The Sum of the Möbius Function on Short Intervals

Martin Brennan

June 2023

Acknowledgements

For his support and encouragement throughout this past year, I would like to thank Professor Adam Harper.

Contents

1 Introduction

The purpose of this dissertation is to study the following question: For which $\theta > 0$ do we have

$$\sum_{x<n\leq x+H} \mu(n) = o(H), \tag{1.1}$$

where $H = x^\theta$ and μ is the Möbius function? For completeness we recall the definition of the Möbius function:

$$\mu(n) = \begin{cases} (-1)^n & \text{if } n = p_1p_2...p_n \text{ with } p_i \neq p_j \text{ for } i \neq j \\ 0 & \text{if } \exists\, p \text{ s.t } p^2 \mid n \end{cases}$$

We shall refer to intervals of the form $[x, x + x^\theta]$ where $\theta < 1$ as short intervals. When the above result is true, it tells us that there are roughly the same number of square-free integers $n \in (x, x + x^\theta]$ with an even number of prime factors as there are with an odd number.

Although the sum by itself is of sufficient interest to warrant further study, the methods covered can also be used to:

- Find estimates for the number of integers in short intervals that can be written as the sum of two squares.
- Find estimates for the number of almost primes in short intervals, where almost primes are integers of the form p_1p_2 and p_1, p_2 are primes.
- Find estimates for sums of the k-fold divisor function, τ_k, over short intervals.

However we will not delve into the details of these applications here [12].

Many of the tools used can also be employed to estimate the number of primes in short intervals which can be used to obtain results of the form $p_{n+1} \leq p_n + p_n^\theta$, where p_n denotes the n-th prime number [6] [8]. It should be noted, however, that the methods we study here can only demonstrate this for $\theta > 0.55$. In 2001, Baker, Harman, and Pintz [1] established the current record for this type of result with the inequality $p_{n+1} \leq p_n + p_n^{0.525}$.

We shall look at three results regarding the sum of the Möbius function over an interval. The first one will be the prime number theorem for the Möbius function. This is a simple application of the ideas originally developed by Riemann [19], applied to the Dirichlet series generated by $1/\zeta(s)$. Unfortunately, due to the limited size of the known zero free regions of the zeta function, we are only able to achieve results for intervals of size $H \geq x \exp(-(\log x)^{3/5})$. If we assume the Riemann hypothesis, we can use the same methods to show that (1.1) is true for $\theta > 1/2$.

The second result we shall discuss is a proof of (1.1) for $\theta > 7/12 \approx 0.58$. This was independently proved by both Motohashi[16] and Ramachandra[17] in 1976. Both of their proofs were based on the work of Huxley[7] and therefore both face the same limitations. We shall focus on Ramachandra's proof as it is the simpler of the two.

Ramachandra's idea was to apply the methods of Riemann; however, instead of moving the entire line of integration to the left using a zero-free region, he utilized Huxley's zero density estimate to shift as much of the line as far left as possible. Huxley's zero density estimate essentially states that

$$N(\alpha, T) \ll T^{A(1-\alpha)}(\log T)^B$$

where $A = 12/5$, $B = 9$, and $N(\alpha, T)$ is the number of zeros of $\zeta(s)$ in the region $\beta + i\gamma :; \alpha \leq \beta \leq 1, ; |\gamma| \leq T$. The exact value of B will not matter for our purposes. Huxley proved this density estimate in order to show his so called prime number theorem, which states that $p_{n+1} \leq p_n + p_n^\theta$, for n sufficiently large and $\theta > 7/12$. When we apply Ramachandra's methods, we observe that (1.1) holds true for $\theta > 1 - 1/A$. Furthermore, for $A = 12/5$, we obtain 1.1 for $\theta > 7/12$. Notice how as A decreases, so does our lower bound for θ. This is expected as a decrease in A implies a smaller number of zeros in the critical strip, enabling us to shift even more of our line of integration to the left.

The density hypothesis asserts that $A \leq 2$, which would give us (1.1) for $\theta > 1/2$. We see that the density hypothesis gives us as good of a result as the Riemann hypothesis, however, the density hypothesis may be much easier to prove. In fact, Huxley's methods show that the density hypothesis is true for $\alpha > 5/6$ and in 2000, Bourgain [2] showed that it was true for $\alpha > 25/32$, whereas the zeta function has only been proven to have no zeros in the Vinogradov-Korobov zero-free region, which is much smaller (see Corollary 3.16).

The proof of Huxley's zero density estimate uses two results. The first, proved by Ingham[8], states that $N(\alpha, T) \ll T^{A_1(\alpha)(1-\alpha)}(\log T)^5$ where $A_1(\alpha) = 2(2-\alpha)^{-1}$. The second, proved by Huxley, states that $N(\alpha, T) \ll T^{A_2(\alpha)(1-\alpha)}(\log T)^9$ for $\alpha \geq 3/4$ where $A_2(\alpha) = (5\alpha - 3)(\alpha^2 + \alpha - 1)^{-1}$. Note that $A_1(\alpha)$ is increasing, $A_2(\alpha)$ is decreasing and $A_1(3/4) = A_2(3/4) = 12/5$, which gives us Huxley's estimate. From this we can immediately see that the limiting factor in Huxley's result is the bound for the number of zeros near the 3/4 line. Indeed, in a recent preprint by Maynard and Pratt[13], they were able to show by studying the vertical distribution of zeros on the 3/4 line, that if all the zeros of ζ in the critical strip lie on finitely many vertical lines then $A \leq 24/11$. This would prove (1.1) for $\theta > 13/24 \approx 0.542$.

Huxley's estimate for when $\alpha = 3/4$ essentially comes down to bounding the product of six Dirichlet polynomials of size $x^{1/6+o(1)}$. Heath-Brown managed to prove Huxley's prime number theorem in [5] using what is now called Heath-Brown's identity (lemma 7.7). This method involves studying large values of Dirichlet polynomials which is what Huxley's work is also based on, so the limitation is essentially the same.

The third result, which most of this dissertation is dedicated to, is due to Matomäki and Teräväinen [12]. It states that

$$\sum_{x<n\leq x+x^{0.55+\epsilon}} \mu(n) \ll_\epsilon \frac{x^{0.55+\epsilon}}{(\log x)^{1/3-\epsilon}}$$

They also applied Heath-Brown's identity, however, they first applied the linear sieve to the sum followed by the introduction of Rameré's identity.

The basic idea of sieve theory is quite simple. Using a form of inclusion exclusion we may show that

$$\pi(x) - \pi(\sqrt{x}) + 1 = \sum_{d|\mathcal{P}(\sqrt{x})} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor$$

where $\mathcal{P}(z) = \prod_{p<z} p$. We would like to now use the estimate $\lfloor n \rfloor = n + O(1)$, however, this will lead to an error term which is too large to be useful. The idea of sieve theory is to restrict the support of the Möbius function in such a way that the main term remains somewhat accurate, while the error term is greatly diminished. More specifically, we form two sequences $(\lambda_d^+)_{d\geq 1}$ and $(\lambda_d^-)_{d\geq 1}$ such that when the first is substituted for the Möbius function, we obtain an upper bound for whatever we are trying to estimate. The combination of the two sequences $(\lambda_d^+)_{d\geq 1}$ and $(\lambda_d^-)_{d\geq 1}$ form what is called a sieve. Similarly, the second gives a lower bound.

There is a general strategy for applying sieve theory to a given sum:

- Use a sieve to eliminate problematic terms.
- Apply a combinatorial identity to the terms that remain in order to transform the sum into something more manageable.
- Apply analytic techniques to the new sum.

In our case we will use a sieve to eliminate all of the terms that do not have a prime factor on the interval $(P, Q]$. The terms which remain will all have a prime factor on the interval $(P, Q]$ and we shall use Rameré's identity to extract this prime factor. As $\mu(p) = -1$ is constant for any $p \in (P, Q]$, when we apply analytic methods, we will find that the Dirichlet polynomial corresponding to this small prime variable is

$$P(s) = \sum_{P<p\leq Q} \frac{1}{p^s}$$

There is a bound for $P(1 + it)$ given in [11], however, much of the detail about how the many error terms are handled is omitted. We will go through the proof of this result in much more detail (lemma 3.19).

The most important property of the Dirichlet polynomial $P(s)$ is that we can make it's length very short. More specifically we shall eventually define $Q = x^{1/(\log\log x)^2} \ll_\epsilon x^\epsilon$. We could not obtain a Dirichlet polynomial of such a length if we instead just used Heath-Brown's identity. The ability to make P very short enables us to handle cases where we have the product of six Dirichlet polynomials of length $x^{1/6+o(1)}$.

We note at this point that although we will follow the arguments of Matomäki and Teräväinen[12] quite closely when applying the above sieve techniques, our illustration looks rather different from the one in the original paper. We have broken up sections 3 and 4 of [12] into more digestible lemmas. There are also many seemingly arbitrary choices of parameters throughout the proof in [12] that we will not make straight away. Instead, we will choose to leave these parameters unspecified so that we are able to see what properties they must satisfy. This means we will have a much messier error term, which we will deal with at the end of Chapter 7. We have also quoted all of the results used and illustrated their applications, some of which were not included in the original proof, e.g., the application of the fundamental lemma of the sieve in Lemma 7.2 and the application of Shiu's bound in Lemma 7.4. Finally we hope to give some extra motivation for the various ideas involved in the proof.

The analytic techniques we will employ involve a comparison principle. In particular we shall look at the difference

$$\sum_{x<n\leq x+x^\theta} \mu(n) - \frac{x^\theta}{y} \sum_{x<n\leq x+y} \mu(n)$$

where y is chosen so that (1.1) is satisfied with $H = y$. In order to prove (1.1) for $H = x^\theta$, it suffices to show that this difference is $o(H)$. Our motivation for looking at this difference is that when we apply Perron's formula to both sums, we shall see that the main parts of the two integrals cancel, leaving us with the tails of the integrals and some error terms. While this theory is covered in many standard textbooks on analytic number theory (see, for example, [4], Ch. 7), we will develop it here for completeness with our own style and presentation.

When we apply these techniques, we will need to introduce additional parameters, namely the upper and lower bounds of the tails of our integrals. In the original paper, these were chosen arbitrarily. However, in this work, we will dedicate a subsection to discussing the required properties of these bounds and their origins.

At the start of section 4 in [12], it is assumed that $\theta = 1/2 + \epsilon$, however, we shall instead assume $\theta = 1/2 + h + \epsilon$ and will only set $h = 0.05$ when we are forced to do so. This will allow us to see how the results which are not contingent on $\theta > 0.55$ generalise.

We shall see that the comparison technique outlined above will only handle the case when we are able to form a Dirichlet polynomial which has a length in the interval $[x^{1/2-h-\epsilon/2}, x^{1/2+h+\epsilon/2}]$. To handle the remaining case we shall use a lemma from Heath-Brown and Iwaniec ([6], lemma 2) to bound sums of the form

$$\sum_{\substack{M<m\leq 2M \\ N<n\leq 2N}} a_m b_n r(x, x^\theta, mn)$$

where

$$r(x, x^\theta, mn) = \left\lfloor \frac{x + x^\theta}{mn} \right\rfloor - \left\lfloor \frac{x}{mn} \right\rfloor - \frac{x^\theta}{mn}$$

and $|a_n|$, $|b_n| \leq \tau_k(n)$ for some $k \geq 1$. The lemma used to bound these sums from [6] is only true for $\theta > 0.55$ and is the reason why we must make this assumption. We note that this lemma was first used in [6] to show that $p_{n+1} - p_n \ll p_n^\theta$.

As this lemma plays such a vital role in the proof of the result from [12], we shall dedicate the final part of this dissertation to its proof. In the original paper it was assumed that $|a_m|, |b_n| \leq 1$, however, as we will show, the proof generalises to the case where a_m and b_n are only divisor bounded and this is required for our application.

The main idea involves examining the difference between two sums by applying Perron's formula to both. However, this time, we observe that both of the main terms from Perron's formula are closely related through a change of variable which will again lead to cancellation in the main terms. We will demonstrate how all the error terms are bounded, with one of them having a minor mistake in the original paper which we have corrected.

The most challenging aspect of this proof involves bounding the sum of a product of three Dirichlet polynomials. The authors of [6] did this by employing six different bounds for the sum.

Three of these bounds were derived using Halasz' method, which gives us a bound for how many times a Dirichlet polynomial can become large given it is a certain length. Additionally, two bounds were obtained through a mean value theorem applied to Dirichlet polynomials.

The final bound was established by deriving an integral formula for a function denoted as $L(s)$, which is a truncation of the zeta function. This integral formula allowed the authors to find a bound for the sum by utilizing the fourth-moment estimate for the zeta function along the critical line.

Many of the explicit details including the application of Halasz' method and the derivation of the integral formula for $L(s)$ were omitted from the original paper and so we have filled them in here.

After they found these six bounds they then subdivided the problem into four cases based on the lengths of the Dirichlet polynomials and the size of the values they are taking. In the first case, to obtain a bound for the sum, they applied Van Der Corput's bound for exponential sums (Theorem 8.3). The details of this application, involving substantial calculations, were omitted from the original paper and so we have provided them here.

In the remaining three cases, they bounded the sum by using powers and products of the six bounds. However, the authors did not provide a specific motivation for the powers they used. Here, we will offer motivation by demonstrating that these powers can be determined by solving sets of linear equations.

In section 8.4 we will find the parameters which make all six bounds described above equal. This will allow us to see that the reason why we must have $\theta > 0.55$ is due to the case when Heath-Brown's identity gives us five variables of size $x^{1/5+o(1)}$, which in turn will give us five Dirichlet polynomials of the same length. While this fact is mentioned in [12], no mathematical details are provided to explain why this is the case, and it is not immediately obvious. We shall therefore shed some light on why this is the new barrier to improvement for θ in (1.1).

2 Preliminary Definitions and Results

Definition 2.1 (Multiplicative function)**.** *Let $f : \mathbb{N} \to \mathbb{C}$. We say that f is multiplicative if for any $m, n \in \mathbb{N}$ with m, n coprime, we have*

$$f(mn) = f(m)f(n)$$

Definition 2.2 (k-fold divisor function)**.** *Let $k \geq 1$ be an integer. We define the k-fold divisor function to be*

$$\tau_k(n) = \sum_{i_1 i_2 \ldots i_k = n} 1$$

We also define $\tau(n) := \tau_2(n)$

We shall often have sums whose terms are bounded by a divisor function. We shall therefore need a bound for sums of divisor functions.

Lemma 2.3. *Let $k, l \in \mathbb{Z}_{\geq 1}$ and $x \geq 2$. Then*

$$\sum_{n \leq x} \tau_k(n)^l \ll x(\log x)^{k^l}$$

and $\forall\ \epsilon > 0$ we have

$$\tau_k(n) \ll_\epsilon n^\epsilon$$

Note that we shall usually not need to know the exact power of $\log x$ in the bound. We may also sometimes use the slightly cruder bound

$$\sum_{n \leq x} \tau_k(n)^l \ll_\epsilon x^{1+\epsilon}$$

Proof. See [10] p.23.

□

In the course of this dissertation we shall encounter multiple instances where we wish to bound the sum of a multiplicative function over a short interval. The following theorem will help us do just that.

Theorem 2.4 (*Shiu's bound*)**.** *Let $0 < a < k$ be positive integers with $(a, k) = 1$ and let $0 < \alpha, \beta < \frac{1}{2}$. Suppose f is a non-negative, multiplicative arithmetic function that satisfies the following two properties:*

- *$\exists A_1 > 0$ such that*
$$f(p^l) \leq A_1^l \quad p \text{ prime}, \quad l \geq 1$$
- *$\forall \epsilon > 0$, $\exists A_2 > 0$ such that*
$$f(n) \leq A_2 n^\epsilon \quad n \geq 1$$

Then as $x \to \infty$ we have

$$\sum_{\substack{x < n \leq x+H \\ n \equiv a \bmod k}} f(n) \ll \frac{H}{\phi(k)} \frac{1}{\log(x+H)} \exp\left(\sum_{\substack{p \leq x+H \\ p \nmid k}} \frac{f(p)}{p} \right)$$

provided that

$$k < H^{1-\alpha}, \quad x^\beta < H \leq x$$

Proof. See [20].

□

3 Dirichlet Series and Polynomials

Definition 3.1 (Dirichlet series)**.** *A Dirichlet series is a function $D : \mathbb{C} \to \mathbb{C}$ of the form*

$$D(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

where $a_n \in \mathbb{C}$.

We call a Dirichlet series a Dirichlet polynomial if the sequence $(a_n)_{n\geq 1}$ is finitely supported. In this section we will see how Dirichlet series any polynomials can be used to estimate sums of sequences.

3.1 Perron's formula

Theorem 3.2 (Truncated Perron formula)**.** *Let $x, \kappa, T > 0$, and suppose that $\sum_{n=1}^{\infty} \frac{|a_n|}{n^\kappa}$ converges. Then*

$$\sum_{n\leq x} a_n = \frac{1}{2\pi i}\int_{\kappa - iT}^{\kappa + iT} \left(\sum_{n=1}^{\infty} \frac{a_n}{n^s}\right) x^s \frac{ds}{s} + O\left(x^\kappa \sum_{n=1}^{\infty} \frac{|a_n|}{n^\kappa} \min\left\{1, \frac{1}{T|\log(x/n)|}\right\}\right)$$

Proof. Follows from Thm 12.1 in [9].

□

Perron's formula is *(probably)* the most important theorem in analytic number theory. It allows us to estimate something algebraic on the left-hand side by something analytic on the right-hand side. Sometimes we may even use it in reverse to estimate something analytic by something algebraic.

In some of our applications we shall have that $|a_n| \leq \psi(n)$ for some non decreasing ψ, and that

$$\sum_{n=1}^{\infty} \frac{|a_n|}{n^\kappa} \ll A(\kappa)$$

where $1 < \kappa$. It therefore makes sense that we find a general formula for this case.

Corollary 3.3. *Let $0 < \kappa < 2$ and $0 < T \ll x$, where T is assumed to be sufficiently large. Suppose $|a_n| \leq \psi(n)$ for some non decreasing ψ and that $\sum_{n=1}^{\infty} \frac{|a_n|}{n^\kappa} \ll A(\kappa)$. Then*

$$\sum_{n\leq x} a_n = \frac{1}{2\pi i}\int_{\kappa - iT}^{\kappa + iT} \left(\sum_{n=1}^{\infty} \frac{a_n}{n^s}\right) x^s \frac{ds}{s} + O\left(\frac{x^\kappa A(\kappa)}{T} + \frac{x\psi(2x)\log x}{T}\right)$$

Proof. The main term is obviously the one from Perron's formula. Our task is to show that the error from Perron's formula is bounded by the one above. To do this we first note that $n \notin (x/2, 2x) \implies |\log(x/n)| \geq \log 2$. Therefore

$$\begin{aligned} x^\kappa \sum_{n=1}^{\infty} \frac{|a_n|}{n^\kappa} \min\left\{1, \frac{1}{T|\log(x/n)|}\right\} &\ll x^\kappa \sum_{n=1}^{\infty} \frac{|a_n|}{Tn^\kappa} + \sum_{n\in(x/2,2x)} \psi(n) \left(\frac{x}{n}\right)^\kappa \min\left\{1, \frac{1}{T|\log(x/n)|}\right\} \\ &\ll \frac{x^\kappa A(\kappa)}{T} + \psi(2x) \sum_{n\in(x/2,2x)} \min\left\{1, \frac{1}{T|\log(x/n)|}\right\}, \end{aligned}$$

where we have used that $\sum_{n=1}^{\infty} \frac{|a_n|}{n^\kappa} \ll A(\kappa)$ and $\psi(n)(x/n)^\kappa \leq 2^\kappa \psi(n) \ll \psi(2x)$ when $n \in (x/2, 2x)$. The reciprocal of $\log(x/n)$ is rather difficult to deal with, however, for $n \in (x/2, 2x)$ we can replace it by the linear term of its Taylor expansion as follows:

$$|\log(x/n)| = |\log(n/x)| = \log\left|\left(1 + \frac{n-x}{x}\right)\right| \asymp \frac{|n-x|}{x}$$

giving us

$$\sum_{n \in (x/2, 2x)} \min\left\{1, \frac{1}{T|\log(x/n)|}\right\} \ll \sum_{n \in (x/2, 2x)} \min\left\{1, \frac{x}{T|n-x|}\right\}$$

Note that

$$1 \leq \frac{x}{T|n-x|} \iff |n-x| \leq \frac{x}{T}.$$

So when $|n-x| \leq x/T$ the minimum will be 1. For the terms with $|n-x| \geq x/T$ we will split the sum into dyadic intervals. i.e. We split the sum into smaller sums over ranges of the form $(2^j x/T, 2^{j+1}x/T]$ where $0 \leq j \leq \lceil \log T/\log 2 \rceil$. This is a common trick we shall make use of a few times throughout this dissertation. The reason why this is often helpful is because on these ranges the dummy variable is roughly of constant size and hence if we replace it by its lower or upper bound we will (hopefully) not lose too much information. In our case $n \asymp 2^j x/T$. Doing this to our sum we see that

$$\begin{aligned}
\sum_{n \in (x/2, 2x)} \min\left\{1, \frac{1}{T|\log(x/n)|}\right\} &\ll \sum_{|n-x| \leq x/T} 1 + \sum_{j=0}^{\lceil \frac{\log T}{\log 2} \rceil} \sum_{2^j x/T < |n-x| \leq 2^{j+1}x/T} \frac{x}{T|n-x|} \\
&\ll \frac{x}{T} + \sum_{j=0}^{\lceil \frac{\log T}{\log 2} \rceil} \sum_{2^j x/T < |n-x| \leq 2^{j+1}x/T} \frac{1}{2^j} \\
&\ll \frac{x}{T} + \sum_{j=0}^{\lceil \frac{\log T}{\log 2} \rceil} \frac{2^j x}{T 2^j} \\
&\ll \frac{x}{T} + \frac{x \log(T)}{T} \\
&\ll \frac{x \log x}{T}
\end{aligned}$$

Therefore

$$x^\kappa \sum_{n=1}^{\infty} \frac{|a_n|}{n^\kappa} \min\left\{1, \frac{1}{T|\log(x/n)|}\right\} \ll \frac{x^\kappa A(\kappa)}{T} + \frac{x\psi(2x)\log x}{T}$$

Proving the result.

□

Remark 3.4. *If instead ψ was non-increasing, then we would have the same result but with $\psi(2x)$ replaced with $\psi(x/2)$. We usually will have $\psi(x/2) \asymp \psi(x) \asymp \psi(2x)$, so we will often just use $\psi(x)$ without comment.*

Remark 3.5. *We shall often use $\kappa = 1 + 1/\log x$ as this implies $x^\kappa = ex \asymp x$ and $(\kappa - 1)^{-1} = \log x$.*

Remark 3.6. *We will also often have that $\psi(n) \leq 1$ for all n. In this case we automatically have*

$$\sum_{n=1}^{\infty} \frac{|a_n|}{n^\kappa} \ll \sum_{n=1}^{\infty} \frac{1}{n^\kappa} \ll \int_1^\infty \frac{dw}{w^\kappa} = (\kappa - 1)^{-1}$$

which implies

$$\sum_{n \leq x} a_n = \frac{1}{2\pi i} \int_{\kappa - iT}^{\kappa + iT} \left(\sum_{n=1}^{\infty} \frac{a_n}{n^s}\right) \frac{x^s}{s} ds + O\left(\frac{x^\kappa}{T(\kappa - 1)}\right) + O\left(\frac{x \log x}{T}\right)$$

As we are looking at sums over $(x, x+y]$, with $y \le x$, we will need to deal with Dirichlet polynomials of the form

$$F(s) = \sum_{n \le x} \frac{f_n}{n^s}$$

It is hence also worth examining this case.

Corollary 3.7. *Let $\kappa > 0$ and $0 < T \ll x$, and suppose that $\sum_{n \le x} \frac{|f_n|}{n^\kappa} \ll A_\kappa(x)$ and $|f_n| \ll \psi(n)$ where ψ is non-decreasing. Then*

$$\sum_{n \le x} f_n = \frac{1}{2\pi i} \int_{\kappa - iT}^{\kappa + iT} \left(\sum_{n \le x} \frac{f_n}{n^s} \right) \frac{x^s}{s} ds + O\left(\frac{x^\kappa A_\kappa(x)}{T} + \frac{x\psi(x) \log x}{T} \right)$$

Proof. Follows along the same lines as 3.3 with some trivial adjustments. □

When we refer to Perron's formula throughout this dissertation, we may be referring to any of the results above depending on the context.

3.2 Perron's Formula on Intervals

We can study sums over intervals using Perron's formula by noting that

$$\sum_{x < n \le x+y} a_n = \sum_{n \le x+y} a_n - \sum_{n \le x} a_n$$

$$= \frac{1}{2\pi i} \int_{\kappa - iT}^{\kappa + iT} \left(\sum_{n=1}^{\infty} \frac{a_n}{n^s} \right) \frac{(x+y)^s - x^s}{s} ds + Error$$

Now the term

$$\frac{(x+y)^s - x^s}{s}$$

is a bit difficult to understand, it would therefore be prudent to form an estimate for it.

Lemma 3.8. *Let $0 < y < x$ and $T_0 > 0$. Suppose that $s = c + it$ and $T_0 \ll x/y$. Then $\forall\, |t| \le T_0$ we have*

$$\frac{(x+y)^s - x^s}{s} = x^{s-1} y + O\left(|s| x^{c-2} y^2 \right)$$

Moreover we have $\forall\, t \in \mathbb{R}$ that

$$\frac{(x+y)^s - x^s}{s} \ll y x^{c-1}.$$

Proof.
Part 1. We first note that

$$\frac{(x+y)^s - x^s}{s} = \frac{x^s((1+y/x)^s - 1)}{s}.$$

Now $y/x < 1$ and therefore

$$\left(1 + \frac{y}{x}\right)^s = \exp\left(s \log\left(1 + \frac{y}{x}\right)\right)$$

$$= \exp\left(\frac{sy}{x} + O\left(\frac{|s| y^2}{x^2} \right)\right)$$

$$= \exp\left(\frac{sy}{x} \right) \exp\left(O\left(\frac{|s| y^2}{x^2} \right)\right)$$

We have that $T_0 \ll x/y \implies yT_0/x \ll 1 \implies |s|y/x \ll 1$. We shall now make use of the fact that if $\alpha \ll 1$ then $\exp(\alpha) = 1 + \alpha + O\left(\alpha^2\right)$. We therefore see that

$$\left(1 + \frac{y}{x}\right)^s = \left(1 + \frac{sy}{x} + O\left(\frac{|s|^2 y^2}{x^2}\right)\right)\left(1 + O\left(\frac{|s|y^2}{x^2}\right)\right)$$

$$= 1 + \frac{sy}{x} + O\left(\frac{|s|y^2}{x^2}\right)$$

Subbing this into our expression we get

$$\frac{(x+y)^s - x^s}{s} = \frac{x^s\left(syx^{-1} + O\left(|s|y^2x^{-2}\right)\right)}{s}$$

$$= x^{s-1}y + O\left(|s|y^2x^{c-2}\right)$$

Part 2. If $|t| < x/y$ then $|s| \ll x/y$ and the proof follows from part 1. If $|t| \geq x/y$ then $|s| \gg x/y$ and therefore

$$\frac{1}{|s|} \ll \frac{y}{x}$$

Combining this with the trivial estimate, $(x+y)^s - x^s \ll x^c$, gives

$$\frac{(x+y)^s - x^s}{s} \ll yx^{c-1}$$

proving the result.

$\square$

Lemma 3.9. *Let $x, y, \kappa, T, T_0 > 0$ and $F(s) = \sum_{n \leq x} \frac{f_n}{n^s}$. Suppose $T_0 \ll x/y$, $T_0 \leq T \ll x$, $y < x$, $\sum_{n \leq x} \frac{|f_n|}{n^\kappa} \ll A_\kappa(x)$, $|f_n| \ll \psi(n)$ and there exists an increasing function in x, $B_\kappa(x,T)$, such that*

$$\int_{\kappa+iT_0}^{\kappa+iT} |F(s)|d|s| \ll B_\kappa(x,T) \qquad \int_{\kappa+iT}^{\kappa-iT_0} |F(s)|d|s| \ll B_\kappa(x,T)$$

Then we have that

$$\sum_{x<n\leq x+y} f_n = \frac{1}{2\pi i}\int_{\kappa-iT_0}^{\kappa+iT_0} F(s)x^{s-1}yds + O\left(x^{\kappa-2}y^2A_\kappa(x)T_0^2 + x^{\kappa-1}yB_\kappa(x,T) + \frac{x^\kappa A_\kappa(x)}{T} + \frac{x\psi(x)\log x}{T}\right)$$

Proof. By Perron's formula we have that

$$\sum_{x<n\leq x+y} f_n = \frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT} F(s)\frac{(x+y)^s - x^s}{s}ds + O\left(\frac{x^\kappa A_\kappa(x)}{T} + \frac{x\psi(x)\log x}{T}\right)$$

$$= \int_{\kappa-iT_0}^{\kappa+iT_0} F(s)\left(x^{s-1}y + O\left(|s|x^{\kappa-2}y^2\right)\right)ds + O\left(\int_{\kappa+iT_0}^{\kappa+iT} |F(s)|x^{\kappa-1}yd|s|\right)$$

$$+ O\left(\int_{\kappa+iT}^{\kappa-iT_0} |F(s)|x^{\kappa-1}yd|s|\right) + O\left(\frac{x^\kappa A_\kappa(x)}{T} + \frac{x\psi(x)\log x}{T}\right)$$

$$= \frac{1}{2\pi i}\int_{\kappa-iT_0}^{\kappa+iT_0} F(s)x^{s-1}yds + O\left(\int_{\kappa-iT_0}^{\kappa+iT_0} A_\kappa(x)x^{\kappa-2}y^2T_0d|s|\right) + O\left(x^{\kappa-1}yB_\kappa(x)\right)$$

$$+ O\left(\frac{x^\kappa A_\kappa(x)}{T} + \frac{x\psi(x)\log x}{T}\right)$$

performing the integral in the 'big Oh' term gives the result.

□

3.3 Comparing two sums

One of the main uses of the previous lemma is that it allows us to compare a sum with a scaled version of a longer sum. Notice that the main terms of

$$\sum_{x<n\leq x+H} f_n - \frac{H}{y}\sum_{x<n\leq x+y} f_n$$

will cancel. Therefore, if we can bound the error terms, we can then estimate the shorter sum by using the longer one.

Corollary 3.10. *With the same assumptions of lemma 3.9 along with the additional assumption that $0 < H \ll y$, we then have that*

$$\sum_{x<n\leq x+H} f_n - \frac{H}{y}\sum_{x<n\leq x+y} f_n \ll \frac{x^\kappa A_\kappa(x)}{T} + \frac{x\psi(x)\log x}{T} + x^{\kappa-2}HyA_\kappa(x)T_0^2 + x^{\kappa-1}HB_\kappa(x,T)$$

Proof. Form lemma 3.9 we have

$$\sum_{x<n\leq x+H} f_n - \frac{H}{y}\sum_{x<n\leq x+y} f_n \ll x^{\kappa-2}H^2A_\kappa(x)T_0^2 + x^{\kappa-1}HB_\kappa(x,T) + \frac{x^\kappa A_\kappa(x)}{T} + \frac{x\psi(x)\log x}{T}$$
$$+ x^{\kappa-2}HyA_\kappa(x)T_0^2 + x^{\kappa-1}HB_\kappa(x,T) + \frac{Hx^\kappa A_\kappa(x)}{yT}$$

Note that $H \ll y$ implies $H/y \ll 1$ and $H^2 \ll Hy$. Picking out the dominant terms proves the statement.

3.4 Zeta Function Theory

The zeta function is the most important function in analytic number theory and we shall need to make use of it several times throughout this dissertation. For completeness we shall now give its definition.

Definition 3.11. *For $Re\{s\} > 1$ we define*

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

Corollary 3.12. *For $Re\{s\} > 1$ we have*

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$$

Proof. Follows from the identity $\mu * 1(n) = 1_{n=1}(n)$ $\square$

We see from this relation that the Möbius function is deeply connected to the reciprocal of the zeta function.

We would like to be able to estimate the zeta function with a truncation of the sum in the definition. For this we have the Hardy-Littlewood formula.

Theorem 3.13 (Hardy-Littlewood formula)**.** *Let $x > 0$. Suppose $0 < \sigma_0 \geq \sigma \leq 2$ and $|t| \leq 2\pi x$, then*

$$\zeta(s) = \sum_{n \leq x} \frac{1}{n^s} + \frac{x^{1-s}}{s-1} + O_{\sigma_0}\left(x^{-\sigma}\right)$$

Proof. See [21] Thm 4.11 $\square$

One of the most important properties of the zeta function is that if we obtain a bound on how big it gets, we also get a bound on how small it gets. This notion is formalized in the following theorem.

Theorem 3.14. *Let $\phi(t)$ and $\theta(t)$ be positive functions which satisfy the following:*

- *Both $\phi(t)$ and $1/\theta(t)$ are non-decreasing for $t \geq 0$.*
- *$\phi(t) \to \infty$ and $\theta(t) \leq 1$*
- *$\frac{\phi(t)}{\theta(t)} = o(e^{\phi(t)})$*

Suppose $\zeta(\sigma + it) = O(e^{\phi(t)})$ when $1 - \theta(t) \leq \sigma \leq 2$. Then there is a constant $A > 0$ such that $\zeta(\sigma + it) \neq 0$ when

$$1 - A\frac{\theta(2t+1)}{\phi(2t+1)} \leq \sigma$$

Moreover we have

$$\frac{1}{\zeta(\sigma + it)} = O\left(\frac{\phi(2t+3)}{\theta(2t+3)}\right)$$

when

$$1 - \frac{A}{4} \cdot \frac{\theta(2t+3)}{\phi(2t+3)} \leq \sigma$$

Proof. See Thm 3.10 and Thm 3.11 in [21]. $\square$

Rather than using two different constants for the zero free region and the region for the the bound of $1/\zeta$, we will use $A_1/4$ for both.

Obviously to use this theorem we need a bound for $\zeta(s)$. Bounds for $\zeta(s)$ are most difficult to obtain inside the critical strip $0 < \sigma < 1$, for this we have Richert's bound.

Theorem 3.15 (Richert's bound)**.** *There exists C such that for $1/2 \leq \sigma \leq 1$ and t large we have*

$$\zeta(\sigma + it) \ll t^{C(1-\sigma)^{3/2}} \log^{2/3} t$$

Proof. See [18]. ☐

Rewriting Richert's bound we see that

$$\zeta(\sigma+it) \ll \exp(C(1-\sigma)^{3/2}\log t + (2/3)\log\log t)$$

so when $\sigma \geq 1-((\log\log t)/\log t)^{2/3}$ we have

$$\zeta(\sigma+it) \ll \exp\left(\left(C+\frac{2}{3}\right)\log\log t\right).$$

Applying theorem 3.14 with $\phi(t) = (C+2/3)\log\log t$ and $\theta(t) = ((\log\log t)/\log t)^{2/3}$ gives the largest known zero free region of the zeta function.

Corollary 3.16. *There exists $c_0 > 0$ small, such that ζ has no zeros in the region*

$$\mathcal{R} = \left\{\sigma+it : \sigma \geq 1 - \frac{c_0}{(\log\log(|t|+3))^{1/3}\log(|t|+2)^{2/3}}\right\}$$

Moreover for $\sigma+it \in \mathcal{R}$ we have

$$\frac{1}{\zeta(\sigma+it)} \ll (\log\log(|t|+3))^{1/3}\log(|t|+2)^{2/3}$$

Remark 3.17. *We may instead use*

$$\mathcal{R}_\epsilon = \left\{\sigma+it : \sigma \geq 1 - \frac{1}{(\log(|t|)^{2/3+\epsilon}}\right\}$$

and the bound

$$\frac{1}{\zeta(\sigma+it)} \ll_\epsilon \log(|t|)^{2/3+\epsilon}$$

Note that $\mathcal{R}_\epsilon \subset \mathcal{R}$ for $|t|$ large enough.

Finally, we state a mean value theorem for the zeta function on the critical line.

Theorem 3.18. *Let $T \geq 1$. Then*

$$\int_1^T \left|\zeta\left(\frac{1}{2}+it\right)\right|^4 dt = O\left(T\log^4 T\right)$$

Proof. See Thm 5.1 in [9]. ☐

3.5 Dirichlet Polynomial Estimates

So far we have showed that sums of sequences can be estimated by integrals of Dirichlet polynomials. We will now state some estimates for various Dirichlet polynomials and integrals of Dirichlet polynomials that we will come across in the course of this dissertation. We shall prove some of these results and provide references for the others.

Lemma 3.19. *Suppose $\exp\left((\log x)^\theta\right) \leq P \leq Q \ll x$, where $\theta > 2/3$, and let*

$$P(s) = \sum_{P<p\leq Q}\frac{1}{p^s}.$$

Then, for any $|t| \leq x$ and $A > 0$ we have

$$|P(1+it)| \ll_A \frac{\log x}{1+|t|} + (\log x)^{-A}.$$

Proof. This proof is from [11]. However, we have added many of the omitted details in regard to the error terms.

When $|t| \leq 10$ then $\frac{1}{1+|t|} \gg 1$ so it suffices to show $P(1+it) \ll \log x$, however this is trivial as by Merten's estimates we have

$$\sum_{P<p\leq Q} \frac{1}{p^{1+it}} \ll \sum_{p\leq Q} \frac{1}{p} \ll \log\log Q \ll \log x.$$

Hence we may assume $|t| \geq 10$. By the Euler product, when $\Re(s) > 1$ we have

$$\begin{aligned}
\log(\zeta(s)) &= -\sum_p \log(1-\frac{1}{p^s}) \\
&= \sum_{k=1}^{\infty}\sum_p \frac{1}{kp^{ks}} \\
\implies \log(\zeta(s+1+it)) &= \sum_{p^k} \frac{1}{kp^{(1+it)k}} \cdot \frac{1}{p^{ks}}
\end{aligned}$$

We see that the $k = 1$ terms gives the series we are interested in and we shall show that the terms given by $k \geq 2$ gives us one of the error terms. Naturally we will now apply Perron's formula. We shall use the version from theorem 3.2 with "x" $= Q$ and "x" $= Q$ where by "x" we mean the x in theorem 3.2 and not the x we have in the statement we wish to prove. If we then take the difference of the two resulting formulas we get

$$\begin{aligned}
\sum_{P<p^k\leq Q} \frac{1}{kp^{(1+it)k}} = \frac{1}{2\pi i}&\int_{c-iT}^{c+iT} \log\zeta(s+1+it)\cdot\frac{Q^s-P^s}{s}ds \\
&+O\left(Q^c\sum_{p^k}\frac{1}{kp^kp^{kc}}\min\left\{1,\frac{1}{T|\log(Q/p^k)|}\right\}\right) \\
&+O\left(P^c\sum_{p^k}\frac{1}{kp^kp^{kc}}\min\left\{1,\frac{1}{T|\log(P/p^k)|}\right\}\right)
\end{aligned}$$

where $c = 1/\log(x)$ and $T = \frac{|t|+1}{2}$. Due to the choice of c we have that $P^c, Q^c \ll 1$ and so it remains to bound the two sums. We shall only calculate the bound for the second sum as the calculation for the first sum is the same, except with P swapped for Q, and when it is calculated it will be seen to be smaller than the second error term. As in the proof of 3.3 we have that $|\log(P/p^k)| \asymp \frac{|P-p^k|}{P}$ when $p^k \in (P/2, 2P)$ and $|\log(P/p^k)| \gg 1$ when $p^k \notin (P/2, 2P)$. Once again, dividing the sum into dyadic intervals, we see that

$$\begin{aligned}
\sum_{p^k}\frac{1}{kp^kp^{kc}}\min\left\{1,\frac{1}{T|\log(P/p^k)|}\right\} \ll& \sum_{p^k\notin(P/2,2P)}\frac{1}{Tp^{(1+c)k}} + \sum_{|p^k-P|\leq P/T}\frac{1}{p^k} \\
&+\sum_{j=0}^{\lceil\frac{\log T}{\log 2}\rceil}\sum_{2^jP/T<|p^k-P|\leq 2^{j+1}P/T}\frac{1}{p^k}\frac{P}{T|p^k-P|}
\end{aligned}$$

We can easily bound the first term by an integral:

$$\sum_{p^k\notin(P/2,2P)}\frac{1}{Tp^{(1+c)k}} \ll \frac{1}{T}\sum_{n=1}^{\infty}\frac{1}{n^{1+c}} \ll \frac{1}{T}\sum_{n=1}^{\infty}\frac{1}{n^{1+c}} \ll \frac{1}{T}\int_1^{\infty}\frac{dw}{w^{1+c}} \ll \frac{1}{cT} = \frac{\log x}{T}$$

In the remaining terms, we shall have sums over integer points on an interval and we must be careful to distinguish between two different cases.

- If the interval is of length greater than or equal to 1, then we shall approximate the sum by its length.
- If the interval is of length less than 1, than we shall approximate the interval by the single term that may still be in the sum.

For the second term, if $T \le P$ we have

$$\sum_{|p^k - P| \le P/T} \frac{1}{p^k} \ll \frac{1}{P} \cdot \frac{P}{T} = \frac{1}{T}$$

where we have used that $p^k \asymp P$ when $|p^k - P| \le P/T$. If, on the other hand, $T > P$, then there will be at most one term in the sum, giving us

$$\sum_{|p^k - P| \le P/T} \frac{1}{p^k} \ll \frac{1}{P}$$

In the remaining term we may again use the estimate $p^k \asymp P$. The contribution of the terms with $2^j P/T \ge 1$ is

$$\begin{aligned}
&\ll \sum_{j=0}^{\lceil \frac{\log T}{\log 2} \rceil} \sum_{2^j P/T < |p^k - P| \le 2^{j+1} P/T} \frac{1}{P} \frac{1}{2^j} \\
&\ll \sum_{j=0}^{\lceil \frac{\log T}{\log 2} \rceil} \frac{1}{P} \frac{1}{2^j} \frac{P 2^j}{T} \\
&\ll \frac{\log x}{T}
\end{aligned}$$

The contribution of the terms with $2^j P/T < 1$ is

$$\begin{aligned}
&\ll \sum_{j=0}^{\lceil \frac{\log T}{\log 2} \rceil} \sum_{2^j P/T < |p^k - P| \le 2^{j+1} P/T} \frac{1}{P} \frac{1}{2^j} \\
&\ll \sum_{j=0}^{\lceil \frac{\log T}{\log 2} \rceil} \frac{1}{P} \frac{1}{2^j} \\
&\ll \frac{1}{P}
\end{aligned}$$

$$\sum_{P < p \le Q} \frac{1}{p^{1+it}} = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \log \zeta(s + 1 + it) \cdot \frac{Q^s - P^s}{s} ds + \sum_{\substack{P < p^k \le Q \\ 2 \le k}} \frac{1}{k p^{(1+it)k}} + O\left(\frac{\log x}{T} + \frac{1}{P} \right)$$

In the sum on the right-hand side we see that $\sqrt{P} \le p$ and hence

$$\sum_{\substack{P < p^k \le Q \\ 2 \le k}} \frac{1}{k p^{(1+it)k}} \ll \sum_{k=2}^{\log Q} \sum_{P^{1/k} < p}^{\infty} \frac{1}{p^k} \ll \sum_{k=2}^{\log Q} \int_{P^{1/k}}^{\infty} \frac{1}{w^k} dw \ll \sum_{k=2}^{\log Q} \frac{1}{P^{1-1/k}} \ll \frac{1}{\sqrt{P}} + \frac{\log Q}{P^{2/3}} \ll \frac{1}{\sqrt{P}}$$

So at this point we may conclude

$$\sum_{P < p \le Q} \frac{1}{p^{1+it}} = \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \log \zeta(s + 1 + it) \cdot \frac{Q^s - P^s}{s} ds + O\left(\frac{\log x}{|t| + 1} + \frac{1}{\sqrt{P}} \right)$$

It remains to estimate the integral. To do this we shall move the line of integration to the left using the Vinogradov-Korobov zero free region. We shall integrate on the contour

$$\mathcal{C} = \{c + iu : \ u \in [-T, T]\} \cup \{v \pm iT : \ v \in [-\sigma_0, c]\} \cup \{-\sigma_0 + iu : \ u \in [-T, T]\}.$$

where

$$\sigma_0 = \frac{c_0}{(\log x)^{2/3} \log\log x}$$

and c_0 is as in 3.16 Let $\mathcal{R}$ be the region contained in the contour $\mathcal{C}$. Note that $s+1+it$ is in the Vinogradov-Korobov zero free region for all $s \in \mathcal{R}$ and hence $\zeta(s+1+it) \neq 0 \ \forall s \in \mathcal{R}$. Also $-it \notin \mathcal{R}$ and therefore $\zeta(s+1+it)$ does not have a pole in $\mathcal{R}$. We conclude that $\log\zeta(s+1+it)$ is defined on $\mathcal{R}$, is analytic and hence the contour integral will be zero. Using Richert's bound (thm. 3.15) and the bound for $1/\zeta$ in corollary 3.16 we can show that for any $s \in \mathcal{R}$

$$\log\zeta(s+1+it) \ll \log\log x$$

On the horizontal sides we have $s \asymp T$, $Q^s - P^s \ll x^c \ll 1$ and the length of the interval is $\ll (\log x)^{-2/3}(\log\log x)^{-1}$. The contribution from the horizontal sides is therefore

$$\int_{\substack{\text{Horizontal}\\ \text{sides}}} \log\zeta(s+1+it) \cdot \frac{Q^s - P^s}{s} ds \ll \frac{\log\log x}{T(\log x)^{2/3}\log\log x} \ll \frac{\log x}{|t|+1}$$

On the vertical sides we have $Q^s - P^s \ = Q^{-\sigma_0} - P^{-\sigma_0} \ll P^{-\sigma_0}$, hence for the vertical integral we have

$$\int_{-\sigma_0 - iT}^{-\sigma_0 + iT} \log\zeta(s+1+it) \cdot \frac{Q^s - P^s}{s} ds \ll P^{-\sigma_0}\log^2 x \int_{-\sigma_0 - iT}^{-\sigma_0 + iT} \frac{d|s|}{|s|}$$

Now we note that $|s| \geq \sigma_0 \gg (\log x)^{-2/3}$ and $s = \sigma + i\tau \asymp |\tau|$ if $|\tau| \geq 1$. Using these estimates we see that

$$\int_{-\sigma_0 - iT}^{-\sigma_0 + iT} \frac{d|s|}{|s|} \ll (\log x)^{2/3} + \int_1^T \frac{d\tau}{\tau} \ll \log x$$

This finally allows us to conclude

$$P(1+it) \ll P^{-\sigma_0}(\log x)^3 + \frac{\log x}{|t|+1}$$

Using the lower bound for P and the definition of σ_0 we see that

$$P^{-\sigma_0}(\log x)^3 \ll \exp\left(-\frac{(\log x)^{\theta - 2/3}}{\log\log x}\right)(\log x)^3 \ll_A \frac{1}{(\log x)^A}$$

when $\theta > 2/3$, proving the result.

□

It is worth emphasising that the last line of the proof is not true when $\theta \leq 2/3$, so the smallest values for P we can pick are of the form $P = \exp\left((\log x)^{2/3+\epsilon}\right)$ for some $\epsilon > 0$. This lower bound for P comes from the boundary of the Vinogradov-Korobov zero free region.

We shall sometimes be integrating Dirichlet polynomials whose coefficients we do not have a good understanding of. For this we have the following estimate, .

Theorem 3.20. *For any Dirichlet polynomial* $\sum_{1 \leq n < N} a_n n^{-it}$ *we have*

$$\int_0^T \left| \sum_{n=1}^{N} \frac{a_n}{n^{it}} \right|^2 dt = (T + O(N)) \cdot \sum_{n=1}^{N} |a_n|^2$$

Proof. See Thm 9.1 from [10]. ◻

Note that the trivial bound for the integral is

$$T\left(\sum_{n=1}^{N}|a_n|\right)^2 = T\sum_{n=1}^{N}|a_n|^2 + T\sum_{\substack{1\le n,m\le N\\ n\neq m}}|a_n a_m|$$

We see that if $N \ll T$ then theorem 3.20 gives us significant saving as it avoids the contribution of the cross terms.

We have an analogous result for sums of arbitrary Dirichlet polynomials.

Theorem 3.21 (Davenport.). *Let $T_0, T \in \mathbb{R}$ with $T > 0$. Given a sequence $T_0 = t_0 < t_1 < .. < t_r < t_{r+1} = T_0 + T$, we have, for an arbitrary complex sequence $(c_n)_{n\ge 1}$, that*

$$\sum_{r=1}^{R}\left|\sum_{n=1}^{N} c_n n^{-it_r}\right|^2 \le (T + O(N\log N))\left(\frac{1}{\delta} + \log N\right)\sum_{n=1}^{N}|c_n|^2$$

where

$$\delta = \min_{0\le r\le R} t_{r+1} - t_r$$

Proof. See theorem 1 from [14]. ◻

The next result shows how we can use information we have on the zeta function to estimate Dirichlet polynomials which are truncations of the zeta function.

Lemma 3.22. *Let $k > 1$ and*

$$L(s) = \sum_{n=L}^{kL}\frac{1}{l^s}$$

Then for $s = c + it \neq 1$ with $|t| \le 2\pi L$ and $c > 0$, we have

$$L(c+it) = \frac{(kL)^{1-s} - L^{1-s}}{1-s} + O(L^{-c})$$

Moreover if c is bounded then

$$L(c+it) \ll \frac{L^{1-c}}{|1-s|}$$

Proof of part 1. To estimate $L(s)$ we use the Hardy Littlewood formula, with $x = L$ and $x = 6L$, to see that for $|t| \le 2\pi L$ we have

$$\zeta(c+it) = \sum_{n\le kL}\frac{1}{n^s} + \frac{(kL)^{1-s}}{s-1} + O\left((kL)^{-c}\right) = \sum_{n\le L}\frac{1}{n^s} + \frac{L^{1-s}}{s-1} + O\left(L^{-c}\right)$$

and hence as $k^{-c} \le 1$ we have

$$\sum_{n\le kL}\frac{1}{n^s} - \sum_{n\le L}\frac{1}{n^s} = \frac{L^{1-s}}{s-1} - \frac{(kL)^{1-s}}{s-1} + O\left(L^{-c}\right)$$

The result follows.

Proof of part 2. By part 1 we have

$$L(c+it) \ll \frac{L^{1-c}}{|1-s|} + L^{-c}$$

Note that the first term will be dominant if $|1-s| \ll L$. However we are assuming $|t| \ll L$ and therefore $|1-s| \ll L$ if c is bounded. The result follows.

□

We shall see later that one method of bounding integrals of Dirichlet polynomials is to take a sum over well spaced points. It will therefore be of interest to understand how often Dirichlet polynomials take large values on these points. The following theorem, due to work from Haslasz and Montgomery, shows a Dirichlet polynomial "can only get so big, so often".

Theorem 3.23. *Let $N_1, N_2 \geq 0$,*

$$F(s) = \sum_{N_1 < n \leq N_1 + N_2} \frac{f_n}{n^s}$$

and

$$G = \sum_{N_1 < n \leq N_1 + N_2} |f_n|^2$$

Suppose $0 \leq t_1 < t_2 ... < t_R \leq T$ satisfy $t_r - t_{r-1} > 1$ and

$$|F(it_r)| \geq B$$

for some $B > 0$. Then

$$R \ll \left(GN_2B^{-2} + G^3N_2TB^{-6}\right)(\log 2T)^6$$

Proof. See Corollary 9.9 from [10].

□

4 Prime Number Theorem for the Möbius Function

The most natural approach to the problem of estimating sums of the mobius function, is to apply Perron's formula to $1/\zeta(s)$ and then shift the contour as far left as possible using the Vinogradov-Korobov zero free region. Doing this gives us the following

Theorem 4.1 (Prime Number Theorem for Möbius Function with Vinogradov-Korobov error term)**.** *Suppose $x, \epsilon > 0$ and $y \geq x\exp(-(\log x)^{3/5-\epsilon})$. Then*

$$\sum_{x<n\leq x+y_1} \mu(n) \ll y\exp(-(\log x)^{3/5-\epsilon}) \ll_A \frac{y}{(\log x)^A}$$

Proof. $1/\zeta(s)$ does not have any poles on the Vinogradov-Korobov zero free region. We therefore have that the series $\sum_{n\geq 1}\mu(n)n^{-s}$ converges in the same region. Note that $|\mu(n)| \leq 1$, and therefore by Perron's formula with $\kappa = 1 + (\log x)^{-1}$ we have for $T \ll x$ that:

$$\sum_{x<n\leq x+y} \mu(n) = \frac{1}{2\pi i}\int_{\kappa-iT}^{\kappa+iT} \frac{1}{\zeta(s)}\frac{(x+y)^s - x^s}{s}ds + O\left(\frac{x\log x}{T}\right)$$

Using contour integration and the Vinograov Korobov zero-free region $\mathcal{R}_\epsilon$ from 3.17 we have:

$$\int_{\kappa-iT}^{\kappa+iT} \frac{1}{\zeta(s)}\frac{(x+y)^s - x^s}{s}ds = \int_{1-\delta-iT}^{1-\delta+iT} \frac{1}{\zeta(s)}\frac{(x+y)^s - x^s}{s}ds - \int_{1-\delta-iT}^{\kappa-iT} \frac{1}{\zeta(s)}\frac{(x+y)^s - x^s}{s}ds$$

$$+ \int_{1-\delta+iT}^{\kappa+iT} \frac{1}{\zeta(s)}\frac{(x+y)^s - x^s}{s}ds$$

where $\delta = (\log T)^{-2/3-\epsilon}$. Using the bound for $1/\zeta$ from remark 3.17, we have:

$$\int_{1-\delta+iT}^{\kappa+iT} \frac{1}{\zeta(s)}\frac{(x+y)^s - x^s}{s}ds \ll \int_{1-\delta+iT}^{\kappa+iT} \frac{x^{1+\kappa}}{T}(\log T)^{2/3+\epsilon}ds \ll \frac{x\log x}{T}$$

where we have used that $x^{1+\kappa} \ll x$ and the length of the interval is $\ll 1$. By symmetry, the lower horizontal side will have the same bound. For the vertical side we use lemma 3.8 to get

$$\int_{1-\delta-iT}^{1-\delta+iT} \frac{1}{\zeta(s)}\frac{(x+y)^s - x^s}{s}ds \ll \int_{1-\delta-iT}^{1-\delta+iT} x^{-\delta}y(\log T)^{2/3+\epsilon}ds$$

$$\ll Tyx^{-\delta}\log x$$

We now have

$$\sum_{x<n\leq x+y} \mu(n) \ll Tyx^{-\delta}\log x + \frac{x\log x}{T}$$

$$= y\log x e^{\log T - \log x(\log T)^{-2/3-\epsilon}} + \frac{x\log x}{T}$$

Letting $T = \exp(3(\log x)^{3/5-\epsilon})$ gives us

$$\sum_{x<n\leq x+y} \mu(n) \ll y\exp(-(\log x)^{3/5}) + x\exp(-2(\log x)^{3/5-\epsilon})$$

The result follows by imposing the bound $y \geq x\exp(-(\log x)^{3/5-\epsilon})$

$\square$

5 The 7/12 result

5.1 The Stratagy

As the main focus of this dissertation is the stronger 0.55 result, we shall not go into too much detail of this proof. We will, however, outline the main ideas and refer the reader to Huxley[7], Ingham[8] and Motohashi[16] for the finer details.

We recall from the proof of the prime number theorem that

$$\sum_{x<n\leq x+y} \mu(n) = \frac{1}{2\pi i}\int_{1-\delta-iT}^{1-\delta+iT} \frac{1}{\zeta(s)} \frac{(x+y)^s - x^s}{s} ds + O\left(\frac{x\log x}{T}\right)$$

where $\delta = (\log T)^{2/3-\epsilon_1}$ for some ϵ_1 arbitrarily small. Although we do not know where the zeros of the zeta function are outside of the zero free region, we can formulate estimates for how many there are. Our strategy will be, instead of moving the entire line of integration to the left, to move as much of the line as possible to the left. We then must find a suitable bound for $1/\zeta(s)$ on this new contour.

5.2 Huxley's Zero Density Estimate

Theorem 5.1. *Let $1/2 \leq \alpha \leq 1$, $T > 0$ and define*

$$N(\alpha, T) = \#\{\beta + i\gamma :\ \alpha \leq \beta \leq 1,\ |\gamma| \leq T,\ \text{and } \zeta(\beta + i\gamma) = 0\}.$$

Then

$$N(\alpha, T) \ll T^{12(1-\alpha)/5}(\log T)^9$$

Sketch of proof. Ingham [8] proved that

$$N(\alpha, T) \ll T^{3(1-\alpha)/(2-\alpha)}(\log T)^5$$

This proves the theorem for $1/2 \leq \alpha \leq 3/4$. We will not go into the details of Ingham's proof, but will instead focus on Huxley's contribution for $3/4 \leq \alpha \leq 1$.

We define

$$M(s) = \sum_{1\leq n\leq X} \frac{\mu(n)}{n^s}$$

for some $X > 0$. Our hope is that this function will approximate $1/\zeta(s)$ away from the zeros of $\zeta(s)$. Using the identity $1 * \mu(n) = 1_{n=1}(n)$, we have

$$M(s)\zeta(s) = 1 + \sum_{X<n} \frac{b_n}{n^s}$$

where

$$b_n = \sum_{\substack{d|n \\ d\leq X}} \mu(d)$$

We now study the integral transform

$$\frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} \zeta(\rho+\omega)M(\rho+\omega)Y^{\omega}\Gamma(\omega)d\omega = e^{-1/Y} + \sum_{m>X} b(m)m^{-\rho}e^{-m/Y}$$

where $Y > 0$ and $\rho = \beta + i\gamma$. This identity can be verified term by term. We now shift the line of integration to $Re\{\omega\} = 1/2 - \beta$. The zero of $\zeta(\rho+\omega)$ will cancel with the pole of $\Gamma(\omega)$ at $\omega = 0$. The only pole of the integrand is at $\omega = 1 - \rho$ and has residue

$$M(1)Y^{1-\rho}\Gamma(1-\rho).$$

We suppose that

$$\log X \leq \log Y \leq 2\log T$$

and we will let $l = \log T$. It can then be shown that

$$\left|\sum_{m>100lY} b(m)m^{-\rho}e^{-m/Y}\right| < 1/10$$

and that for $|\gamma| > 100l$, we have

$$\left|M(1)Y^{1-\rho}\Gamma(1-\rho)\right| < 1/10.$$

Note that by theorem 9.2 from [21] there is $\ll l^2$ zeros of $\zeta(s)$ with imaginary part of size less than $100l$ which, in view of the result we wish to prove, is a negligible amount.

We can therefore show that if $|\gamma| > 100l$ then

$$\frac{1}{2\pi i}\int_{1/2-i\infty}^{1/2+i\infty} \zeta(\rho+\omega)M(\rho+\omega)Y^{\omega}\Gamma(\omega)d\omega \geq \frac{2}{3} + \sum_{X<m\leq 100lY} b(m)m^{-\rho}e^{-m/Y}$$

Note that this means if the Dirichlet polynomial takes a value close to zero, then the integral cannot also take a value close to zero. This allows us to classify the zeros into two types.

- Type (i) zeros will be those for which our Dirichlet polynomial takes relatively large values.
- Type (ii) zeros will be those for which our integral takes relatively large values.

There may be zeros which are both type (i) and type (ii), however, all that matters is that every zero will be at least one of the two. We will first split the range of summation into dyadic intervals. Let

$$n_s = \left\lceil \frac{\log X - \log Y}{\log 2} \right\rceil \qquad n_f = \left\lfloor \frac{\log(100\ell)}{\log 2} \right\rfloor$$

We have chosen n_s and n_f so that $2^{n_s-1}Y \leq X \leq 2^{n_s}Y < 2^{n_s} < ... < 2^{n_f-1} < 2^{n_f} \leq 100lY \leq 2^{n_f+1}$. We now define $I(n) := [2^{n-1}Y, 2^nY]$ for $n_s < n < n_f$, $I(n_s) = [X, 2^{n_s-1}]$ and $I(n_f) = [2^{n_f}Y, 100lY]$

Class (i, n) : zeros ρ at which

$$\left|\sum_{m\in I(n)} b(m)m^{-\rho}e^{-m/Y}\right| > (6l)^{-1},$$

Class (ii, n) : zeros ρ at which

$$\max_{|\gamma-t|\leqq 2^n}\left|\zeta\left(\frac{1}{2}+it\right)M\left(\frac{1}{2}+it\right)\right| > c_2 2^n Y^{\alpha-\frac{1}{2}}$$

where c_2 is a fixed constant which is sufficiently small (see [7]).

We will focus on the type (i) zeros as the type (ii) zeros are more of a byproduct of the method we have used and are not of fundamental importance in fact modern treatments of the problem avoid the appearance of type (ii) zeros completely (see Ch.10 [10]). Applying Haslasz' method (theorem 3.23) to

$$\left(\sum_{m\in I(n)} b(m)m^{-\rho}e^{-m/Y}\right)^a$$

for $a \in \mathbb{Z}_{\geq 1}$, giving us

$$\left(2^nY\right)^{2a(1-\alpha)} e^{-a2^n} H(a) + T\left(2^nY\right)^{2a(2-3\alpha)} e^{-3a2^n} H(a)^3$$

where

$$H(a) = (3al)^{4a^2}(6l)^{2a}l$$

The question is which a should we use to bound the number of type (i, n) zeros. To decide this we define the decreasing sequence, $Y_1 > Y_2 > ... > Y_A$, by the relation

$$H(a+1)Y_a^{2(a+1)(1-\alpha)}l = H^3(a)Y_a^{2a(2-3\alpha)}T$$

We note that there was a typo in Huxley's paper[7] where he suggests that this sequence is increasing in a.

Huxley then showed that if we use the $a'th$ estimate for the number of class (i, n) zeros when $Y_a \leq 2^n Y \leq T_{a-1}$ we will get the required bound. Again, we refer to [7] for the rest of the details.

□

The main point we wish to illustrate here is that the bound for the number of zeros of the zeta function relies on bounding products of Dirichlet polynomials. Specifically in the case when $\alpha = 3/4$ we find that the worst bound is due to when $a = 3$. When $a = 3$ we need to bound the cue of a Dirichlet polynomial. However Halasz' method involves squaring this polynomial so we in fact end up with the product of six polynomials.

5.3 Application to Sum of Möbius Function

Theorem 5.2. *Let $x > 0$ and $H = x^{7/12+\epsilon}$. Then*

$$\sum_{x<n\leq x+H} \mu(n) = o(H)$$

Sketch of proof. We continue the argument outlined in section 5. We shall divide the area to the left of the line of integration into rectangles of area 1. More specifically we let

$$\Delta(j,k) = \{s = \sigma + it : \ \sigma_j \leq \sigma < \sigma_{j+1}, k(\log T) \leq t < (k+1)\log T\}$$

where $\sigma_j = 1/2 + j(\log T)^{-1}$ and j, k are integers. The range of j and k will be $[0, J]$ and $[-K, K]$ respectively, where $J = \lfloor (1/2 - \delta)\log T \rfloor$ and $K = \lfloor T(\log T)^{-1} \rfloor$. These rectangles will tile the area between our line of integration and the 1/2 line. We now divide these tiles into two classes W and Y. We may think of W as the set of *"bad tiles"* and Y as the set of *"good tiles"*. More specifically we let $\epsilon > 0$ be small and assume T is large enought to ensure the $1 - \epsilon < 1 - \delta$. If $\sigma_j \leq 1 - \epsilon$ then

$$\begin{aligned}
\Delta(j,k) \in W &\iff \Delta(j,k) \text{ contains a zero of } \zeta(s) \\
\Delta(j,k) \in Y &\iff \Delta(j,k) \text{ does not contains a zero of } \zeta(s)
\end{aligned}$$

If instead $\sigma_j > 1 - \epsilon$ then

$$\begin{aligned}
\Delta(j,k) \in W &\iff \exists\, s \in \Delta(j,k) \text{ such that } \zeta(s)G_j(s) > 1/2 \\
\Delta(j,k) \in Y &\iff \forall\, s \in \Delta(j,k) \text{ we have } \zeta(s)G_j(s) \geq 1/2
\end{aligned}$$

where

$$G_j(s) = \sum_{1\leq n\leq X_j} \frac{\mu(n)}{n^s}$$

for some $X_j \geq 1$. Using Huxleys zero density estimate we have for $\sigma_j \leq 1 - \epsilon$ that

$$\#\{k : \Delta(j,k) \in W\} \ll T^{12(1-\sigma_j)/5}(\log T)^9$$

Using an argument of Montgomery ([15] pp. 110-112) we can choose X_j so that for $\sigma_j > 1 - \epsilon$ we have

$$\#\{k : \Delta(j,k) \in W\} \ll T^{c(1-\sigma_j)^{3/2}}(\log T)^{16}$$

where c is a constant which comes from the application of Richert's bound (theorem 3.15).

For $k \in [-K, K]$ let

$$j_k = \mathrm{Max}\{j : \Delta(j,k) \in W\}$$

This means that $\Delta(j_k, k)$ are the *"bad tiles"* which are furthest to the right. We now define

$$\mathcal{D} = \bigcup_{j=0}^{J} \bigcup_{k=-K}^{K} \Delta(j,k),$$

$$\mathcal{D}' = \bigcup_{k=-K}^{K} \bigcup_{j \leqq j_k} \Delta(j,k)$$

and $\mathcal{D}_0 = \mathcal{D} - \mathcal{D}'$. We define our contour of integration, denoted as L, to consist of vertical and horizontal segments. The horizontal segments are chosen to maintain a distance of $\log\log T$ from $\mathcal{D}'$, while the vertical segments are selected to maintain a distance of ϵ^2 if $\sigma \le 1-\epsilon$ and a distance of $(\log T)^{-1}$ if $1-\epsilon < \sigma$. It remains to bound $1/\zeta(s)$ on L. If $s \in L$ and $\sigma \leqq 1-\epsilon+\epsilon^2$ we can use Borel-Carathodory and Hadamard's three circle theorems to find a bound for $1/\zeta(s)$ (See [21]; pp. 282-283). If, on the other hand, we have $s \in L$ and $1-\varepsilon < \sigma_j \leqq \sigma < \sigma_{j+1}$, then we use that $|1/\zeta(s)| \ll |G_j(s)|$ in order to obtain a bound for $1/\zeta(s)$. Again we refer the reader to [12] for details. Following this procedure we can show that

$$\frac{1}{\zeta(s)} \ll \exp(c\sqrt{\epsilon}(1-\sigma)\log T)(\log T)^4$$

for any $s = \sigma + it \in L$.

The integral over the horizontal segments will be $\ll$ the integral over the vertical segments. Each vertical segment will either have real part $\sigma_j + \epsilon^2$ if $\sigma_j \le 1-\epsilon$ or real part $\sigma_j + (\log T)^{-1}$ if $\sigma_j > 1-\epsilon$. Using our estimates for the numbers of $\Delta(j,k) \in W$ we can see that the length of the contour with real part $\sigma_j + \epsilon^2$ is

$$\ll (\log T) T^{12(1-\sigma_j)/5} (\log T)^9$$

if $\sigma_j \le 1-\epsilon$. Similarly the length of the contour with real part $\sigma_j + (\log T)^{-1}$ is

$$\ll (\log T) T^{c(1-\sigma_j)^{3/2}} (\log T)^{16}$$

$$\ll T^{c\sqrt{\epsilon}(1-\sigma_j)} (\log T)^{17}$$

if $\sigma_j > 1-\epsilon$. We therefore have that

$$\begin{aligned}
\sum_{x<n\le x+H} \mu(n) &\ll H \int_L \left|\frac{1}{\zeta(s)}\right| x^{\sigma-1} |ds| + \frac{x\log x}{T} \\
&\ll H(\log T)^{14} \sum_{j=0}^{J_1} \exp(c\sqrt{\epsilon}(1-\sigma_j-\epsilon^2)\log T)\exp(-(1-\sigma_j-\epsilon^2)\log x)\exp(c\sqrt{\epsilon}(1-\sigma_j)\log T) \\
&\quad + H(\log T)^{23} \sum_{j=J_1+1}^{J} \exp(c\sqrt{\epsilon}(1-\sigma_j-\epsilon^2)\log T)\exp(-(1-\sigma_j-(\log T)^{-1}))\log x)\exp(c\sqrt{\epsilon}(1-\sigma_j)\log T) \\
&\quad + \frac{x\log x}{T} \\
&\ll Hx^{\varepsilon^2}(\log T)^{14} \sum_{j=0}^{J_1} \exp\left((1-\sigma_j)\left((12/5 + c\sqrt{\varepsilon})\log T - \log x\right)\right) \\
&\quad + H(\log T)^{23} \sum_{j=J_1+1}^{J} \exp\left((1-\sigma_j)\left(2c\sqrt{\varepsilon}\log T - \log x\right)\right) \\
&\quad + \frac{x\log x}{T}
\end{aligned}$$

where $J_1 = \lfloor (1/2 - \epsilon) \log T \rfloor$. If we now let

$$T = \exp((5/12)(1 - c\sqrt{\epsilon}/12) \log x) = x^{5/12 - c\sqrt{\epsilon}/144}$$

we can see that

$$\sum_{x < n \leq x+H} \mu(n) = o(H) + O(x^{7/12 + c\sqrt{\epsilon}/144} \log x) = o(H)$$

provided $H \geq x^{7/12+\epsilon}$.

□

Note that $7/12 = 1 - 5/12$ and the "5/12" comes from Huxley's zero density estimate. In general if

$$N(\alpha, T) \ll T^{A(1-\alpha)} (\log T)^B$$

then we will have

$$\sum_{x < n \leq x+H} \mu(n) = o(H)$$

for $H \geq x^{1-1/A+\epsilon}$. This shows that if the so called "density hypothesis" is true and $A = 2$, then our result will hold true for $H \geq x^{1/2+\epsilon}$.

6 Some Sieve Theory

6.1 Background

Consider the following naive approach to the problem of counting primes less than a given number x:

Given a composite number $n \leq x$, there must exist a prime $p \leq \sqrt{x}$ such that $p \mid n$. On the other hand, given a prime p there are $\lfloor x/p \rfloor$ positive integers $n \leq x$ which it divides. Our idea for counting the number of primes less than x will be to subtract 1 from x for every $n \leq x$ which is divisible by $p \leq \sqrt{x}$. This would give us

$$\pi(x) - \pi(\sqrt{x}) + 1 = x - \sum_{p \leq \sqrt{x}} \sum_{\substack{n \leq x \\ p \mid n}} 1 = x - \sum_{p \leq \sqrt{x}} \left\lfloor \frac{x}{p} \right\rfloor$$

where the +1 on the left had side comes from the fact that this method will count 1 as prime. Unfortunately this is not correct as any number less than x with exactly two distinct prime factors will cause us to subtract twice as much as we should have. We shall therefore add 1 to our guess for every n which is divisible by the product of two distinct primes, giving us

$$\pi(x) - \pi(\sqrt{x}) + 1 = x - \sum_{p \leq \sqrt{x}} \left\lfloor \frac{x}{p} \right\rfloor + \sum_{\substack{p_1, p_2 \leq \sqrt{x} \\ p_1 \neq p_2}} \left\lfloor \frac{x}{p_1 p_2} \right\rfloor$$

However, again this will not be correct as now for every number with exactly 3 prime factors we will have subtracted 3 in the first sum and added 3 in the second sum. We therefore must now subtract 1 for every $n \leq x$ divisible by three distinct primes, and so on. Continuing like this we will eventually arrive at the correct formula of

$$\pi(x) - \pi(\sqrt{x}) + 1 = \sum_{d \mid \mathcal{P}(\sqrt{x})} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor$$

where $\mathcal{P}(z) = \prod_{p<z} p$. We shall now apply the obvious estimate of $\lfloor n \rfloor = n + O(1)$, giving us

$$\pi(x) - \pi(\sqrt{x}) + 1 = x \sum_{d \mid \mathcal{P}(\sqrt{x})} \frac{\mu(d)}{d} + O\left(\sum_{d \mid \mathcal{P}(\sqrt{x})} 1 \right)$$

$$= x \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p}\right) + O\left(2^{\pi(\sqrt{x})}\right)$$

By Merten's products we have that

$$x \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p}\right) \asymp \frac{x}{\log x}$$

We see that the main term is roughly what we would expect due to Merten's prime number theorem. However, the major problem with this approach is the exponentially growing error term which is much larger than any reasonable estimate for the primes.

The previous argument, known as the sieve of Eratosthenes, is not a very effective method for counting primes due to the large error. However, similar techniques can be employed to study sets which are smaller or less dense than the primes; e.g. twin primes, primes of the form n^2+1 and primes on short intervals. The main idea of sieve theory is to restrict the support of $\mu(d)$ in a way which reduces the error while keeping the main term relatively accurate. If we do this carefully we achieve the fundamental lemma of the sieve (lemma 6.1). For more details on sieve theory see [4], [3].

6.2 Fundamental Lemma

Lemma 6.1 (Fundamental Lemma of the Sieve)**.** *$\forall\, \kappa > 0$, $y > 1$ $\exists$ $\Lambda^+ = (\lambda_d^+)_{d \geq 1}$, $\Lambda^- = (\lambda_d^-)_{d \geq 1}$ sequences of real numbers such that:*

- $\lambda_1^{\pm} = 1$
- $|\lambda_d^{\pm}| \leq 1$
- $\lambda_d^{\pm} = 0 \quad \forall d \geq y$
- $\forall n > 1$
$$\sum_{d|n} \lambda_d^- \leq 0 \leq \sum_{d|n} \lambda_d^+ \iff \sum_{d|n} \lambda_d^- \leq \sum_{d|n} \mu(d) \leq \sum_{d|n} \lambda_d^+$$

Moreover, for any multiplicative function $g(d)$ with $0 \leq g(p) < 1$ and

$$\prod_{w \leq p < z} (1 - g(p))^{-1} \leq \left(\frac{\log z}{\log w} \right)^{\kappa} \left(1 + \frac{K}{\log w} \right)$$

we have that $\forall z, w$ such that $2 \leq w < z \leq y$,

$$\sum_{d|P(z)} \lambda_d^{\pm} g(d) = \left(1 + O_{\kappa} \left(e^{-s} \left(1 + \frac{K}{\log z} \right)^{10} \right) \right) \prod_{p<z} (1 - g(p))$$

where

$$P(z) = \prod_{p<z} p \qquad s = \frac{\log y}{\log z}$$

Proof. See lemma 6.3 from [10].

□

Here κ is called the dimension of the sieve. When $\kappa = 1$ we call the combination of Λ^+ and Λ^- "the linear sieve". The linear sieve is the only sieve we shall need in this dissertation. More specifically we will need the following corollary:

Corollary 6.2. *Let $\lambda_d^{\pm}$ be as in 6.1. Then*

$$\sum_{d|n} \mu(d) = \sum_{d|n} \lambda_d^+ + O \left(\sum_{d|n} \lambda_d^+ - \sum_{d|n} \lambda_d^- \right)$$

Proof. Trivial.

□

The usefulness of this statement comes from the fact that

$$1_{(m,n)=1} = \sum_{d|(m,n)} \mu(d) = \sum_{d|(m,n)} \lambda_d^+ + O \left(\sum_{d|(m,n)} \lambda_d^+ - \sum_{d|(m,n)} \lambda_d^- \right).$$

This means that we can estimate the coprimality condition with a sum whose support we can restrict. On its own, it is not obvious how estimating a function that takes the values of 0 and 1 with something that seemingly has an error of $O(1)$ is in anyway useful. However, when $m = \mathcal{P}(P,Q) := \prod_{P<p\leq Q} p$, if we sum the error term over $(x, x+H]$ we see that

$$\sum_{x<n\leq x+H} \sum_{d|(\mathcal{P}(P,Q),n)} \lambda_d^+ - \lambda_d^- = \sum_{d|\mathcal{P}(P,Q)} \sum_{x<kd\leq x+H} \lambda_d^+ - \lambda_d^- \ll H \sum_{d|\mathcal{P}(P,Q)} \left(\frac{\lambda_d^+}{d} - \frac{\lambda_d^-}{d} \right) + O(y)$$

This is something that we can use the linear sieve to bound, as seen in the following corollary:

Corollary 6.3. *Let* $0 < P \leq Q$ *and* $y > 1$*. Then*

$$\sum_{d \mid \mathcal{P}(P,Q)} \left(\frac{\lambda_d^+}{d} - \frac{\lambda_d^-}{d} \right) \ll e^{-s}$$

where

$$s = \frac{\log y}{\log Q}$$

and $\lambda_d^{\pm}$ *are the linear sieve coefficients which are supported on* $[0, y)$*.*

Proof. Application of the fundamental lemma of the sieve with $\kappa = 1$, $z = Q$ and $g(d) = (1/d) \cdot 1_{d \mid \mathcal{P}(P,Q)}$. $\square$

7 The 0.55 result

As we have stated is the introduction, the following argument is due to Matomäki and Teräväinen [12].

Sections $7.1-7.4$ is based on section 3.1 in [12]. We have split this section up in order to illustrate where various error terms come from. We have also filled in many of the details which were omitted in [12] such as the application of the fundamental lemma of the sieve and Shiu's bound.

In section 7.5 we prove Heath-Brown's identity as is done in section 4.1 of [12]. We shall also give a motivating discussion on why Heath-Brown's identity is useful here.

In sections 7.6-7.9 we shall apply the analytic techniques we described in section 3.2. This is outlined in section 4.2 of [12], however, many of the details were omitted.

In section 7.10 we have demonstrated how the lemma from Heath-Brown and Iwaniec is applied to our sum. Most of these details were also omitted in the original paper.

Finally in sections 7.10 and 7.11, we shall bring all of the results together to prove the desired statement. We shall see in these sections why the reasons for many of the parameter choices made in [12]

7.1 Ramaré's Identity

Ramaré's identity will play a key role in the proof of the 0.55 result. It allows us to replace the truncated Möbius function with a useful double sum.

Lemma 7.1 (*Ramaré's identity*)**.** *Suppose* $0 < P \leq Q$ *and let* $n \in \mathbb{Z}_{>0}$*. Then*

$$\mu(n)1_{\left(n,\Pi_{P<p\leq Q}p\right)>1} = \sum_{P<p\leq Q}\sum_{pm=n}\frac{\mu(p)\mu(m)}{\omega_{(P,Q]}(m)+1} + O\left(1_{\exists p\in(P,Q]:\, p^2|n}\right) \tag{7.1}$$

where $\omega_{(P,Q]}(m)$ *is the number of distinct prime divisors of* m *on* $(P,Q]$

Proof. ([12], Section 3.1)If $(n,\Pi_{P<p\leq Q}p)=1$ then the *left-hand side* is zero, and the sum on the *right-hand side* is empty; hence, the result trivially follows.
Assume $(n,\Pi_{P<p\leq Q}p)>1$. Suppose $\forall p\in(P,Q],\ p^2\nmid n$. Then, if $n=pm$ for some $p\in(P,Q]$, we must have $(p,m)=1$ so that $\mu(n)=\mu(p)\mu(m)$ and $\omega_{(P,Q]}(n)=\omega_{(P,Q]}(m)+1$. Thus

$$\mu(n) = \sum_{P<P\leqslant Q}\sum_{pm=n}\frac{\mu(n)}{\omega_{(P,Q]}(n)} = \sum_{P<p\leq Q}\sum_{pm=n}\frac{\mu(p)\mu(m)}{\omega_{(P,Q]}(m)+1}$$

which is what we want to show. If $\exists p\in(P,Q]\colon p^2\mid n$, then $\mu(n)=0$ and the error term is $O(1)$. For each m in the sum we have $\omega_{(P,Q]}(m)\geq\omega_{(P,Q]}(n)-1$ as m has at most one fewer distinct prime divisors then n. Hence

$$\left|\sum_{P<p\leq Q}\sum_{pm=n}\frac{\mu(p)\mu(m)}{\omega_{(P,Q]}(m)+1}\right| \leq \sum_{P<p\leq Q}\sum_{pm=n}\frac{1}{\omega_{(P,Q]}(n)} = 1$$

as there are $\omega_{(P,Q]}(n)$ terms in the sum. Therefore the identity we wish to prove reads as

$$0 = 1 + O(1)$$

which is trivially true.

□

Note that $\mu(p)=-1$ for all p, hence the summand is independent of p. We therefore see that if we sum over $x<n\leq x+H$ and interchange the order of summation, we shall achieve terms of the form $\sum_{A<p\leq B}1$. These sums' corresponding Dirichlet series will be of the same form as $P(s)$ in 3.19. As a result, when we eventually use Perron's formula to estimate our sum, we shall get a factor of $(\log x)^{-A}$ where we can make A as large as we want. This shall allow us to eliminate any factors of $\log x$ that arise in our estimates by simply increasing A.

We would now like to apply Ramaré's identity to (1.1). Unfortunately we must first deal with the problematic terms when $\mu(n)\neq\mu(n)1_{\left(n,\Pi_{P<p\leq Q}p\right)>1}$. We will deal with these terms by sieving them out.

Lemma 7.2. *Suppose* $0 < P \leq Q$ *and* $H, x > 0$. *Then*

$$\sum_{x<n\leq x+H} \mu(n) = \sum_{x<n\leq x+H} \mu(n) 1_{(n, \Pi_{P<p\leq Q} p)>1} + \left(\frac{H \log(P)}{\log(Q)} \right) + O(Q) \tag{7.2}$$

Proof. ([12], Section 3.1) We first split the sum in two,

$$\sum_{x<n\leq x+H} \mu(n) = \sum_{\substack{x<n\leq x+H \\ \exists p \in (P,Q]: p|n}} \mu(n) + \sum_{\substack{x<n\leq x+H \\ p|n \implies p \notin (P,Q]}} \mu(n)$$

The first sum is the same as the main term in (7.2). The second sum can be bounded using the fundamental lemma of the sieve with $y = z = Q$ as follows:

$$\sum_{\substack{x<n\leq x+H \\ p|n \implies p \notin (P,Q]}} \mu(n) \ll \sum_{\substack{x<n\leq x+H \\ p|n \implies p \notin (P,Q]}} 1$$

$$= H - \sum_{\substack{x<n_1p_1\leq x+H \\ p_1 \in (P,Q]}} 1 + \sum_{\substack{x<n_1p_1, n_2p_2\leq x+H \\ p_1, p_2 \in (P,Q]}} 1 - ...$$

$$= H - \left(\sum_{p_1 \in (P,Q]} \left\lfloor \frac{x+H}{p_1} \right\rfloor - \left\lfloor \frac{x}{p_1} \right\rfloor \right) + \left(\sum_{p_1, p_2 \in (P,Q]} \left\lfloor \frac{x+H}{p_1p_2} \right\rfloor - \left\lfloor \frac{x}{p_1p_2} \right\rfloor \right) - ...$$

$$= \sum_{d | P((P,Q])} \mu(d) \left(\left\lfloor \frac{x+H}{d} \right\rfloor - \left\lfloor \frac{x}{d} \right\rfloor \right)$$

$$\ll \sum_{d | P((P,Q])} \lambda_d^+ \left(\left\lfloor \frac{x+H}{d} \right\rfloor - \left\lfloor \frac{x}{d} \right\rfloor \right)$$

$$\ll \sum_{d | P((P,Q])} \lambda_d^+ \left(\frac{H}{d} + 1 \right)$$

$$= H \sum_{d \leq Q} \frac{\lambda_d^+ 1_{d|P((P,Q]}(d)}{d} + O(Q)$$

$$= H \prod_{p \in (P,Q]} \left(1 - \frac{1_{p|P((P,Q]}(p)}{p} \right) + O(Q)$$

$$= \frac{H \prod_{p\leq Q} \left(1 - \frac{1}{p}\right)}{\prod_{p \leq P} \left(1 - \frac{1}{p}\right)} + O(Q)$$

Where $P((P,Q])$ is a product of all the primes on the interval $(P,Q]$. However from Mertens' estimates we know that

$$\prod_{p\leq n} \left(1 - \frac{1}{p}\right) \asymp \frac{1}{\log(n)}$$

hence,

$$\sum_{\substack{x<n\leq x+H \\ p|n\implies p\notin(P,Q]}} \mu(n) = O\left(\frac{H\log(P)}{\log(Q)}\right) + O(Q)$$

proving the result.

□

We can now apply Ramaré's identity to the main term in (7.2).

Lemma 7.3. *Let $H, x > 0$ and suppose $0 < P \leq Q \ll H^{1/2}$. Then*

$$\sum_{x<n\leq x+H} \mu(n)1_{(n,\Pi_{P<p\leq Q}p)>1} = \sum_{\substack{x<pm\leq x+H \\ P<p\leq Q}} \frac{\mu(p)\mu(m)}{\omega_{(P,Q]}(m)+1} + O\left(\frac{H}{P}\right) \tag{7.3}$$

Proof. ([12], Section 3.1) Applying Ramaré's identity we see that

$$\sum_{x<n\leq x+H} \mu(n)1_{(n,\Pi_{P<p\leq Q}p)>1} = \sum_{\substack{x<pm\leq x+H \\ P<p\leq Q}} \frac{\mu(p)\mu(m)}{\omega_{(P,Q]}(m)+1} + \sum_{x<n\leq x+H} O\left(1_{\exists p\in(P,Q]:\, p^2|n}\right).$$

All that remains is to bound the error term, which we can do as follows

$$\begin{aligned}\sum_{x<n\leq x+H} 1_{\exists p\in(P,Q]:\, p^2|n} &= \sum_{\substack{x/p^2<m\leq(x+H)/p^2 \\ p\in(P,Q]}} 1 \\ &\ll \sum_{p\in(P,Q]} \frac{H}{p^2} \\ &\ll \frac{H}{P}\end{aligned}$$

where for the final inequality, we have bounded the sum by the integral. Note that $Q \ll H^{1/2} \implies H/p^2 \gg 1$ which meant we were able to bound the number of integer points in $x/p^2 < m \leq (x+H)/p^2$ by the length of the interval H/p^2. □

7.2 Examining $\omega_{(P,Q]}(m)$

One of the disadvantages of the sum we have on the right-hand side of (7.3) is that the function $\omega_{(P,Q]}(m)$ is pretty difficult to understand. The following lemma will focus on restricting the influence of this function on our sum.

Lemma 7.4. *Let $0 < P \leq Q$, $0 < x, M$ and $0 < \epsilon < 1$. Suppose $x^\epsilon < H \leq x$. Then*

$$\sum_{\substack{x<pm\leq x+H \\ P<p\leq Q}} \frac{\mu(p)\mu(m)}{\omega_{(P,Q]}(m)+1} = \sum_{P<p\leq Q} \sum_{\substack{x<pm_1m_2\leq x+H \\ p'|m_1\implies p'\in(P,Q] \\ p''|m_2\implies p''\notin(P,Q] \\ m_1\leq M}} \frac{\mu(p)\mu(m_1)\mu(m_2)}{\omega_{(P,Q]}(m_1)+1} + O\left(H\log^2(x)2^{-\log M/\log Q}\right) \tag{7.4}$$

Proof. Again, the objective of this lemma is to reduce the influence of $\omega_{(P,Q]}(m)$. To do this we will factor m as $m = m_1m_2$ so that each of m_1's prime factors are in $(P,Q]$ and none of m_2's prime factors are in $(P,Q]$.

Note that this means $\omega_{(P,Q]}(m) = \omega_{(P,Q]}(m_1)$ and hence

$$\sum_{\substack{x<pm\leq x+H \\ P<p\leq Q}} \frac{\mu(p)\mu(m)}{\omega_{(P,Q]}(m)+1} = \sum_{P<p\leq Q} \sum_{\substack{x<pm_1m_2\leq x+H \\ p'|m_1\Longrightarrow p'\in(P,Q] \\ p''|m_2\Longrightarrow p''\notin(P,Q]}} \frac{\mu(p)\mu\left(m_1\right)\mu\left(m_2\right)}{\omega_{(P,Q]}\left(m_1\right)+1}$$

We now shall restrict the support of m_1, and hence the influence of $\omega_{(P,Q]}(m_1)$ on our sum. We shall assume that both m_1 and m_2 are squarefree as otherwise $\mu(m_1)\mu(m_2) = 0$.

Let $\omega(k)$ denote the number of distinct prime factors of k. Now if $p \mid m_1$ then $p < Q$. Hence if $m_1 > M$ we have that

$$M < m_1 \leq Q^{\omega(k)}$$

$$\Longrightarrow \frac{\log M}{\log Q} < \omega(k)$$

For each $x < k \leq x + H$, there are at most $\omega(k)$ ways to write $k = pm_1m_2$. Note that the assumption that m_1 and m_2 implies k is cubefree. Hence, since all the terms in the sum have an absolute value at most 1, we can see that the sum of the terms with $m_1 > M$ is bounded by

$$\sum_{\substack{x<k\leq x+H \\ \omega(k)\geq\log M/\log Q \\ k \text{ is cubefree}}} \omega(k)$$

Note that $2^{\omega(k)} \leq k$ and hence

$$\omega(k) \leq \frac{\log(k)}{\log(2)} \ll \log(x).$$

Therefore

$$\sum_{\substack{x<k\leq x+H \\ \omega(k)\geq\log M/\log Q \\ k \text{ is cubefree}}} \omega(k) \ll \log(x) \sum_{\substack{x<k\leq x+H \\ \omega(k)\geq\log M/\log Q \\ k \text{ is cubefree}}} 1$$

$$\ll \log(x) \sum_{\substack{x<k\leq x+H \\ k \text{ is cubefree}}} 2^{\omega(k)-\log M/\log Q}$$

$$= \log(x) \cdot 2^{-\log M/\log Q} \cdot \sum_{\substack{x<k\leq x+H \\ k \text{ is cubefree}}} 2^{\omega(k)}$$

We shall now apply theorem 2.4 with $a = 1$, $k = 2$, $\alpha = 0.4$, $\beta = \epsilon$ and

$$f(n) = 2^{\omega(n)}.$$

It is easy to see the requirements on a, k, α, and β are trivially satisfied. Clearly f is non negative, multiplicative and

$$f(p^l) = 2^{\omega(p^l)} = 2 \leq 2^l.$$

To prove $f(n) \ll_\epsilon n^\epsilon$ recall

$$\omega(n) \ll \frac{\log n}{\log\log n}$$

$$\implies 2^{\omega(n)} \ll 2^{\log n/(\log\log n)}$$

$$= n^{\log 2/(\log\log n)}$$

$$\ll_\epsilon n^\epsilon$$

which proves f satisfies all required properties. Hence we may conclude

$$\sum_{\substack{x<n\leq x+H \\ n\equiv 1 \bmod 2}} 2^{\omega(n)} \ll H \frac{1}{\log(x+H)} \exp\left(\sum_{\substack{p\leq x+H \\ p\nmid 2}} \frac{2}{p}\right)$$

$$\ll H \frac{1}{\log(x+H)} \exp\left(2\log\log(x+H)\right)$$

$$\ll H \frac{1}{\log(x+H)} \log^2(x+H)$$

$$\ll H\log(x)$$

Although this only bounds the odd terms, we can use it to bound the entire sum as follows:

$$\sum_{\substack{x<k\leq x+H \\ k \textit{ is cubefree}}} 2^{\omega(k)} \ll \sum_{\substack{x<k\leq x+H \\ k\equiv 1 \bmod 2}} 2^{\omega(k)} + \sum_{\substack{x/2<k\leq (x+H)/2 \\ k\equiv 1 \bmod 2}} 2^{\omega(2k)} + \sum_{\substack{x/4<k\leq (x+H)/4 \\ k\equiv 1 \bmod 2}} 2^{\omega(4k)}$$

$$\ll \sum_{m=0}^{2} \sum_{\substack{x/2^m<k\leq (x+H)/2^m \\ k\equiv 1 \bmod 2}} 2^{1+\omega(k)}$$

$$\ll H\log(x) \sum_{m=0}^{2} \frac{1}{2^m}$$

$$\ll H\log(x)$$

So we can finally conclude

$$\sum_{P<p\leq Q} \sum_{\substack{x<pm_1m_2\leq x+H \\ p'|m_1 \implies p'\in(P,Q] \\ p''|m_2 \implies p''\notin(P,Q] \\ m_1>M}} \frac{\mu(p)\mu(m_1)\mu(m_2)}{\omega_{(P,Q]}(m_1)+1} \ll \log(x)\cdot 2^{-\log M/\log Q} H\log(x).$$

The result follows.

$\square$

7.3 Simplifying the sum over m_2

The sum on the right-hand side of (7.4) can be written in the form

$$\sum_{x<pm_1m_2\leq x+H} f(p)g(m_1)h(m_2)$$

where

$$f(p)=\mu(p) \qquad g(m_1)=\frac{\mu(m_1)\,1_{p|m_2\Longrightarrow p\notin(P,Q]}}{\omega_{(P,Q]}(m_1)+1} \qquad h(m_2)=\mu(m_2)1_{\left(m_2,\Pi_{P<p\leq Q}p\right)=1}$$

We see that f is relatively simple, however both g and h are quite complex. It would obviously be better if instead of having two complicated functions we just had one. As we shall see in the next section we have methods for handling this kind of sum when $h(m_2)=\mu(m_2)$. We would therefore like to remove the coprimality condition on m_2. In the proceeding lemma we shall do this by estimating the coprimality condition using the fundamental lemma of the sieve. This will inevitably result in more terms in our sum, and these terms themselves may be challenging to handle. However, we can incorporate these terms into the definition of 'g.' We will then aim to choose the parameters in our sum so that the remaining complex function, 'g,' has limited size and support, thereby minimizing its impact on our sum.

Lemma 7.5. *Let $0<P\leq Q$, $0<x,M$, $y>1$ and $0<\epsilon<1$. Suppose $x^\epsilon<H\leq x$ and $Q^2x^\beta<H$ for some $0<\beta<1/2$. Then*

$$\sum_{P<p\leq Q}\ \sum_{\substack{x<pm_1m_2\leq x+H\\ p'|m_1\Longrightarrow p'\in(P,Q]\\ p''|m_2\Longrightarrow p''\notin(P,Q]\\ m_1\leq M}} \frac{\mu(p)\mu(m_1)\mu(m_2)}{\omega_{(P,Q]}(m_1)+1} = -\sum_{\substack{x<prn\leq x+H\\ P<p\leq Q\\ r\leq My}} a_r\mu(n)+O(E_1) \tag{7.5}$$

where

$$E_1=H\log(Q)\log(M)e^{-s}+yQM+\frac{H\log^3(x)}{P}$$

$$s=\frac{\log y}{\log Q}$$

and a_r satisfies $|a_r|\leq\tau(r)$.

The definition of a_r will be given in the proof however all that matters is its size and support.

Proof. ([12], Section 3.1)From 6.2 we have that

$$1_{(m_2,\mathcal{P}(P,Q))=1}=\sum_{d|(m_2,\mathcal{P}(P,Q))}\mu(d)=\sum_{d|(m_2,\mathcal{P}(P,Q))}\lambda_d^+ +O\left(\sum_{d|(m_2,\mathcal{P}(P,Q))}\lambda_d^+-\sum_{d|(m_2,\mathcal{P}(P,Q))}\lambda_d^-\right)$$

Subbing this into the left-hand side of (7.5) gives us

$$\sum_{P<p\leq Q}\ \sum_{\substack{x<pm_1m_2\leq x+H\\ p'|m_1\Longrightarrow p'\in(P,Q]\\ p''|m_2\Longrightarrow p''\notin(P,Q]\\ m_1\leq M}} \frac{\mu(p)\mu(m_1)\mu(m_2)}{\omega_{(P,Q]}(m_1)+1} = \sum_{\substack{x<pm_1dn\leq x+H\\ P<p\leq Q,m_1\leq M\\ p'|dm_1\Longrightarrow p'\in(P,Q]}} \lambda_d^+\frac{\mu(p)\mu(m_1)\mu(dn)}{\omega_{(P,Q]}(m_1)+1}+O\left(\sum_{\substack{x<pm_1dn\leq x+H\\ P<p\leq Q,m_1\leq M\\ p'|dm_1\Longrightarrow p'\in(P,Q|}}\left(\lambda_d^+-\lambda_d^-\right)\right).$$

To bound the error term we first sum over n and then apply corollary 6.3 as follows

$$\sum_{\substack{x<pm_1dn\leq x+H\\P<p\leq Q,m_1\leq M\\p'|dm_1\Longrightarrow p'\in(P,Q|}}(\lambda_d^+-\lambda_d^-)=\sum_{P<p\leq Q}\sum_{\substack{d\leq y\\m_1\leq M\\p'|dm_1\Longrightarrow p'\in(P,Q|}}(\lambda_d^+-\lambda_d^-)\left(\frac{H}{pdm_1}+O(1)\right)$$

$$\ll\sum_{P<p\leq Q}\sum_{m_1\leq M}\frac{H}{pm_1}\left(\sum_{d|\mathcal{P}(P,Q)}\frac{\lambda_d^+}{d}-\sum_{d|\mathcal{P}(P,Q)}\frac{\lambda_d^-}{d}\right)+\sum_{\substack{P<p\leq Q\\m_1\leq M\\d<y}}1$$

$$\ll\sum_{P<p\leq Q}\sum_{m_1\leq M}\frac{He^{-s}}{pm_1}\quad+\quad yQM$$

$$\ll H\log(Q)\log(M)e^{-s}+yQM$$

If it was the case that $\mu(dn)=\mu(d)\mu(n)$, we would have achieved our goal of splitting the sum into three terms with two simple and one complex. However this is not necessarily true when $(d,n)>1$. We shall proceed to show that the terms for this case are negligible.

To do this we note that if $q\mid d$ then $P<q\leq Q$. Therefore if $k=pm_1dn$ and $(d,n)>1$ then there exists $P<q\leq Q$ such that $q^2\mid k$. Each k appears in the sum at most $\tau_4(k)$ times and we again may assume that k is cubefree. Therefore the contribution to the sum of the terms with $(d,n)>1$ is bounded by

$$\sum_{P<q\leq Q}\sum_{\substack{x<k\leq x+H\\q^2|k\\k\text{ is cubefree}}}\tau_4(k)\ll\sum_{P<q\leq Q}\sum_{\substack{x/q^2<n\leq(x+H)/q^2\\n\text{ is cubefree}}}\tau_4(q^2n)$$

$$\ll\sum_{P<q\leq Q}\sum_{\substack{x/q^2<n\leq(x+H)/q^2\\n\text{ is cubefree}}}\tau_4(n)$$

We now note that τ_4 is a multiplicative function,

$$\tau_4(p^l)\leq l^4\leq 5^l,\quad l\geq 1$$

and it can be shown that

$$\tau_4(n)<<_\epsilon n^\epsilon$$

So we can once again use theorem 2.4 to see that for $L\leq x$ we have

$$\sum_{\substack{x<n\leq x+L\\n\equiv 1\bmod 2}}\tau_4(n)\ll\frac{L}{\log(x+L)}\exp\left(\sum_{p\leq x+L}\frac{4}{p}\right)$$

$$\ll\frac{L}{\log(x+L)}\exp\left(4\log\log(x+L)\right)$$

$$=\frac{L\log^4(x+L)}{\log(x+L)}$$

$$\ll L\log^3 x$$

Similar to the previous application of Shiu's bound, this gives us

$$\sum_{\substack{x<n\leq x+L \\ n \text{ is cubefree}}} \tau_4(n) \ll L\log^3 x$$

We therefore have that

$$\sum_{P<q\leq Q}\ \sum_{x/q^2<n\leq (x+H)/q^2} \tau_4(n) \ll \sum_{P<q\leq Q} \frac{H}{q^2}\log^3(x) \ll \frac{H\log^3(x)}{P}.$$

Note that we were able to apply Shiu's bound here as $x^\beta < H/q^2$. All together we may conclude that

$$\begin{aligned}
&\sum_{P<p\leq Q}\ \sum_{\substack{x<pm_1m_2\leq x+H \\ p'|m_1\implies p'\in(P,Q] \\ p''|m_2\implies p''\notin(P,Q] \\ m_1\leq M}} \frac{\mu(p)\mu(m_1)\mu(m_2)}{\omega_{(P,Q]}(m_1)+1} \\
&= \sum_{\substack{x<pm_1dn\leq x+H \\ P<p\leq Q, m_1\leq M \\ p'|dm_1\implies p'\in(P,Q]}} \lambda_d^+ \frac{\mu(p)\mu(m_1)\mu(d)\mu(n)}{\omega_{(P,Q]}(m_1)+1} + O(E_1) \\
&= -\sum_{P<p\leq Q}\ \sum_{x<prn\leq x+H} \mu(n) \sum_{m_1d=r} \lambda_d^+\mu(d) \frac{\mu(m_1)1_{p|m_1 \implies p\in(p,Q]}1_{m_1\leq M}}{\omega_{(P,Q]}(m_1)+1} + O(E_1)
\end{aligned}$$

where we have defined

$$E_1 = H\log(Q)\log(M) + yQM + \frac{H\log^3(x)}{P}$$

We now let

$$a_r = \sum_{m_1d=r} \lambda_d^+\mu(d)\frac{\mu(m_1)1_{p|m_1 \implies p\in(p,Q]}1_{m_1\leq M}}{\omega_{(P,Q]}(m_1)+1} = (\lambda^+\mu * \mu w)(r)$$

where

$$w(m) = \frac{\mu(m)1_{p|m \implies p\in(p,Q]}1_{m\leq M}}{\omega_{(P,Q]}(m)+1}$$

It remains to examine the support of a_r. Suppose $m_1d = r \geq My$. If $d > y$ then $\lambda_d^+ = 0$, otherwise $d \leq Q^{100\log\log x}$ in which case

$$m_1 = \frac{r}{d} \geq M \implies w(m_1) = 0$$

We conclude $a_r = 0$ when $r \geq My$. Finally we note that

$$|a_r| \leq \sum_{mn=r} |\lambda_m^+\mu(m)\mu(n)w(n)| \leq \sum_{mn=r} 1 = \tau(r)$$

$\square$

Note that we can control the support of a_r by adjusting the size of M and y. This gives us a lot of control over the influence of these complicated terms on our sum.

7.4 Combining the Results

Theorem 7.6. *Let $0 < x, M$, $0 < P \leq Q$, $y > 1$ and $0 < \epsilon < 1$. Suppose $x^\epsilon < H \leq x$ and $Q \ll H^{1/2}$. Then*

$$\sum_{x<n\leq x+H} \mu(n) = -\sum_{\substack{x<prn\leq x+H \\ P<p\leq Q \\ r\leq My}} a_r \mu(n) + O(E)$$

where

$$E = \frac{H\log(P)}{\log(Q)} + Q + \frac{H}{P} + H\log^2(x)2^{-\log M/\log Q} + H\log(Q)\log(M)e^{-s} + yQM + \frac{H\log^3(x)}{P}$$

$$s = \frac{\log y}{\log Q}$$

and a_r satisfies $|a_r| \leq \tau(r)$

Proof. Trivial consequence of the previous four lemmas.

□

Although the error is quite unwieldy, we have plenty of parameters which we can use to control it. We shall forget about it for now and come back to it once we know more about what properties we need our parameters to satisfy.

7.5 Heath-Brown's Identity

Lemma 7.7 (Heath-Brown's identity)**.** *Let $k \geq 1$ be an integer. Then for $x \geq 2^k$ and $n \leq 2x$ we have*

$$\mu(n) = \sum_{1\leq j\leq k} (-1)^{j-1}\binom{k}{j} 1^{(*)(j-1)} * \left(\mu 1_{\left[1,(2x)^{1/k}\right]}\right)^{(*)j}(n)$$

$$= \sum_{1\leq j\leq k} (-1)^{j-1}\binom{k}{j} \sum_{\substack{n_1n_2\ldots n_{2j-1}=n \\ i\geq j \implies n_i\leq(2x)^{1/k}}} \mu(n_j)\mu(n_{j+1})\ldots\mu(n_{2j-1})$$

Proof. ([12], Section 4.1) First we let

$$M(s) = \sum_{m\leq(2x)^{1/2k}} \frac{\mu(m)}{m^s},$$

then by the binomial theorem we have that

$$(1-\zeta(s)M(s))^k = \sum_{j=0}^{k}\binom{k}{j}(-\zeta(s)M(s))^j = 1 - \sum_{j=1}^{k}(-1)^{j-1}\binom{k}{j}\zeta(s)^j M(s)^j.$$

Dividing across by $\zeta(s)$ and moving the sum to the other side gives

$$\frac{1}{\zeta(s)} = \sum_{j=1}^{k}(-1)^{j-1}\binom{k}{j}\zeta(s)^{j-1}M(s)^j + \frac{1}{\zeta(s)}(1-\zeta(s)M(s))^k.$$

We now look at the Dirichlet series expansion of both sides of this expression when $\Re(s) > 1$. The n'th coefficient of $\frac{1}{\zeta(s)}$ is $\mu(n)$ and the n'th coefficient of the sum is the expression on the right-hand side of the identity we are trying to prove. It remains to show that the $n'th$ coefficient of the last term is zero when $n \leq 2x$. To do this we note that

$$\zeta(s)M(s) = \sum_{n\geq 1} \frac{(1 * \mu 1_{[1,(2x)^{1/k}]})(n)}{n^s}$$

$$= \sum_{1\leq n\leq (2x)^{1/k}} \frac{(1*\mu)(n)}{n^s} + \sum_{(2x)^{1/k}<n} \frac{(1 * \mu 1_{[1,(2x)^{1/k}]})(n)}{n^s}$$

$$= 1 + \sum_{(2x)^{1/k}<n} \frac{(1 * \mu 1_{[1,(2x)^{1/k}]})(n)}{n^s},$$

hence we see that $(1 - \zeta(s)M(s))^k$ is of the form

$$\sum_{2x<n} \frac{a_n}{n^s}$$

finishing the proof.

□

The advantage of Heath-Brown's identity is that it replaces $\mu(n)$ with a sum of products of Möbius functions whose support is restricted to $n \leq (2x)^{1/k}$. This means when we estimate our sum using Dirichlet polynomials, they will have a shorter length. In practice we will take products of subsets of the $2k-1$ Dirichlet polynomials given to us by Heath-Brown's identity, which will give us Dirichlet polynomials of certain required lengths. In summary, Heath-Brown's identity is useful, as it will give rise to Dirichlet polynomials over whose length we shall have a certain amount of control.

Inserting Heath-Brown's identity into 7.6 gives us the following theorem.

Theorem 7.8. *Let everything be as in theorem 7.6. Then*

$$\sum_{x<n\leq x+H} \mu(n) = \sum_{j=1}^{k} (-1)^{j-1}\binom{k}{j} \sum_{\substack{x<prn_1\cdots n_{2j-1}\leq x+H \\ P<p\leq Q \\ i\geq j \Longrightarrow n_i \leq (2x)^{1/k}}} a_r \mu(n_j)\cdots\mu(n_{2j-1}) + O(E)$$

$$= \sum_{j=1}^{k} (-1)^{j-1}\binom{k}{j} \sum_{\substack{x<prn_1\cdots n_{2j-1}\leq x+H \\ P<p\leq Q}} a_r a_1(n_1)a_2(n_2)\cdots a_{2k-1}(n_{2k-1}) + O(E)$$

where

$$a_i(n) = \begin{cases} 1 & 1 \leq i < j \\ 1_{n=1}(n) & j \leq i \leq k \\ \mu(n)1_{n\leq (2x)^{1/k}}(n) & k < i \leq k+j \\ 1_{n=1}(n) & k+j < i \leq 2k-1 \end{cases}$$

Proof. Result follows by inserting Heath-Brown's identity into 7.6 and then introducing some dummy variables.

□

We have introduced dummy variables in order to keep the number of variables in each summand the same for each value of j. Note that technically we should write $a_{i,j}(n)$ however the form of $a_i(n)$ will not matter too much and hence we shall treat each summand the same regardless of the value of j.

7.6 A Note on Notation

In order to keep our sums somewhat tidy we shall introduce a small bit of notation. Specifically we define

$$\sum_{n\sim N} x_n := \sum_{AN<n\le BN} x_n$$

where $0 < A < B$ are fixed positive constants or are bounded by a fixed positive constant. If A and B ever seem to depend on variables it is because these variables will be fixed in the future. Mainly we shall find A and B might depend on k, however we shall fix $k = 20$ later so this does not matter. We do this as the size of A and B will not matter for our purposes so there is no point in calculating them explicitly.

7.7 Strategy for Applying Analytic Techniques

We wish to form an estimate for

$$\sum_{x<n\le x+x^\theta} \mu(n)$$

where θ is close to $1/2$. To do this we shall compare it to

$$\sum_{x<n\le x+y_1} \mu(n)$$

where $y_1 = x\exp\left(-3(\log x)^{1/3}\right)$. We could choose y_1 to be smaller as in 4, however we have decided to stick with the choice of the authors in We would like to apply (3.10) to the two sums, however, if we attempted to do this directly we would find that we do not have adequate control over the Dirichlet series and polynomials which appear in the calculation. Instead we shall apply 7.8 with $H = x^\theta$ and $H = y_1$. We shall then apply (3.10) to the two resulting sums. We shall not however apply (3.3) to the entire sum at once. We shall first split each sum into dyadic ranges. We shall then use (3.10) to estimate the difference of each sum over their corresponding dyadic ranges.

We proceed by dividing the sum into the following $\ll (\log x)^{2k+1}$ dyadic ranges

$$\sum_{\substack{x<prn_1\cdots n_{2k}-1\le x+H\\ p\in(P_1,2P_1],r\in(R,2R],n_i\in(N_i,2N_i]\\ p\le Q}} a_r a_1(n_1)\cdots a_{2k-1}(n_{2k-1})$$

where

$$\begin{aligned} &P_1\in[P,Q],\quad R\in[1/2,yM],\quad N_1,\ldots,N_{k-1}\in[1/2,x],\\ &N_k,\ldots,N_{2k-1}\in\left[1/2,(2x)^{1/k}\right],\quad P_1RN_1\cdots N_{2k-1}\asymp x\end{aligned}$$

Additionally we note that $P_1RN_1N_2...N_{j-1}N_kN_{k+1}...N_{2j-1}\asymp x$. Therefore $N_i\asymp 1$ when $j<i<k$ or $2j-1<i\le 2k-1$. We shall, however, not need this fact until section 7.10

For convenience we will define

$$D(H,y_1,P_1,R,N_1,...,N_{2k-1}) := \sum_{\substack{x<prn_1\cdots n_{2k-1}\le x+H\\ p\in(P_1,2P_1],r\in(R,2R],n_i\in(N_i,2N_i]\\ p\le Q}} a_r a_1(n_1)\cdots a_{2k-1}(n_{2k-1}) - \frac{H}{y_1}\sum_{\substack{x<prn_1\cdots n_{2k-1}\le x+y_1\\ p\in(P_1,2P_1],r\in(R,2R],n_i\in(N_i,2N_i]\\ p\le Q}} a_r a_1(n_1)\cdots a_{2k-1}(n_{2k-1})$$

It will be our objective to show that

$$D(H,y_1,P_1,R,N_1,...,N_{2k-1}) \ll_A \frac{H}{(\log x)^A}$$

7.8 Applying Perron's formula

First we define

$$P(s) = \sum_{P_1 < p \le \min\{2P_1, Q\}} p^{-s} \qquad R(s) = \sum_{R < r \le 2R} a_r r^{-s}$$

$$N_i(s) = \sum_{N_i < n \le 2N_i} a_i(n) n^{-s}$$

Furthermore, if $I \subset [2k-1]$, then we define $N_I = \prod_{i \in I} N_i$ and similarly $N_I(s) = \prod_{i \in I} N_i(s)$. We would like to apply Corollary 3.10; however, to do this, we need an estimate for the tail of the integral. Therefore, we must find suitable estimates for the Dirichlet polynomials above.

Lemma 7.9. *$\forall\, , t \in \mathbb{R}$, $\kappa > 0$ and $R > 2$ we have that*

$$|R(\kappa + it)| \ll R^{1-\kappa} \log x$$

Proof.

$$|R(\kappa + it)| \ll \sum_{r \in (R, 2R]} \frac{|a_r|}{r^\kappa} \ll \frac{1}{R^\kappa} \sum_{r \in (R, 2R]} \tau(r) \ll \frac{R \log R}{R^\kappa} \ll R^{1-\kappa} \log x$$

□

We shall from now on assume $P \ge \exp\left((\log x)^{2/3+\epsilon_1}\right)$ for some $\epsilon_1 > 0$. We do this so we can apply lemma 3.19 to bound $P(\kappa + it)$.

Lemma 7.10. *Let $T_0 \gg_A (\log x)^A$. If $T_0 \le |t| \le x$ then for any $\kappa > 0$ we have*

$$|P(\kappa + it)| \ll_A \frac{P_1^{1-\kappa}}{(\log x)^A}$$

Proof. If $|t| \gg_A (\log x)^A$ and $S \ge \exp\left((\log x)^{2/3+\epsilon_1}\right)$ then

$$\sum_{S < p \le \min\{Q, 2P_1\}} \frac{1}{p^{1+it}} \ll_A \frac{\log x}{1 + |t|} + (\log x)^{-A} \ll \frac{\log x}{(\log x)^A} + (\log x)^{-A} \ll (\log x)^{-A},$$

where we have dropped a factor of $\log x$ by adjusting A and the implied constant. Then by partial summation we have

$$P(\kappa + it) = \sum_{P_1 < p \le \min\{Q, 2P_1\}} \frac{p^{1-\kappa}}{p^{1+it}} \ll P_1^{1-\kappa} \max_{p_1 \le S \le \min\{Q, 2P_1\}} \left\{ \left| \sum_{S < p \le \min\{Q, 2P_1\}} \frac{1}{p^{1+it}} \right| \right\} \ll_A \frac{P_1^{1-\kappa}}{(\log x)^A}$$

□

Lemma 7.11. *For any $I \subset [2k-1]$ and $0 < T_0 < T$, we have that*

$$\int_{\kappa + iT_0}^{\kappa + iT} |N_I(s)|^2 \, ds \ll (T + N_I) N_I^{1-2\kappa} (\log x)^B$$

where $B > 0$ is a constant that only depends on k.

Proof. First note that

$$N_I(\kappa + it) = \sum_{n = N_I}^{2^{2k-1} N_I} \frac{c_n n^{-\kappa}}{n^{it}} \qquad \text{where} \qquad |c_n| \le \tau_{2k-1}(n)$$

Applying the mean value theorem (3.20) gives us

$$\int_{\kappa+iT_0}^{\kappa+iT} |N_I(s)|^2 \, ds \ll (T+N_I) \sum_{n=N_I}^{2^{2k-1}N_I} \frac{|c_n|^2}{n^{2\kappa}}$$

$$\ll (T+N_I) N_I^{-2\kappa} \sum_{n=N_I}^{2^{2k-1}N_I} \tau_{2k-1}(n)^2$$

$$\ll (T+N_I) N_I^{1-2\kappa} (\log x)^B$$

□

Theorem 7.12. *Let $\kappa \in \mathbb{R}$, $0 < T_0 \le T$ and $[n] = I \cup J$ is a partition of $[n]$. Suppose that $\forall\ A > 0$ we have $T_0 \gg_A (\log x)^A$. Then*

$$\int_{T_0}^{T} |P(\kappa+it)\, R(\kappa+it)\, N_I(\kappa+it)\, N_J(\kappa+it)| \, dt \ll_A \frac{x^{1-\kappa}}{(\log x)^A} \left(\frac{T}{N_I} \cdot \frac{T}{N_J} + \frac{T}{N_J} + \frac{T}{N_I} + 1 \right)^{1/2}$$

Proof. ([12], Lemma 4.3) As $[n] = I \cup J$ is a partition, we have that $P_1 R N_I N_J \asymp x$. In what follows we shall use this fact, along with our Dirichlet polynomial estimates and the Cauchy Schwarz inequality, to prove the result.

$$\int_{T_0}^{T} |P(\kappa+it)\, R(\kappa+it)\, N_I(\kappa+it)\, N_J(\kappa+it)| \, dt$$

$$\ll_A \frac{(P_1R)^{1-\kappa}\log x}{(\log x)^A} \int_{T_0}^{T} |N_I(\kappa+it)\, N_J(\kappa+it)\,|$$

$$\ll_A \frac{(P_1R)^{1-\kappa}}{(\log x)^{A-1}} \left(\int_{T_0}^{T} |N_I(\kappa+it)|^2 dt \int_{T_0}^{T} |N_J(\kappa+it)|^2 dt \right)^{1/2}$$

$$\ll_A \frac{(\log x)^B}{(\log x)^{A-1}} \left((P_1R)^{2-2\kappa} \left((T+N_I) N_I^{1-2\kappa} \right) \left((T+N_J) N_J^{1-2\kappa} \right) \right)^{1/2}$$

$$\ll_A \frac{x^{(1/2)-\kappa}}{(\log x)^A} \left(P_1RT^2 + P_1RTN_I + P_1RTN_J + P_1RN_IN_J \right)^{1/2}$$

$$\ll_A \frac{x^{(1/2)-\kappa}}{(\log x)^A} \left(\frac{xT^2}{N_IN_J} + \frac{xT}{N_J} + \frac{xT}{N_I} + x \right)^{1/2}$$

$$\ll_A \frac{x^{1-\kappa}}{(\log x)^A} \left(\frac{T}{N_I} \cdot \frac{T}{N_J} + \frac{T}{N_J} + \frac{T}{N_I} + 1 \right)^{1/2}$$

Note that here we have replaced $A - 1 - B$ with A and then adjusted the implied constant appropriately.

□

All of the Dirichlet polynomials above have real coefficients and therefore

$$|P(s)R(s)N_I(s)N_J(s)| = |P(\overline{s})R(\overline{s})N_I(\overline{s})N_J(\overline{s})|$$

This means integral over $[-T, -T_0]$ will have the same bound as the integral over $[T_0, T]$. Therefore we have a formula for $B_\kappa(x, T)$ in corollary 3.10, however, we still need to find a formula for the bound $A_\kappa(x)$. This is relatively easy after we note that we may write

$$P(\kappa + it)\, R(\kappa + it)\, N_I(\kappa + it)\, N_J(\kappa + it) = \sum_{n \sim x} \frac{f_n}{n^{\kappa + it}}$$

for some $|f_n| \leq \tau_{2k+1}(n)$. Therefore

$$\sum_{n \sim x} \frac{|f_n|}{n^\kappa} \leq x^{-\kappa} \sum_{n \sim x} \tau_{2k+1}(n) \ll x^{1-\kappa} (\log x)^B$$

for some $B > 0$. We can now apply corollary 3.10 to get the following.

Theorem 7.13. *Let $\kappa \in \mathbb{R}$, $0 < T_0 \leq T \ll x$ and $[n] = I \cup J$ is a partition of $[n]$. Suppose that $\forall\ A > 0$ we have $(\log x)^A \ll T_0 \ll x/y_1$. Then*

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) \ll_{A,\eta} \frac{x^{1+\eta}}{T} + H(\log x)^B T_0^2 \exp(-3(\log x)^{1/3})$$

$$+ \frac{H}{(\log x)^A} \left(\frac{T}{N_I} \cdot \frac{T}{N_J} + \frac{T}{N_J} + \frac{T}{N_I} + 1 \right)^{1/2}$$

for some $B > 0$.

Proof. ([4], Lemma 7.2) We apply corollary 3.10 with $A_\kappa(x) = x^{1-\kappa}(\log x)^B$, $\psi(n) = \tau_{2k+1}(n) \ll_\eta n^\eta$ and

$$B_\kappa(x) = \frac{x^{1-\kappa}}{(\log x)^A} \left(\frac{T}{N_I} \cdot \frac{T}{N_J} + \frac{T}{N_J} + \frac{T}{N_I} + 1 \right)^{1/2}.$$

This gives us

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) \ll_A \frac{x^\kappa x^{1-\kappa} (\log x)^B}{T} + x^{\kappa-2} H y_1 x^{1-\kappa} (\log x)^B T_0^2 + \frac{x \tau_{2k+1}(x) \log x}{T}$$

$$+ x^{\kappa-1} H \frac{x^{1-\kappa}}{(\log x)^A} \left(\frac{T}{N_I} \cdot \frac{T}{N_J} + \frac{T}{N_J} + \frac{T}{N_I} + 1 \right)^{1/2}$$

$$\ll_{A,\eta} \frac{x(\log x)^B}{T} + H y_1 x^{-1} (\log x)^B T_0^2 + \frac{x^{1+\eta} \log x}{T}$$

$$+ \frac{H}{(\log x)^A} \left(\frac{T}{N_I} \cdot \frac{T}{N_J} + \frac{T}{N_J} + \frac{T}{N_I} + 1 \right)^{1/2}$$

Note that the third term dominates the first term and we can drop the factor of $\log x$ in the third term by adjusting η. The result follows after subbing in $y_1 = x \exp(-3(\log x)^{1/3})$ and then simplifying.

$\square$

Note that κ does not feature in our final result. We therefore could use any vertical line of integration to obtain the bound.

7.9 Choosing T and T_0

From now on we shall assume that $H = x^{1/2+h+\epsilon}$, where we think of h as being a small positive constant (in our case it will be 0.05) and $\epsilon > 0$ will be arbitrarily small. The result for larger H follows by splitting the sum into shorter sums.

At the very least our bound for D must be $o(H)$. Therefore

$$H(\log x)^B T_0^2 \exp(-3(\log x)^{1/3}) = o(H)$$

$$\implies T_0 = o\left((\log x)^{B/2} \exp\left(\frac{3}{2}(\log x)^{1/3}\right)\right)$$

We must also have that $T_0 \ll x/y_1$, however, the above bound is stronger. Finally we have the lower bound $T_0 \gg_A (\log x)^A$. Note that the lower bound for T_0 comes from the lower bound for P which in turn comes from the Vinogradov-Korobov zero free region. The upper bound for T_0 comes from y_1, which also comes from the Vinogradov-Korobov zero free region. Increasing the size of the zero free regions will simultaneously increase both the lower and upper bounds of T_0.
If $T_0 = \exp((\log x)^{1/3})$ then for any $A > 0$ we shall have $T_0 \gg_A (\log x)^A$ and

$$H(\log x)^B T_0^2 \exp(-3(\log x)^{1/3}) = (\log x)^B \exp(-(\log x)^{1/3}) \ll_A \frac{H}{(\log x)^A}.$$

In lieu of the previous bound we would like to have that $\forall\ A > 0$,

$$\frac{x^{1+\eta}}{T} \ll_A \frac{H}{(\log x)^A} \iff \frac{x^{1+\eta}(\log x)^A}{H} \ll_A T$$

where as usual we have adjusted A appropriately. Note that the remaining error term is increasing in T and so we would like T to be as small as possible and therefore we shall set

$$T = \frac{x^{1+2\eta}}{H}$$

The remaining error term will be $\ll H(\log x)^{-A}$ if

$$\frac{T}{N_I}, \frac{T}{N_J} \ll 1 \iff \frac{x^{1+2\eta}}{HN_I}, \frac{x^{1+2\eta}}{HN_J} \ll 1$$

$$\iff \frac{x^{1+2\eta}}{H} \ll N_I, N_J$$

$$\iff x^{(1/2)-h+2\eta-\epsilon} \ll N_I, N_J$$

The restriction on N_I, N_J implies that $N_I \ll x/N_J \ll x^{(1/2)+h-2\eta+\epsilon}$ and therefore $x^{(1/2)-h+2\eta-\epsilon} \ll N_I \ll x^{(1/2)+h+\epsilon'}$. This condition is necessary but not sufficient. It would be nice if we could find a similar kind of condition that was sufficient as then, instead of finding a partition with the required property above, we could just find a subset with a similar property. We can achieve this by again recalling that $N_I N_J P_1 R \asymp x$. If we therefore limit the size of P_1, R and N_I we will have a lower bound for N_J. More specifically we shall assume $P_1 R \ll x^{\epsilon/4}$ and $N_I \ll x^{(1/2)+h+\epsilon/2}$ giving us

$$N_J \asymp \frac{x}{P_1 R N_I} \gg x^{(1/2)-h-3\epsilon/4}.$$

We therefore let $2\eta - \epsilon = -3\epsilon/4 \iff \eta = \epsilon/8$. In total we have proved the following theorem:

Theorem 7.14. *Suppose $\exists\ I \subset [2k-1]$ such that $N_I \in [x^{(1/2)-h-\epsilon/2}, x^{(1/2)+h+\epsilon/2}]$. Then, $\forall\ A > 0$, we have*

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) \ll_A \frac{H}{(\log x)^A}$$

Proof. Above.

□

7.10 The 0.55 lemma

In light of theorem 7.14, we let $N_i = x^{\alpha_i}$. Due to the assumptions $P_1 R \ll x^{\epsilon/4}$ and $P_1 R N_1 N_2 ... N_{2k-1} \asymp x$, we have that

$$\sum_{i=1}^{2k-1} \alpha_i \in \left[1 - \frac{3\epsilon}{4}, 1\right] \subset (1-\epsilon, 1].$$

If there is a subset $I \subset [n]$ with $\sum_{i \in I} \alpha_i \in [1/2 - h - \epsilon/2, 1/2 + h + \epsilon/2]$, then we have a bound for D. It is therefore worth asking *"what can we say when no such subset exists?"*. The following combinatorial lemma will answer this question.

Lemma 7.15. *Let $0 \leq h \leq 1/2$ and let $n \in \mathbb{Z}_{>0}$. If $\alpha_1 \cdots \alpha_n \in [0,1]$ satisfies:*

$$\sum_{i=1}^{n} \alpha_i \in [1-\epsilon, 1] \subset [1/2 - h, 1]$$

for some $\epsilon > 0$. Then one of the following holds:

(1) $\exists I \subset [n]$ *s.t* $\displaystyle\sum_{i \in I} \alpha_i \in [1/2 - h, 1/2 + h]$

(2) *There is a partition* $[n] = I_1 \cup I_2 \cup \{l\}$ *where* $\displaystyle\sum_{i \in I_j} \alpha_j < 1/2 - h$ *for* $j = 1, 2,$ *and* $\alpha_l > 2h - \epsilon$.

Proof. ([12], Lemma 4.5) We will assume (1) is not true and show this implies (2).
Let $\mathcal{I} = \left\{ I \subset [n] : \sum_{i \in I} \alpha_i \leq 1/2 + h \right\}$ and let $I_1 \in \mathcal{I}$ be a maximal element of $\mathcal{I}$ (Note that if $\mathcal{I} = \emptyset$, then there must be only one non-zero α_i, in which case the result is trivial). Now since we have assumed (1) is not true and $\sum_{i \in I_1} \alpha_i \leq 1/2 + h$ we must have that $\sum_{i \in I_1} \alpha_i < 1/2 - h$.

Now pick any $l \in [n] \setminus I_1$.(We can do this as if no such l existed then we would have $I_1 = [n]$, in which case $\sum_{i=1}^{n} \alpha_i < 1/2 - h$, giving a contradiction.) Due to the maximality of I_1 we have that $\sum_{i \in I_1} \alpha_i + \alpha_l > 1/2 + h$.

Now let $I_2 = [n] \setminus (I_1 \cup \{l\})$. As $\sum_{i=1}^{2k+1} \alpha_i \leq 1$ we have that

$$\sum_{i \in I_2} \alpha_i \leq 1 - \sum_{i \in I_1} \alpha_i - \alpha_l < 1 - 1/2 - h = 1/2 - h$$

Finally, we have

$$\alpha_l \geq 1 - \epsilon - \sum_{i \in I_1} \alpha_i - \sum_{i \in I_2} \alpha_i > 1 - \epsilon - (1/2 - h) - (1/2 - h) = 2h - \epsilon$$

which proves the result.

□

From this we can see that if we are unable to apply theorem 7.14, then we can assume there exists a partition $[n] = I \cup J \cup \{l\}$ such that $N_I, N_J \leq x^{1/2-h}$ and $N_l \geq x^{2h-\epsilon}$. Recall that if $i \geq k$ then $N_i \ll x^{1/k}$. Therefore if $1/k < 2h - \epsilon$ then $l < k$ which would be nice as $a_l(n)$ is simpler when $l < k$. We shall therefore require that $k \geq h^{-1}$ so that

$$\frac{1}{k} \leq h < 2h - \epsilon$$

for ϵ small enough. We must now recall the definition of $a_l(n)$ for $l < k$:

$$a_l(n) = \begin{cases} 1 & 1 \leq l < j \\ 1_{n=1}(n) & j \leq l < k \end{cases}$$

Recall that $N_l \asymp 1$ when $j \le l < k$ and therefore $l < j$ and $a_l(n) = 1$. Given this information about N_l see that D is of the form

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) = \sum_{\substack{x < prn_I n_J n_l \le x+H \\ p \sim P_1,\ r \sim R,\ n_l \sim N_l \\ n_I \sim N_I,\ n_J \sim N_J}} a_r a_I(n_I) a_J(n_J) - \frac{H}{y_1} \sum_{\substack{x < prn_I n_J n_l \le x+y_1 \\ p \sim P_1,\ r \sim R,\ n_l \sim N_l \\ n_I \sim N_I,\ n_J \sim N_J}} a_r a_I(n_I) a_J(n_J)$$

We may drop the condition $n_l \sim N_l$ as this is implied by the other conditions. We now let $U = P_1 N_I$ and $V = RN_J$. Recall that we have assumed $P_1 R \ll x^{\epsilon/4}$, we shall now make the stronger assumption $P_1, R \le x^{\epsilon/8}$. Therefore we have that $U, V \le x^{1/2-h+\epsilon/8}$ giving us

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) = \sum_{\substack{x < uvn_l \le x+H \\ u \sim U,\ v \sim V}} b_u c_v - \frac{H}{y_1} \sum_{\substack{x < uvn_l \le x+y_1 \\ u \sim U,\ v \sim V}} b_u c_v$$

If we sum over the n_l variable first, we shall have for both $y = H$ and $y = y_1$ that

$$\sum_{\substack{x < uvn_l \le x+y \\ u \sim U,\ v \sim V}} b_u c_v = \sum_{u \sim U,\ v \sim V} b_u c_v \left(\left\lfloor \frac{x+y}{uv} \right\rfloor - \left\lfloor \frac{x}{uv} \right\rfloor \right)$$

$$= y \sum_{u \sim U,\ v \sim V} \frac{b_u c_v}{uv} + \sum_{u \sim U,\ v \sim V} b_u c_v r(x, y, uv)$$

where

$$r(x, y, uv) = \left\lfloor \frac{x+y}{uv} \right\rfloor - \left\lfloor \frac{x}{uv} \right\rfloor - \frac{y}{uv} = O(1)$$

On substituting this into our formula for D, we see that the main terms cancel, giving us

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) = \sum_{u \sim U,\ v \sim V} b_u c_v r(x, H, uv) - \frac{H}{y_1} \sum_{u \sim U,\ v \sim V} b_u c_v r(x, y_1, uv)$$

Note that $b_u \le \tau_{2k}(u)$ and $c_v \le \tau_{2k}(v)$, which implies $b_u \ll_{\delta_0} U^{\delta_0}$ and $c_v \ll_{\delta_0} V^{\delta_0}$ for any $\delta_0 > 0$. We therefore have that

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) \ll_{\delta_0} (UV)^{\delta_0} \sum_{u \sim U,\ v \sim V} r(x, H, uv) + \frac{H(UV)^{\delta_0}}{y_1} \sum_{u \sim U,\ v \sim V} r(x, y_1, uv)$$

We can bound the second term by the trivial estimate:

$$\frac{H(UV)^{\delta_0}}{y_1} \sum_{u \sim U,\ v \sim V} r(x, y_1, uv) \ll \frac{H(UV)^{1+\delta_0}}{y_1} = H \frac{x^{2(1/2-h+\epsilon/8)(1+\delta_0)}}{x \exp(-3(\log x)^{1/3})} \ll_A \frac{H}{(\log x)^A}$$

where for the last inequality we have used the fact that we can make both δ_0 and ϵ arbitrarily small w.r.t h, giving us a power saving which is clearly $\ll_A H(\log x)^{-A}$.

The trivial estimate will not be enough for the first term, instead we shall use a lemma from Heath-Brown and Iwaniec [6].

Lemma 7.16. *Let $\eta > 0$, $0.55 + 3\eta < \Theta < 7/12$ and $0 \le \phi \le (6\Theta - 1)/5 - 3\eta$. Suppose $U, V \le x^\phi$, then*

$$\sum_{\substack{U < u \le 2U \\ V < v \le 2V}} b_u c_v r(x, x^\theta, uv) \ll_\delta x^{\Theta - \delta}$$

where $|b_u| \le \tau_k(u)$, $|c_v| \le \tau_k(v)$, for some $k \ge 1$,and $\delta = \delta(\eta) > 0$.

We have made some slight modifications to the original statement. We have replaced $0.55 + 2\eta$ with $0.55 + 3\eta$ as on examination of the proof it seems this was a slight, albeit insignificant, error by the authors. The original statement required that $|b_u|, |c_v| \leq 1$ however the proof generalises to the case in which they are only divisor bounded which we need for our purposes.

Theorem 7.17. *Let $H = x^\theta$ where $\theta = 0.55 + \epsilon$. Suppose there exists a partition $[n] = I \cup J \cup \{l\}$ such that $N_I,\ N_J \leq x^{0.45}$. Then, $\forall\ A > 0$ we have*

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) \ll_{A,\epsilon} \frac{H}{(\log x)^A}.$$

Proof. ([12], Lemma 4.4) We apply lemma 7.16 with $\eta = \epsilon/4$ and $\Theta = \theta = 0.55 + \epsilon$. With these values we see that $0.55 + 3\eta = 0.55 + 3\epsilon/4 < \theta$ and $(6\theta - 1)/5 - 3\eta = 0.46 + 9\epsilon/20 > 0.46$. If $N_I, N_J \leq x^{0.45}$ then $U, V \leq x^{0.45+\epsilon/8} \leq x^{0.46}$, so we have

$$\sum_{u\sim U,\ v\sim V} r(x, x^\theta, uv) \ll_\delta x^{\theta-\delta}$$

where $\delta = \delta(\epsilon) > 0$.

We note that although lemma 7.16 has ranges of summation $U < u \leq 2U$ and $V < v \leq 2V$, we may still apply it on the ranges $u \sim U$ and $v \sim V$ as we can split the sum into a finite number of dyadic ranges.

We may conclude

$$D(H, y_1, P_1, R, N_1, ..., N_{2k-1}) \ll_{A,\epsilon,\delta_0} (UV)^{\delta_0} x^{\theta-\delta(\epsilon)} + \frac{H}{(\log x)^A}.$$

Finally we note that $UV \ll x$, so if we set $\delta_0 \leq \delta$ we shall get a power saving in the first term which clearly is $\ll_A H(\log x)^{-A}$, proving the result. $\square$

Note that every result we have proved prior to the previous one works for $\theta < 0.55$. The previous theorem is the limiting factor in our result.

Corollary 7.18. *Let $H = x^\theta$ where $\theta = 0.55 + \epsilon$. then for any $P_1, R, N_1, ...N_{2k-1}$ as in section 7.7 we have*

$$D(H, y_1, P_1, R, N_1, ...N_{2k-1}) \ll_{A,\epsilon} \frac{H}{(\log x)^A}$$

Proof. By the combinatorial lemma with $h = 0.05$, if we cannot apply theorem 7.14 with $h = 0.05$, we can instead apply corollary 7.17. $\square$

7.11 Putting the Sum Back Together

Recall that we had the requirement $k \geq h^{-1}$. We are now assuming $h = 0.05$, so we will take $k = 20$. We now have for $H = x^{0.55+\epsilon}$ that

$$\sum_{j=1}^{k} (-1)^{j-1} \binom{k}{j} \sum_{\substack{x < prr_1 \cdots n_{2j-1} \leq x+H \\ P < p \leq Q \\ i \geq j \Longrightarrow n_i \leq (2x)^{1/k}}} a_r \mu(n_j) \cdots \mu(n_{2j-1}) - \frac{H}{y_1} \sum_{\substack{x < prr_1 \cdots n_{2j-1} \leq x+y_1 \\ P < p \leq Q \\ i \geq j \Longrightarrow n_i \leq (2x)^{1/k}}} a_r \mu(n_j) \cdots \mu(n_{2j-1})$$

$$\ll \sum_{j=1}^{k} \binom{k}{j} \sum_{P_1, R, N_1, \ldots, N_{2k+1}} D(H, y_1, P_1, R, N_1, \ldots N_{2k-1})$$

$$\ll_{A,\epsilon} \sum_{j=1}^{k} \binom{k}{j} \sum_{P_1, R, N_1, \ldots, N_{2k+1}} \frac{H}{(\log x)^A}$$

$$\ll_{A,\epsilon} \sum_{j=1}^{k} \binom{k}{j} \frac{H (\log x)^{2k+1}}{(\log x)^A}$$

$$\ll_{A,\epsilon} \frac{H}{(\log x)^A}$$

Where we have adjusted A and used the fact that k is a fixed constant, and so the sum of the binomials is also just a fixed constant.

Theorem 7.19. *Let $\epsilon > 0$ and $H = x^{0.55+\epsilon}$. Then for any $A \geq 1$*

$$\sum_{x < n \leq x+H} \mu(n) - \frac{H}{y_1} \sum_{x < n \leq x+y_1} \mu(n) \ll_{\epsilon,A} E + \frac{H}{(\log x)^A}$$

where

$$E = \frac{H \log(P)}{\log(Q)} + Q + \frac{H}{P} + H \log^2(x) 2^{-\log M / \log Q} + H \log(Q) \log(M) e^{-s} + yQM + \frac{H \log^3(x)}{P}$$

$$s = \frac{\log y}{\log Q}$$

Proof. Follows from combining the argument above with theorem 7.8 □

Corollary 7.20. *Let ϵ, H and E be as in theorem 7.19. Then for any $A \geq 1$ we have*

$$\sum_{x < n \leq x+H} \mu(n) \ll_A \frac{H}{(\log x)^A} + E$$

Proof. By the prime number theorem for the Möbius function we have that for any $A \geq 1$

$$\frac{H}{y_1} \sum_{x < n \leq x+y_1} \mu(n) \ll_A \frac{H}{y_1} \cdot \frac{y_1}{(\log x)^A} = \frac{H}{(\log x)^A}$$

The result follows by substituting this into theorem 7.19. □

7.12 Handling the Error

In order to apply lemma 3.19, we required that $P \geq \exp\left((\log x)^{2/3+\epsilon_1}\right)$ for some $\epsilon_1 > 0$. this implies that

$$\frac{H}{P} + \frac{H(\log x)^3}{P} \ll \frac{H}{(\log x)^{1/3-\epsilon}}$$

The only term that remains which depends on P is

$$\frac{H\log(P)}{\log(Q)}$$

which is increasing in P. We therefore take P as small as possible, i.e. $P = \exp\left((\log x)^{2/3+\epsilon_1}\right)$.

We have also required that $R \leq x^{\epsilon/4}$ and therefore, as yM is the upper bound for R, we shall require $y, M \leq x^{\epsilon/8}$. Similarly as we assumed $P_1 \leq x^{\epsilon/4}$ and Q is the upper bound for P_1, so we shall require $Q \leq x^{\epsilon/4}$. With these bounds we easily have

$$yQM + Q \ll_\epsilon \frac{H}{(\log x)^{1/3-\epsilon}}$$

The last term in the error that depends on y is e^{-s} where $s = \log y/\log Q$. This is decreasing in y and so we will set $y := x^{\epsilon/8}$. Now we note that if we also set $M = x^{\epsilon/8}$ we get

$$H\log^2(x)2^{-\log M/\log Q} + H\log(Q)\log(M)e^{-s} \ll H(\log x)^2 e^{-(\epsilon/8)\log x/\log Q}$$

We would like

$$H(\log x)^2 e^{-(\epsilon/8)\log x/\log Q} \ll_\epsilon \frac{H}{(\log x)^{1/3-\epsilon}}$$

To obtain this it suffices to have

$$\frac{-\epsilon\log x}{8\log Q} < \left(\epsilon - \frac{1}{3}\right)\log\log x \iff \log Q < \frac{(\epsilon/8)\log x}{(1/3-\epsilon)\log\log x}$$

We also would like

$$\frac{H\log P}{\log Q} = \frac{H(\log x)^{2/3+\epsilon_1}}{\log Q} \ll_\epsilon \frac{H}{(\log x)^{1/3-\epsilon}}$$

$$\iff (\log x)^{1+\epsilon_1-\epsilon} \ll_\epsilon \log Q$$

We can easily see that if we set $\epsilon_1 = \epsilon/2$ and $Q = x^{1/(\log\log x)^2}$, both bounds will be satisfied when x is large enough, giving us the following result

Theorem 7.21. *Let $H = x^{0.55+\epsilon}$ for some $\epsilon > 0$. Then*

$$\sum_{x<n\leq x+H} \mu(n) \ll_\epsilon \frac{H}{(\log x)^{1/3-\epsilon}}$$

8 Proof of the 0.55 Lemma

8.1 The Statement

For posterity we shall restate the lemma in terms of the variables used in the original paper.[6]

Lemma 8.1. *Let $\eta > 0$, $0.55 + 3\eta < \Theta < 7/12$ and $0 \leq \phi \leq (6\Theta - 1)/5 - 3\eta$. Suppose $M, N \leq x^\phi$, then*

$$\sum_{\substack{M<m\leq 2M \\ N<n\leq 2N}} a_m b_n r(x, x^\theta, mn) \ll_\delta x^{\Theta-\delta}$$

where $|a_m| \leq \tau_k(u)$, $|b_n| \leq \tau_k(m)$, for some $k \geq 1$,and $\delta = \delta(\eta) > 0$.

Again we note that in the original paper the lemma is only stated for $|a_m|, |b_n| \leq 1$, however, the proof generalises to when the sequences are only divisor bounded.

Our proof is significantly longer than the proof in the original paper. This is because we have elaborated on how the many error terms are bounded and also derived some of the formulas which are stated without proof, e.g., the integral formula for the function $L(s)$. We have also showed how all of the results cited in the original paper were applied, such as the details of the application of Van Der Corput's bound which was originally omitted.

Finally, at the end of the proof we will have a choice of various exponents. These are chosen seemingly arbitrarily in the original paper, however we have given motivation on why these specific exponents were chosen.

8.2 Prerequisite Results

In order to prove the lemma we shall need a small proposition that allows us to bound minimums by products.

Proposition 8.2. *Let $\alpha_1, ...\alpha_r > 0$ and $e_1, ..., e_r \geq 0$ with $\sum_{i=1}^r e_i = 1$. Then*

$$\min\{\alpha_1, ..., \alpha_r\} \leq \alpha_1^{e_1}\alpha_2^{e_2}...\alpha_r^{e_r}$$

Proof. Let $\alpha = \min\{\alpha_1, ..., \alpha_r\}$. Then we have

$$\min\{\alpha_1, ..., \alpha_r\} = \alpha^1 = \alpha^{e_1+e_2+...+e_r} = \alpha^{e_1}\alpha^{e_2}...\alpha^{e_r} \leq \alpha_1^{e_1}\alpha_2^{e_2}...\alpha_r^{e_r}$$

□

We shall also need Van Der Corput's estimate for exponential sums.

Theorem 8.3. *If $f(x)$ is real and twice differentiable on $[a, b]$ with $b \geq a + 1$, and $\forall x \in [a, b]$ we have*

$$0 < \lambda_2 \leqslant f''(x) \leqslant h\lambda_2$$

then

$$\sum_{a<n\leqslant b} e^{2\pi i f(n)} = O\left(h(b-a)\lambda_2^{1/2}\right) + O\left(\lambda_2^{-1/2}\right)$$

Proof. The restriction on the second derivative allows us to approximate the sum with the corresponding integral. (See theorem 5.9 in [21])

8.3 Proof.

We define $a_m = 0$ and $b_n = 0$ for m and n not in our range of summation. Furthermore let $y = x^\Theta$. We shall only assume that $1/2 < \Theta$ until we are forced to assume $0.55 < \Theta$ towards the end of the proof. We are interested in the difference

$$\mathcal{R}(x, y, M, N) = \sum_{x<mnl\leq x+y} a_m b_n - y \sum_{m,n\geq 1} \frac{a_m}{m}\frac{b_n}{n}$$

where $y \leq x/2$. Note that the range of l is $(x/(4MN), 3x/(2MN))$, hence, in order to study the first sum we will use the Dirichlet polynomial $f(s) = M(s)N(s)L(s)$ where

$$M(s) = \sum_{M<m\leq 2M} \frac{a_m}{m^s}, \qquad N(s) = \sum_{N<n\leq 2N} \frac{b_n}{n^s},$$

$$L(s) = \sum_{L<l\leq 6L} \frac{1}{l^s}.$$

Note that here we have let $L = x/(4MN)$. The Dirichlet polynomial for the second sum will be $M(s+1)N(s+1)$ and this indicates the strategy which we will use to prove the result. We shall estimate both the sums using Perron's formula, then for the first sum we shall use an estimate $L(s)$ so that the integral will be independent of l. We shall then use the change of variable $s \mapsto s-1$ and compare the result with the main term given by Perron's formula applied to the second sum.

We first let $0 < T_0 < T \ll x$ and $T_0 \ll xy^{-1}$. We can see from the proof of lemma 3.9, in conjunction with the bound $|a_m b_n| \ll \tau_k(mn)$ for some $k > 0$, that

$$\sum_{x < mnl \leq x+y} a_m b_n = \frac{1}{2\pi i} \int_{c-iT_0}^{c+iT_0} f(s) x^{s-1} y ds + E_1 + E_{tails} + E_{perron}$$

where

$$E_1 \ll \int_{c-iT_0}^{c+iT_0} |f(s)||s| x^{c-2} y^2 d|s|$$

$$E_{tails} \ll yx^{c-1} \left(\int_{c-iT}^{c-iT_0} f(s) ds + \int_{c+iT_0}^{c+iT} f(s) ds \right)$$

$$E_{perron} \ll \frac{x^c}{T} \sum_{1 \leq lmn \leq x+y} \frac{|a_m b_n|}{(lmn)^c} + \frac{x \tau_k(x) \log x}{T}$$

To bound E_1 we shall need a bound for $M(s)N(s)L(s)$. For $M(s)$ we may use the trivial estimate:

$$|M(c+it)| \leq \sum_{M < m \leq 2M} \frac{1}{m^c} \ll \sum_{M < m \leq 2M} \frac{1}{M^c} \ll M^{1-c}.$$

Similarly we have $N(c+it) \ll N^{1-c}$. Finally, if we assume $T_0 \leq 2\pi L$, by lemma 3.22 we have $L(s) \ll L^{1-c}/|1-s|$, giving us

$$f(s) \ll \frac{M^{1-c} N^{1-c} L^{1-c}}{|1-c|} \asymp \frac{x^{1-c}}{|1-s|}$$

Inserting this into our bound for E_1 gives us

$$E_1 \ll y^2 x^{-1} \int_{c-iT_0}^{c+iT_0} \frac{|s|}{|1-s|} d|s|$$

If $|s-1| > 1$ we have that

$$\frac{|s|}{|1-s|} \ll 1$$

On the other hand, if $|s-1| \leq 1$, we have

$$\frac{|s|}{|1-s|} \ll \frac{1}{|1-c|}$$

As always c will be bounded and therefore the latter will be the larger of these two bounds. Using this bound in our estimate for E_1 we get

$$E_1 \ll y^2 x^{-1} \int_{c-iT_0}^{c+iT_0} \frac{ds}{|1-c|} \ll \frac{T_0 y^2 x^{-1}}{|1-c|}$$

We shall now bound E_{perron}. Note that if $a_m b_n \neq 0$ then $L < l \leq 6L$, $M < m \leq 2M$ and $N < n \leq 2N$.

Using this in combination with the bound $\tau_k(mn) \ll_\eta x^\eta$, we have

$$E_{perron} \ll \frac{x^c}{T} \sum_{1\le lmn\le x+y} \frac{\tau_k(mn)}{(LMN)^c} + \frac{x\tau_k(x)\log x}{T}$$

$$\ll \frac{x^c}{T} \sum_{1\le lmn\le x+y} \frac{x^\eta}{x^c} + \frac{x^{1+\eta}\log x}{T}$$

$$\ll_\eta \frac{x^\eta}{T} \sum_{1\le r\le x+y} \tau_3(r) + \frac{x^{1+\eta}}{T}$$

$$\ll_\eta \frac{x^{1+\eta}}{T}$$

Where we have adjusted η and the implied constant multiple times. We shall bound E_{tails} later, for now we shall return to the main term. From lemma 3.22 we have:

$$L(c+it) = \frac{(6L)^{1-s} - L^{1-s}}{1-s} + O(L^{-c})$$

Subbing this into the main term we get

$$\int_{c-iT_0}^{c+iT_0} M(s)N(s)L(s)x^{s-1}yds = \int_{c-iT_0}^{c+iT_0} M(s)N(s)\cdot\frac{(6L)^{1-s} - L^{1-s}}{1-s}x^{s-1}yd+O\left(\int_{c-iT_0}^{c+iT_0} |M(s)N(s)|L^{-c}x^{c-1}yd|s|\right)$$

To bound the error term we can use the trivial estimates we had before for $M(s)$ and $N(s)$ to get

$$\int_{c-iT_0}^{c+iT_0} |M(s)N(s)|L^{-c}x^{c-1}yd|s| \ll \int_{c-iT_0}^{c+iT_0} M^{1-c}N^{1-c}L^{-c}x^{c-1}yd|s| \ll T_0yL^{-1}$$

We now have the following estimate for the first sum:

$$\sum_{x<mnl\le x+y} a_mb_n = \frac{1}{2\pi i}\int_{c-iT_0}^{c+iT_0} M(s)N(s)\cdot\frac{(6L)^{1-s} - L^{1-s}}{1-s}x^{s-1}yds+O_\eta\left(T_0yL^{-1} + \frac{T_0y^2x^{-1}}{|1-c|}` + E_{tails} + \frac{L^{c-1}x^{1+\eta}}{T}\right)$$

We now shall apply Perron's formula to the second sum. Before we do this we note that $a_mb_n \neq 0$ implies $x/6L < mn \le x/L$. Therefore for $\kappa > 0$ we have

$$y\sum_{m,n}\frac{a_mb_n}{mn} = y\sum_{x/6L<mn\le x/L}\frac{a_mb_n}{mn}$$

$$= \frac{y}{2\pi i}\int_{\kappa-iT_0}^{\kappa+iT_0}\sum_{mn}\frac{a_mb_n}{mn}\cdot\frac{1}{(mn)^s}\cdot\left(\left(\frac{x}{L}\right)^r - \left(\frac{x}{6L}\right)^r\right)\frac{dr}{r} + O_\eta\left(\frac{yx^\kappa}{T_0}\sum_{1\le lmn\le x+y}\frac{|a_nb_m|}{(mn)^{1+\kappa}} + \frac{yx^\eta\log x}{T_0}\right)$$

Similarly to E_{perron} we have the bound

$$\frac{yx^\kappa}{T_0}\sum_{1\le lmn\le x+y}\frac{|a_nb_m|}{(mn)^{1+\kappa}} + \frac{yx^\eta\log x}{T_0} \ll_\eta \frac{yx^\eta}{T_0}$$

where we have assumed $\kappa < 1$. Note that we can see the main term for the estimates of our two sums is the same if we use the substitution $r = s-1$ and we assume $c = 1+\kappa$.
We therefore have that

$$\mathcal{R}(x,y,M,N) \ll_\eta T_0yL^{-1} + \frac{T_0y^2x^{-1}}{|1-c|}` + E_{tails} + \frac{L^{c-1}x^{1+\eta}}{T} + \frac{yx^\eta}{T_0}$$

We shall set $c = 1 + (\log x)^{-1}$ so that $c > 1$, $x^c \ll x$ and $x^\kappa \ll 1$. Setting $T_0 = L^{1/2}$, in order to balance the first and last terms, gives us

$$\mathcal{R}(x, y, M, N) \ll_\eta \frac{y}{L^{1/2}} + \frac{L^{1/2} y^2 \log x}{x} + E_{tails} + \frac{x^{1+\eta}}{T} + \frac{yx^\eta}{L^{1/2}}$$

We shall define $T = x^{1+2\eta} y^{-1}$ so that $x^{1+\eta}/y \ll x^{\Theta-\eta}$. Note that we may assume $MN \gg x^{\Theta-\eta}$ as otherwise the result is trivial. Therefore $L \asymp x/MN \ll x^{1-\Theta+\eta} \implies L^{1/2} \ll x^{(1/2)(1-\Theta+\eta)} \ll xy^{-1}$. Similarly we have $MN \ll x^{2\phi} \implies L^{-1/2} \ll x^{\phi-1/2}$. Using our bound for ϕ we have that

$$\phi - \frac{1}{2} \leq \frac{6\Theta - 1}{5} - 2\eta - \frac{1}{2} = \frac{12\Theta - 7}{10} - 2\eta \leq -2\eta$$

when $\Theta \leq 7/12$. Note that this is the reason we have made the assumption $\Theta \leq 7/12$. We therefore have that $L^{-1/2} \ll x^{-2\eta}$. Subbing in these bounds for $L^{1/2}$ and $L^{-1/2}$ into our bound for $\mathcal{R}$ gives

$$\mathcal{R}(x, y, M, N) \ll_\eta x^{\Theta-\eta} + E_{tails}$$

We note that in the original paper, when bounding one of the error terms, it was stated that

$$T_0 L^{1-c} (MN)^{1-c} y^2 (\log x) x^{c-2} \ll T_0 MN y x^{-1}$$

However as $LMN \asymp x$ we can see this statement is equivalent to $y \log x \ll MN$ which is not necessarily true as we have only assumed $x^{\Theta-\eta} \ll MN$.

It remains to show that for some $\delta > 0$ we have

$$E_{tails} \ll yx^{c-1} \left(\int_{c-iT}^{c-iT_0} f(s)ds + \int_{c+iT_0}^{c+iT} f(s)ds \ll x^{\Theta-\delta} \right)$$

By symmetry it suffices to estimate the integral over $[T_0, T]$. To do this we shall again use dyadic intervals; i.e. we shall split the interval $[T_0, T]$ into $\ll \log x$ smaller intervals of the form $[T_1, T_2]$ where $T_2 \leq 2T_1$. On each of these intervals we then pick well spaced points $T_1 \leq t_1 \leq t_2 ... \leq t_R \leq T_2$ where $t_r - t_{r-1} \geq 1$. Note that due to the spacing of the points that there are at most T_1 of them. We now estimate the integral on $[T_1, T_2]$ by taking the sum if f on these points as follows

$$\int_{c+iT_1}^{c+iT_2} f(s)ds \ll \sum_{r=1}^{R} |f(c + it_r)|$$

There are many ways we could choose the points so that the integral is bounded by the sum, for example we could split $[T_1, T_2]$ into R smaller intervals of length at most 2, and then let t_i be the value on the ith interval which maximises f on that interval. The precise method we use, however, will not matter. All that matters is that it is possible. From now on $[T_1, T_2]$ will be the dyadic interval on which the previous integral is maximised. This means that

$$yx^{c-1} \int_{c+iT_0}^{c+iT} f(s)ds \ll \log(x) yx^{c-1} \sum_{r=1}^{R} |f(c + it_r)|$$

It therefore remains to bound this sum. To do this we shall again use a form of dyadic intervals, however, this time rather than splitting the ranges of L, M and N, we shall instead split up their domains and then count how many t_r's are in each element of our partition. Before we begin this we note that by the trivial bounds we have

$$|L(c + it)|, |M(c + it)|, |N(c + it)| \leq 8.$$

We picked 8 as it is a power of 2, which will make a calculation a little cleaner later. First we show that the points where one of L, M or R are sufficiently small contribute a negligible amount to our sum. Let

$$\mathcal{R} = \{t_r \; L(c + it_r) \leq x^{-2} \text{ or } M(c + it_r) \leq x^{-2} \text{ or } N(c + it_r) \leq x^{-2}\}$$

As $|t_i - t_j| \geq 1$ for $i \neq j$, we must have that $R \leq T_1$ and therefore

$$\log(x) y x^{c-1} \sum_{r \in \mathcal{R}} f(c+it_r) \leq \log(x) y x^{c-1} \sum_{r=1}^{R} 8^2 \cdot \frac{1}{x^2} \ll T_1 \log(x) y x^{c-3} \ll x^{\Theta - \eta}$$

proving the set $\mathcal{R}$ contributes a negligible amount to our sum. We now split our remaining t_r's into $\ll \log(x)^3$ sets of the form

$$S(U,V,W) = \{U \leq L^{c-\frac{1}{2}} |L(c+it_r)| < 2U, V \leq M^{c-\frac{1}{2}} |M(c+it_r)| < 2V, W \leq N^{c-\frac{1}{2}} |N(c+it_r)| < 2W\}$$

where

$$x^{-2} \leq L^{\frac{1}{2}-c} U = 2^{-u} \leq 8$$

for some integer $u \geq -3$, and similarly for V and W. If $t_r \in S(U,V,W)$ then

$$f(c+it_r) \asymp (LMN)^{1/2-c} UVW \asymp x^{1/2-c} UVW,$$

giving us

$$\log(x) y x^{c-1} \sum_{r=1}^{R} |f(c+it_r)| \ll \log(x) y x^{c-1} \sum_{U_i, V_i, W_i} x^{1/2-c} U_i V_i W_i |S(U_i, V_i, W_i)|$$

$$\ll \log(x)^4 y x^{-1/2} UVW |S(U,V,W)|$$

where U, V and W will from now on denote the quantities which maximise $U_i V_i W_i |S(U_i, V_i, W_i)|$. In order to prove the lemma it suffices to show $UVW|S(U,V,W)| \ll x^{1/2-\delta_0(\eta)}$. To do this we will need different estimates for different sizes of U, V and W.

For our first estimate we shall apply theorem 3.21 from Davenport with $\delta = 1$ to get

$$\sum_{t \in S(U,V,W)} |M(c+it)|^2 \leq (T_1 + O(2M \log M))(1 + \log 2M) \sum_{m=1}^{2M} \frac{|a_m|^2}{m^{2c}}$$

$$\ll (T_1 + M) \log(M)^2 \sum_{M < m \leq 2M} \frac{\tau_k(m)^2}{m^{2c}}$$

$$\ll (T_1 + M) \log(x)^2 M^{-2c} \sum_{M < m \leq 2M} \tau_k(m)^2$$

$$\ll (T_1 + M) M^{1-2c} (\log x)^{A_1}$$

for some $A_1 > 0$. We also have that

$$\sum_{t \in S(U,V,W)} |M(c+it)|^2 \geq \sum_{t \in S(U,V,W)} (M^{1/2-c} V)^2 = |S(U,V,W)| M^{1-2c} V^2.$$

So we may conclude that

$$|S(U,V,W)| \ll (T_1 + M) V^{-2} \log(x)^{A_1}.$$

Similarly we may apply the theorem to $N(c+it)$ to give us

$$|S(U,V,W)| \ll (T_1 + N) W^{-2} \log(x)^{A_1}.$$

For our next estimate we apply theorem 3.23 with

$$D(s) = M(s+c) = \sum_{M<m\leq 2M} \frac{a_m m^{-c}}{m^{it}}$$

We therefore have that $N_1 \asymp M$,

$$G = \sum_{M<m\leq 2M} \frac{|a_m|^2}{m^{2c}} \ll M^{-2c} \sum_{M<m\leq 2M} \tau_k(m)^2 \ll M^{1-2c}(\log x)^{A_2}$$

for some $A_2 > 0$, and

$$|D(it_r)| > M^{1/2-c}V$$

for $t_r \in S(U,V,W)$. Using the theorem we may conclude

$$\begin{aligned}|S(U,V,W)| \ll & [M^{1-2c}M(M^{1/2-c}V)^{-2} + (M^{1-2c})^3(M^{1/2-c}V)^{-6}MT_1](\log x)^{A_3} \\ = & \left(V^{-2}M + V^{-6}MT_1\right)(\log x)^{A_3}\end{aligned}$$

For some $A_3 > 0$. Similarly we have

$$|S(U,V,W)| \ll \left(W^{-2}N + W^{-6}NT_1\right)(\log x)^{A_3}$$

For our fifth bound we apply theorem 3.23 with

$$D(s) = L(s+c)^2 = \sum_{L^2<l\leq 36L^2} \frac{c_l l^{-2c}}{l^s}$$

where $c_l \leq \tau(l)$. Therefore this time we have $N_1 \asymp L^2$,

$$G = \sum_{L^2<l\leq 36L^2} \frac{|c_l|^2}{l^{4c}} = L^{-4c} \sum_{L^2<l\leq 36L^2} \tau(l)^2 \ll L^{2-4c}(\log x)^{A_4}$$

for some $A_4 > 0$, and

$$|D(it_r)| = |L(c+it_r)|^2 > L^{1-2c}U^2$$

Theorem 3.23 now gives us

$$\begin{aligned}|S(U,V,W)| \ll & [L^{2-4c}L^2(L^{1-2c}U^2)^{-2} + (L^{2-4c})^3L^2T_1(L^{1-2c}U^2)^{-6}](\log x)^{A_5} \\ = & [U^{-4}L^2 + U^{-12}L^2T_1](\log x)^{A_5}\end{aligned}$$

for some $A_5 > 0$.

To obtain our final bound we again shall take advantage of the fact that $L(s)$ is a truncation of the zeta function. More specifically we are going to construct a formula for $L(s)$ in terms of an integral involving the zeta function on the critical line. By the Hardy-Litlewood formula (theorem 3.13) we have

$$\zeta(c+it+z) = \sum_{1\leq n\leq X} \frac{1}{n^{c+it}} \cdot \frac{1}{n^z} + \frac{X^{1-c-it-z}}{c+it+z-1} + O\left(X^{-c-\Re(z)}\right)$$

and therefore

$$\sum_{1\leq n\leq X} \frac{1}{n^{c+it}} \cdot \frac{1}{n^z} = \zeta(c+it+z) - \frac{X^{1-c-it-z}}{c+it+z-1} + O\left(X^{-c-\Re(z)}\right)$$

where $X \geq 2\pi T_0 = 2\pi L^{1/2}$. We shall apply Perron's formula, to the above Dirichlet polynomial, in the form of corollary 3.7 along the line $\Re(z) = 1/2 - c$. Note first that

$$\sum_{n \leq X} \frac{|n^{-c-it}|}{n^{1-c+\kappa}} = \sum_{n \leq X} \frac{1}{n^{1+\kappa}} \ll \frac{1}{\kappa} = A_\kappa(X)$$

and $|n^{-c-it}| = n^{-c} = \psi(n)$.

$$\sum_{1 \leq n \leq X} \frac{1}{n^{c+it}} = \int_{1/2-c+i(1/2)T_1}^{1/2-c+i(1/2)T_1} \zeta(c+it+z)\frac{X^z}{z} - \frac{X^{1-c-it}}{z(c+it+z-1)} + O\left(X^{-1/2}\right)\frac{X^z}{z}dz$$

$$+ O\left(\frac{L^{1-c+\kappa}\log x}{T_1} + \frac{L^{1-c}\log x}{T_1}\right)$$

Recalling that $T_1 \geq T = \sqrt{L}$ and $L^\kappa \ll 1$, we have

$$\frac{L^{1-c+\kappa}\log x}{T_1} + \frac{L^{1-c}\log x}{T_1} \ll L^{1/2-c}\log x$$

Now we note that as $t \in [T_1, 2T_1]$ we have $c + it + z - 1 \asymp T_1 \geq L^{1/2}$ and therefore

$$\int_{1/2-c+i(1/2)T_1}^{1/2-c+i(1/2)T_1} \frac{X^{1-c-it}}{z(c+it+z-1)}ds \ll \frac{X^{1-c}}{L^{1/2}}\int_0^{(1/2)T_1} \frac{d\tau}{1+\tau} \ll \frac{X^{1-c}\log x}{L^{1/2}}.$$

For the error term in the integral we have

$$\int_{1/2-c+i(1/2)T_1}^{1/2-c+i(1/2)T_1} O\left(X^{-1/2}\right)\frac{X^z}{z}dz \ll X^{-c}\int_0^{(1/2)T_1}\frac{d\tau}{1+\tau} \ll X^{-c}\log x$$

We conclude:

$$\sum_{1 \leq n \leq X} \frac{1}{n^{c+it}} = \int_{1/2-c+i(1/2)T_1}^{1/2-c+i(1/2)T_1} \zeta(c+it+z)\frac{X^z}{z} + O\left(\frac{X^{1-c}\log x}{L^{1/2}}\right)$$

Applying this formula twice with $X = 6L$ and $X = L$, then taking the difference of the two resulting formulas, we have

$$L(s) = \frac{1}{2\pi i}\int_{1/2-c-i(1/2)T_1}^{1/2-c+i(1/2)T_1} \zeta(s+z)\left((6L)^z - L^z\right)\frac{dz}{z} + O\left(L^{\frac{1}{2}-c}\log x\right)$$

We shall now use the estimate from theorem 3.18 for the fourth order moment of the zeta function to estimate $L(s)$.In order to do this we will need the following formulation of Hölders inequality

$$\left(\int |f(z)g(z)|dz\right)^4 \leqslant \left(\int |f(z)|^4 dz\right)\left(\int |g(z)|^{4/3}dz\right)^3$$

We shall apply this with $f(z) = \zeta(s+z)z^{-1/4}$ and $g(z) = z^{-3/4}$ to get:

$$\begin{aligned}
L^{4c-2} \sum_{r \in S(U,V,W)} |L(c+it_r)|^4 &\ll L^{4c-2} \sum_r \left\{ \int_{1/2-c-i(1/2)T_1}^{1/2-c+i(1/2)T_1} |\zeta(c+it_r+z)| \, L^{1/2-c} \frac{d|z|}{|z|} \right\}^4 + |S(U,V,W)|(\log x)^4 \\
&\ll L^{4c-2} \sum_r \left\{ \int_{1/2-c-i(1/2)T_1}^{1/2-c+i(1/2)T_1} \left| \zeta\left(\frac{1}{2} + i(v+t_r)\right) \right| L^{1/2-c} \frac{dv}{1+|v|} \right\}^4 + T(\log x)^4 \\
&\ll \sum_r \left\{ \int_{-2T_1}^{2T_1} \left| \zeta\left(\frac{1}{2} + i\tau\right) \right| \frac{d\tau}{1+|\tau - t_r|} \right\}^4 + T(\log x)^4 \\
&\ll \sum_r \left\{ \int_{-2T_1}^{2T_1} \left| \zeta\left(\frac{1}{2} + i\tau\right) \right|^4 \frac{d\tau}{1+|\tau - t_r|} \right\} \left\{ \int_{-2T_1}^{2T_1} \frac{d\tau}{1+|\tau - t_r|} \right\}^3 + T(\log x)^4 \\
&\ll (\log x)^3 \int_{-2T_1}^{2T_1} \left| \zeta\left(\frac{1}{2} + i\tau\right) \right|^4 \left(\sum_r \frac{1}{1+|\tau - t_r|} \right) d\tau + T(\log x)^4 \\
&\ll (\log x)^4 \int_{-2T_1}^{2T_1} \left| \zeta\left(\frac{1}{2} + i\tau\right) \right|^4 d\tau + T(\log x)^4 \\
&\ll (\log x)^8 T
\end{aligned}$$

Where in the last line we have applied theorem 3.18. On the other hand we also have

$$U^4 |S(U,V,W)| \le L^{4c-2} \sum_{r \in S(U,V,W)} |L(c+it_r)|^4$$

Putting these two estimates together gives us

$$|S(U,V,W)| \ll U^{-4} T (\log x)^8$$

Now we shall let

$$\begin{aligned}
F = \min \Big\{ & V^{-2}(M+T),\ W^{-2}(N+T),\ V^{-2}M + V^{-6}MT, \\
& W^{-2}N + W^{-6}NT,\ U^{-4}L^2 + U^{-12}L^2T,\ U^{-4}T \Big\}
\end{aligned}$$

Note that by our six bounds for $|S(U,V,W)|$ we have

$$UVW|S(U,V,W)| \ll UVWF(\log x)^A$$

for some $A > 0$. Recall that we wish to show that $UVW|S(U,V,W)| \ll x^{1/2-\delta_0(\eta)}$ and so it clearly suffices to show that $UVWF(\log x)^A \ll x^{1/2-\delta_0(\eta)}$ which will be true if $UVWF \ll x^{1/2-2\delta_0(\eta)}$. In order to prove this we must look at 4 different cases. In these cases we shall make repeated use of the proposition about minimums.

Case 1. $F \le 2V^{-2}M,\ 2W^{-2}N$
We have

$$UVWF \le UVW \min\{2V^{-2}M,\ 2W^{-2}N\} \le UVW(2V^{-2}M)^{1/2}(2W^{-2}N)^{1/2} = 2U(MN)^{1/2}$$

To bound U we pick some $t_r \in S(U,V,W)$ to see that

$$U \leq L^{c-1/2} \left| L\left(c+it_r\right)\right| = L^{c-1/2} \left| \sum_{L<l\leq 6L} \frac{1}{l^{c+it_r}} \right| \ll L^{c-1/2} \frac{1}{L^c} \max_{1\leq Y\leq 5L} \left\{ \sum_{L<l\leq L+Y} \exp\left(2\pi i \left(\frac{-\log l}{2\pi}\right) t_r\right)\right\}$$

Now we apply Van Der Corputs bound (theorem 8.3) with $f(x) = -t_r \log(x)/(2\pi)$, in which case

$$f''(x) = \frac{t_r}{2\pi x^2}$$

So for $x \in [L, 6L]$ we have

$$\frac{t_r}{2\pi(6L)^2} \leq f''(x) \leq \frac{t_r}{2\pi L^2}$$

Hence we see that $\lambda_2 = t_r(2\pi(6L)^2)^{-1}$ and h does not matter as it is just a constant. Applying Van Der Corput's bound we find that

$$\max_{1\leq Y\leq 5L} \left\{ \sum_{L<l\leq L+Y} \exp\left(2\pi i \left(\frac{-\log l}{2\pi}\right) t_r\right)\right\} \ll (6L-L)\left(\frac{t_r}{2\pi(6L)^2}\right)^{1/2} + \left(\frac{2\pi(6L)^2}{t_r}\right)^{1/2}$$

$$\ll t_r^{1/2} + L t_r^{-1/2}$$

$$\ll x^{1/2+\eta-\Theta/2} + L^{3/4}.$$

Where we have used that $L^{1/2} = T_0 \leq t_r \leq T = x^{1+2\eta-\Theta}$. Applying this to our bound for U we get

$$U \ll L^{-1/2} x^{1/2+\eta-\Theta/2} + L^{1/4}$$

It suffices to prove that $U \ll L^{1/2}x^{-\eta}$. We will first show $L^{1/4} \ll L^{1/2}x^{-\eta}$ which is equivalent to $x^{\eta} \ll L^{1/4}$. Recall that from the assumptions in the statement we have $MN \ll x^{2\phi}$ and

$$\phi < \frac{6\Theta - 1}{5} - 2\eta < \frac{1}{2} - 2\eta.$$

We therefore have that $L \gg x^{1-2\phi} \gg x^{4\eta}$ which implies $x^{\eta} \ll L^{1/4}$. It remains to show that

$$L^{-1/2} x^{1/2+\eta-\Theta/2} \ll L^{1/2} x^{-\eta}$$

$$\iff x^{1/2+2\eta-\Theta/2} \ll L$$

Given the lower bound $L \gg x^{1-2\phi}$, it suffices to show that

$$1/2 + \eta - \Theta/2 \leq 1 - 2\phi$$

$$\iff \phi \leq \frac{1+\Theta}{4} - \frac{\eta}{2}$$

By the upper bound for ϕ this will be true if

$$\frac{1+\Theta}{4} - \frac{\eta}{2} \leq \frac{6\Theta-1}{5} - 2\eta$$

$$\iff \frac{9}{19} + \frac{30}{19}\eta \leq \Theta$$

Which is true for η small because $9/19 < 1/2 < \Theta$. We can therefore conclude $U \ll L^{1/2}x^{-\eta}$. Subbing this into our bound for $UVWF$ gives

$$UVWF \ll (LMN)^{1/2}x^{-\eta} \ll x^{1/2-\eta}$$

Case 2. $F > 2V^{-2}M,\ 2W^{-2}N$
The sums inside the minimum in the definition of F are problematic, however, our assumed bound on F will help us deal with these sums. From the definition of F we have that $F \leq V^{-2}M + V^{-2}T < F/2 + V^{-2}$. Solving this inequality for F gives us

$$F < 2V^{-2}T$$

Similar arguments using the other bounds from the definition of F gives us

$$F < 2W^{-2}T, \qquad F < 2V^{-6}MT, \qquad F < 2W^{-6}NT.$$

Combining these bounds with the definition of F gives us

$$F \leq \min\{2V^{-2}T,\ 2W^{-2}T,\ 2V^{-6}MT,\ 2W^{-6}NT,\ U^{-4}L^2 + U^{-12}L^2T,\ U^{-4}T\}$$

We have one remaining sum in this bound for F. To deal with this we will use the trivial inequality $\min\{a, b+c\} \leq \min\{a, 2b\} + \min\{a, 2c\}$. Applying this, we get

$$\begin{aligned}F \leq\ & 2\min\{V^{-2}T, W^{-2}T,\ V^{-6}MT,\ W^{-6}NT,\ U^{-4}L^2,\ U^{-4}T\} \\ & + 2\min\{V^{-2}T,\ W^{-2}T,\ V^{-6}MT,\ W^{-6}NT,\ U^{-12}L^2T,\ U^{-4}T\}\end{aligned}$$

Where we have replace $U^{-4}T$ by the loser bound $2U^{-4}T$ and then we have proceeded to factor the 2 out of the minimum. We shall now use the proposition which allows us to bound the minimums by products. Applying the proposition to the first term and recalling that $L \asymp x(MN)^{-1}$ gives

$$\begin{aligned}&\min\{V^{-2}T,\ W^{-2}T,\ V^{-6}MT,\ W^{-6}NT,\ U^{-4}L^2,\ U^{-4}T\} \\ &\ll (V^{-2}T)^{e_1}(W^{-2}T)^{e_2}(V^{-6}MT)^{e_3}(W^{-6}NT)^{e_4}(U^{-4}x^2M^{-2}N^{-2})^{e_5}(U^{-4}T)^{e_6} \\ &= V^{-2e_1-6e_3}W^{-2e_2-6e_4}U^{-4e_5-4e_6}M^{e_3-2e_5}N^{e_4-2e_5}x^{2e_5}T^{e_1+e_2+e_3+e_4+e_6}\end{aligned}$$

where $e_1 + ... + e_6 = 1$. Now recall that we know what T and x are, however, we do not know much about U, V, W, M and N. We wish to find a bound for $UVWF$ that obviously does not depend on U, V or W. We can easily achieve this by demanding that the exponent of U, V and W is -1. To deal with M and N we shall simply demand that their exponents are zero. This gives us the system of equations

$$\begin{aligned}e_1 + e_2 + e_3 + e_4 + e_5 + e_6 &= 1 \\ 2e_1 + 6e_3 &= 1 \\ 2e_2 + 6e_4 &= 1 \\ 4e_5 + 4e_6 &= 1 \\ e_3 - 2e_5 &= 0 \\ e_4 - 2e_5 &= 0\end{aligned}$$

Solving these gives us the solution

$$e_1 = e_2 = \frac{5}{16} \qquad e_3 = e_4 = \frac{1}{16}$$

$$e_5 = \frac{1}{32} \qquad e_6 = \frac{7}{32}$$

Subbing these values in gives us

$$\min\{V^{-2}T,\ W^{-2}T,\ V^{-6}MT,\ W^{-6}NT,\ U^{-4}L^2,\ U^{-4}T\} \ll (UVW)^{-1}x^{1/16}T^{31/32}$$

We now repeat this process with the second minimum:

$$\min\{V^{-2}T,\ W^{-2}T,\ V^{-6}MT,\ W^{-6}NT,\ U^{-12}L^2T,\ U^{-4}T\}$$

$$\ll V^{-2e_1-6e_3}W^{-2e_2-6e_4}U^{-12e_5-4e_6}M^{e_3-2e_5}N^{e_4-2e_5}x^{2e_5}T^{e_1+e_2+e_3+e_4+e_5+e_6}$$

which gives us the system of equations

$$\begin{aligned} e_1+e_2+e_3+e_4+e_5+e_6 &= 1 \\ 2e_1+6e_3 &= 1 \\ 2e_2+6e_4 &= 1 \\ 12e_5+4e_6 &= 1 \\ e_3-2e_5 &= 0 \\ e_4-2e_5 &= 0 \end{aligned}$$

Solving the system gives us

$$e_1 = e_2 = \frac{7}{20} \qquad e_3 = e_4 = \frac{1}{20}$$

$$e_5 = \frac{1}{40} \qquad e_6 = \frac{7}{40}$$

giving us the bound

$$\min\{V^{-2}T,\ W^{-2}T,\ V^{-6}MT,\ W^{-6}NT,\ U^{-12}L^2T,\ U^{-4}T\} \ll (UVW)^{-1}Tx^{1/20}$$

Using these bounds for F and subbing in $T = x^{1+2\eta}/y \asymp x^{1+2\eta-\Theta}$, we get

$$\begin{aligned} UVWF &\ll x^{1/16}T^{31/32} + Tx^{1/20} \\ &\ll x^{1/16+31/32+(62/32)\eta-(31/32)\Theta} + x^{1/20+1+2\eta-\Theta} \\ &= x^{33/32+(62/32)\eta-(31/32)\Theta} + x^{21/20+2\eta-\Theta} \\ &\ll x^{21/20+2\eta-\Theta} \end{aligned}$$

as $33/32 + (62/32)\eta - (31/32)\Theta \le 21/20 + 2\eta - \Theta \iff \Theta \le 3/5 + 2\eta$ which is true as $\Theta \le 7/12 < 3/5$. We would like $UVWF \ll x^{1/2-\eta}$. This will be true if

$$21/20 + 2\eta - \Theta \le 1/2 - \eta \iff 0.55 + 3\eta \le \Theta$$

which we shall henceforth assume.

We stress here that everything in this proof and the the proof of Matomäki and Teräväinen's result [12], works for any $\Theta > 1/2$ or $\theta > 1/2$ respectively. This is the first time we have been forced to make the stronger assumption of $\Theta > 0.55$.

Case 3. $F > 2V^{-2}M,\ F \le 2W^{-2}N$

As in the previous case, the assumption $F > 2V^{-2}M$ gives us the bounds $F < 2V^{-2}T$ and $F < 2V^{-6}MT$.

Combining these bounds with the assumption that $F \leq 2W^{-2}N$ along with the definition of F gives us

$$
\begin{aligned}
F &\leq \min\{2V^{-2}T,\ 2V^{-6}MT,\ 2W^{-2}N,\ U^{-4}L^2 + U^{-12}L^2T,\ U^{-4}T\} \\
&\leq 2\min\{V^{-2}T,\ V^{-6}MT,\ W^{-2}N,\ U^{-4}L^2,\ U^{-4}T\} \\
&\qquad + 2\min\{V^{-2}T,\ V^{-6}MT,\ W^{-2}N,\ U^{-12}L^2T,\ U^{-4}T\} \\
&\ll (V^{-2}T)^{e_1}(V^{-6}MT)^{e_2}(W^{-2}N)^{e_3}(U^{-4}x^2M^{-2}N^{-2})^{e_4}(U^{-4}T)^{e_5} \\
&\qquad + (V^{-2}T)^{f_1}(V^{-6}MT)^{f_2}(W^{-2}N)^{f_3}(U^{-12}x^2M^{-2}N^{-2}T)^{f_4}(U^{-4}T)^{f_5} \\
&= V^{-2e_1-6e_2}W^{-2e_3}U^{-4e_4-4e_5}x^{2e_4}M^{e_2-2e_4}N^{e_3-2e_4}T^{e_1+e_2+e_5} \\
&\qquad + V^{-2f_1-6f_2}W^{-2f_3}U^{-12f_4-4f_5}x^{2f_4}M^{f_2-2f_4}N^{f_3-2f_4}T^{f_1+f_2+f_4+f_5}
\end{aligned}
$$

where $e_1 + e_2 + e_3 + e_4 + e_5 = f_1 + f_2 + f_3 + f_4 + f_5 = 1$. Again, in both terms, we shall require that the exponent of U, V and W is -1 giving us three equations. In total, including the requirement that the exponents add to 1, we have four equations in five variables. This means, that unlike in case 2, we cannot require both the exponents of M and N be zero. We shall instead just require that the exponent of M is zero and then use the estimate $N \ll x^{\phi}$. These requirements give us the following systems of equations

$$
\begin{aligned}
e_1 + e_2 + e_3 + e_4 + e_5 &= 1 \\
2e_1 + 6e_2 &= 1 \\
2e_3 &= 1 \\
4e_4 + 4e_5 &= 1 \\
e_2 - 2e_4 &= 0
\end{aligned}
$$

and

$$
\begin{aligned}
f_1 + f_2 + f_3 + f_4 + f_5 &= 1 \\
2f_1 + 6f_2 &= 1 \\
2f_3 &= 1 \\
12f_4 + 4f_5 &= 1 \\
f_2 - 2f_4 &= 0
\end{aligned}
$$

Solving these gives us

$$
e_1 = e_2 = \frac{1}{8} \qquad e_3 = \frac{1}{2}
$$

$$
e_4 = \frac{1}{16} \qquad e_5 = \frac{3}{16}
$$

and

$$
f_1 = \frac{1}{4} \qquad f_2 = \frac{1}{12} \qquad f_3 = \frac{1}{2}
$$

$$
f_4 = \frac{1}{24} \qquad f_5 = \frac{1}{8}
$$

Subbing these values in to our bound for F will give us

$$UVWF \ll x^{1/8}T^{7/16}N^{3/8} + x^{1/12}N^{5/12}T^{1/2}$$

$$\ll x^{1/8+(7/16)(1+2\eta-\Theta)+3\phi/8} + x^{1/12+(5/12)\phi+(1/2)(1+2\eta-\Theta)}$$

$$= x^{9/16+(7/16)(2\eta-\Theta)+3\phi/8} + x^{7/12+(1/2)(2\eta-\Theta)+5\phi/12}$$

We would like $UVWF \ll x^{1/2-\eta/8}$ (η is a little too big this time where as $\eta/8$ is just small enough). The first term will be $\ll x^{1-\eta/8}$ if

$$\frac{9}{16} + \frac{7}{16}(2\eta - \Theta) + \frac{3\phi}{8} \leq \frac{1}{2} - \frac{\eta}{8}$$

$$\iff \phi \leq \frac{7\Theta - 1}{6} - \frac{16\eta}{6}$$

which is true if

$$\frac{6\Theta - 1}{5} \leq \frac{7\Theta - 1}{6}$$

$$\iff \Theta \leq 1$$

which we know to be the case.

The second term will be $\ll x^{1/2-\eta/8}$ if

$$\frac{7}{12} + \frac{1}{2}(2\eta - \Theta) + \frac{5\phi}{12} \leq \frac{1}{2} - \frac{\eta}{8}$$

$$\iff \phi \leq \frac{6\Theta - 1}{5} - \frac{14\eta}{5}$$

which is fundamentally why we make the assumption that

$$\phi \leq \frac{6\Theta - 1}{5} - 3\eta.$$

We have therefore proved case 3.

Case 4. $F \leq 2V^{-2}M,\ F > 2W^{-2}N$

This is the exact same as case 3, except M is interchanged with N, and V is interchanged with W.

□

8.4 The New Barrier

We wish to find out what choice of parameters maximise $UVWF$. Consider the following lower bound for F which can be easily seen from its definition:

$$F \geq \min\{TV^{-2}, TW^{-2}, V^{-6}MT, W^{-6}NT, U^{-12}L^2T, U^{-4}T\}$$

We therefore have that

$$UVWF \geq T\min\{UV^{-1}W, UVW^{-1}, UV^{-5}WM, UVW^{-5}N, U^{-11}VWL^2, U^{-3}VW\}$$

This lower bound will be maximised when each argument in the minimum is the same. Letting the first two arguments equal, we find

$$UV^{-1}W = UVW^{-1} \implies V = W$$

Letting the third and fourth arguments equal we now find

$$UV^{-5}WM = UVW^{-5}N \implies V^{-4}M = V^{-4}N \implies M = N$$

Equality in the first and third arguments gives us

$$UV^{-1}W = UV^{-5}WM \implies UV^{-1}V = UV^{-5}VM \implies M = V^4$$

Setting the first and sixth bound equal we get

$$UV^{-1}W = U^{-3}VW \implies V^2 = U^4 \implies V = U^2$$

By the relation $L \asymp x/MN = xV^{-8} = xU^{-16}$, we see that if we let the first and fifth bounds equal, then:

$$UV^{-1}W = U^{-11}VWL^2$$

$$\implies UU^{-2}U^2 \asymp U^{-11}U^2U^2x^2U^{-32}$$

$$\implies U^{40} \asymp x^2$$

$$\implies U \asymp x^{0.05}$$

Substituting in $U = x^{0.05}$ to our relations for M, N, L, V and W we get:

$$M = N = x^{0.4}, \qquad L \asymp x^{0.2}, \qquad V = W = x^{0.1}$$

Using these values in our bound for $UVWF$ we get

$$UVWF \gg T \min\{x^{0.05}, x^{0.05}, x^{0.05}, x^{0.05}, x^{0.05}, x^{0.05}\}$$

$$= Tx^{0.05}$$

$$= x^{1.05-\Theta+2\eta}$$

Recall that we required that

$$UVWF \ll x^{\Theta-\eta}$$

We therefore must have that

$$x^{1.05-\Theta+2\eta} \ll x^{1/2-\eta}$$

which means we must again have that $0.55 + 3\eta \leq \Theta$.

This tells us that having $0.55+3\eta \leq \Theta$ is not only sufficient but also necessary, using only the information provided by our six bounds for F.

Recall that in our proof of 7.21, M corresponded to N_I and N corresponded to N_J. In light of the previous discussion we can see that if we could always choose N_I and N_J so that neither was close to $x^{0.4}$, then we may be able to take $\theta < 0.55$ in theorem 7.21. The problem is that if all but five of the $N_i's$ in Heath-Brown's identity were of size $x^{o(1)}$ and the remaining five were of size $x^{1/5+o(1)}$, then we would be forced into the case where $N_I \asymp N_J = x^{0.4+o(1)}$.

This shows us that while Rameré's identity helped us handle the case with six variables of size $x^{1/6+o(1)}$, we have now encountered a new barrier due to five variables of size $x^{1/5+o(1)}$.

References

[1] Roger C. Baker, Harman Glyn, and János Pintz. The difference between consecutive primes, ii. *Proceedings of the London Mathematical Society*, 83(3):532–562, 2001.

[2] Jean Bourgain. On large values estimates for Dirichlet polynomials and the density hypothesis for the Riemann zeta function. *International Mathematics Research Notices*, 2000(3):133–146, 2000.

[3] John B. Friedlander and Henryk Iwaniec. *Opera De Cribro*, volume 57. American Mathematical Soc., 2010.

[4] Glyn Harman. *Prime-Detecting Sieves.(LMS-33)*. Princeton University Press, 2012.

[5] David R Heath-Brown. Prime numbers in short intervals and a generalized Vaughan identity. *Canadian Journal of Mathematics*, 34(6):1365–1377, 1982.

[6] David R. Heath-Brown and Henryk Iwaniec. On the difference between consecutive primes. 1979.

[7] Martin N. Huxley. On the difference between consecutive primes. *Inventiones mathematicae*, 15:164–170, 1971.

[8] Albert E. Ingham. On the estimation of $N(\sigma, T)$. *The Quarterly Journal of Mathematics*, (1):201–202, 1940.

[9] Aleksandar Ivić. *The Riemann zeta-function: theory and applications.* Courier Corporation, 2012.

[10] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53. American Mathematical Soc., 2021.

[11] Kaisa Matomäki and Maksym Radziwiłł. A note on the Liouville function in short intervals. *arXiv preprint arXiv:1502.02374*, 2015.

[12] Kaisa Matomäki and Joni Teräväinen. On the Möbius function in all short intervals. *Journal of the European Mathematical Society*, 25(4):1207–1225, 2022.

[13] James Maynard and Kyle Pratt. Half-isolated zeros and zero-density estimates. *arXiv preprint arXiv:2206.11729*, 2022.

[14] Hugh L. Montgomery. Mean and large values of Dirichlet polynomials. *Inventiones mathematicae*, 8(4):334–345, 1969.

[15] Hugh L. Montgomery. *Topics in multiplicative number theory*, volume 227. Springer, 2006.

[16] Yoichi Motohashi. On the sum of the Möbius function in a short segment. *Proceedings of the Japan Academy*, 52(9):477–479, 1976.

[17] Kanakanahalli Ramachandra. Some problems of analytic number theory. *Acta Arithmetica*, 31(4):313–324, 1976.

[18] H. E. Richert. Zur abschätzung der Riemannschen zetafunktion in der nähe der vertikalen $\sigma=1$. *Mathematische Annalen*, 169:97–101, 1967.

[19] Bernhard Riemann. On the number of prime numbers less than a given quantity. 1859.

[20] Peter Shiu. A Brun-Titschmarsh theorem for multiplicative functions. 1980.

[21] Edward Charles Titchmarsh and David Rodney Heath-Brown. *The theory of the Riemann zeta-function.* Oxford university press, 1986.