

Abstract

We show that

$$\sum_{\substack{p_n \leq x \\ p_{n+1} - p_n \geq \sqrt{p_n}}} (p_{n+1} - p_n) \ll_{\varepsilon} x^{3/5+\varepsilon}$$

or any fixed $\varepsilon > 0$. This improves a result of Matomäki, in which the exponent was $2/3$

The Differences Between Consecutive Primes. V

D.R. Heath-Brown
Mathematical Institute, Oxford

October 3, 2019

1 Introduction

In this paper we shall continue our investigations into the sum

$$\sum_{\substack{p_n \leq x \\ p_{n+1} - p_n \geq \sqrt{p_n}}} (p_{n+1} - p_n). \quad (1)$$

It was shown by Wolke [17] that the sum is $O(x^{1-\delta})$ (with $\delta = 1/30$), thereby answering a question of Erdős. The exponent was improved firstly to $5/6 + \varepsilon$ for any fixed $\varepsilon > 0$, and then to $3/4 + \varepsilon$, by the author [2], [3]. A further reduction, to $25/36 + \varepsilon$, was achieved by Peck [13], and the present record is held by Matomäki [10], with exponent $2/3$.

One would conjecture that the sum (1) only contains at most the terms with $p_n = 3, 7, 13, 23, 31$ and 113 , and hence is bounded. However we are far from proving this, even under the Riemann Hypothesis. The latter assumption allows for an estimate $O(x^{1/2}(\log x)^2)$, as was proved by Selberg [14]. The Lindelöf Hypothesis similarly implies that the sum is $O_\varepsilon(x^{1/2+\varepsilon})$, for any positive ε , as was shown in the fourth paper of this series, [6].

Our goal is to improve the unconditional estimates as follows.

Theorem 1 *For any fixed $\varepsilon > 0$ we have*

$$\sum_{\substack{p_n \leq x \\ p_{n+1} - p_n \geq \sqrt{p_n}}} (p_{n+1} - p_n) \ll_\varepsilon x^{3/5+\varepsilon}.$$

We should view the exponents $2/3$ and $3/5$ as being

$$1/2 + 1/6 \quad \text{and} \quad 1/2 + 1/10$$

respectively, so that we have reduced the excess over $1/2$ by 40%, from $1/6$ down to $1/10$. (For comparison, Heath-Brown [2] gives roughly a 29% improvement over Wolke [17]; Heath-Brown [3] sharpens [2] by 25%; Peck [13] improves on [3] by about 22%; and Matomäki reduces the excess in Peck's exponent by some 14%.)

In fact we prove a stronger result than Theorem 1.

Theorem 2 *For any fixed $\varepsilon > 0$ the measure of the set of $y \in [0, x]$ such that*

$$\max_{0 \leq h \leq \sqrt{y}} \left| \pi(y+h) - \pi(y) - \int_y^{y+h} \frac{dt}{\log t} \right| \geq \frac{\sqrt{y}}{(\log y)(\log \log y)}$$

is $O_\varepsilon(x^{3/5+\varepsilon})$.

If $p_n \leq y \leq p_{n+1} - \frac{1}{2}\sqrt{p_{n+1}}$ then $\pi(y + \frac{1}{2}\sqrt{y}) - \pi(y) = 0$. Thus each $p_n \leq x$ for which $p_{n+1} - p_n \geq \sqrt{p_n}$ contributes an interval whose length is at least $\frac{1}{3}(p_{n+1} - p_n)$, say, to the set in Theorem 2, provided that p_n is large enough. Thus Theorem 1 is a corollary of Theorem 2.

We remark that the analysis in this paper would be very considerably simplified if our question had been about gaps $p_{n+1} - p_n$ of size at least $p_n^{1/2+\varepsilon}$, rather than $p_n^{1/2}$. There would be no “bad” ranges to be handled by sieve upper bounds, and one could merely have used the generalized Vaughan identity, rather than the Buchstab formula. The situation is analogous to that in the author's papers [4] and [5], proving $\pi(x+y) - \pi(x) \sim y/\log x$ for Huxley's range $x^{7/12+\varepsilon} \leq y \leq x$, and for $x^{7/12} \leq y \leq x$ respectively.

With the exception of Matomäki's work, all previous results on the sum (1) could have been adapted to prove a corresponding version of Theorem 2. Matomäki uses sieve methods in an essential way, so that her method shows the sparsity of values y where $\pi(y + \sqrt{y}) - \pi(y) \leq cy/\log y$ for some small positive constant c . In contrast, our approach only uses sieve methods to handle relatively minor contributions to $\pi(y + \sqrt{y}) - \pi(y)$. Theorem 1 could undoubtedly be improved further by deploying sieve methods in the same way that Matomäki does. We have decided against doing this largely from laziness, but partly so as to demonstrate more clearly the power of our primary new tool, Proposition 1, described below.

Our approach to the sum (1) uses the standard mean and large values estimates for Dirichlet polynomials, which arise naturally in this context when one applies a sieve decomposition to the problem. The Dirichlet polynomials one encounters are typically products of shorter polynomials of the form $\sum_{N < n \leq 2N} p^{-s}$, the sum being over primes.

The mean value theorem for Dirichlet polynomials (Montgomery [12, Theorem 6.1]) shows that

$$\int_0^T \left| \sum_{m \leq M} a_m m^{-it} \right|^2 dt \ll (T + M) \sum_{m \leq M} |a_m|^2. \quad (2)$$

This is quite efficient when $M \ll T$ and the coefficients a_m are fairly even in size. Our main new tool is the following quite different mean value estimate, which remains useful for certain longer Dirichlet polynomials.

Proposition 1 *Let $T \geq 1$ and let \mathcal{M} be a set of distinct integers m in $(0, T]$, of cardinality $\#\mathcal{M} \leq R$, with associated complex coefficients ζ_m of modulus 1. Suppose we are given a positive integer N and complex coefficients q_1, \dots, q_N . Then we have*

$$\begin{aligned} & \int_0^T \left| \sum_{m \in \mathcal{M}} \zeta_m m^{-it} \right|^2 \left| \sum_{n \leq N} q_n n^{-it} \right|^2 dt \\ & \ll_{\varepsilon} \left(N^2 R^2 + (NT)^{\varepsilon} \{NRT + NR^{7/4} T^{3/4}\} \right) \max_n |q_n|^2 \end{aligned}$$

for any fixed $\varepsilon > 0$.

For the proof we refer the reader to the author's paper [7]. To see the strength of this result we observe that the term $N^2 R^2$ corresponds to the (square of the) maximum value that the product of our two Dirichlet polynomials could attain, while the term NRT is what one would get if one had square root cancellation throughout the range $[0, T]$. Thus the bound is sub-optimal largely because of the term $NR^{7/4} T^{3/4}$. When $R \leq T^{1/3}$ one has $NR^{7/4} T^{3/4} \leq NRT$, so that our result is essentially best possible in this situation. Indeed, this is what happens in the critical case for our theorems, in which it transpires that $T = x^{1/2+o(1)}$ and $R = x^{1/10+o(1)}$.

Estimates of the type in Proposition 1 originate with the work of Yu [18], who gave a bound $O_{\varepsilon}((N^2 R^2 + NRT)(NT)^{\varepsilon})$ subject to the Lindelöf hypothesis, and used it to show that

$$\sum_{p_n \leq x} (p_{n+1} - p_n)^2 \ll_{\varepsilon} x^{1+\varepsilon} \quad (3)$$

under the same assumption. Since we only obtain an optimal estimate in Proposition 1 when $R \leq T^{1/3}$ it turns out that we are unable to say anything useful about gaps $p_{n+1} - p_n$ shorter than $p_n^{1/3}$. This does not preclude a new unconditional result for the sum in (3), but in the present paper we will restrict our attention to gaps of size $p_n^{1/2}$ or more.

Acknowledgements. This work was partly supported by EPSRC grant number EP/K021132X/1. For a further part of the preparation of this paper the author was supported by the NSF under Grant No. DMS-1440140, while in residence at the *Mathematical Sciences Research Institute* in Berkeley, California, during the Spring 2017 semester.

It is a pleasure also to acknowledge the work of the anonymous referees, whose efforts have removed various misprints and improved the exposition of this paper.

2 Structure of the Proof, and the Choice of Parameters

Theorem 2 compares the number of primes in an interval $(y, y + h]$ with its expected main term, for variable h . Our argument begins with some preliminary steps to replace the interval $(y, y + h]$ by $(y, y + y\delta_0]$ for a suitably small $\delta_0 > 0$ independent of y . Rather than compare the number of primes in $(y, y + y\delta_0]$ with its expected main term, we find it convenient to compare with the number of primes in a longer interval $(y, y + y\delta_1]$.

The argument then goes on to apply sieve methods to both $(y, y + y\delta_0]$ and $(y, y + y\delta_1]$. This has two effects. Firstly it enables us to remove certain short ranges of variables that would otherwise be awkward to handle. Secondly it allows us to translate the problem into one involving products of Dirichlet polynomials. At this point we use the key idea from Yu [18], coding the points y for which $(y, y + y\delta_0]$ does not have the expected number of primes into a Dirichlet polynomial. This process produces Proposition 2 in Section 6, which is a major waypoint in our argument.

Proposition 2 requires us to estimate the mean value of a product of Dirichlet polynomials, and we do this using a variety of well established techniques, in combination with Proposition 1. In particular we use Vinogradov’s zero-free region, various forms of the “Large Values” estimate for Dirichlet polynomials, and the classical mean value estimate (2). This stage of the argument requires us to examine several separate cases. For these estimates to produce a suitable saving it is crucial that certain critical ranges for the lengths of the Dirichlet polynomials are avoided, and these are the ranges that the initial sieve argument eliminates.

The ranges we will avoid take the shape $[x^{1/\ell-\eta}, x^{1/\ell+\eta}]$ for certain positive integers ℓ , and by removing these we are able to make savings of factors of order $x^{c\eta}$ for certain constants $c > 0$. Here $\eta = \eta(x)$ is a small function of x which we will specify in a moment.

At other points in the argument the saving we obtain is related to the available zero-free region for the Riemann zeta-function, which allows us to improve on the trivial bound by factors of the type $\exp((\log x)^\theta)$ for certain constants $\theta \in (0, 1)$. With this in mind we define

$$\mathbf{S} = \mathbf{S}(x) = \exp((\log \log x)^{11}) \quad (4)$$

with a view to saving at least a positive power of \mathbf{S} in the various key arguments. This means, conversely, that we can afford to lose factors $\mathbf{S}^{o(1)}$, since they will be more than compensated for by the gain of a power of \mathbf{S} .

We set

$$\nu = \nu(x) = (\log \log x)^5, \quad (5)$$

$$z_1 = z_1(x) = (4x)^{1/\nu} = x^{o(1)}, \quad (6)$$

$$\eta = \eta(x) = (\log \log x)^{-12000}, \quad (7)$$

$$\varpi = \varpi(x) = (\log \log x)^2, \quad (8)$$

and

$$z_2 = z_2(x) = z_1^\varpi = x^{o(1)}. \quad (9)$$

Then

$$\log x \ll \mathbf{S}^{o(1)}, \quad \text{and} \quad \nu^\nu \ll \mathbf{S}^{o(1)}, \quad (10)$$

while

$$\mathbf{S} \ll \exp((\log x)^\theta) \ll x^{o(\eta)} \ll z_1 \quad \text{for any } \theta \in (0, 1). \quad (11)$$

We shall use all these bounds repeatedly in our argument.

For the entirety of the paper we will assume without further comment that x is sufficiently large.

3 Preliminary Steps

To prove Theorem 2 it suffices to establish the corresponding bound when y varies over a dyadic range $(x, 2x]$. Theorem 2 requires us to estimate the measure of a set of real numbers y , defined using the maximum over intervals $(y, y+h]$ for varying h . We begin by showing how to replace the real variable y by an integer variable m , and how to use an interval whose length is a fixed fraction δ_0 of its left-hand endpoint. It will be convenient to write $\mathcal{I}(x)$ for the set of $y \in (x, 2x]$ such that

$$\max_{0 \leq h \leq \sqrt{y}} \left| \pi(y+h) - \pi(y) - \int_y^{y+h} \frac{dt}{\log t} \right| \geq \frac{\sqrt{y}}{(\log y)(\log \log y)},$$

so that our goal is to estimate $\text{Meas}(\mathcal{I}(x))$.

Lemma 1 *Let*

$$\delta_0 = x^{-1/2}(\log \log x)^{-2} \quad (12)$$

and

$$H = x^{1/2}(\log \log x)^{-4}. \quad (13)$$

Then there is a set of R_0 distinct integers $m_1, \dots, m_{R_0} \in [x/H, 3x/H]$ for which

$$\left| \pi(mH(1 + \delta_0)) - \pi(mH) - \int_{mH}^{mH(1 + \delta_0)} \frac{dt}{\log t} \right| \geq \frac{\delta_0 x}{3(\log x)(\log \log x)},$$

and such that

$$\text{Meas}(\mathcal{I}(x)) \ll x^{1/2} R_0.$$

If $0 \leq h \leq \sqrt{y}$ then the interval $(y, y + h]$ is a union of at most $\delta_0^{-1} y^{-1/2}$ disjoint subintervals $(y_1, y_1(1 + \delta_0)]$, together with a shorter interval at the end, of length $O(\delta_0 x)$. By the Brun–Titchmarsh theorem this last interval contains $O(\delta_0 x / \log x)$ primes. It follows that

$$\begin{aligned} & \left| \pi(y + h) - \pi(y) - \int_y^{y+h} \frac{dt}{\log t} \right| \\ & \leq \delta_0^{-1} y^{-1/2} \left| \pi(y_1(1 + \delta_0)) - \pi(y_1) - \int_{y_1}^{y_1(1 + \delta_0)} \frac{dt}{\log t} \right| + O(\delta_0 x / \log x) \end{aligned}$$

for some y_1 with $y \leq y_1 \leq y + \sqrt{y}$. Thus if $x < y \leq 2x$ and

$$\left| \pi(y + h) - \pi(y) - \int_y^{y+h} \frac{dt}{\log t} \right| \geq \frac{\sqrt{y}}{(\log y)(\log \log y)} \quad (14)$$

then there is some y_1 as above with

$$\begin{aligned} \left| \pi(y_1(1 + \delta_0)) - \pi(y_1) - \int_{y_1}^{y_1(1 + \delta_0)} \frac{dt}{\log t} \right| & \geq \frac{\delta_0 y}{2(\log y)(\log \log y)} \\ & \geq \frac{\delta_0 x}{2(\log x)(\log \log x)}. \end{aligned}$$

For each value y satisfying (14) we choose the smallest such y_1 and define $m = 1 + [y_1/H]$, so that $x/H < m \leq 3x/H$. In this way we produce a collection of distinct integers m_1, \dots, m_{R_0} . Each such integer m_i may arise from a range of values for y , satisfying $y = m_i H + O(\sqrt{x})$. It therefore follows that the measure we have to estimate in Lemma 1 is $O(x^{1/2} R_0)$,

as required. Moreover, if $m = 1 + [y_1/H]$ then $\pi(y_1(1 + \delta_0))$ differs from $\pi(mH(1 + \delta_0))$ by $O(H/\log x)$, by the Brun–Titchmarsh Theorem. Similarly $\pi(y_1) = \pi(mH) + O(H/\log x)$. On the other hand,

$$\int_{y_1}^{y_1(1+\delta_0)} \frac{dt}{\log t}$$

differs from the corresponding integral between mH and $mH(1 + \delta_0)$ by $O(H/\log x)$. Since $H/\log x = o(\delta_0 x/(\log x)(\log \log x))$ we therefore have

$$\left| \pi(mH(1 + \delta_0)) - \pi(mH) - \int_{mH}^{mH(1+\delta_0)} \frac{dt}{\log t} \right| \geq \frac{\delta_0 x}{3(\log x)(\log \log x)},$$

for large enough x , and the lemma follows.

For each $m \in [x/H, 3x/H]$ we will locate the primes in the interval

$$\mathcal{A} := \mathcal{A}(m) = \{n \in \mathbb{Z} : mH < n \leq mH(1 + \delta_0)\}$$

by sieving. Rather than compare the number of primes with the integral

$$\int_{mH}^{mH(1+\delta_0)} \frac{dt}{\log t}$$

it will be more convenient to work with the number of primes in a long interval

$$\mathcal{B} := \mathcal{B}(m) = \{n \in \mathbb{Z} : mH < n \leq mH(1 + \delta_1)\},$$

where

$$\delta_1 := \exp\{-(\log x)^{1/2}\}. \quad (15)$$

We write $\mathcal{A}^{(k)} = \mathcal{A}^{(k)}(m)$ for the weighted set in which $n \in \mathcal{A}(m)$ has weight $\tau_k(n)$, and similarly for $\mathcal{B}^{(k)}$. In what follows we will re-number the integers m_i in Lemma 1 as necessary. We proceed to show the following.

Lemma 2 *There are integers $k = 1, 2$ or 3 and $m_1, \dots, m_{R_1} \in [x/H, 3x/H]$ with $R_1 \geq R_0/3$, satisfying*

$$|S(\mathcal{A}^{(k)}(m_i), (4x)^{1/4}) - \delta_0 \delta_1^{-1} S(\mathcal{B}^{(k)}(m_i), (4x)^{1/4})| \geq \frac{\delta_0 x}{25(\log x)(\log \log x)}.$$

We begin the proof by using the prime number theorem with Vinogradov's error term, whence

$$\begin{aligned} \pi(mH(1 + \delta_1)) - \pi(mH) &= \int_{mH}^{mH(1+\delta_1)} \frac{dt}{\log t} + O(\delta_1 x (\log x)^{-2}) \\ &= \delta_1 \delta_0^{-1} \int_{mH}^{mH(1+\delta_0)} \frac{dt}{\log t} + O(\delta_1 x (\log x)^{-2}), \end{aligned}$$

so that if x is large enough we have

$$\begin{aligned} & \left| \pi(mH(1 + \delta_0)) - \pi(mH) - \delta_0 \delta_1^{-1} \{ \pi(mH(1 + \delta_1)) - \pi(mH) \} \right| \\ & \geq \frac{\delta_0 x}{4(\log x)(\log \log x)} \end{aligned}$$

for $m = m_1, \dots, m_{R_0}$.

We now consider the weighted set \mathcal{A}_* , consisting of elements $n \in \mathcal{A}$ weighted by

$$3 - \frac{3}{2}\tau(n) + \frac{1}{3}\tau_3(n).$$

One may check that this takes the value 1 when n is prime, and that it vanishes for square-free n having 2 or 3 prime factors. It follows that $S(\mathcal{A}_*, (4x)^{1/4})$ differs from

$$\pi(mH(1 + \delta_0)) - \pi(mH)$$

by the contribution from integers which have a factor $p^2 \in [(4x)^{1/2}, 4x]$. When $(4x)^{1/4} \leq p \leq x^{1/3}$ there are $O(1 + x\delta_0 p^{-2})$ multiples of p^2 , producing total contribution $O(x^{1/3})$. On the other hand, if $p^2 t \in \mathcal{A}$ with $p \geq x^{1/3}$, then $t \leq 4x^{1/3}$. For each such t the prime p is restricted to an interval $(\sqrt{mH/t}, \sqrt{(mH(1 + \delta_0))/t}]$ of length $O(1)$, so that the total contribution in this case is also $O(x^{1/3})$. It follows that

$$S(\mathcal{A}_*, (4x)^{1/4}) = \pi(mH(1 + \delta_0)) - \pi(mH) + o\left(\frac{\delta_0 x}{(\log x)(\log \log x)}\right).$$

We define \mathcal{B}_* analogously, and find this time that the error is

$$\ll \sum_{(4x)^{1/4} \leq p \leq (4x)^{1/2}} (1 + x\delta_1 p^{-2}) \ll x^{3/4} = o\left(\frac{\delta_1 x}{(\log x)(\log \log x)}\right).$$

We may then deduce that

$$\left| S(\mathcal{A}_*, (4x)^{1/4}) - \delta_0 \delta_1^{-1} S(\mathcal{B}_*, (4x)^{1/4}) \right| \geq \frac{\delta_0 x}{5(\log x)(\log \log x)},$$

for large enough x . Lemma 2 then follows.

The integer k appearing in Lemma 2 will be fixed for the rest of the proof.

4 The First Sieve Stage

In this section we introduce our first sieve process, and show that terms in which certain variables lie in awkward ranges make a negligible contribution to $S(\mathcal{A}^{(k)}(m_i), (4x)^{1/4})$ and $S(\mathcal{B}^{(k)}(m_i), (4x)^{1/4})$.

We begin by noting that $S(\mathcal{A}^{(k)}(m), (4x)^{1/4})$ counts products $n_1 \dots n_k$ in $\mathcal{A}(m)$ for which each factor n_i has no prime divisor below $(4x)^{1/4}$. We will use the parameters ν and z_1 given by (5) and (6). We now define

$$\Pi_1 := \prod_{p < z_1} p, \quad \text{and} \quad \Pi_2 = \prod_{z_1 \leq p < (4x)^{1/4}} p,$$

so that

$$S(\mathcal{A}^{(k)}(m), (4x)^{1/4}) = \sum_{q_1, \dots, q_k | \Pi_2} \mu(q_1) \dots \mu(q_k) N_k(\mathcal{A}, q_1 \dots q_k),$$

where

$$N_k(\mathcal{A}, q) := \#\{h_1, \dots, h_k : qh_1 \dots h_k \in \mathcal{A}(m), (h_1 \dots h_k, \Pi_1) = 1\}, \quad (16)$$

and similarly for \mathcal{B} . Each q_i is composed of various prime factors p , which belong to dyadic intervals of the type $(2^s, 2^{s+1}]$. Similarly, we can decompose the range for the variables h_i into dyadic intervals.

We proceed to show that there is a negligible contribution from terms with a divisor close to a reciprocal power $x^{1/\ell}$, say.

Lemma 3 *Let η be given by (7) and let $\mathcal{A}_\dagger(m)$ be the set of integers in $\mathcal{A}(m)$ having a divisor in the range $[x^{1/\ell-2\eta}, x^{1/\ell+2\eta}]$ for some integer $\ell \in [4, \nu + 2]$. Then the number of $2k$ -tuples $(q_1, \dots, q_k, h_1, \dots, h_k)$ with $q_1 \dots q_k h_1 \dots h_k$ in $\mathcal{A}_\dagger(m)$, and such that $q_1, \dots, q_k \mid \Pi_2$ and $(h_1 \dots h_k, \Pi_1) = 1$ is*

$$\ll \delta_0 x (\log x)^{-1} (\log \log x)^{-3/2}.$$

If we define \mathcal{B}_\dagger similarly then we get an analogous bound with δ_1 in place of δ_0 .

If the prime divisors p_i of $q_1 \dots q_k$ lie in dyadic intervals I_1, \dots, I_t , say, and the h_i lie in dyadic intervals J_1, \dots, J_k , then $t + k \leq k\nu + k$. The lemma then shows that there is a negligible contribution from those collections of intervals for which any product from $I_1, \dots, I_t, J_1, \dots, J_k$ lies in $[x^{1/\ell-\eta}, x^{1/\ell+\eta}]$. Here we use the fact that $2^{k\nu+k} \leq x^\eta$ for large x , by (5) and (7). In particular the lemma shows that we can reduce the sieving range, restricting our attention to divisors q of Π_2 whose prime factors satisfy $p < x^{1/4-\eta}$.

For the proof of the lemma we begin by observing that an integer $n \in \mathcal{A}$ arises in at most $\tau_6(n) \leq \tau(n)^5$ ways as $n = h_1 q_1 \dots h_k q_k$, so that for each ℓ the contribution we have to consider is

$$\ll \sum_{x^{1/\ell-2\eta} \leq d \leq x^{1/\ell+2\eta}} \sum_{\substack{n \in \mathcal{A} \\ (n, \Pi_1)=1, d|n}} \tau(n)^5.$$

By Cauchy's inequality this is at most $\Sigma_1^{1/2} \Sigma_2^{1/2}$, with

$$\Sigma_1 = \sum_d \sum_{\substack{n \in \mathcal{A} \\ (n, \Pi_1)=1, d|n}} \tau(n)^{10} \leq \sum_{n \in \mathcal{A}: (n, \Pi_1)=1} \tau(n)^{11}$$

and

$$\Sigma_2 = \sum_{x^{1/\ell-2\eta} \leq d \leq x^{1/\ell+2\eta}} \sum_{\substack{n \in \mathcal{A} \\ (n, \Pi_1)=1, d|n}} 1.$$

We now apply the following lemma, which is an immediate corollary of the theorem of Shiu [15].

Lemma 4 *Suppose that $X, Y, z \geq 2$ are real numbers such that $X^c \leq Y \leq X$, for some constant $c > 0$, and let N be a positive integer. Then*

$$\sum_{\substack{X < n \leq X+Y \\ p|n \Rightarrow p \geq z}} \tau(n)^N \ll_{c,N} \frac{Y}{\log X} \left(\frac{\log X}{\log z} \right)^{2^N}.$$

This produces

$$\Sigma_1 \ll \frac{\delta_0 x}{\log x} \nu^{2^{11}}.$$

Moreover, since $d \leq x^{1/\ell+2\eta} \leq x^{1/3}$, a simple sieve upper bound yields

$$\sum_{\substack{n \in \mathcal{A} \\ (n, \Pi_1)=1, d|n}} 1 \ll \frac{\delta_0 x}{d \log z_1},$$

and

$$\sum_{\substack{x^{1/\ell-2\eta} \leq d \leq x^{1/\ell+2\eta} \\ (d, \Pi_1)=1}} d^{-1} \ll \frac{\eta \log x}{\log z_1},$$

whence

$$\Sigma_2 \ll \frac{\delta_0 x \eta (\log x)}{(\log z_1)^2}.$$

We therefore conclude that each value for ℓ contributes

$$\ll (\Sigma_1 \Sigma_2)^{1/2} \ll \frac{\delta_0 x}{\log z_1} \nu^{2^{10}} \eta^{1/2}.$$

Since $\ell \ll \nu$ the total is

$$\ll \delta_0 x (\log x)^{-1} \nu^{1026} \eta^{1/2} \ll \delta_0 x (\log x)^{-1} (\log \log x)^{-3/2},$$

in view of (5) and (7). This proves Lemma 3 for \mathcal{A} , and the treatment of \mathcal{B} is similar.

Other terms we wish to remove are those in which $q_1 \dots q_k$ has two or more prime factors p_1, p_2 lying in the same dyadic interval $(2^s, 2^{s+1}]$.

Lemma 5 *We have*

$$\sum_{\substack{z_1 \leq p_1, p_2 < x^{1/4-\eta} \\ p_1/2 \leq p_2 \leq p_1}} \sum_{\substack{q_1, \dots, q_k | \Pi_2 \\ p_1 p_2 | q_1 \dots q_k}} N_k(\mathcal{A}, q_1 \dots q_k) \ll \delta_0 x (\log x)^{-1} (\log \log x)^{-3/2},$$

and

$$\sum_{\substack{z_1 \leq p_1, p_2 < x^{1/4-\eta} \\ p_1/2 \leq p_2 \leq p_1}} \sum_{\substack{q_1, \dots, q_k | \Pi_2 \\ p_1 p_2 | q_1 \dots q_k}} N_k(\mathcal{B}, q_1 \dots q_k) \ll \delta_1 x (\log x)^{-1} (\log \log x)^{-3/2}.$$

Arguing as for Lemma 3 we see that the total contribution from the first sum is

$$\sum_{\substack{z_1 \leq p_1, p_2 < x^{1/4-\eta} \\ p_1/2 \leq p_2 \leq p_1}} \sum_{\substack{n \in \mathcal{A} \\ (n, \Pi_1)=1, p_1 p_2 | n}} \tau_6(n).$$

This time Cauchy's inequality gives a bound $\ll (\Sigma_1 \Sigma_3)^{1/2}$, with Σ_1 as before, and

$$\Sigma_3 = \sum_{\substack{z_1 \leq p_1, p_2 < x^{1/4-\eta} \\ p_1/2 \leq p_2 \leq p_1}} \sum_{\substack{n \in \mathcal{A} \\ (n, \Pi_1)=1, p_1 p_2 | n}} 1.$$

A simple sieve upper bound shows that the inner sum above is

$$\ll \frac{\delta_0 x / p_1 p_2}{\log \min(z_1, \delta_0 x / p_1 p_2)} \ll \frac{\delta_0 x / p_1 p_2}{\log z_1} + \frac{\delta_0 x / p_1 p_2}{\log \delta_0 x / p_1 p_2},$$

whence

$$\begin{aligned} \Sigma_3 &\ll \sum_{p_1} \left\{ \frac{\delta_0 x}{p_1 (\log p_1) (\log z_1)} + \frac{\delta_0 x}{p_1 (\log p_1) (\log \delta_0 x / p_1^2)} \right\} \\ &\ll \frac{\delta_0 x}{(\log z_1)^2} + \frac{\delta_0 x \log \log x}{(\log x)^2} \\ &\ll \frac{\delta_0 x}{(\log z_1)^2}, \end{aligned}$$

by (5) and (6). It then follows that the total contribution from terms where two prime factors lie in the same dyadic interval is

$$\ll (\Sigma_1 \Sigma_3)^{1/2} \ll \left(\frac{\delta_0 x}{\log x} \nu^{2^{11}} \right)^{1/2} \left(\frac{\delta_0 x}{(\log z_1)^2} \right)^{1/2} \ll \frac{\delta_0 x}{(\log x)^{3/2}} \nu^{2^{10}+1}.$$

By the choice of ν in (5) this will be suitably small, which completes the treatment of \mathcal{A} . The proof for \mathcal{B} is similar.

5 A Second Sieve Operation

The variables h in (16) are now constrained to lie in dyadic ranges which avoid the intervals $[x^{1/\ell-\eta}, x^{1/\ell+\eta}]$. We now write $\xi(h) = 1$ if $(h, \Pi_1) = 1$, and $\xi(h) = 0$ otherwise. This is satisfactory when $h \leq x^{1/4}$, but for larger h we need to pick out values satisfying $(h, \Pi_1) = 1$ by using a simple Fundamental Lemma sieve. For this we use the parameters ϖ and z_2 given by (8) and (9). We then define $\xi_0(h) = \xi(h)$ if $h < x^{1/4}$ and

$$\xi_0(h) = \sum_{\substack{d|(h, \Pi_1) \\ d < z_2}} \mu(d) \quad (17)$$

otherwise. Our immediate goal will then be the following result.

Lemma 6 *Let*

$$\Delta(n) = \Delta(n, m) = \chi_{\mathcal{A}(m)}(n) - \delta_0 \delta_1^{-1} \chi_{\mathcal{B}(m)}(n),$$

where $\chi_{\mathcal{A}(m)}$ is the characteristic function of $\mathcal{A}(m)$, for example. Then

$$\begin{aligned} & S(\mathcal{A}^{(k)}, (4x)^{1/4}) - \delta_0 \delta_1^{-1} S(\mathcal{B}^{(k)}, (4x)^{1/4}) \\ &= \sum_{q_1, \dots, q_k | \Pi_2} \mu(q_1) \dots \mu(q_k) \sum_{h_1, \dots, h_k} \xi_0(h_1) \dots \xi_0(h_k) \Delta(q_1 h_1 \dots q_k h_k) \\ & \quad + O(\delta_0 x (\log x)^{-1} (\log \log x)^{-3/2}), \end{aligned} \quad (18)$$

with suitably restricted sums over the q_i and h_i . Specifically the q_i are composed of prime factors which run over disjoint dyadic intervals, and the h_i also run over dyadic intervals; and no product from a subset of these dyadic intervals falls in any of the ranges $[x^{1/\ell-\eta}, x^{1/\ell+\eta}]$ for $4 \leq \ell \leq \nu + 2$.

The claim in the lemma is that replacing

$$\sum_{\substack{h_1, \dots, h_k \\ q h_1 \dots h_k \in \mathcal{A}}} \xi(h_1) \dots \xi(h_k)$$

by

$$\sum_{\substack{h_1, \dots, h_k \\ qh_1 \dots h_k \in \mathcal{A}}} \xi_0(h_1) \dots \xi_0(h_k)$$

produces a negligible error, and similarly for \mathcal{B} . The key result which facilitates this is the following, which is an immediate deduction from the author's work [5, Lemma 15].

Lemma 7 *We have*

$$|\xi(h) - \xi_0(h)| \leq \left| \sum_{\substack{d|(h, \Pi_1) \\ d \geq z_2}} \mu(d) \right| \leq \sum_{\substack{d|(h, \Pi_1) \\ z_2 \leq d < z_1 z_2}} 1 \leq \tau(h).$$

Using the lemma we then see that

$$|\xi(h_1) \dots \xi(h_k) - \xi_0(h_1) \dots \xi_0(h_k)| \ll \max_j \left(|\xi(h_j) - \xi_0(h_j)| \prod_{i; i \neq j} \tau(h_i) \right).$$

Then, writing $h = h_1 \dots h_k$, we have

$$\begin{aligned} |\xi(h_1) \dots \xi(h_k) - \xi_0(h_1) \dots \xi_0(h_k)| &\ll \tau(h)^2 \max_j |\xi(h_j) - \xi_0(h_j)| \\ &\ll \tau(h)^2 \max_j \sum_{\substack{d|(h_j, \Pi_1) \\ z_2 \leq d < z_1 z_2}} 1 \\ &\ll \tau(h)^2 \sum_{\substack{d|(h, \Pi_1) \\ z_2 \leq d < z_1 z_2}} 1. \end{aligned}$$

Since $\tau_3 \mu^2 * \tau_3 \tau^2 \leq \tau^5$ the error we have to control is then

$$\ll \sum_{q|\Pi_2} \tau_k(q) \sum_{h: qh \in \mathcal{A}} \tau_k(h) \tau(h)^2 \sum_{\substack{d|(h, \Pi_1) \\ z_2 \leq d < z_1 z_2}} 1 \ll \sum_{\substack{d|\Pi_1 \\ z_2 \leq d < z_1 z_2}} \sum_{\substack{n \in \mathcal{A} \\ d|n}} \tau(n)^5.$$

We have $z_1 z_2 \leq x^{1/4}$ say, by (6) and (9). Moreover, if $d \mid n$, then $\tau(n) \leq \tau(d) \tau(n/d)$, whence Lemma 4 with $z = 1$ yields

$$\sum_{\substack{n \in \mathcal{A} \\ d|n}} \tau(n)^5 \ll \tau(d)^5 \frac{\delta_0 x}{d} (\log x)^{31}.$$

It follows that the overall error on replacing $\xi(h)$ by $\xi_0(h)$ for $\mathcal{A}(m)$ is

$$\ll \delta_0 x (\log x)^{31} \sum_{\substack{d|\Pi_1 \\ z_2 \leq d < z_1 z_2}} \tau(d)^5/d.$$

We bound this last sum using Rankin's trick, as follows. For any $\theta > 0$ we have

$$\begin{aligned} \sum_{\substack{d|\Pi_1 \\ z_2 \leq d < z_1 z_2}} \tau(d)^5/d &\leq z_2^{-\theta} \sum_{\substack{d|\Pi_1 \\ z_2 \leq d < z_1 z_2}} \tau(d)^5 d^{-1+\theta} \\ &\leq z_2^{-\theta} \sum_{\substack{d=1 \\ d|\Pi_1}}^{\infty} \tau(d)^5 d^{-1+\theta} \\ &= z_2^{-\theta} \prod_{p < z_1} (1 + 32p^{-1+\theta}) \\ &\leq z_2^{-\theta} \exp \left\{ 32 \sum_{p < z_1} p^{-1+\theta} \right\}. \end{aligned}$$

We choose $\theta = 1/\log z_1$, so that

$$\sum_{p < z_1} p^{-1+\theta} \leq 3 \log \log z_1,$$

(for x large enough) and hence

$$\sum_{\substack{d|\Pi_1 \\ z_2 \leq d < z_1 z_2}} \tau(d)^5/d \ll z_2^{-1/\log z_1} (\log x)^{96} = e^{-\varpi} (\log x)^{96},$$

by (9). Thus the error induced by replacing $\xi(h)$ with $\xi_0(h)$ is $O(\delta_0 x / (\log x)^2)$, say, if ϖ satisfies (8). As in the previous section, although we have presented the argument as it applies to \mathcal{A} , it applies in the same way for \mathcal{B} , and the lemma then follows.

There is one final step that belongs in this section. Thus far our adjustments have affected \mathcal{A} and \mathcal{B} separately. However, if h_1 , say, is large we will show that the corresponding average of $\Delta(q_1 h_1 \dots q_k h_k)$ must be negligibly small. To be specific, we will show the following.

Lemma 8 *Let V be the largest power of 2 such that $V \leq x^{5/8}$. Then terms with $h_i > V$ contribute $O(\delta_0 x (\log x)^{-2})$ in (18).*

For the proof we consider the union of dyadic ranges for h_1 covering the interval (V, ∞) . If we set $f = (q_1 h_1 \dots q_k h_k)/h_1$ we will trivially have $\Delta(q_1 h_1 \dots q_k h_k) = 0$ unless $f \leq 8x^{3/8}$. In this latter case we find that the overall contribution is

$$\ll \sum_{f \leq 8x^{3/8}} \tau(f)^6 \left| \sum_{h > V} \xi_0(h) \Delta(fh) \right| \leq \sum_{f \leq 8x^{3/8}} \tau(f)^6 \sum_{d < z_2} \left| \sum_{g > V/d} \Delta(fdg) \right|.$$

We now recall that

$$\mathcal{A} = \mathbb{Z} \cap (mH, mH(1 + \delta_0)], \quad \text{and} \quad \mathcal{B} = \mathbb{Z} \cap (mH, mH(1 + \delta_1)],$$

with $\delta_0 \leq \delta_1$. Thus if $Vf > mH(1 + \delta_1)$ we have $\Delta(fdg) = 0$ for all $g > V/d$. On the other hand, if $Vf \leq mH$ then

$$\sum_{g > V/d} \Delta(fdg) = \left\{ \frac{mH\delta_0}{fd} + O(1) \right\} - \delta_0 \delta_1^{-1} \left\{ \frac{mH\delta_1}{fd} + O(1) \right\} = O(1).$$

In the remaining case $mH < Vf \leq mH(1 + \delta_1)$, and

$$\sum_{g > V/d} |\Delta(fdg)| \leq \left\{ \frac{mH\delta_0}{fd} + O(1) \right\} + \delta_0 \delta_1^{-1} \left\{ \frac{mH\delta_1}{fd} + O(1) \right\} \ll \delta_0 x / fd + 1.$$

We therefore see that the overall contribution from terms with $h_1 > V$ is

$$\ll \sum_{f \leq 8x^{3/8}} \sum_{d < z_2} \tau(f)^6 + \sum_{\substack{f \\ Vf \in \mathcal{B}}} \sum_{d < z_2} \tau(f)^6 \frac{\delta_0 x}{fd}.$$

The first sum is $\ll z_2 x^{3/8+o(1)} = x^{3/8+o(1)}$, by (9). In the second sum we have $f \gg x/V$ and

$$\sum_{\substack{f \\ Vf \in \mathcal{B}}} \tau(f)^6 \ll \delta_1 x V^{-1} (\log x)^{63}$$

by Lemma 4 with $z = 1$. It follows that our bound is

$$\ll x^{3/8+o(1)} + \delta_0 \delta_1 x (\log x)^{64} \ll \delta_0 x (\log x)^{-2},$$

say, by (15). This is satisfactory for the lemma.

6 Introducing Dirichlet Polynomials

We are now interested only in the case in which $q_1 \dots q_k$ is square-free. Suppose that

$$q_i = p_{i,1} \dots p_{i,t_i}, \quad (1 \leq i \leq k).$$

Each prime $p_{i,j}$ runs over a corresponding dyadic interval $I_{i,j}$, and the number of possible intervals is $O(\log x)$. Since $t_i \leq \nu$ for $i \leq k$ the total number of choices for these dyadic intervals is at most $(C \log x)^{3+3\nu}$ for some absolute constant C . In the same way the variables h_1, \dots, h_k belong to dyadic interval J_1, \dots, J_k , and there are at most $(C \log x)^3$ choices for these intervals. By (4) and (5) there are therefore $O(\mathbf{S}^{o(1)})$ choices for the entire collection of intervals. Since we have arranged that the intervals $I_{i,j}$ are distinct, each relevant product $q = q_1 \dots q_k$ arises exactly once as $p_{i,j}$ runs over $I_{i,j}$, and each q corresponds to $\tau_k(q)$ choices for the k -tuple q_1, \dots, q_k . We note that

$$\tau_k(q) \leq \tau_3(q) \leq 3^{\omega(q)} \leq 3^{3\nu} \ll \mathbf{S}^{o(1)}.$$

It will be convenient to re-label the intervals $I_{i,j}$ as I_1, \dots, I_t , where $t = t_1 + \dots + t_k$, and to replace I_j by $I_j \cap [z_1, (4x)^{1/4})$. Then

$$\begin{aligned} & \left| \sum_{q_1, \dots, q_k | \Pi_2} \mu(q_1) \dots \mu(q_k) \sum_{h_1, \dots, h_k} \xi_0(h_1) \dots \xi_0(h_k) \Delta(q_1 h_1 \dots q_k h_k) \right| \\ & \ll \mathbf{S}^{o(1)} \sum_{I_1, \dots, I_t} \sum_{J_1, \dots, J_k} \left| \sum_{\substack{p_i \in I_i, h_j \in J_j \\ 1 \leq i \leq t, 1 \leq j \leq k}} \xi_0(h_1) \dots \xi_0(h_k) \Delta(p_1 \dots p_t h_1 \dots h_k) \right|. \end{aligned}$$

We also observe that $\xi_0(h) = 0$ for $1 < h < z_1$ so that we may replace J_j by $J_j \cap [z_1, x^{5/8}]$. When $h_j = 1$ for some index j we may omit J_j altogether, reducing k by 1. As a result we may have to allow for the possibility that $k = 0$. Referring to Lemma 2 and (18) we see that

$$\delta_0 x \ll \mathbf{S}^{o(1)} \sum_{I_1, \dots, I_t} \sum_{J_1, \dots, J_k} \left| \sum_{\substack{p_j \in I_j \\ 1 \leq j \leq t}} \sum_{\substack{h_j \in J_j \\ 1 \leq j \leq k}} \xi_0(h_1) \dots \xi_0(h_k) \Delta(p_1 \dots p_t h_1 \dots h_k, m_i) \right|$$

for $i = 1, \dots, R_1$, where we now omit collections of dyadic intervals any subset of which contains a product from any of the ranges $[x^{1/\ell-\eta}, x^{1/\ell+\eta}]$ for $4 \leq \ell \leq \nu + 2$.

Thus there is a subset of the m_i , with cardinality at least $R_1 \mathbf{S}^{-1}$, on which the contribution from some specific set of intervals is large. We can therefore conclude as follows.

Lemma 9 *Suppose all parameters are as previously defined. Then there are:-*

(i) *Integers $k = 0, 1, 2$ or 3 and $t \in [0, 3\nu]$;*

(ii) *Disjoint intervals $I_j = (A_j, B_j] \subseteq [z_1, (4x)^{1/4}]$ for $1 \leq j \leq t$, with $B_j \leq 2A_j$, and intervals $J_j = (C_j, D_j] \subseteq [z_1, x^{5/8}]$ for $1 \leq j \leq k$ with $D_j \leq 2C_j$, with the following property. For any subsets $\mathcal{J}_1 \subseteq \{1, \dots, t\}$ and $\mathcal{J}_2 \subseteq \{1, \dots, k\}$, and for any integer in the range $4 \leq \ell \leq \nu + 2$, we have either*

$$\prod_{j \in \mathcal{J}_1} A_j \prod_{j \in \mathcal{J}_2} C_j \geq x^{1/\ell + \eta} \quad \text{or} \quad \prod_{j \in \mathcal{J}_1} B_j \prod_{j \in \mathcal{J}_2} D_j < x^{1/\ell - \eta};$$

and

(iii) *Distinct integers $m_1, \dots, m_R \in [x/H, 3x/H]$;*

such that

$$\mathbf{S}^{o(1)} \left| \sum_{\substack{p_j \in I_j \\ 1 \leq j \leq t}} \sum_{\substack{h_j \in J_j \\ 1 \leq j \leq k}} \xi_0(h_1) \dots \xi_0(h_k) \Delta(p_1 \dots p_t h_1 \dots h_k, m_i) \right| \gg x^{1/2}$$

for $m = m_1, \dots, m_R$, and with

$$\text{Meas}(\mathcal{I}(x)) \ll x^{1/2 + o(1)} R.$$

Here we have used the estimate $\mathbf{S} \leq x^{o(1)}$ in estimating $\text{Meas}(\mathcal{I}(x))$.

Clearly we must have

$$(\prod A_j)(\prod C_j) \leq 4x \quad \text{and} \quad (\prod B_j)(\prod D_j) \geq x$$

in order for there to be any overlap with \mathcal{A} or \mathcal{B} , and we therefore assume henceforth that

$$2^{-3-3\nu} x \leq (\prod A_j)(\prod C_j) \leq 4x. \quad (19)$$

We now define Dirichlet polynomials

$$P_j(s) = \sum_{p \in I_j} p^{-s}, \quad (1 \leq j \leq t),$$

$$F_j(s) = \sum_{h \in J_j} \xi_0(h) h^{-s}, \quad (1 \leq j \leq k),$$

and

$$D(s) = P_1(s) \dots P_t(s) F_1(s) \dots F_k(s). \quad (20)$$

Thus $D(s)$ has coefficients supported on $[2^{-3-3\nu}x, 2^{5+3\nu}x]$.

At this point it is convenient to establish a general result on the coefficients of products of these Dirichlet polynomials.

Lemma 10 *The coefficients of any product of distinct factors $P_j(s)$ take values 0 and 1 only. The coefficients c_n of any sub-product of $D(s)$ satisfy $|c_n| \leq \tau_7(n)$. Provided that one excludes factors F_j for which $C_j \geq x^{1/4}$, any product of powers of the polynomials P_j and F_j will have coefficients c_n satisfying $|c_n| \leq \nu^\nu \ll \mathbf{S}^{o(1)}$ for $n \leq x$.*

The first assertion follows from the fact that the polynomials P_j have coefficients supported on primes in disjoint intervals I_j . For the second claim we observe that the coefficients of F_j have size at most $\tau(n)$, so that the sub-product in question will have coefficients dominated by those of $\zeta(s) (\zeta(s)^2)^k$, giving us the required bound $\tau_7(n)$. For the final assertion, we observe that if our product of Dirichlet polynomials contains a term with $n \leq x$ it can have at most $h = [\nu]$ factors, since A_j and C_j are at least z_1 . It follows that $|c_n| \leq \tau_h(n)$. Moreover the c_n are supported on products of primes $p \geq z_1$, since we are excluding the case $C_j \geq x^{1/4}$. We then have $\Omega(n) \leq \nu$, in light of the assumption that $n \leq x$. However $\tau_h(n)$ is at most the number of ways that a set of $\Omega(n)$ primes (distinct or not) can be partitioned into h subsets, whence $\tau_h(n) \leq h^{\Omega(n)} \leq \nu^\nu$. This completes the proof of the lemma.

We are now ready to state the main result of this section.

Proposition 2 *Let*

$$T = x^{1/2} \mathbf{S}^2 \quad (21)$$

and

$$T_0 = \exp(\tfrac{1}{3} \sqrt{\log x}). \quad (22)$$

Then, in the situation of Lemma 9, there are complex coefficients ζ_j of modulus 1 for which the function

$$M(s) = \sum_{j=1}^R \zeta_j m_j^{-s}$$

satisfies

$$Rx \ll \mathbf{S}^{o(1)} \int_{T_1}^{2T_1} |D(it)M(it)| dt \quad (23)$$

for some $T_1 \in [T_0, T]$.

As explained in connection with the definition (4) we think of this final bound as involving a loss of a factor $\mathbf{S}^{o(1)}$. The integral on the right should suggest the use of Proposition 1, although much work must be done first. However we observe at this point that the integers m_j satisfy $0 < m_j \leq 3x/H \leq T$, as required for Proposition 1, by virtue of (13) and (21).

We begin the proof of Proposition 2 by following the usual analysis of Perron's formula, as in Titchmarsh [16, Sections 3.12 and 3.19] for example. This produces

$$\begin{aligned} & \sum_{p_j \in I_j, (1 \leq j \leq t)} \sum_{\substack{h_j \in J_j (1 \leq j \leq k) \\ p_1 \dots p_t h_1 \dots h_k \in \mathcal{A}(m)}} \xi_0(h_1) \dots \xi_0(h_k) \\ &= \frac{1}{2\pi i} \int_{-iT}^{iT} D(s) \frac{(1 + \delta_0)^s - 1}{s} (Hm)^s ds + O(E), \end{aligned} \quad (24)$$

where the error E is given by

$$\sum_{2^{-3-3\nu}x \leq n \leq 2^{5+3\nu}x} \tau_7(n) \min \left\{ \frac{T^{-1}}{|\log(mH(1 + \delta_0)/n)|} + \frac{T^{-1}}{|\log mH/n|}, \log T \right\}.$$

Terms with $n < x/2$ or $n > 5x$ contribute $O(\tau_7(n)/T)$ each, and hence produce $O(2^{3\nu}x(\log x)^6/T)$ in total. When $J < |mH(1 + \delta_0) - n| \leq 2J$ with $x^{1/4} \leq J \ll x$ we have

$$\frac{T^{-1}}{|\log(mH(1 + \delta_0)/n)|} \ll (JT)^{-1}x$$

and the corresponding terms contribute $O((JT)^{-1}x(J \log^6 x))$, by Lemma 4, taking $z = 1$. Summing over dyadic ranges for J produces $O(x(\log x)^7/T)$, and similarly for the contribution from $T^{-1}/|\log mH/n|$. Finally, for terms with $|mH(1 + \delta_0) - n| \leq x^{1/4}$ or $|mH - n| \leq x^{1/4}$ we bound the minimum in E by $\log T$, obtaining a contribution $O(x^{1/4}(\log x)^6(\log T))$, by a further application of Lemma 4. It therefore follows that

$$E \ll 2^{3\nu}x(\log x)^6/T + x(\log x)^7/T + x^{1/4}(\log x)^6(\log T).$$

Our choice (21) ensures that $E \ll x^{1/2}\mathbf{S}^{-1}$, by (4) and (10).

A similar analysis applies to $\mathcal{B}(m)$, leading to the estimate

$$\begin{aligned} & \sum_{p_j \in I_j, (1 \leq j \leq t)} \sum_{\substack{h_j \in J_j (1 \leq j \leq k) \\ p_1 \dots p_t h_1 \dots h_k \in \mathcal{B}(m)}} \xi_0(h_1) \dots \xi_0(h_k) \\ &= \frac{1}{2\pi i} \int_{-iT}^{iT} D(s) \frac{(1 + \delta_1)^s - 1}{s} (Hm)^s ds + O(x^{1/2}\mathbf{S}^{-1}). \end{aligned}$$

It then follows that

$$\begin{aligned} & \sum_{p_j \in I_j, (1 \leq j \leq t)} \sum_{h_j \in J_j, (1 \leq j \leq k)} \xi_0(h_1) \dots \xi_0(h_k) \Delta(p_1 \dots p_t h_1 \dots h_k, m) \\ &= \frac{1}{2\pi i} \int_{-iT}^{iT} D(s) G(s) m^s ds + O(x^{1/2} \mathbf{S}^{-1}), \end{aligned}$$

with

$$G(s) = \left(\frac{(1 + \delta_0)^s - 1}{s} - \delta_0 \delta_1^{-1} \frac{(1 + \delta_1)^s - 1}{s} \right) H^s.$$

Now if $0 \leq \mu \leq 1$ and t is real, we have

$$\frac{(1 + \mu)^{it} - 1}{it} = \mu + \int_0^\mu \int_0^\nu (1 + \lambda)^{it-2} (it - 1) d\lambda d\nu = \mu + O(\mu^2(1 + |t|)),$$

whence $G(it) \ll \delta_0 \delta_1 (1 + |t|)$. Moreover $|P_j(it)| \leq A_j$ and $F_j(it) \ll C_j(\log x)$. We therefore deduce from (19) that

$$\int_{-iT_0}^{iT_0} |D(it)G(it)| dt \ll \delta_0 \delta_1 x (\log x)^3 T_0^2.$$

The choice (22) shows that the above bound is $O(x^{1/2} \mathbf{S}^{-1})$, by (12), (15), and (11). This allows us to conclude from part (iii) of Lemma 9 that

$$\mathbf{S}^{o(1)} \left| \int_{T_0 < |t| \leq T} D(it)G(it) m_j^{it} dt \right| \gg x^{1/2}$$

for $m_j = m_1, \dots, m_R$, when x is large enough.

We have now reached an important stage in the argument. By choosing suitable complex coefficients ζ_j of modulus 1 we can write

$$\overline{\zeta_j} \int_{T_0 < |t| \leq T} D(it)G(it) m_j^{it} dt = \left| \int_{T_0 < |t| \leq T} D(it)G(it) m_j^{it} dt \right|$$

for $1 \leq j \leq R$, whence

$$\mathbf{S}^{o(1)} \int_{T_0 < |t| \leq T} D(it)G(it) \overline{M(it)} dt \gg R x^{1/2},$$

with $M(s)$ as in Proposition 2. Indeed, since

$$|G(it)| = \left| \int_1^{1+\delta_0} v^{-it-1} dv - \delta_0 \delta_1^{-1} \int_1^{1+\delta_1} v^{-it-1} dv \right| \leq 2\delta_0,$$

we have

$$\begin{aligned} Rx^{1/2} &\ll \delta_0 \mathbf{S}^{o(1)} \int_{T_0 < |t| \leq T} |D(it) \overline{M(it)}| dt \\ &\ll x^{-1/2} \mathbf{S}^{o(1)} \int_{T_0 < |t| \leq T} |D(it) M(it)| dt \end{aligned}$$

for large enough x , by (12). Moreover, by dyadic subdivision there will be a value $T_1 \in [T_0, T]$ such that

$$\begin{aligned} Rx^{1/2} &\ll (\log x) x^{-1/2} \mathbf{S}^{o(1)} \int_{T_1 \leq |t| \leq 2T_1} |D(it) M(it)| dt \\ &\ll x^{-1/2} \mathbf{S}^{o(1)} \int_{T_1 \leq |t| \leq 2T_1} |D(it) M(it)| dt. \end{aligned}$$

The contribution from negative t has the same shape as that for positive t , but with ζ_j replaced by its conjugate, so that it suffices to consider $T_1 \leq t \leq 2T_1$. The proposition then follows.

It is the introduction of the Dirichlet polynomial $M(s)$, and the estimation of mean-values involving it, via Proposition 1, which are the most significant features of this paper.

7 Extremely Large Values of Dirichlet Polynomials

The next stage in the argument is to show that $P_j(it)$ and $F_j(it)$ cannot be extremely large.

Lemma 11 *We have*

$$|P_j(it)| \leq A_j \exp(-(\log x)^{1/5}), \quad (25)$$

if x is large enough. Similarly, we have

$$|F_j(it)| \leq C_j \exp(-(\log x)^{1/5}), \quad (26)$$

if x is large enough.

For $P_j(s)$ this follows by the argument used for Lemma 19 of Heath-Brown [5], which handled Dirichlet polynomials evaluated at $\frac{1}{2} + it$ rather than it . Since $A_j \geq z_1$ the argument shows that

$$P_j(it) \ll A_j (z_1^{-\beta(T_1)} + T_1^{-1}) (\log x)^2$$

with $\beta(T_1)$ of order $(\log T_1)^{-2/3}(\log \log T_1)^{-1/3}$, so that $\beta(T_1) \geq (\log x)^{-3/4}$ for large x . Thus (6) and (22) yield

$$\begin{aligned} P_j(it) &\ll A_j(z_1^{-(\log x)^{-3/4}} + T_0^{-1})(\log x)^2 \\ &\ll A_j\{\exp(-(\log x)^{1/4}/\nu) + \exp(-\frac{1}{3}(\log x)^{1/2})\}(\log x)^2, \end{aligned}$$

and (25) follows by (5), if x is large enough.

When $C_j \leq x^{1/4}$ we have

$$F_j(it) = \sum_{C_j < n \leq D_j} \xi(n)n^{-it}.$$

We write n as pm where $p = P^+(n)$ is the largest prime factor of n . This allows us to classify terms according to the value of m , giving

$$|F_j(it)| \ll \sum_m \left| \sum_{\substack{C_j/m < p \leq D_j/m \\ p \geq \max(P^+(m), z_1)}} p^{-it} \right|.$$

The inner sum is empty unless $D_j/m \geq z_1$, and then the previous argument shows that

$$\sum_{\substack{C_j/m < p \leq D_j/m \\ p > \max(P^+(m), z_1)}} p^{-it} \ll (C_j/m) \exp(-2(\log x)^{1/5}),$$

say. The required estimate then follows on summing over m .

In the remaining range $C_j \geq x^{1/4}$ we have

$$F_j(it) = \sum_{C_j < n \leq D_j} \xi_0(n)n^{-it} \ll \sum_{d \leq z_2} \left| \sum_{C_j/d < n \leq D_j/d} n^{-it} \right|, \quad (27)$$

by (17). When $T_1 \leq t \leq 2T_1$ the inner sum is

$$\ll (C_j/d)^{1/2} T_1^{1/6} + (C_j/d) T_1^{-1/6}$$

by the van der Corput third derivative estimate (see Titchmarsh [16, Theorem 5.11]). It follows that

$$F_j(it) \ll z_2^{1/2} C_j^{1/2} T_1^{1/6} + C_j T_1^{-1/6} \log x \ll z_2^{1/2} C_j x^{-1/8} T_1^{1/6} + C_j T_1^{-1/6} \log x,$$

since $C_j \geq x^{1/4}$. However $T_1 \leq T \ll x^{1/2+o(1)}$ by (21), whence (9) yields

$$z_2^{1/2} C_j x^{-1/8} T_1^{1/6} \ll C_j x^{-1/24+o(1)};$$

and $T_1 \geq T_0$, whence

$$C_j T_1^{-1/6} \log x \ll C_j \exp\{-\frac{1}{4}\sqrt{\log x}\},$$

say, by (22). The bound required for the lemma then follows.

We have already shown that we can take $C_j \leq x^{5/8}$, and we now reduce this bound further.

Lemma 12 *If $C_j \geq \max(x^{1/4}, T_1 z_2)$ then $R = 0$.*

Thus we will assume henceforth that $C_j \leq \max(x^{1/4}, T_1 z_2)$.

For the proof we apply (27). According to Titchmarsh [16, Theorem 4.11] we have

$$\sum_{N < n \leq M} n^{-1/2-it} \ll M^{1/2}/|t|$$

uniformly for $M \geq N \geq |t|/2$, say. In our situation we have

$$C_j/d \geq C_j/z_2 \geq T_1 \geq |t|/2,$$

and it follows by partial summation that

$$\sum_{C_j/d < n \leq D_j/d} n^{-it} \ll C_j d^{-1} T_1^{-1}.$$

We therefore conclude that $F_j(it) \ll C_j(\log x)/T_1$ uniformly for $t \in [T_1, 2T_1]$. Since $C_j \leq x^{5/8}$ the product $D(it)$ must contain at least one other factor apart from $F_j(it)$. We therefore see from Lemma 11 that

$$\begin{aligned} D(it) &\ll T_1^{-1}(\log x) \exp\{-(\log x)^{1/5}\} \prod_j A_j \prod_j C_j \\ &\ll T_1^{-1}(\log x) \exp\{-(\log x)^{1/5}\} x, \end{aligned}$$

by (19). Since $M(it) \ll R$ we then deduce from Proposition 2 that

$$Rx \ll \mathbf{S}^{o(1)}(\log x) \exp\{-(\log x)^{1/5}\} Rx.$$

This then shows that we must have $R = 0$, in view of (11).

8 The Fourth Moment of $F_j(it)$

Our next goal is the following estimate.

Lemma 13 *Suppose that $C_j \geq x^{1/4}$. Let $t_1, \dots, t_M \in [T_1, 2T_1]$, and assume that $|t_m - t_n| \geq 1$ for $m \neq n$. Then*

$$\sum_{m=1}^M |F_j(it_m)|^4 \ll \mathbf{S}^{o(1)} T_1 C_j^2.$$

Moreover

$$\int_{T_1}^{2T_1} |F_j(it)|^4 dt \ll \mathbf{S}^{o(1)} T_1 C_j^2.$$

The second claim clearly follows from the first. The lemma would still be true when $C_j < x^{1/4}$, but we only need the lemma for $C_j \geq x^{1/4}$. It would be quite easy to establish an estimate of the above form with an additional factor z_2^4 , say, using the classical fourth moment estimate for the Riemann zeta-function; but unfortunately $z_2 \neq O(\mathbf{S}^{o(1)})$. The key input for the proof is therefore the following estimate, which is an immediate corollary of Theorem 1 of Bettin, Chandee and Radziwiłł [1].

Lemma 14 *Let*

$$A(s) = \sum_{n \leq N} a_n n^{-s}$$

with $|a_n| \leq \tau_3(n)$ and $N \leq T^{1/2+1/67}$. Then

$$\int_T^{2T} |\zeta(\tfrac{1}{2} + it) A(\tfrac{1}{2} + it)|^2 dt \ll T \log T \sum_{m,n} \frac{|a_m a_n|}{[m, n]} + T.$$

Indeed

$$\begin{aligned} \sum_{m,n \leq N} \frac{\tau_3(m) \tau_3(n)}{[m, n]} &\leq \sum_{d \leq N} \sum_{\substack{m,n \leq N \\ d|m,n}} \frac{\tau_3(m) \tau_3(n)}{mn/d} \\ &\leq \sum_{d \leq N} \frac{\tau_3(d)^2}{d} \left\{ \sum_{u \leq N/d} \frac{\tau_3(u)}{u} \right\}^2 \\ &\ll \sum_{d \leq N} \frac{\tau_3(d)^2}{d} \{(\log N)^3\}^2 \\ &\ll (\log N)^{15}, \end{aligned}$$

whence

$$\int_T^{2T} |\zeta(\tfrac{1}{2} + it) A(\tfrac{1}{2} + it)|^2 dt \ll T(\log T)^{16}. \quad (28)$$

We begin our proof of Lemma 13 by writing

$$Z(s) = \sum_{\substack{d < z_2 \\ d | \Pi_1}} \mu(d) d^{-s} \quad \text{and} \quad F(s) = Z(s) \zeta(s).$$

It then follows from (17) that the coefficients of $F_j(s)$ and $F(s)$ agree for $C_j < n \leq D_j$. Using Perron's formula (Titchmarsh [16, Sections 3.12 and 3.19] for example) we then deduce that

$$F_j(it) = \frac{1}{2\pi i} \int_{5/4 - iT_1/2}^{5/4 + iT_1/2} F(s + it) \frac{D_j^s - C_j^s}{s} ds + O(E(C_j)) + O(E(D_j))$$

for $T_1 \leq t \leq 2T_1$, with

$$E(A) = \sum_{n=1}^{\infty} \tau(n) (A/n)^{5/4} \min \left\{ \frac{T_1^{-1}}{|\log(A/n)|}, \log T_1 \right\}.$$

Since $C_j \leq \max(x^{1/4}, T_1 z_2)$ we have $C_j \leq T_1 z_2$, so that $T_1 \geq C_j^{7/8}$, say. An analysis similar to that used for (24) then shows that the error terms satisfy

$$E(C_j) + E(D_j) \ll C_j^{1/2}.$$

We proceed to move the line of integration to $\operatorname{Re}(s) = \frac{1}{2}$, incurring an error

$$\left(\frac{C_j^{5/4}}{T_1} + \frac{C_j^{1/2} T_1^{1/4}}{T_1} \right) z_2 \ll C_j^{1/2},$$

say, whence

$$F_j(it) \ll C_j^{1/2} \left(1 + \int_{-T_1/2}^{T_1/2} |F(\tfrac{1}{2} + i(\tau + t))| \frac{d\tau}{1 + |\tau|} \right).$$

We then find via Hölder's inequality that

$$\sum_{m=1}^M |F_j(it_m)|^4 \ll C_j^2 \left\{ T_1 + (\log T_1)^3 \sum_{m=1}^M \int_{-T_1/2}^{T_1/2} |F(\tfrac{1}{2} + i(\tau + t_m))|^4 \frac{d\tau}{1 + |\tau|} \right\}.$$

However

$$\begin{aligned} \int_{-T_1/2}^{T_1/2} |F(\tfrac{1}{2} + i(\tau + t_m))|^4 \frac{d\tau}{1 + |\tau|} &= \int_{t_m - T_1/2}^{t_m + T_1/2} |F(\tfrac{1}{2} + i\tau)|^4 \frac{d\tau}{1 + |\tau - t_m|} \\ &\leq \int_{T_1/2}^{5T_1/2} |F(\tfrac{1}{2} + i\tau)|^4 \frac{d\tau}{1 + |\tau - t_m|}, \end{aligned}$$

and

$$\sum_{m=1}^M \frac{1}{1 + |\tau - t_m|} \ll \log T_1,$$

in view of the spacing condition on the points t_m . We therefore conclude that

$$\sum_{m=1}^M |F_j(it_m)|^4 \ll C_j^2 \left\{ T_1 + (\log T_1)^4 \int_{T_1/2}^{5T_1/2} |F(\tfrac{1}{2} + i\tau)|^4 d\tau \right\}.$$

Our task is now to estimate the fourth power moment of the function $F(\tfrac{1}{2} + i\tau) = Z(\tfrac{1}{2} + i\tau)\zeta(\tfrac{1}{2} + i\tau)$, using Lemma 14. For this we will employ the approximate functional equation for $\zeta(s)$ as given by Titchmarsh [16, Theorem 4.13]. This yields

$$\zeta(\tfrac{1}{2} + i\tau) \ll |Z_1(\tfrac{1}{2} + i\tau)| + |Z_2(\tfrac{1}{2} + i\tau; \tau)| + 1,$$

with

$$Z_1(s) = \sum_{n \leq X} n^{-s} \quad (X = T_1^{1/2+1/68}) \quad (29)$$

and

$$Z_2(s; \tau) = \sum_{n \leq 2\pi\tau/X} n^{-s}.$$

We apply this to two factors $\zeta(s)$, whence

$$\sum_{m=1}^M |F_j(it_m)|^4 \ll C_j^2 \{T_1 + (\log T_1)^4 (I_1 + I_2 + I_3)\}, \quad (30)$$

with

$$\begin{aligned} I_1 &= \int_{T_1/2}^{5T_1/2} |\zeta(\tfrac{1}{2} + i\tau) Z_1(\tfrac{1}{2} + i\tau) Z(\tfrac{1}{2} + i\tau)^2|^2 d\tau, \\ I_2 &= \int_{T_1/2}^{5T_1/2} |\zeta(\tfrac{1}{2} + i\tau) Z_2(\tfrac{1}{2} + i\tau; \tau) Z(\tfrac{1}{2} + i\tau)^2|^2 d\tau, \end{aligned}$$

and

$$I_3 = \int_{T_1/2}^{5T_1/2} |\zeta(\tfrac{1}{2} + i\tau) Z(\tfrac{1}{2} + i\tau)^2|^2 d\tau.$$

The first and third of these can be handled immediately by (28), giving bounds $O(\mathbf{S}^{o(1)} T_1)$.

For I_2 we use Cauchy's inequality together with the usual fourth moment estimate for the Riemann zeta-function (Titchmarsh [16, (7.6.1)] for example) to show that

$$I_2^2 \leq \left\{ \int_{T_1/2}^{5T_1/2} |\zeta(\tfrac{1}{2} + i\tau)|^4 d\tau \right\} I_4 \ll T_1 (\log T_1)^4 I_4, \quad (31)$$

with

$$I_4 = \int_{T_1/2}^{5T_1/2} |Z_2(\tfrac{1}{2} + i\tau; \tau) Z(\tfrac{1}{2} + i\tau)^2|^4 d\tau.$$

If we expand $Z_2^2 Z^4 \overline{Z_2}^2 \overline{Z}^4$ we get a sum over 12-tuples $(m_1, \dots, m_4, n_1, \dots, n_8)$ in which $m_j \leq 2\pi\tau/X$ and $n_j < z_2$. We then have to examine

$$\int_{\tau \in [T_1/2, 5T_1/2], \tau \geq (X/2\pi) \max(m_j)} \left(\frac{U}{V} \right)^{i\tau} d\tau,$$

where $U = m_1 m_2 n_1 n_2 n_3 n_4$ and $V = m_3 m_4 n_5 n_6 n_7 n_8$. Since the range for τ is a sub-interval of $[T_1/2, 5T_1/2]$ the integral is of order $|\log U/V|^{-1}$ whenever $U \neq V$, and is of order T_1 otherwise. Each value of U occurs at most $\tau_6(U)$ times, and similarly for V . Moreover since $\tau \leq 5T_1/2$ we have $m_i \ll T_1/X$, whence $U, V \ll (T_1/X)^2 z_2^4 \ll T_1$ via our choice (29) for X . It follows that

$$I_4 \ll T_1 \sum_{U \ll T_1} \frac{\tau_6(U)^2}{U} + \sum_{U \neq V \ll T_1} \frac{\tau_6(U) \tau_6(V)}{(UV)^{1/2}} |\log U/V|^{-1}.$$

The first sum is $O((\log T_1)^{36})$. For the second we note that $2ab \leq a^2 + b^2$ for real a, b whence

$$\begin{aligned} & \sum_{U \neq V \ll T_1} \frac{\tau_6(U) \tau_6(V)}{(UV)^{1/2}} |\log U/V|^{-1} \\ & \leq \frac{1}{2} \sum_{U \neq V \ll T_1} \left(\frac{\tau_6(U)^2}{U} + \frac{\tau_6(V)^2}{V} \right) |\log U/V|^{-1} \\ & = \sum_{U \neq V \ll T_1} \frac{\tau_6(U)^2}{U} |\log U/V|^{-1}, \end{aligned}$$

by symmetry. However

$$\sum_{\substack{V \ll T_1 \\ V \neq U}} |\log U/V|^{-1} \ll T_1 \log T_1,$$

whence $I_4 \ll T_1 (\log T_1)^{37}$. Thus $I_2 \ll \mathbf{S}^{o(1)} T_1$ by (31), so that (30) yields the required estimate for Lemma 13.

9 Large Values of Dirichlet Polynomials

In this section we handle moderately large values of $P_j(it)$ and $F_j(it)$. For this section only, it will be convenient to define

$$P_{t+j}(s) = F_j(s) \text{ and } A_{t+j} = C_j, B_{j+t} = D_j \text{ for } 1 \leq j \leq k.$$

We proceed by covering the range $[T_1, 2T_1]$ with unit intervals $[n, n+1]$, and examining the contribution from those where one has

$$\sup_{[n, n+1]} |P_j(it)| > A_j^{4/5} \text{ for some } j \leq t+k. \quad (32)$$

Consider the set \mathcal{N} of integers n with $[T_1] \leq n \leq [2T_1]$ for which (32) holds, and such that

$$V_0 < \sup_{[n, n+1]} |D(it)| \leq 2V_0$$

for a given $V_0 \geq 1$, where $D(s)$ is given by (20). Our goal is to prove the following bound.

Lemma 15 *We have*

$$\#\mathcal{N} \ll x V_0^{-1} \exp\{-(\log x)^{1/6}\}.$$

Before proceeding to the proof of the lemma we note the following consequence. By using a dyadic subdivision into values of V_0 which are powers of 2, we will have

$$\int |D(it)| dt \ll x (\log x) \exp\{-(\log x)^{1/6}\}$$

where the integral is over relevant intervals $[n, n+1]$ such that (32) holds. (Note that intervals where $V_0 \leq 1$ contribute a total $O(T_1)$, which is satisfactory.) We may compare this with the lower bound in Proposition 2. Since $|M(it)| \leq R$ for all t , and

$$x (\log x) \exp\{-(\log x)^{1/6}\} \ll x \mathbf{S}^{-1},$$

by (11), we see that intervals $[n, n+1]$ where (32) holds make a negligible contribution in Proposition 2. Thus, in subsequent work we will be able to assume that for all relevant t we have $|P_j(it)| \leq A_j^{4/5}$ and $|F_j(it)| \leq C_j^{4/5}$.

For the proof of the lemma we write

$$x_0 = \prod_{j=1}^{t+k} A_j \leq 4x,$$

by (19). For each j we choose $V_j = V_j(n)$ to be a power of 2 with

$$V_j/2 < \sup_{[n, n+1]} |P_j(it)| \leq V_j,$$

and we choose $j = j_0 = j_0(n)$ such that

$$\sigma = \frac{\log V_j}{\log A_j}$$

is maximal. In particular we have $\sigma > 4/5$ by (32). Moreover

$$A_j^{1-\sigma} \geq \frac{1}{2} \exp((\log x)^{1/5})$$

by Lemma 11. Thus

$$x^{1-\sigma} \geq \exp((\log x)^{1/5}),$$

since $A_j \leq x^{5/8}$ by part (ii) of Lemma 9. It follows that

$$1 - \sigma \geq (\log x)^{-4/5}. \quad (33)$$

We also have

$$V_0 \leq \sup_{[n, n+1]} |D(it)| \leq \prod_j \sup_{[n, n+1]} |P_j(it)| \leq \prod_j V_j \leq \left(\prod_j A_j \right)^\sigma = x_0^\sigma,$$

so that

$$x^\sigma \gg x_0^\sigma \geq V_0. \quad (34)$$

We subdivide the integers $n \in \mathcal{N}$ according to the values of $j_0(n)$ and V_{j_0} , producing $O((\log x)^2)$ subsets. There is thus a choice of j_0 and V_{j_0} for which the corresponding subset, \mathcal{N}_1 say, satisfies $\#\mathcal{N} \ll (\log x)^2 \#\mathcal{N}_1$, and such that

$$\sup_{[n, n+1]} |P_{j_0}(it)| \geq \frac{1}{2} A_{j_0}^\sigma$$

for $n \in \mathcal{N}_1$. It will be typographically convenient to drop the subscript j_0 , and to write

$$\sup_{[n, n+1]} |P(it)| \geq \frac{1}{2} A^\sigma.$$

We then get a succession of points $t_n \in [n, n+1]$ where the suprema are attained, and by restricting either to even n in \mathcal{N}_1 or to odd n , and then re-labeling, we obtain a set of points $t_1, \dots, t_K \in [T_1, 2T_1]$, with $\#\mathcal{N} \ll (\log x)^2 K$, such that $|t_i - t_j| \geq 1$ for $i \neq j$, and with

$$|P(it_m)| \geq \frac{1}{2} A^\sigma \quad (m = 1, \dots, K).$$

We first dispose of the case in which $P(s) = F_j(s)$ with $A = C_j > x^{1/4}$. In this situation Lemma 13 shows that $KA^{4\sigma} \ll \mathbf{S}^{o(1)} T A^2$. Since $\sigma > 4/5$ and $C_j \geq x^{1/4+\eta}$ by part (ii) of Lemma 9 we deduce via (21) that

$$\#\mathcal{N} \ll \mathbf{S}^{o(1)} T A^{2-4\sigma} \ll \mathbf{S}^3 x^{1/2+(2-4\sigma)(\frac{1}{4}+\eta)} \ll \mathbf{S}^3 x^{1-\sigma-\eta}.$$

This is enough for Lemma 15, in view of (11).

We turn now to the case in which (32) holds for a polynomial with $A < x^{1/4}$, so that $z_1 \leq A \leq x^{1/4-\eta}$, by part (ii) of Lemma 9. We will use the standard theory of large values estimates for Dirichlet polynomials, considering two separate sub-cases. Suppose firstly that

$$A \leq x^{1/4-\eta}; \text{ and either } A \geq x^{3/14} \text{ or } \sigma \geq \frac{9}{10}.$$

We choose a non-negative integer w such that

$$A^w \leq x^{1/2-2\eta} < A^{w+1}.$$

Thus $w \geq 2$, since $A \leq x^{1/4-\eta}$, and therefore

$$A^w \geq (x^{1/2-2\eta})^{w/(w+1)} \geq x^{1/3-2\eta}. \quad (35)$$

Since $A \geq z_1 = (4x)^{1/\nu}$ we will have $w \leq \nu$. The Dirichlet polynomial $P(s)^w$ has coefficients c_n supported on integers $n \leq (2A)^w \leq x$, and Lemma 10 shows that $|c_n| \leq \nu^\nu$.

We will now apply the following ‘‘Large Values Estimate’’.

Lemma 16 *Let $t_1, \dots, t_K \in [0, T]$ with $|t_i - t_j| \geq 1$ for $i \neq j$. Suppose we have complex coefficients c_n such that*

$$\left| \sum_{n=1}^N c_n n^{-it_j} \right| \geq V$$

for $1 \leq j \leq K$. Then

$$K \ll GNV^{-2} + G^3NTV^{-6}(\log NT)^2,$$

where

$$G = \sum_{n \leq N} |c_n|^2.$$

This follows from Huxley [8, (2.9)].

In our situation we take $\sum c_n n^{-s} = P(s)^w$, with $N = (2A)^w$, $V = 2^{-w}A^{w\sigma}$, and $G \leq (2A)^w \nu^{2\nu}$. This leads to

$$K \ll \{A^{w(2-2\sigma)} + TA^{w(4-6\sigma)}\} 2^{10w} \nu^{6\nu} (\log x)^2 \ll \mathbf{S}^{o(1)} \{A^{w(2-2\sigma)} + TA^{w(4-6\sigma)}\},$$

by (10). It should be emphasized that this holds uniformly with respect to w . If $A \geq x^{3/14}$ then since $\sigma > 4/5$ we see that (21) yields

$$A^{w(4\sigma-2)} \geq A^{6w/5} \geq A^{12/5} \geq x^{18/35} \geq T,$$

whence $A^{w(2-2\sigma)} \geq TA^{w(4-6\sigma)}$. On the other hand, if $\sigma \geq 9/10$ then

$$A^{w(4\sigma-2)} \geq A^{8w/5} \geq x^{\frac{8}{5}(\frac{1}{3}-2\eta)} \geq x^{8/15-4\eta} \geq T,$$

by (35), and again we find that $A^{w(2-2\sigma)} \geq TA^{w(4-6\sigma)}$. Thus, under our current assumptions, we have

$$K \ll \mathbf{S}^{o(1)} A^{w(2-2\sigma)},$$

whence

$$\#\mathcal{N} \ll \mathbf{S}^{o(1)} x^{1-\sigma-\eta(2-2\sigma)}.$$

However $x^{1-\sigma} \ll xV_0^{-1}$ by (34), and (33) yields

$$x^{-\eta(2-2\sigma)} \ll \exp(-2\eta(\log x)^{1/5}).$$

Here $\eta(\log x)^{1/5} \geq (\log x)^{1/6}$, say, by (7). Thus (11) yields

$$\mathbf{S}^{o(1)} x^{-\eta(2-2\sigma)} \ll \exp(-(\log x)^{1/6}),$$

and we obtain the bound required for Lemma 15 in the current case.

The final situation we examine is that in which we have $A \leq x^{3/14}$ and $4/5 \leq \sigma \leq 9/10$. We begin as before, but now choosing w so that

$$A^w \leq x^{15/31} < A^{w+1}.$$

Thus $w \geq 2$ and hence $A^w > x^{10/31}$. Instead of Lemma 16 we use the following estimate.

Lemma 17 *Let $t_1, \dots, t_J \in [\tau_0, \tau_0 + \tau]$ with $|t_i - t_j| \geq 1$ for $i \neq j$. Then for any complex coefficients a_m and any fixed $\varepsilon > 0$ we have*

$$\left\{ \sum_{j \leq J} \left| \sum_{M < m \leq 2M} a_m m^{-it_j} \right| \right\}^2 \ll_{\varepsilon} \tau^{\varepsilon} (JM + J^{11/6} \tau^{1/2} + J^{23/12} \tau^{1/12} M^{1/2}) \sum_{M < m \leq 2M} |a_m|^2.$$

This follows from the analysis in Section 3 of Jutila [9], taking $k = 3$. We cover the range $[T_1, 2T_1]$ with $O(1 + T_1/\tau)$ subintervals of length at most τ , whence some such subinterval contains J points t_j , with

$$K \ll (1 + T_1/\tau)J.$$

We proceed to split the sum $P(s)^w = \sum c_m m^{-s}$ into dyadic ranges, and deduce that there is some $M \leq (2A)^w$ such that

$$(\tfrac{1}{2}A^{\sigma})^w J \ll (\log x) \sum_{j \leq J} \left| \sum_{M < m \leq 2M} c_m m^{-it_j} \right|.$$

Lemma 17 then shows that

$$\begin{aligned} & (\tfrac{1}{2}A^{\sigma})^{2w} J^2 \\ & \ll_{\varepsilon} (\log x)^2 x^{\varepsilon} (J(2A)^w + J^{11/6} \tau^{1/2} + J^{23/12} \tau^{1/12} (2A)^{w/2}) \sum_{n \leq (2A)^w} |c_n|^2. \end{aligned}$$

Since $|c_n| \leq \nu^{\nu} = x^{o(1)}$ by (5) this simplifies to give

$$A^{2w\sigma} J^2 \ll x^{o(1)} (JA^w + J^{11/6} \tau^{1/2} + J^{23/12} \tau^{1/12} A^{w/2}) A^w,$$

and hence

$$J \ll x^{o(1)} (A^{w(2-2\sigma)} + \tau^3 A^{w(6-12\sigma)} + \tau A^{w(18-24\sigma)}).$$

However

$$\{A^{w(2-2\sigma)}\}^{2/3} \{\tau^3 A^{w(6-12\sigma)}\}^{1/3} = \tau A^{w(10-16\sigma)/3} \geq \tau A^{w(18-24\sigma)}$$

for $\sigma \geq 4/5$, and so the final term may be dropped. Now, since

$$\#\mathcal{N} \ll (\log x)^2 K \ll (\log x)^2 (1 + T_1/\tau) J \ll x^{o(1)} (1 + x^{1/2}/\tau) J$$

we have

$$\#\mathcal{N} \ll x^{o(1)} \left\{ 1 + \frac{x^{1/2}}{\tau} \right\} (A^{w(2-2\sigma)} + \tau^3 A^{w(6-12\sigma)}).$$

We choose

$$\tau = A^{w(10\sigma-4)/3},$$

whence

$$\#\mathcal{N} \ll x^{o(1)} \{ A^{w(2-2\sigma)} + x^{1/2} A^{w(10-16\sigma)/3} \}.$$

Since $A^w \leq x^{15/31}$ and $\sigma \leq 9/10$ we have

$$A^{w(2-2\sigma)} \leq x^{1-\sigma-(1-\sigma)/31} \leq x^{1-\sigma-1/310}.$$

Moreover, since $A^w \geq x^{10/31}$ and $\sigma \geq 4/5$ we have

$$x^{1/2} A^{w(10-16\sigma)/3} \leq x^{1/2+10(10-16\sigma)/93} \leq x^{1-\sigma-1/930},$$

on noting that

$$\frac{1}{2} + \frac{10(10-16\sigma)}{93} \leq 1 - \sigma - \frac{1}{930}$$

on $[4/5, 1]$, with equality at the lower endpoint. These estimates give suitable bounds for $\#\mathcal{N}$ in this final case. This completes the proof of Lemma 15.

10 Factors of Length Below $x^{1/4}$ — The Key Proposition

This section will be devoted to the proof of a general estimate which will be used to handle a number of different cases. We suppose that we have arranged the factors of $D(s)$ into three groups, so that

$$D(s) = P_1(s) \dots P_t(s) F_1(s) \dots F_k(s) = A(s) B(s) C(s).$$

We suppose further that any factor $F_j(s)$ of $A(s)$ has length at most $x^{1/4}$. We write our Dirichlet polynomials as

$$A(s) = \sum_{A < n \leq 2^{3\nu+3}A} a_n n^{-s}, \quad B(s) = \sum_{B < n \leq 2^{3\nu+3}B} b_n n^{-s},$$

and

$$C(s) = \sum_{C < n \leq 2^{3\nu+3}C} c_n n^{-s},$$

where $|a_n| \leq \nu^\nu$ and $|b_n| \leq \tau_7(n) \leq \tau(n)^6$, by Lemma 10. We may assume that $2^{-3-3\nu}x \leq ABC \leq 4x$, as in (19).

We now have the following result.

Proposition 3 Suppose that $C \notin [x^{1/\ell-\eta}, x^{1/\ell+\eta}]$ for every integer ℓ in the range $4 \leq \ell \leq \nu + 2$, and that

$$A \leq Bx^{o(\eta)}, \quad B \leq x^{1/2}x^{o(1)} \quad \text{and} \quad C^2 \leq Ax^{o(\eta)}. \quad (36)$$

Then if $Rx \ll \mathbf{S}^{o(1)}I$ with

$$I = \int_{T_1 \leq t \leq 2T_1; |C(it)| \leq C^{4/5}} |A(it)B(it)C(it)M(it)| dt$$

we have

$$R \ll_{\varepsilon} x^{1/10+\varepsilon},$$

for any fixed $\varepsilon > 0$.

The reader will see that the result holds under somewhat weaker but more complicated conditions. However the above suffices for our needs. Moreover one sees that the exponent $1/10$ corresponds to the situation in which A, B and C are roughly $x^{2/5}, x^{2/5}$ and $x^{1/5}$. This is the critical case for our theorem. Clearly (36) implies that $C < x^{1/4}$, since $ABC \leq 4x$.

We start by using Cauchy's inequality to show that $I \leq (I_1 I_2)^{1/2}$, where

$$\begin{aligned} I_1 &= \int_0^T |A(it)M(it)|^2 dt \\ &\ll_{\varepsilon} \mathbf{S}^{o(1)} (A^2 R^2 + (AT)^{\varepsilon} \{ART + AR^{7/4}T^{3/4}\}) \max_n |a_n|^2, \end{aligned} \quad (37)$$

for any fixed $\varepsilon > 0$, by Proposition 1, and

$$I_2 = \int_{T_1 \leq t \leq 2T_1; |C(it)| \leq C^{4/5}} |B(it)C(it)|^2 dt.$$

In (37) we have $|a_n|^2 \leq \nu^{2\nu} \ll \mathbf{S}^{o(1)}$. Thus depending on which of the three terms in (37) dominates we find that

$$\{Rx\}^2 \ll_{\varepsilon} \mathbf{S}^{o(1)} A^2 R^2 I_2,$$

or

$$\{Rx\}^2 \ll_{\varepsilon} \mathbf{S}^{o(1)} (AT)^{\varepsilon} ART I_2,$$

or

$$\{Rx\}^2 \ll_{\varepsilon} \mathbf{S}^{o(1)} (AT)^{\varepsilon} AR^{7/4}T^{3/4} I_2.$$

Rearranging these leads to

$$x^2 \ll_{\varepsilon} \mathbf{S}^{o(1)} A^2 I_2, \quad (38)$$

or

$$R \ll_{\varepsilon} x^{-2+2\varepsilon} AT I_2,$$

or

$$R^{1/4} \ll_{\varepsilon} x^{-2+2\varepsilon} AT^{3/4} I_2.$$

We will show that (38) cannot happen, for large x , and that

$$AI_2 \ll_{\varepsilon} x^{8/5+3\varepsilon}. \quad (39)$$

Thus if we have the second of the above three alternatives, then (21) produces $R \ll_{\varepsilon} x^{1/10+6\varepsilon}$. This is satisfactory for our theorem, on re-defining ε . Similarly if we have the third alternative, then $R \ll_{\varepsilon} x^{-1/10+21\varepsilon}$, which is more than satisfactory.

Our estimate for I_2 is given in the following lemma.

Lemma 18 *Let $B(s)$ and $C(s)$ be as above, and assume that the integer $w \geq 2$ is chosen so that*

$$C^{2w-1} \leq x < C^{2w+1}. \quad (40)$$

Then

$$\begin{aligned} I_2 &= \int_{T_1 \leq t \leq 2T_1; |C(it)| \leq C^{4/5}} |B(it)C(it)|^2 dt \\ &\ll \mathbf{S}^{o(1)} \{B^2 C^{8/5} + T B^{1+1/2w} C + T^{1-1/w} B^{1+1/2w} C^2\}. \end{aligned}$$

We might hope for a bound $O(\mathbf{S}^{o(1)} \{B^2 C^{8/5} + TBC\})$ for such an integral. indeed something like this would follow from a suitable form of Montgomery's Large Values Conjecture. Because we have a product of Dirichlet polynomials there is some flexibility in the way that the usual mean and large values estimates can be applied. One sees something similar in the work of Matomäki and Radziwiłł [11, Section 8]. However our analysis is geared to gaining a good power of x , while they are interested in making a smaller saving but with one very short polynomial.

We proceed to establish Lemma 18, before going on to apply it to (38) and (39). To estimate I_2 we cover the range $[T_1, 2T_1]$ with intervals $[n, n+1]$ and focus attention either on even values of n , or on odd values, depending on which case makes the larger contribution. For each such interval we choose a point t_n for which $|B(it)C(it)|$ is maximal, subject to the condition that $|C(it)| \leq C^{4/5}$. Intervals in which $B(it_n)$ or $C(it_n)$ is of order x^{-1} , say, contribute at most $O(T_1)$ to I_2 . We subdivide the remaining points further into $O(\log^2 x)$ classes according to the dyadic ranges

$$U_1 < |B(it_n)| \leq 2U_1, \quad U_2 < |C(it_n)| \leq 2U_2, \quad (41)$$

in which $|B(it_n)|$ and $|C(it_n)|$ lie. After renumbering the points t_n we find that

$$I_2 \ll T_1 + U_1^2 U_2^2 K \log^2 x, \quad (42)$$

where (41) holds for $1 \leq n \leq K$.

Our task now is to estimate K , for which we will use Huxley's large values estimate, given by Lemma 16, and the mean value estimate of Montgomery [12, Theorem 7.3], taking $Q = 1$, $\chi = 1$, $\delta = 1$. This latter result produces the following bound.

Lemma 19 *Under the assumptions of Lemma 16 we have*

$$K \ll G(N + T)V^{-2} \log N.$$

We may apply Lemmas 16 and 19 to $B(it)$, noting that

$$\sum |b_n|^2 \ll \sum \tau(n)^{12} \ll (\log x)^{4095} 2^{3\nu+3} B \ll \mathbf{S}^{o(1)} B,$$

say, by (10), to show that

$$K \ll \{B^2 U_1^{-2} + \min(BTU_1^{-2}, B^4 TU_1^{-6})\} \mathbf{S}^{o(1)}. \quad (43)$$

When the term $B^2 U_1^{-2}$ in (43) dominates the estimate (42) becomes

$$I_2 \ll T + B^2 U_2^2 \mathbf{S}^{o(1)} \ll T + B^2 C^{8/5} \mathbf{S}^{o(1)}, \quad (44)$$

since $U_2 \leq C^{4/5}$. This is clearly acceptable for Lemma 18. For the remainder of the proof we may therefore assume that (43) reduces to

$$K \ll \min(BTU_1^{-2}, B^4 TU_1^{-6}) \mathbf{S}^{o(1)}. \quad (45)$$

In addition to considering mean and large values of $B(s)$ we can use the Dirichlet polynomial

$$C(s)^w = \sum_{n \leq N} c_n n^{-s},$$

where the integer $w \geq 2$ is chosen to satisfy (40). Then $2w - 1 \leq \nu$, by (6), and $N \leq 2^{w(3\nu+3)} C^w = 2^{O(\nu^2)} C^w$. Moreover $|c_n| \leq \nu^\nu$ by Lemma 10. Here we have

$$\sum |c_n|^2 \leq 2^{O(\nu^2)} C^w,$$

whence Lemma 19 yields

$$K \ll \{C^{2w} + C^w T\} U_2^{-2w} \mathbf{S}^{o(1)}.$$

We have assumed that C does not fall in any interval $[x^{1/\ell-\eta}, x^{1/\ell+\eta}]$ with $4 \leq \ell \leq \nu + 2$. Hence (40) implies that we have

$$x^{1/(2w+1)+\eta} \ll C \ll x^{1/(2w-1)-\eta}. \quad (46)$$

Moreover (36) yields

$$C = (ABC)^{1/5} (C^2/A)^{2/5} (A/B)^{1/5} \ll x^{1/5+o(\eta)},$$

and it follows that $w \neq 2$.

By virtue of (45) we have

$$K \ll \min(BTU_1^{-2}, B^4TU_1^{-6}, \{C^{2w} + C^wT\}U_2^{-2w}) \mathbf{S}^{o(1)},$$

and therefore

$$\begin{aligned} K &\ll (BTU_1^{-2})^{1-3/2w} (B^4TU_1^{-6})^{1/2w} (\{C^{2w} + C^wT\}U_2^{-2w})^{1/w} \mathbf{S}^{o(1)} \\ &\ll \{T^{1-1/w} + TC^{-1}\} B^{1+1/2w} C^2 U_1^{-2} U_2^{-2} \mathbf{S}^{o(1)}. \end{aligned}$$

Thus (42) becomes

$$\begin{aligned} I_2 &\ll T + \{T^{1-1/w} + TC^{-1}\} B^{1+1/2w} C^2 \mathbf{S}^{o(1)} \\ &\ll \{T^{1-1/w} + TC^{-1}\} B^{1+1/2w} C^2 \mathbf{S}^{o(1)}. \end{aligned} \quad (47)$$

This is satisfactory for Lemma 18, thereby completing the proof.

It remains to use Lemma 18 to handle (38) and (39). When (38) holds Lemma 18 produces

$$\begin{aligned} x^2 &\ll_{\varepsilon} \mathbf{S}^{o(1)} A^2 \{B^2 C^{8/5} + TB^{1+1/2w} C + T^{1-1/w} B^{1+1/2w} C^2\} \\ &\ll_{\varepsilon} x^{o(\eta)} A^2 \{B^2 C^{8/5} + x^{1/2} B^{1+1/2w} C + x^{1/2-1/2w} B^{1+1/2w} C^2\} \end{aligned} \quad (48)$$

by (21). Since $ABC \leq 4x$ we have

$$x^{o(\eta)} A^2 B^2 C^{8/5} \ll x^{2+o(\eta)} C^{-2/5} \ll x^{2+o(\eta)} z_1^{-2/5}$$

so that the first term of (48) is $o(x^2)$, by (11). Since $A \leq Bx^{o(\eta)}$ the inequalities (46) yield

$$\begin{aligned} A^2 B^{1+1/2w} C &\leq (AB)^{3/2+1/4w} C x^{o(\eta)} \\ &= (ABC)^{3/2+1/4w} C^{-(2w+1)/4w} x^{o(\eta)} \\ &\ll x^{3/2+1/4w} \cdot x^{-1/4w-(2w+1)\eta/4w} x^{o(\eta)}, \end{aligned}$$

so that the overall contribution of this term to (48) is

$$\ll x^{o(\eta)} \cdot x^{1/2} \cdot x^{3/2-\eta/2} = o(x^2).$$

Similarly, we find that

$$\begin{aligned} A^2 B^{1+1/2w} C^2 &\leq (AB)^{3/2+1/4w} C^2 x^{o(\eta)} \\ &= (ABC)^{3/2+1/4w} C^{(2w-1)/4w} x^{o(\eta)} \\ &\ll x^{3/2+1/4w} \cdot x^{1/4w-(2w-1)\eta/4w} x^{o(\eta)}, \end{aligned}$$

so that the corresponding contribution to (48) is again $o(x^2)$. We therefore see that (38) cannot hold.

We remark that this would fail for $A = B = x^{2/5}$, $C = x^{1/5}$. It is crucial that C should not be close to $x^{1/5}$, for example, and this is the reason for the removal of such ranges in Section 4.

We now examine (39). Using Lemma 18 and (21) we have

$$\begin{aligned} AI_2 &\ll \mathbf{A}^{o(1)} \{B^2 C^{8/5} + T B^{1+1/2w} C + T^{1-1/w} B^{1+1/2w} C^2\} \\ &\ll_{\varepsilon} \mathbf{S}^{o(1)} \{AB^2 C^{8/5} + x^{3/2} B^{1/2w} + x^{3/2-1/2w} B^{1/2w} C\}, \end{aligned} \quad (49)$$

since $ABC \leq 4x$. However our assumptions (36) give

$$\begin{aligned} AB^2 C^{8/5} &\ll_{\varepsilon} AB^2 A^{1/5} C^{6/5} x^{\varepsilon} \\ &= (ABC)^{6/5} B^{4/5} x^{\varepsilon} \\ &\ll_{\varepsilon} x^{6/5} (x^{1/2+\varepsilon})^{4/5} x^{\varepsilon} \\ &\leq x^{8/5+2\varepsilon}, \end{aligned}$$

so that $AI_2 \ll_{\varepsilon} x^{8/5+3\varepsilon}$ when the first term of (49) dominates.

We noted above firstly that $w \geq 2$, and then that $w \neq 2$. It therefore follows that

$$B^{1/2w} \leq B^{1/6} \leq x^{1/12+\varepsilon}$$

via (36). Thus $AI_2 \ll_{\varepsilon} x^{19/12+2\varepsilon} \ll_{\varepsilon} x^{8/5+2\varepsilon}$ when the second term of (49) dominates.

Finally, when $w = 3$ we note that

$$\begin{aligned} x^{3/2-1/2w} B^{1/2w} C &= x^{4/3} B^{1/6} C \\ &= x^{4/3} (ABC)^{4/15} (C^2/A)^{11/30} (A/B)^{1/10} \\ &\ll_{\varepsilon} x^{4/3+4/15+\varepsilon} \\ &= x^{8/5+\varepsilon}, \end{aligned}$$

by (36). On the other hand, if $w \geq 4$ then

$$x^{3/2-1/2w} B^{1/2w} C \leq x^{3/2-1/2w+1/4w+1/(2w-1)+\varepsilon} \leq x^{177/112+\varepsilon} \ll x^{8/5+\varepsilon},$$

by (46). We therefore obtain (39) whichever term of (49) dominates. This completes our treatment of Proposition 3.

11 Factors of Length Below $x^{1/4}$

In this section we handle the various cases in which every factor $F_j(s)$ of $D(s)$ has length $C_j \leq x^{1/4}$. In this situation any factor, whether of type $P_j(s)$ or $F_j(s)$, will have length at least z_1 and at most $x^{1/4-\eta}$, as shown by part (ii) of Lemma 9. It will be convenient to combine factors $P_j(s)$ and $F_j(s)$ of $D(s)$ as far as possible, subject to the lengths of the resulting Dirichlet polynomials being at most $x^{1/4-\eta}$. Such products will no longer run over dyadic intervals, but they will be of the form

$$Q(s) = \sum_{A < n \leq 2^{3+3\nu}A} q_n n^{-s},$$

and we will refer to A as being the “length” of $Q(s)$. Thus the procedure described above involves multiplying any two Dirichlet polynomials whose lengths A_1 and A_2 have $A_1 A_2 \leq x^{1/4-\eta}$ recursively, until no further polynomials can be combined. We may therefore assume that $A_i \leq x^{1/4-\eta}$ for any $Q_i(s)$ and that $A_i A_j > x^{1/4-\eta}$ for any two distinct factors $Q_i(s)$ and $Q_j(s)$. If there are m factors $Q_i(s)$ altogether, we deduce from (19) that

$$2^{-3-3\nu}x \leq \prod_{i=1}^m A_i \leq 4x. \quad (50)$$

We therefore see that $5 \leq m \leq 8$. We will index the polynomials with $A_1 \geq A_2 \geq \dots$

We begin by considering the case in which $m = 5$. In view of our ordering of the $Q_i(s)$ we will have

$$A_3 A_4 \leq A_1 A_2 \leq x^{1/2-2\eta} \leq x^{1/2}$$

and $A_5^2 \leq A_3 A_4$. It follows that we can apply Proposition 3 with

$$A(s) = Q_3(s)Q_4(s), \quad B(s) = Q_1(s)Q_2(s), \quad \text{and} \quad C(s) = Q_5(s).$$

We then have $R \ll x^{1/10+o(1)}$ when $m = 5$.

For $m = 6$ we note that $A_1 A_3 \leq x^{1/2-2\eta} \leq x^{1/2}$ and

$$A_2 A_4 A_6 \leq \{A_1 A_2 A_3 A_4 A_5 A_6\}^{1/2} \leq (4x)^{1/2}$$

by (50). Moreover $A_5^2 \leq A_1 A_3$ and $A_5^2 \leq A_2 A_4 A_6$. We may therefore apply Proposition 3 with $C(s) = Q_5(s)$ and either

$$A(s) = Q_1(s)Q_3(s), \quad B(s) = Q_2(s)Q_4(s)Q_6(s)$$

or vice-versa, depending on which of A_1A_3 or $A_2A_4A_6$ is smaller.

When $m = 7$ we consider two cases. Suppose firstly that

$$A_1A_2A_6 \leq x^{1/2}.$$

Then $A_3A_4A_7 \leq A_1A_2A_6 \leq x^{1/2}$. Moreover $A_5^2 \leq A_3A_4A_7$. Thus we may successfully apply Proposition 3 with

$$A(s) = Q_3(s)Q_4(s)Q_7(s) \quad B(s) = Q_1(s)Q_2(s)Q_6(s), \quad \text{and} \quad C(s) = Q_5(s).$$

In the alternative case we have $A_1A_2A_6 \geq x^{1/2}$, whence $A_3A_4A_5A_7 \leq 4x^{1/2}$, by (50). We also know that $A_1A_2 \leq x^{1/2-2\eta} \leq x^{1/2}$. Moreover $A_6^2 \leq A_1A_2$ and $A_6^2 \leq A_3A_4A_5A_7$. It follows in this alternative case that we may apply Proposition 3 with $C(s) = Q_6(s)$ and either

$$A(s) = Q_1(s)Q_2(s), \quad B(s) = Q_3(s)Q_4(s)Q_5(s)Q_7(s)$$

or vice-versa, depending on which of A_1A_2 or $A_3A_4A_5A_7$ is smaller.

There remains the case $m = 8$. Here we have

$$A_2A_4A_6A_8 \leq \{A_1A_2A_3A_4A_5A_6A_7A_8\}^{1/2} \leq (4x)^{1/2}$$

by (50). Moreover $A_1 \leq x^{1/4-\eta} \leq A_6A_8$, whence

$$A_1A_3A_5 \leq A_6A_8A_3A_5 \leq A_6A_8A_2A_4.$$

We can therefore apply Proposition 3 with

$$A(s) = Q_1(s)Q_3(s)Q_5(s), \quad B(s) = Q_2(s)Q_4(s)Q_6(s)Q_8(s)$$

and $C(s) = Q_7(s)$ to show that $R \ll x^{1/10+o(1)}$ in this final case.

12 Factors of Length at Least $x^{1/4}$

In this section we consider the case in which $D(s)$ has one or more factors $F_j(s)$ with $C_j \geq x^{1/4}$. As in the previous section we combine factors to produce Dirichlet polynomials $Q(s)$, but this time we omit from the procedure any factors $F_j(s)$ for which $C_j > x^{1/4}$. Thus any factor $Q_j(s)$ will have length $A_j \leq x^{1/4-\eta}$, and we will have $A_iA_j > x^{1/4-\eta}$ for any distinct polynomials $Q_i(s), Q_j(s)$.

We begin by teating the case in which $D(s)$ has precisely two factors, F_1 and F_2 say, for which $C_j > x^{1/4}$. According to part (ii) of Lemma 9 we then

have $C_j \geq x^{1/4+\eta}$. We now write $D(s) = F_1(s)F_2(s)H(s)$, so that the length A of $H(s)$ satisfies

$$A \ll 4x/C_1C_2 \ll x^{1/2-2\eta}.$$

Moreover the coefficients of $H(s)$ will have order $\mathbf{S}^{o(1)}$ by Lemma 10.

We then deduce from (23) that

$$Rx \ll \mathbf{S}^{o(1)} \int_{T_1}^{2T_1} |F_1(it)F_2(it)H(it)M(it)| dt.$$

By Hölder's inequality we therefore have

$$\{Rx\}^2 \ll \mathbf{S}^{o(1)} I_1^{1/2} I_2^{1/2} \int_{T_1}^{2T_1} |H(it)M(it)|^2 dt,$$

with

$$I_j = \int_{T_1}^{2T_1} |F_j(it)|^4 dt.$$

We now apply Lemma 13 together with Proposition 1 to deduce that

$$\{Rx\}^2 \ll_{\varepsilon} \mathbf{S}^{o(1)} C_1 C_2 T (A^2 R^2 + (AT)^{\varepsilon} \{ART + AR^{7/4} T^{3/4}\}),$$

for any fixed $\varepsilon > 0$. We then find that either

$$x^2 \ll_{\varepsilon} \mathbf{S}^{o(1)} C_1 C_2 A^2 T,$$

or

$$R \ll_{\varepsilon} C_1 C_2 A T^2 x^{-2+2\varepsilon},$$

or

$$R \ll_{\varepsilon} C_1^4 C_2^4 A^4 T^7 x^{-8+8\varepsilon}.$$

Since $C_1 C_2 A \leq 4x$ the definition (21) of T allows us to deduce that either

$$x^{1/2} \ll_{\varepsilon} \mathbf{S}^3 A,$$

or

$$R \ll_{\varepsilon} x^{3\varepsilon},$$

or

$$R \ll_{\varepsilon} x^{-1/2+9\varepsilon}.$$

The first of these is impossible by (11), since $A \ll x^{1/2-2\eta}$, while the other options are more than enough to give $R \ll x^{1/10+o(1)}$. This completes our treatment of the case in which exactly two of the factors $F_j(s)$ have length at least $x^{1/4}$.

We turn now to the case in which there are three factors $F_j(s)$ with corresponding lengths $C_j \geq x^{1/4}$, for which we use a variant of the previous method. Writing $D(s) = F_1(s)F_2(s)F_3(s)H(s)$ we find this time that $H(s)$ has length A satisfying

$$A \leq 4xC_1^{-1}C_2^{-1}C_3^{-1} \ll x^{1/4-3\eta}.$$

We then find via Hölder's inequality that

$$Rx \ll \mathbf{S}^{o(1)}\{I_1 I_2 I_3\}^{1/4} \left\{ \int_{T_1}^{2T_1} |H(it)|^4 |M(it)|^4 dt \right\}^{1/4},$$

with I_j as before. To estimate the remaining integral we observe that

$$|M(it)|^4 \leq R^2 |M(it)|^2.$$

We may then apply Proposition 1 with

$$H(it)^2 = \sum_{n \leq N} q_n n^{-it}.$$

We will have $N \ll 2^{6\nu} A^2$ and $q_n \ll \mathbf{S}^{o(1)}$, by Lemma 10. A similar calculation to before then shows that either

$$x \ll_{\varepsilon} \mathbf{S}^{o(1)}(C_1 C_2 C_3)^{1/2} A T^{3/4},$$

or

$$R \ll_{\varepsilon} x^{2\varepsilon},$$

or

$$R \ll_{\varepsilon} x^{-1/2+8\varepsilon}.$$

The first of these is impossible when $A \ll x^{1/4-3\eta}$, and the other alternatives yield $R \ll x^{1/10+o(1)}$.

Finally in this section we examine the situation in which there is exactly one factor F_j with $C_j > x^{1/4}$. Here we shall use the following result.

Lemma 20 *Suppose $D(s)$ factors as $F_1(s)A(s)B(s)$ with*

$$A \leq 2x^{1/2-\eta}, \quad B \leq x^{7/20} \quad \text{and} \quad AB \leq 4x^{3/4-\eta},$$

and where $A(s)$ and $B(s)$ have no factors $F_j(s)$ of length $C_j > x^{1/4}$. Then $R \ll x^{1/10+o(1)}$.

From (23) we deduce that

$$Rx \ll \mathbf{S}^{o(1)} \int_{T_1}^{2T_1} |F_1(it)A(it)B(it)M(it)|dt,$$

whence Hölder's inequality yields

$$Rx \ll \mathbf{S}^{o(1)} I_1^{1/4} \left\{ \int_{T_1}^{2T_1} |B(it)|^4 dt \right\}^{1/4} \left\{ \int_{T_1}^{2T_1} |A(it)M(it)|^2 dt \right\}^{1/2}.$$

We estimate I_1 via Lemma 13, noting that $C_1 \gg 2^{-3\nu}x/AB \gg x^{1/4+\eta/2}$. To handle the second integral we use the mean value theorem (2) coupled with the bound $O(\mathbf{S}^{o(1)})$ for the coefficients of $B(s)$ given by Lemma 10. The final integral can be dealt with via Proposition 1, again using Lemma 10 to estimate the coefficients. We conclude that

$$\begin{aligned} Rx &\ll \mathbf{S}^{o(1)} \{C_1^2 T\}^{1/4} \{(T+B^2)B^2\}^{1/4} \left\{ \int_{T_1}^{2T_1} |A(it)M(it)|^2 dt \right\}^{1/2} \\ &\ll \mathbf{S}^{o(1)} C_1^{1/2} (T^{1/2} B^{1/2} + T^{1/4} B) \{R^2 A^2 + x^\varepsilon (RAT + R^{7/4} AT^{3/4})\}^{1/2}. \end{aligned}$$

Thus either

$$Rx \ll \mathbf{S}^{o(1)} C_1^{1/2} (T^{1/2} B^{1/2} + T^{1/4} B) RA \ll \mathbf{S}^2 C_1^{1/2} (x^{1/4} B^{1/2} + x^{1/8} B) RA,$$

by (21), or

$$Rx^2 \ll x^{\varepsilon+o(1)} C_1 (TB + T^{1/2} B^2) AT \ll x^{2\varepsilon} C_1 (xB + x^{3/4} B^2) A,$$

or

$$Rx^8 \ll x^{4\varepsilon+o(1)} C_1^4 (T^4 B^4 + T^2 B^8) A^4 T^3 \ll x^{5\varepsilon} C_1^4 (x^{7/2} B^4 + x^{5/2} B^8) A^4.$$

The first alternative is impossible, since

$$C_1^{1/2} B^{1/2} A \leq (4x)^{1/2} A^{1/2} \leq 4x^{3/4-\eta/2}$$

and

$$C_1^{1/2} BA \leq (4x)^{1/2} (AB)^{1/2} \leq 4x^{7/8-\eta/2}.$$

The second option yields

$$R \ll x^{2\varepsilon} (1 + Bx^{-1/4}) \ll x^{1/10+2\varepsilon}.$$

Finally, the third case produces

$$R \ll x^{5\varepsilon} (x^{-1/2} + x^{-3/2} B^4) \ll 1.$$

The lemma therefore follows.

We are now ready to complete our treatment of the case in which

$$D(s) = F_1(s) \dots F_k(s) P_1(s) \dots P_t(s),$$

with $1 \leq k \leq 3$, where $C_1 \geq x^{1/4+\eta}$ and $C_j \leq x^{1/4-\eta}$ for $j \neq 1$. We combine all factors other than $F_1(s)$ as far as possible into Dirichlet polynomials $Q_i(s)$ of length $A_i \leq x^{1/4-\eta}$. We may then write $D(s) = F_1(s) Q_1(s) \dots Q_m(s)$ with $x^{1/4-\eta} \geq A_1 \geq A_2 \geq \dots$, and $A_i A_j \geq x^{1/4-\eta}$ whenever $i \neq j$. We therefore see that we must have $m \leq 6$. Indeed, since $C_1 \leq x^{5/8}$ we must also have $m \geq 2$. Moreover we will have

$$A_1 \dots A_m \leq \frac{4x}{C_1} \leq 4x^{3/4-\eta}.$$

Lemma 20 immediately handles the cases $m = 2$ and $m = 3$, by taking $B(s) = Q_1(s)$ and $A(s) = Q_2(s)$ for $m = 2$, and $B(s) = Q_1(s)$ and $A(s) = Q_2(s) Q_3(s)$ for $m = 3$. When $m = 4$ the choice $B(s) = Q_2(s)$ and $A(s) = Q_1(s) Q_3(s) Q_4(s)$ works similarly if $A_1 A_3 A_4 \leq 2x^{1/2-\eta}$. On the other hand, if $m = 4$ and $A_1 A_3 A_4 \geq 2x^{1/2-\eta}$ we will have

$$C_1 A_2 \leq 4x / (A_1 A_3 A_4) \leq x^{1/2+O(\eta)}$$

and $A_4^2 \leq A_1 A_3 \leq C_1 A_2$. Thus Proposition 3 applies, with $A(s) = Q_1(s) Q_3(s)$, $B(s) = F_1(s) Q_2(s)$ and $C(s) = Q_4(s)$.

When $m = 5$ we apply Lemma 20, taking $A(s) = Q_1(s) Q_3(s) Q_5(s)$ and $B(s) = Q_2(s) Q_4(s)$. Since $A_1 \geq A_2 \geq \dots$ we have

$$(A_1 A_3 A_5)^2 \leq C_1 A_1 A_2 A_3 A_4 A_5 \frac{A_1}{C_1} \leq 4x \frac{x^{1/4-\eta}}{x^{1/4+\eta}} = 4x^{1-2\eta},$$

whence $A \leq 2x^{1/2-\eta}$. Moreover

$$(A_2 A_4)^{3/2} \leq A_1 A_2 A_4 = \frac{C_1 A_1 A_2 A_3 A_4 A_5}{C_1 A_3 A_5} \leq \frac{4x}{x^{1/4+\eta} \cdot x^{1/4-\eta}} = 4x^{1/2},$$

so that $B = A_2 A_4 \leq x^{7/20}$ for large x . The conditions of the lemma are therefore satisfied, whence $R \ll x^{1/10+o(1)}$.

There remains the case $m = 6$. Since $C_1 \geq x^{1/4+\eta}$ and $A_i A_j \geq x^{1/4-\eta}$ whenever $i \neq j$ it follows from (19) that $C_1 = x^{1/4+O(\eta)}$ and $A_j = x^{1/8+O(\eta)}$ for every index j . We may then apply Proposition 3, with

$$A(s) = Q_3(s) Q_4(s) Q_5(s), \quad B(s) = F_1(s) Q_1(s) Q_2(s)$$

and $C(s) = Q_6(s)$, again concluding that $R \ll x^{1/10+o(1)}$.

We have now covered all the relevant cases and have thus completed the proof of Theorem 2.

References

- [1] S. Bettin, V. Chandee, and M. Radziwiłł, The mean square of the product of the Riemann zeta-function with Dirichlet polynomials, *J. Reine angew. Math.*, 729 (2017), 51–79.
- [2] D.R. Heath-Brown, The differences between consecutive primes, *J. London Math. Soc.*, 18 (1978), 7–13.
- [3] D.R. Heath-Brown, The differences between consecutive primes, III, *J. London Math. Soc.* 20, (1979), 177–178.
- [4] D.R. Heath-Brown, Prime numbers in short intervals and a generalized Vaughan identity, *Can. J. Math.*, 34 (1982), 1365–1377.
- [5] D.R. Heath-Brown, The number of primes in a short interval, *J. Reine angew. Math.*, 389 (1988), 22–63.
- [6] D.R. Heath-Brown, The differences between consecutive primes, IV, *A tribute to Paul Erdős*, (Cambridge University Press, 1990), 277–287.
- [7] D.R. Heath-Brown, The differences between consecutive smooth numbers, *Acta Arithmetica*, 184 (2018), 267–285.
- [8] M.N. Huxley, On the difference between consecutive primes, *Invent. Math.*, 15 (1972), 164–170.
- [9] M. Jutila, Zero density estimates for L -functions, *Acta Arith.*, 32 (1977), 55–62.
- [10] K. Matomäki, Large differences between consecutive primes. *Q. J. Math.* 58 (2007), 489–518.
- [11] K. Matomäki and M. Radziwiłł, Multiplicative functions in short intervals. *Ann. of Math.* 183 (2016), 1015–1056.
- [12] H.L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math. 227, (Springer, Berlin-Heidelberg- New York), 1971.
- [13] A.S. Peck, Differences between consecutive primes, *Proc. London Math. Soc.*, 76 (1998), 33–69.
- [14] A. Selberg, On the normal density of primes in small intervals and the difference between consecutive primes, *Arch. Math. Naturvid.*, t. 47, (1943), no. 6, 87–105.

- [15] P. Shiu, A Brun–Titchmarsh theorem for multiplicative functions, *J. Reine angew. Math.*, 313 (1980), 161–170.
- [16] E.C. Titchmarsh, *The theory of the Riemann zeta-function*. Second edition. (Oxford University Press, New York, 1986).
- [17] D. Wolke, Grosse Differenzen aufeinanderfolgender Primzahlen, *Math. Ann.*, 218 (1975), 269–271.
- [18] G. Yu, The differences between consecutive primes, *Bull. London Math. Soc.* 28 (1996), 242–248.

Mathematical Institute,
 Radcliffe Observatory Quarter,
 Woodstock Road,
 Oxford
 OX2 6GG
 UK

`rhb@maths.ox.ac.uk`