\zeta(1+w) = 1$.

Lemma 18.9. (Gallagher) *For any integer* $k \leq \sqrt{\log\log N}$ *we have*

$$\frac{1}{\binom{N}{k}} \sum_{\substack{\mathcal{H}\subset\{1,\dots,N\}\\ |\mathcal{H}|=k}} \mathfrak{S}(\mathcal{H}) = 1 + O\left(\frac{k^2}{\log\log N}\right).$$

Proof. Let $y = c\log N$ for some constant $c < 1$, and define

$$\mathfrak{S}_y(\mathcal{H}) = \prod_{p\leq y} \left(1-\frac{1}{p}\right)^{-k}\left(1-\frac{\nu_{\mathcal{H}}(p)}{p}\right),$$

so that

$$\mathfrak{S}(\mathcal{H})/\mathfrak{S}_y(\mathcal{H}) = \prod_{p>y} \left(1-\frac{1}{p}\right)^{-k}\left(1-\frac{k}{p}\right)\left(1+\frac{k-\nu_{\mathcal{H}}(p)}{p-k}\right).$$

Therefore $\log(\mathfrak{S}(\mathcal{H})/\mathfrak{S}_y(\mathcal{H}) \ll \sum_{p>y} k^2/p^2 + \sum_{i<j}\sum_{p|b_j-b_i, p>y} \frac{1}{p}$. Now $0 < |b_j - b_i| \leq N^2$ so that there are $\leq 2\log N/\log y$ such primes p for each pair i and j. Thus, in total, we have $\ll \frac{k^2}{y\log y} + k^2\frac{\log N}{y\log y} \ll \frac{k^2}{\log\log N}$.

Let $m = \prod_{p\leq y} p$ which is $= N^{c+o(1)}$ by the prime number theorem; the value of $\mathfrak{S}_y(\mathcal{H})$ depends only upon the value of the elements of $\mathcal{H} \pmod m$. Therefore

$$\begin{aligned} k! \sum_{\substack{\mathcal{H}\subset\{1,\dots,N\}\\ |\mathcal{H}|=k}} \mathfrak{S}_y(\mathcal{H}) &= \sum_{h_1,\dots,h_k\in\{1,\dots,N\}} \mathfrak{S}_y(\mathcal{H}) - \sum_{\substack{h_1,\dots,h_k\in\{1,\dots,N\}\\ h_i=h_j \text{ for some } i\neq j}} \mathfrak{S}_y(\mathcal{H}) \\ &= \left(\frac{N}{m}+O(1)\right)^k \sum_{h_1,\dots,h_k\in\{1,\dots,m\}} \mathfrak{S}_y(\mathcal{H}) + O(k^2N^{k-1}(2\log y)^k) \end{aligned}$$

since each $S_y(\mathcal{H}) \ll (2\log y)^k$. By the Chinese Remainder Theorem we have

$$\frac{1}{m^k}\sum_{h_1,\dots,h_k\in\{1,\dots,N\}} \mathfrak{S}_y(\mathcal{H}) = \prod_{p\leq y}\left\{\frac{1}{p^k}\sum_{h_1,\dots,h_k\in\{1,\dots,p\}} \left(1-\frac{1}{p}\right)^{-k}\left(1-\frac{\nu_{\mathcal{H}}(p)}{p}\right)\right\}.$$

Now (for the sum over all possible $h_1,\dots,h_k \in \{1,\dots,p\}$)

$$\sum_h (p-\nu_{\mathcal{H}}(p)) = \sum_h \sum_{\substack{a \pmod p\\ a\neq h_i \text{ for all } i}} = \sum_{a \pmod p} \sum_{\substack{h\\ h_i\neq a \text{ for all } i}} 1 = p(p-1)^k.$$

Combining the above estimates yields the result.

Exercises

18.3a. Prove Lemma 18.7.

18.3b. Show that the contributions of the integrands on the final contours in Lemma 18.8 are indeed acceptable.

18.3c. Do the "combining" at the end of the proof of Lemma 18.9

18.4. Proof of Proposition 18.4. We will write "$\approx$" to indicate an error term that will be considered a little later. In the first part we have

$$\begin{aligned}
\sum_{n=N+1}^{2N} \Lambda_R(n;\mathcal{H},\ell)^2 &= \frac{1}{(k+\ell)!^2} \sum_{d,e\le R} \mu(d)\mu(e) \left(\log \frac{R}{d}\right)^{k+\ell} \left(\log \frac{R}{e}\right)^{k+\ell} \sum_{\substack{N<n\le 2N \\ [d,e]|P_{\mathcal{H}}(n)}} 1 \\
&\approx \frac{1}{(k+\ell)!^2} \sum_{d,e\le R} \mu(d)\mu(e)\nu_{[d,e]}(\mathcal{H}) \frac{N}{[d,e]} \left(\log \frac{R}{d}\right)^{k+\ell} \left(\log \frac{R}{e}\right)^{k+\ell} \\
&= \frac{N}{(2\pi i)^2} \int_{(1)}\int_{(1)} \sum_{d,e\ge 1} \frac{\mu(d)}{d^{s_1}} \frac{\mu(e)}{e^{s_2}} \frac{\nu_{[d,e]}(\mathcal{H})}{[d,e]} \frac{R^{s_1}}{s_1^{k+\ell+1}} \frac{R^{s_2}}{s_2^{k+\ell+1}} ds_1 ds_2 \\
&= \frac{N}{(2\pi i)^2} \int_{(1)}\int_{(1)} \prod_p \left(1 - \frac{\nu_p(\mathcal{H})}{p}\left(\frac{1}{p^{s_1}} + \frac{1}{p^{s_2}} - \frac{1}{p^{s_1+s_2}}\right)\right) \frac{R^{s_1+s_2}}{(s_1 s_2)^{k+\ell+1}} ds_1 ds_2 \\
&= \frac{N}{(2\pi i)^2} \int_{(1)}\int_{(1)} G(s_1,s_2) \frac{\zeta(1+s_1+s_2)^k}{\zeta(1+s_1)^k \zeta(1+s_2)^k} \frac{R^{s_1+s_2}}{(s_1 s_2)^{k+\ell+1}} ds_1 ds_2
\end{aligned}$$

where

$$G(s_1,s_2) = \prod_p \frac{\left(1 - \frac{\nu_p(\mathcal{H})}{p}\left(\frac{1}{p^{s_1}} + \frac{1}{p^{s_2}} - \frac{1}{p^{s_1+s_2}}\right)\right)\left(1 - \frac{1}{p^{1+s_1+s_2}}\right)^k}{\left(1 - \frac{1}{p^{1+s_1}}\right)^k \left(1 - \frac{1}{p^{1+s_2}}\right)^k},$$

which is absolutely convergent when $\mathrm{Re}(s_1)$, $\mathrm{Re}(s_2) > -1/2$. The result follows from lemma 18.8 since $G(0,0) = \mathfrak{S}(\mathcal{H})$, after we have justified the "$\approx$": By lemma 18.6(a), the error term here is

$$\begin{aligned}
&\ll \frac{(\log R)^{2(k+\ell)}}{(k+\ell)!^2} \sum_{d,e\le R} \mu(d)^2\mu(e)^2 \nu_{[d,e]}(\mathcal{H}) \le \frac{(\log R)^{2(k+\ell)}}{(k+\ell)!^2} \sum_{d,e\le R} \mu(d)^2\mu(e)^2 \nu_{[d,e]}(\mathcal{H}) \frac{R}{d}\frac{R}{e} \\
&\le R^2 \frac{(\log R)^{2(k+\ell)}}{(k+\ell)!^2} \prod_{p\le R} \left(1 + \nu_p(\mathcal{H})\left(\frac{2}{p} + \frac{1}{p^2}\right)\right) \ll_k R^2 (\log R)^{2(k+\ell)} \prod_{p\le R}\left(1 + \frac{2k}{p}\right) \\
&\ll R^2 (\log R)^{4k+2\ell}
\end{aligned}$$

and this is negligible provided $R \le N^{1/2}/(\log N)^{2k}$.

For the second part we have, proceeding analogously to the above,

$$\sum_{n=N+1}^{2N} \Lambda_R(n;\mathcal{H},\ell)^2 \vartheta(n+h_0)$$

$$= \frac{1}{(k+\ell)!^2} \sum_{d,e\le R} \mu(d)\mu(e) \left(\log \frac{R}{d}\right)^{k+\ell} \left(\log \frac{R}{e}\right)^{k+\ell} \sum_{\substack{N<n\le 2N \\ [d,e] \mid P_{\mathcal{H}}(n)}} \theta(n+h_0)$$

$$\approx \frac{1}{(k+\ell)!^2} \sum_{d,e\le R} \mu(d)\mu(e)\nu^*_{[d,e]}(\mathcal{H}) \frac{N}{\phi([d,e])} \left(\log \frac{R}{d}\right)^{k+\ell} \left(\log \frac{R}{e}\right)^{k+\ell}$$

$$= \frac{N}{(2\pi i)^2} \int_{(1)}\int_{(1)} \prod_p \left(1 - \frac{(\nu_p(\mathcal{H})-1)}{(p-1)} \left(\frac{1}{p^{s_1}} + \frac{1}{p^{s_2}} - \frac{1}{p^{s_1+s_2}}\right)\right) \frac{R^{s_1+s_2}}{(s_1 s_2)^{k+\ell+1}} ds_1 ds_2$$

$$= \frac{N}{(2\pi i)^2} \int_{(1)}\int_{(1)} G^*(s_1,s_2) \frac{\zeta(1+s_1+s_2)^{k-1}}{\zeta(1+s_1)^{k-1}\zeta(1+s_2)^{k-1}} \frac{R^{s_1+s_2}}{(s_1 s_2)^{k+\ell+1}} ds_1 ds_2$$

where

$$G^*(s_1,s_2) = \prod_p \frac{\left(1 - \frac{(\nu_p(\mathcal{H})-1)}{(p-1)} \left(\frac{1}{p^{s_1}} + \frac{1}{p^{s_2}} - \frac{1}{p^{s_1+s_2}}\right)\right) \left(1 - \frac{1}{p^{1+s_1+s_2}}\right)^{k-1}}{\left(1 - \frac{1}{p^{1+s_1}}\right)^{k-1} \left(1 - \frac{1}{p^{1+s_2}}\right)^{k-1}},$$

which is absolutely convergent when $\mathrm{Re}(s_1)$, $\mathrm{Re}(s_2) > -1/2$. The result follows from lemma 18.8 (with $k-1$ in place of k, and $\ell+1$ in place of ℓ) since $G^*(0,0) = \mathfrak{S}(\mathcal{H})$, after we have justified the "$\approx$": By lemma 18.6(b), the error term here is

$$\le \frac{(\log R)^{2(k+\ell)}}{(k+\ell)!^2} \sum_{d,e\le R} \mu(d)^2\mu(e)^2\nu^*_{[d,e]}(\mathcal{H}) \max_{a:(a,[d,e])=1} \left|\theta(N;[d,e],a) - \frac{N}{\phi([d,e])}\right|$$

$$\le \frac{(\log R)^{2(k+\ell)}}{(k+\ell)!^2} \sum_{m\le R^2} \mu(m)^2\tau(m)\nu^*_m(\mathcal{H}) \max_{a:(a,m)=1} \left|\theta(N;m,a) - \frac{N}{\phi(m)}\right|$$

The square of this sum is, using Cauchy-Schwartz and the trivial upper bound $\theta(N;m,a) \ll (N/m)\log N$,

$$\le \sum_{m\le R^2} \mu(m)^2 \frac{\tau(m)^2\nu^*_m(\mathcal{H})^2}{m} \cdot \sum_{m\le R^2} \max_{a:(a,m)=1} m \left|\theta(N;m,a) - \frac{N}{\phi(m)}\right|^2$$

$$\ll \prod_{p\le R^2} \left(1 + \frac{\tau(p)^2\nu^*_p(\mathcal{H})^2}{p}\right) \cdot \sum_{m\le R^2} \max_{a:(a,m)=1} \left|\theta(N;m,a) - \frac{N}{\phi(m)}\right| N\log N$$

$$\ll (\log R)^{4(k-1)^2} N\log N \sum_{m\le R^2} \max_{a:(a,m)=1} \left|\theta(N;m,a) - \frac{N}{\phi(m)}\right|.$$

Inserting this above we get that our error term is negligible provided

$$\sum_{m \le R^2} \max_{a:(a,m)=1} \left| \theta(N; m, a) - \frac{N}{\phi(m)} \right| \ll \frac{N}{(\log N)^A}$$

for some $A > 4k^2 - 6k + 3$. This holds by the Bombieri–Vinogradov theorem for $R \le N^{1/4}/(\log N)^{B(k)}$; and by the Elliott-Halberstam conjecture for $R \le N^{1/2-\epsilon}$.