ON THE LIOUVILLE FUNCTION IN SHORT INTERVALS

JAKE CHINIS

Abstract. Let \( \lambda \) denote the Liouville function. Assuming Riemann’s Hypothesis, we’ll prove that
\[
\frac{1}{X} \int_X^{2X} \left| \sum_{x \leq n \leq x+h} \lambda(n) \right|^2 \, dx \ll \frac{(\log X)^8}{h},
\]
as \( X \to \infty \), provided \( h = h(X) \leq \exp \left( \sqrt{\frac{1}{2} - o(1)} \log X \log \log X \right) \). The proof uses a simple variation of the methods developed by Matomäki and Radziwiłł in their work on multiplicative functions in short intervals, as well as some standard results concerning smooth numbers.

1. Introduction

Let \( \lambda \) denote the Liouville function; that is, the completely multiplicative function defined by \( \lambda(p) := -1 \), for all primes \( p \). It’s well-known that the Prime Number Theorem (PNT) is equivalent to the fact that \( \lambda \) exhibits some cancellation in its partial sums; more precisely, that
\[
\sum_{n \leq x} \lambda(n) = o(x),
\]
as \( x \to \infty \).

The Möbius randomness principle (see [IK04]) tells us that \( \{\lambda(n)\}_n \) should actually behave like a sequence of independent random variables, taking on the values \( \pm 1 \) with equal probability. As a result, we expect “square-root cancellation” in the partial sums for the Liouville function; that is, for any \( \epsilon > 0 \),
\[
\sum_{n \leq x} \lambda(n) \ll x^{\frac{1}{2} + \epsilon},
\]
as \( x \to \infty \). In fact, the above estimate is equivalent to the Riemann Hypothesis (RH); see Theorem 14.25(C) of [THB86], for example. Furthermore, it’s possible to quantify \( \epsilon \) in terms of \( x \); see [Sou09].

That being said, the Liouville function also exhibits cancellation in short intervals, provided that the length of the interval is sufficiently large. For example, Motohashi [Mot76] and Ramachandra [Ram76], independently, proved that
\[
\sum_{x \leq n \leq x+h} \lambda(n) = o(h),
\]
provided \( h > x^{\frac{11}{20} + \epsilon} \). Assuming RH, this can be improved to \( h > x^{\frac{1}{2}} (\log x)^A \), for some suitable constant \( A > 0 \); see [MM09]. Recently, Matomäki and Teräväinen [MT19] improved Motohashi and Ramachandra’s result with a small saving and in the larger range \( h > x^{\frac{11}{20} + \epsilon} \).

By relaxing the condition that our estimates hold for all short intervals, we can get results that hold for smaller values of \( h \). In unpublished work of Gao [1], it’s shown that
\[
(1.1) \quad \int_X^{2X} \left| \sum_{x \leq n \leq x+h} \lambda(n) \right|^2 \, dx = o(Xh^2),
\]
\footnote{Strictly speaking, Gao’s result, and those stated above his, were initially proven for the M"{o}bius function, \( \mu \), but the proofs extend to the Liouville function with little effort.}
assuming RH and provided \( h > (\log X)^A \), for some large constant \( A > 0 \).

Now, the preceding results should be compared with what is known for primes in short intervals. More precisely, an equivalent form of the PNT is given by
\[
\sum_{n \leq x} \Lambda(n) = x + o(x),
\]
where \( \Lambda \) denotes the von Mangoldt function (defined to be \( \log p \) if \( n \) is a power of the prime \( p \) and 0 otherwise). Furthermore, the Riemann Hypothesis is equivalent to the following estimate:
\[
\sum_{n \leq x} \Lambda(n) = x + O(x^{1/2}(\log x)^2);
\]
in particular, RH implies that
\[
\sum_{x \leq n \leq x+h} \Lambda(n) = h + o(h),
\]
provided \( h > x^{1/2}(\log x)^2 + \epsilon \).

Again restricting ourselves to results that hold almost everywhere, Selberg \cite{Sel43} proved that if RH holds and \( h > (\log X)^{2+\epsilon} \), then
\[
\int_X^{2X} \left| \sum_{x \leq n \leq x+h} \Lambda(n) - h \right|^2 \, dx = o(Xh^2),
\]
so that almost all short intervals contain the correct number of primes.

It’s important to note that both Gao and Selberg obtain square-root cancellation in their estimates. In fact, their proofs are quite similar to one another, although Gao’s proof is more involved. In both cases, they relate the short sum to a contour integral of some Dirichlet series and then shift this contour to the edge of the critical strip, picking up any poles along the way. For Selberg, the corresponding Dirichlet series is \( \zeta'(s)/\zeta(s) \) and, assuming RH, the only pole will be at \( s = 1 \). For Gao, the corresponding Dirichlet series is \( \zeta(2s)/\zeta(s) \) and you have poles at the non-trivial zeros, \( \rho \), of \( \zeta(s) \), with residues \( \zeta(2\rho)/\zeta'\rho) \). Unfortunately, very little is known about \( 1/\zeta'(\rho) \), so we need to proceed in a slightly indirect manner. The key idea is to use a sum over primes to approximate \( \zeta(s) \) and then avoid “clusters” of zeros of \( \zeta(s) \) on the half-line: near these regions, \( 1/\zeta(s) \) is large and the contour is chosen so that \( 1/\zeta(s) \) is not too large, at least in some sense; see \cite{Sou09}, to get an idea of Gao’s proof.

With all that in mind, we now turn our attention to the breakthrough work of Matomäki and Radziwiłł, where they relate the average value of 1-bounded multiplicative functions in short intervals to the corresponding average value in large intervals:

**Theorem 1.1** \cite{MR16}. Let \( f : \mathbb{N} \to [-1, 1] \) be a multiplicative function and let \( h = h(X) \to \infty \) arbitrarily slowly as \( X \to \infty \). Then, for almost all \( x \in [X, 2X] \),
\[
\frac{1}{h} \sum_{x \leq n \leq x+h} f(n) = \frac{1}{X} \sum_{X \leq n \leq 2X} f(n) + o(1),
\]
with \( o(1) \) not depending on \( f \).

In the case of the Liouville function, Theorem 1.1 then implies that
\[
\sum_{x \leq n \leq x+h} \lambda(n) = o(h),
\]
for almost all $x \in [X, 2X]$; in particular,

$$\int_X^{2X} \left| \sum_{x \leq n \leq x + h} \lambda(n) \right|^2 dx = o(Xh^2), (1.2)$$

which goes beyond the work of Gao, by removing the assumption on RH and by extending the range of $h$ to include all $h = h(X) \to \infty$ as $X \to \infty$. Furthermore, the work of Matomäki and Radziwiłł avoids the complex analytic approach used by both Gao and Selberg; instead, they employ a clever decomposition of the corresponding Dirichlet polynomials. For a detailed account of the work in [MR16] restricted to the Liouville function, see [Son16].

In this paper, we apply a simple variation of the methods developed in [MR16] in order to improve, conditionally, the bounds in (1.2); more precisely, we prove the following:

**Theorem 1.2.** Assume RH. Then,

$$\frac{1}{X} \int_X^{2X} \left| \frac{1}{h} \sum_{x \leq n \leq x + h} \lambda(n) \right|^2 dx \ll \frac{(\log X)^8}{h},$$

provided $h \leq \exp \left( \sqrt{\left( \frac{1}{4} - o(1) \right) \log X \log \log X} \right)$. \(\sum\)

**Remark 1.1.** Note that our result recovers (1.1): for larger $h$, we can simply cover $[x, x + h]$ by smaller intervals and, since we obtain cancellation on each of these subintervals, this is enough to show a bound of $o(Xh^2)$, provided $h > (\log X)^{\Omega(1)}$, say.

That being said, we’d like to emphasize that Theorem 1.2 shows square-root cancellation for the Liouville function in almost all short intervals, provided $h > (\log X)^{\Omega(1)}$. For $h < (\log X)^{\epsilon}$, our estimate is quite poor; in this case, the trivial bounds are better. Note also that Theorem 1.2 gives an upper bound on the exceptional set of $x \in [X, 2X]$ for which square-root cancellation does not hold (via Chebyshev’s Inequality). Finally, although we don’t have the full range of $h$ (as in (1.2)), our method avoids the complex analytic approach of both Gao and Selberg, and uses a much simpler variation of the methods developed in [MR16].

1.1. **Outline of the Proof.** Before we get into the details of our proof, we’ll briefly discuss the ideas needed from [MR16]. We also refer the reader to [MR15], where Matomäki and Radziwiłł study the average of $\lambda$ on intervals of length $h = X^\delta$, for $0 < \delta < 1$. The latter contains many of the same ideas as the former, without all of the technical complexities.

After converting to Dirichlet polynomials, our problem reduces to that of bounding the mean square of $\sum_{n \sim X} \lambda(n)/n^{1/2 + it}$, where $n \sim X$ is shorthand for $X \leq n \leq 2X$. More precisely, it will suffice to show that

$$\frac{1}{T} \int_X^{X + T} \left| \sum_{n \sim X} \frac{\lambda(n)}{n^{1/2 + it}} \right|^2 dt \ll \frac{(\log X)^8}{h},$$

for $X/h \leq T \leq X$, where we’ve removed the contribution from the small values of $t$ via a pointwise bound on the integrand. For larger $t$, our main tool is the Mean Value Theorem (Lemma 2.2), but this is only winning if we shorten the length our sum: we should think of $T \approx X/h$ and note that the MVT is best possible if the length of your Dirichlet polynomial is of size $N \approx T$; see Remark 2.1. Our Dirichlet polynomial has length $X$ and our goal is to split the sum over $n \sim X$ into two sums, one of which has length $\approx X/h$ and the other which needs to be handled separately.
In section 4, we deal with the integers \( n \sim X \) which have at least one prime factor \( p > h \). Using a variant of Ramaré’s identity, we can write the sum over such integers as

\[
\sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} = \sum_{h < p \leq 2X} \frac{\lambda(p)}{p^{\frac{1}{2} + it}} \sum_{m \sim X / p} \frac{\lambda(m)}{m^{\frac{1}{2} + it}} \sum_{\# \{q \text{ prime} : q > h, q | m \}} \frac{1}{\mathbb{1}_{p|m}},
\]

where \( \# \{q \text{ prime} : q > h, q | m\} + \mathbb{1}_{p|m} \) represents the number of ways \( n \sim X \) can be written as \( n = mp \) with \( p > h \). From here, we split the sum over \( p \) into dyadic intervals and use a contour integral to separate the variables \( p \) and \( m \). Note that the sum over \( m \) is of length \( X/h \); in particular, we can then take a pointwise bound on the sum over \( p \) and apply the Mean Value Theorem to the sum over \( m \), and this is enough to deal with the integers which have a large prime factor \( p > h \).

In section 5, we deal with the remaining integers, all of whose primes factors are \( \leq h \). Fortunately for us, there are few of these so-called \( h \)-smooth integers, at least for small \( h \), so that the MVT can be applied directly to give us what we want. Getting cancellation for larger \( h \) will require some new ideas; this is ongoing work.

1.2. Further results. The proof of Theorem 1.2 can easily be adapted to a more general setting. For an arbitrary multiplicative function \( f \), all we need is an analogue to Lemma 2.4. Essentially, we’re looking for square-root cancellation to the corresponding sum over primes:

\[
\sum_{p \sim P} f(p) p^{it} \ll P^\varepsilon \left( \sum_{p \sim P} (f(p))^2 \right)^{\frac{1}{2}}.
\]

For example, the above estimate is known to hold, assuming RH, for coefficients of automorphic forms and for multiplicative functions of the form \( \mu(n) \lambda_\pi(n) \) or \( \lambda_\pi(n) \), where \( \mu \) is the Möbius function and where the \( \lambda_\pi(n) \)’s are the coefficients of an automorphic representation \( \pi \).

2. Preliminaries

We present here a collection of standard results, which we use freely throughout our paper. We’ll begin with an effective version of Perron’s Formula, which serves as an approximation to the indicator function on \((1, \infty)\):

**Lemma 2.1** (Effective version of Perron’s Formula). Fix \( \kappa > 0 \). Then,

\[
\frac{1}{2\pi i} \int_{\kappa-iT}^{\kappa+iT} \frac{y^s}{s} \frac{1}{\log y} \, ds = \begin{cases} 1 & \text{if } y > 1 \\ \frac{1}{2} & \text{if } y = 1 \\ 0 & \text{if } y < 1 \end{cases} + O\left( \frac{y^\kappa}{\max\{1, T\log y\}} \right),
\]

uniformly for both \( y > 0 \) and \( T > 0 \).

*Proof*. See [Ten15, Theorem II.2.3].

Before getting into the specifics of our problem, we make note of another result, which is our main tool in proving Theorem 1.2.

---

2In the published version of [MR16], the term \( \mathbb{1}_{p|m} \) appears as the constant 1, but this was corrected in later versions of their paper. In any case, we’ll see that this misprint does not affect their argument.
Lemma 2.2 (Mean Value Theorem). For any sequence of complex numbers \( \{a_n\}_n \), we have that

\[
\int_0^T \left| \sum_{1 \leq n \leq N} a_n n^t \right|^2 dt = (T + \mathcal{O}(N)) \sum_{1 \leq n \leq N} |a_n|^2.
\]

Proof. See [IK04, Theorem 9.1] \( \square \)

Remark 2.1. Notice that the main term in Lemma 2.2 corresponds to the contribution from the diagonal terms (as seen by expanding the square and integrating). Notice further that the Mean Value Theorem (MVT) is exceptionally powerful when \( N \ll T \): in this case, the integral is bounded above by the contribution from the diagonal terms and this is best possible.

We need two more preliminary results. The first is a pointwise bound on \( \sum_{n \sim X} \lambda(n)/n^{\frac{1}{2}+it} \), which allows us to remove the contribution from the small values of \( t \) in our average value (this can be thought of as the analogous result to Lemma 1 in [MR15]). The second is the analogue to Lemma 2 in [MR16]: we’ll need a pointwise bound on sums of the form \( \sum_{h \ll N} 1/p^{\frac{1}{2}+it} \), for large values of \( t \).

Lemma 2.3. Assuming RH,

\[
\sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} \ll \epsilon (1 + |t|)^\epsilon X^\epsilon,
\]

for all \( \epsilon > 0 \) and all \( t \in \mathbb{R} \).

Proof. By Lemma 3.12 of [THB86], we have that

\[
\sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} = \frac{1}{2\pi i} \int_{2- iT}^{2+ iT} \frac{\zeta(2s + 1 + 2it)}{\zeta(s + \frac{1}{2} + it)} \frac{(2X)^s - X^s}{s} ds + \mathcal{O}\left(\frac{X^2}{T}\right).
\]

Given \( \epsilon > 0 \) and assuming RH, the function \( \zeta(2(s + \frac{1}{2} + it))/\zeta(s + \frac{1}{2} + it) \) is analytic for \( \Re(s) \geq \epsilon \); as a result, we may shift the contour to the edge of this region and get that

\[
\sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} = -\frac{1}{2\pi i} \left( \int_{2+ iT}^{\epsilon+ iT} + \int_{\epsilon - iT}^{2- iT} + \int_{2- iT}^{\epsilon - iT} \right) \frac{\zeta(2s + 1 + 2it)}{\zeta(s + \frac{1}{2} + it)} \frac{(2X)^s - X^s}{s} ds + \mathcal{O}\left(\frac{X^2}{T}\right).
\]

Then, using the facts that \( \zeta(s) \ll 1/(\Re(s) - 1) \), for \( \Re(s) > 1 \), and \( 1/\zeta(s) \ll |\Im(s)|^\epsilon \), for \( \Re(s) \geq 1/2 + \epsilon \), (see 14.2.6 in [THB86]), we have that

\[
\sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} \ll \epsilon (|t| + T)^\epsilon X^\epsilon \int_{-T}^T \frac{dy}{\sqrt{y^2 + \epsilon^2}} + (|t| + T)^\epsilon \frac{X^2}{T}
\]

\[
\ll \epsilon (|t| + T)^\epsilon (X^\epsilon \log T + \frac{X^2}{T}).
\]

Taking \( T = X^2 \), this boils down to

\[
\sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2}+it}} \ll \epsilon (|t| + X^2)^\epsilon X^\epsilon \ll (1 + |t|)^\epsilon X^\epsilon,
\]

as claimed. \( \square \)

Lemma 2.4. Let \( P \leq X \). Assuming RH,

\[
\sum_{p \sim P} \frac{1}{p^{\frac{1}{2}+it}} \ll (\log X)^2,
\]

uniformly for \( X^\frac{1}{5} \leq |t| \leq X \).
Lemma 3.1. We'll begin with the following lemma, which essentially follows from Perron's Formula (together with a few other tricks):

\[
\sum_{p \sim x} \frac{1}{p^{\frac{1}{2} + \frac{\log x}{2} + it}} = \frac{1}{2\pi i} \int_{\frac{1}{2} - iT}^{\frac{1}{2} + iT} \log \zeta(s + \frac{1}{2} + \frac{1}{\log X} + it) \frac{(2P)^s - P_s}{s} ds + O\left(\frac{X^{3/2} \log X}{T} + \log X\right).
\]

Now, given \( \epsilon > 0 \) and assuming RH, \( \log \zeta(s + \frac{1}{2} + \frac{1}{\log X} + it) \) is well-defined for \( \Re(s) \geq \epsilon \), so long as \( \Im(s) + t \) is bounded away from 0 (this is required so that we stay away from the pole of \( \zeta(s) \) at \( s = 1 \); furthermore, this condition is satisfied if \( T \leq X^{1/2}/2 \), as \( |t| \geq X^{1/2} \), by hypothesis). Taking \( \epsilon = 1/\log X \) and shifting the contour to the edge of this region, we have that

\[
\sum_{p \sim x} \frac{1}{p^{\frac{1}{2} + \frac{\log x}{2} + it}} = -\frac{1}{2\pi i} \left( \int_{\frac{1}{2} - iT}^{\frac{1}{2} + iT} \log \zeta(s + \frac{1}{2} + \frac{1}{\log X} + it) \frac{(2P)^s - P_s}{s} ds \right) + O\left(\frac{X^{3/2} \log X}{T} + \log X\right).
\]

Finally, we use the bound \( |\log \zeta(s + \frac{1}{2} + \frac{1}{\log X} + it)| \ll \log X \) for all \( s \) in the region described above (see §14.3 of [THB86], for example), to obtain

\[
\sum_{p \sim x} \frac{1}{p^{\frac{1}{2} + \frac{\log x}{2} + it}} \ll (\log X) \int_T^{-T} \frac{dt}{\sqrt{\log x^2 + t^2}} + \frac{X^{3/2} \log X}{T} \log X \ll (\log X)^2,
\]

which follows by recalling that \( P \leq X \) and by taking \( T = X^{1/2}/2 \).

\[\square\]

3. Initial Reductions

In this section, we'll reduce our problem to that of bounding the mean square of a Dirichlet polynomial. We'll begin with the following lemma, which essentially follows from Perron's Formula (together with a few other tricks):

Lemma 3.1. For \( 1 \leq h \leq X \),

\[
\frac{1}{X} \int_X^{2X} \left| \frac{1}{h} \sum_{x \leq n \leq x + h} \lambda(n) \right|^2 dx \ll \frac{1}{X} \int_0^{X/h} \left| \sum_{X \leq n \leq 4X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} \right|^2 dt + \max_{T > X/h} \frac{1}{hT} \int_T^{2T} \left| \sum_{X \leq n \leq 4X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} \right|^2 dt.
\]

Proof. See [MR16 Lemma 14] (equivalently, [MR15 Lemma 4]). \[\square\]

Remark 3.1. By splitting the Dirichlet polynomial in Lemma 3.1 as

\[
\sum_{X \leq n \leq 2X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} + \sum_{2X \leq n \leq 4X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}},
\]

and using the fact that \( |a + b|^2 \ll |a|^2 + |b|^2 \), it suffices to consider the first sum over \( n \sim X \) alone (in order to deal with the second sum, we simply replace \( X \) by \( 2X \) in everything that follows).

With the help of Lemma 3.1, we can now remove the contribution from the small values of \( t \):

\[
\frac{1}{X} \int_0^{T_0} \left| \sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} \right|^2 dt \ll \frac{T_0}{X} (1 + T_0)^{X^\epsilon} \ll \delta (X^{1/2})^{\delta - 1}.
\]

The choice of \( \kappa \) is made so that the error incurred from Lemma 2.1 converges. To get a bound of \( (X^{1/2} \log X)/T + \log X \), we follow the same line of reasoning as in the proof of Lemma 3.12 in [THB86], recalling that \( \log \zeta(s) = \sum_{p, m} 1/mp^{ms} \), for \( \Re(s) > 1 \).
as long as $T_0 \leq X^{\frac{1}{2}}$, and for some suitable choice of $\epsilon = \epsilon(\delta)$. If we want this to be bounded above by something like $h^{\delta - 1}$, then we simply restrict ourselves to $h \leq X^{\frac{1}{2}}$.

Putting all of this together, Theorem 1.2 follows from Lemma 3.1 once we show that

$$\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{n \sim X} n^{\frac{1}{2} + it} \right|^2 dt \ll \frac{(\log X)^8}{h},$$

for $X/h \leq T \leq X$, recalling that $h \leq X^{\frac{1}{2}}$. For $T > X$, the Mean Value Theorem (Lemma 2.2) immediately gives the desired bound:

$$\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{n \sim X} n^{\frac{1}{2} + it} \right|^2 dt \ll \frac{1}{hT} (T + X) \sum_{n \sim X} \frac{1}{n} \ll \frac{1}{h}.$$

In the next two sections, we’ll consider the integers $n \sim X$ which have at least one prime factor $p > h$ and those integers $n \sim X$ all of whose prime factors are $\leq h$, respectively.

4. INTÉGERS WITH LARGE PRIME FACTORS

For the integers $n \sim X$ which have at least one prime factor $p > h$, we have the following:

**Proposition 4.1.** If $h \leq X^{\frac{1}{2}}$ and $X/h \leq T \leq X$, then

$$\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{p > h; p|n} \lambda(n) \right|^2 dt \ll \frac{(\log X)^8}{h},$$

where $\frac{h}{p} \leq \frac{T}{X^{\frac{1}{2}}}$. We have the following:

$$\sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} = \sum_{h \leq x \sim X} \frac{\lambda(p)}{p^{\frac{1}{2} + it}} \sum_{m \sim X/p} \frac{1}{m^{\frac{1}{2} + it} \cdot \# \{ \text{prime factor} p > h, q|m \} \cdot \mathbb{1}_{p|m}}.$$

Proof. To begin, note that

$$\sum_{\frac{h}{p} \leq \frac{T}{X^{\frac{1}{2}}}} \frac{\lambda(p)}{p^{\frac{1}{2} + it}} \sum_{m \sim X/p} \frac{\lambda(m)}{m^{\frac{1}{2} + it} \cdot \# \{ \text{prime factor} p > h, q|m \} \cdot \mathbb{1}_{p|m}}$$

where $\frac{h}{p} \leq \frac{T}{X^{\frac{1}{2}}}$ and $X/h \leq T \leq X$, then

$$\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{p > h; p|n} \lambda(n) \right|^2 dt \ll \frac{(\log X)^8}{h}.$$

Our goal now is to remove the dependence on $\mathbb{1}_{p|m}$, which is done by splitting the inner sum into those $m$ for which $p \nmid m$ and $p|m$, respectively:

$$\sum_{h \leq x \sim X} \frac{\lambda(p)}{p^{\frac{1}{2} + it}} \sum_{m \sim X/p} \frac{\lambda(m)}{m^{\frac{1}{2} + it} \cdot \# \{ \text{prime factor} p > h, q|m \} \cdot \mathbb{1}_{p|m}}$$

with $q$ varying over the set of primes. This can be simplified further by adding and subtracting all $m$ for which $p|m$ to the first term and setting

$$a_m := \frac{-\lambda(m)}{\# \{ \text{prime factor} p > h, q|m \} + 1}, b_m := \frac{-\lambda(m)}{\# \{ \text{prime factor} p > h, q|m \} (\# \{ \text{prime factor} p > h, q|m \} + 1)}.$$
which yields
\[
\sum_{h < p \leq 2X} \frac{1}{p^{\frac{1}{2} + it}} \sum_{m \sim X/p} \frac{\lambda(m)}{m^{\frac{1}{2} + it}} \frac{1}{\# \{ q : h \mid q \} + 1} \frac{1}{p|m}
\]
\[
= \sum_{h < p \leq 2X} \frac{1}{p^{\frac{1}{2} + it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2} + it}} + \sum_{h < p \leq 2X} \frac{1}{p^{\frac{1}{2} + it}} \sum_{m \sim X/p} \frac{b_m}{m^{\frac{1}{2} + it}}
\]
\[
= \sum_{h < p \leq 2X} \frac{1}{p^{\frac{1}{2} + it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2} + it}} + \sum_{h < p \leq 2X} \frac{1}{p^{1 + 2it}} \sum_{m \sim X/p^2} \frac{b_m}{m^{\frac{1}{2} + it}},
\]
where the last line follows by writing \( m = mp \).

That being said,
\[
\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} \right|^2 dt \ll \frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{h < p \leq 2X} \frac{1}{p^{\frac{1}{2} + it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2} + it}} \right|^2 dt
\]
\[
+ \frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{h < p \leq 2X} \frac{1}{p^{1 + 2it}} \sum_{m \sim X/p^2} \frac{b_m}{m^{\frac{1}{2} + it}} \right|^2 dt.
\]
Applying the Mean Value Theorem to the second integral, we see that
\[
\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{h < p \leq 2X} \frac{1}{p^{1 + 2it}} \sum_{m \sim X/p^2} \frac{b_m}{m^{\frac{1}{2} + it}} \right|^2 dt
\]
\[
\ll \frac{1}{hT} (T + X) \sum_{h < p \leq 2X} \frac{1}{p^{1 + 2it}} \sum_{m \sim X/p^2} \frac{1}{m} \ll \frac{1}{h},
\]
recalling that \( X/h \leq T \leq X \).

For the remaining integral, we wish to separate the variables \( p \) and \( m \), so that we may apply a pointwise bound to the sum over \( p \). In [MR16], this is done by splitting the sum over \( p \) into shorter intervals and over-counting. We’ll also split the sum over \( p \) into shorter intervals, but we’ll see that it suffices to consider \( p \) in some dyadic interval (as opposed to an interval of the form \([e^{j/H}, e^{(j+1)/H}]\), with \( H \) depending on the decomposition of their Dirichlet polynomial; see either Lemma 5 in [MR15] or Lemma 12 in [MR16]). Furthermore, we’ll see that using a contour integral produces sharper bounds than over-counting: over-counting does not yield square-root cancellation, but it does win by a factor of \( 1/H \). This is enough to get a bound of \( o(Xh^2) \), but getting something like \( \ll Xh^{1+\delta} \) requires taking \( H \) as large as \( h \), which requires \( h \log X \) intervals (and this is now losing).

**Remark 4.1.** Following the comments above, we can also use a dyadic decomposition in [MR16] Lemma 12 (resp. [MR15] Lemma 5), provided we assume RH: using a contour integral to separate the variables requires that we remove the contribution from small \( |t| \) all the way up to \( X^{\frac{1}{2}} \); see Lemma 2.3.
So, let’s begin by splitting the sum over $p$ into dyadic intervals $[2^j, 2^{j+1}]$, with $\log h / \log 2 \leq j \leq \log X / \log 2$:

$$
\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \sum_{h < p \leq 2X} \frac{1}{p^{\frac{1}{2}+it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \left| \sum_{j=\lceil \log_2 h \rceil}^{\lfloor \log_2 2X \rfloor} \sum_{2^j < p \leq 2^{j+1}} \frac{1}{p^{\frac{1}{2}+it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \right|^2 \, dt
$$

$$
= \frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{j=\lceil \log_2 h \rceil}^{\lfloor \log_2 2X \rfloor} \sum_{2^j < p \leq 2^{j+1}} \frac{1}{p^{\frac{1}{2}+it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \right|^2 \, dt
$$

$$
\ll \frac{(\log X)^2}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{j=\lceil \log_2 h \rceil}^{\lfloor \log_2 2X \rfloor} \sum_{2^j < p \leq 2^{j+1}} \frac{1}{p^{\frac{1}{2}+it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \right|^2 \, dt
$$

$$
+ \frac{(\log X)^2}{h} \left( \frac{X^{\frac{1}{2}} \log X}{Y} \right)^2,
$$

for some $\lceil \log_2 h \rceil \leq j < \log_2 2X$, where $\log_2$ denotes the base 2 logarithm function and where the last line follows by taking the absolute value inside the sum over $j$ and noting that there are $\ll \log X$ such dyadic intervals.

Now, we’ll use Lemma 2.1 with $\kappa = 1 / \log X$, to remove the condition that $mp \sim X$:

$$
1_{X^m \sim X} = \frac{1}{2\pi i} \int_{\kappa-iY}^{\kappa+iY} \frac{(2X)^s - X^s}{s} \, ds + O\left( \frac{1/(mp)^\kappa}{\max\{1,Y | \log X/mp|\}} \right),
$$

which yields:

$$
\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{h < p \leq 2X} \frac{1}{p^{\frac{1}{2}+it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \right|^2 \, dt
$$

$$
\ll \frac{(\log X)^2}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{j=\lceil \log_2 h \rceil}^{\lfloor \log_2 2X \rfloor} \sum_{2^j < p \leq 2^{j+1}} \frac{1}{p^{\frac{1}{2}+it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \right|^2 \, dt
$$

$$
+ \frac{(\log X)^2}{h} \left( \frac{X^{\frac{1}{2}} \log X}{Y} \right)^2.
$$

We can now use Minkowski’s Inequality for integrals ([Ste70, Section A.1]) to change the order of integration:

$$
\int_{X^{\frac{1}{2}}}^{T} \left| \int_{\kappa-iY}^{\kappa+iY} \frac{(2X)^s - X^s}{s} \, ds \right|^2 \, dt
$$

$$
\ll \left( \int_{\kappa-iY}^{\kappa+iY} \left( \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{2^j < p \leq 2^{j+1}} \frac{1}{p^{\frac{1}{2}+it}} \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \right|^2 \, dt \right)^\frac{1}{2} \, ds \right)^2.
$$

Then, by taking $Y = X^{\frac{1}{2}}/2$, we can apply Lemma 2.4 to bound the sum over $p$; this yields the upper bound

$$
\ll (\log X)^4 \left( \int_{\kappa-iY}^{\kappa+iY} \frac{1}{s} \left( \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{m \sim X/p} \frac{a_m}{m^{\frac{1}{2}+it}} \right|^2 \, dt \right)^\frac{1}{2} \, ds \right)^2
$$

$$
\ll (\log X)^6 (T + X/h) \sum_{X^{2j+1} \leq m \leq X^{2j-1}} \frac{1}{m}
$$

$$
\ll (\log X)^6 (T + X/h),
$$
where the second to last line follows from the Mean Value Theorem and from the fact that \( j \geq \log h / \log 2 \) and where the additional powers of \( \log X \) come from the bound
\[
\int_{\kappa - Y}^{\kappa + Y} \frac{1}{|s|} \, ds \ll \log X,
\]
recalling that \( \kappa = 1 / \log X \) with \( Y = X^{\frac{1}{2}} / 2 \).

Therefore,
\[
\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{n \sim X} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} \right|^2 \, dt \ll \frac{(\log X)^8}{h},
\]
provided that \( X/h \leq T \leq X \) and \( h \leq X^{\frac{1}{2}} \), which is the desired result. \( \square \)

To complete the proof of Theorem \ref{thm:main} it remains to consider the \( h \)-smooth integers. The next section is dedicated to this task.

### 5. Smooth integers

For the integers \( n \sim X \) all of whose prime factors are \( \leq h \), our goal is to obtain the following estimate:
\[
\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{n \sim X \atop p | n \Rightarrow p \leq h} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} \right|^2 \, dt \ll \frac{1}{h},
\]
for \( X/h \leq T \leq X \) and \( h \leq X^{\frac{1}{2}} \).

We’ll begin with the following lemma, which allows us to count smooth numbers:

**Lemma 5.1.** Let \( \Psi(x, y) \) denote the number of \( y \)-smooth integers up to \( x \). Write \( x = y^u \), so that \( u = \log x / \log y \), and fix \( \epsilon > 0 \). Then,
\[
\Psi(x, y) = xu^{-\left(1+o(1)\right)u},
\]
uniformly in the range \( u \leq y^{1-\epsilon} \), as both \( y \) and \( u \) tend to infinity.

**Proof.** See \cite[Corollary 1.3]{HT93}. \( \square \)

Using Lemma \ref{lem:smooth-integers} in conjunction with the Mean Value Theorem, we then obtain the following:
\[
\frac{1}{hT} \int_{X^{\frac{1}{2}}}^{T} \left| \sum_{n \sim X \atop p | n \Rightarrow p \leq h} \frac{\lambda(n)}{n^{\frac{1}{2} + it}} \right|^2 \, dt \ll \frac{1}{hT} (T + X) \sum_{p \sim X} \frac{1}{p} \ll u^{-\left(1+o(1)\right)u},
\]
which follows after some simplification, recalling that \( X/h \leq T \leq X \), and where we’ve taken \( u = \log X / \log h \).

To complete our proof, we wish to choose the largest possible \( h \) for which \( u^{-\left(1+o(1)\right)u} \leq h^{-1} \).
To this end, we can optimize our choice of $h$ with the following argument: let’s write $u = C \sqrt{\log X \log \log X}$ for some constant $C > 0$, so that $\log h = C \sqrt{\log X \log \log X}$ and

$$\log u^{-(1+o(1))u} = -(1+o(1))u \log u$$

$$= -\frac{C}{2} (1 + o(1)) \sqrt{\log X \log \log X}$$

$$= -\frac{C^2}{2} (1 + o(1)) \log h;$$

in particular,

$$h = \exp \left( \frac{1}{C} \sqrt{\log X \log \log X} \right)$$

and we’re looking for the smallest possible $C > 0$ which allows

$$u^{-(1+o(1))u} = h^{-\frac{C^2}{2} (1 + o(1))} \leq h^{-1}.$$ 

Taking $C = \sqrt{2} + \epsilon$, for some $\epsilon > 0$, does the trick and we’ve proven the following:

**Proposition 5.1.** Suppose $h = h(X) \to \infty$ as $X \to \infty$, then

$$\frac{1}{hT} \int_{X^{1/2}}^{T} \left| \sum_{n \sim X \atop p | n \Rightarrow p \leq n} \lambda(n) \frac{t}{n^{1/2} + t} \right|^2 dt \ll \frac{1}{h},$$

provided $h \leq \exp \left( \sqrt{\left( \frac{1}{2} - o(1) \right) \log X \log \log X} \right)$.

That being said, we can actually state Theorem 1.2 in terms of the Dickman-de Bruijn function, $\rho$, which is the (unique) continuous solution to the following differential-delay equation:

$$\rho(u) = 1 \quad \text{if} \quad 0 \leq u \leq 1$$

$$u \rho'(u) + \rho(u - 1) = 0 \quad \text{if} \quad u > 1.$$

Since $\Psi(x, y) = (1 + o(1)) x \rho(u)$, uniformly for $u \leq y^{1-\epsilon}$, we have the following:

**Theorem 5.1.** Assume Riemann’s Hypothesis holds and suppose $2 \leq h \leq X$. Then,

$$\frac{1}{X} \int_{X}^{2X} \left| \frac{1}{h} \sum_{x \leq n \leq x+h} \lambda(n) \right|^2 dx \ll \delta \left( \frac{\log X}{h} \right)^8 + \rho \left( \frac{\log X}{\log h} \right) + \left( X^{1/2} \right)^{\delta - 1}.$$

as $X \to \infty$.

**Proof.** To see this, simply note that the term $u^{-(1+o(1))u}$ in Lemma 5.1 corresponds to an asymptotic estimate for $\rho(u)$ (see [HT93 Theorem 1.2]), while the term $(X^{1/2})^{\delta - 1}$ corresponds to the contribution from the small values of $t$, which is dominated by $\left( \frac{\log X}{h} \right)^8$ for smaller values of $h$. \qed

6. Acknowledgements

The author would like to thank Maksym Radziwiłł, for suggesting this problem and for his continuous support throughout the research phase of this paper. The author would also like to thank Andrew Granville, for optimizing the argument presented in Section 5. In addition to this, the author thanks Renaud Alie, Alison Inglis, Oleksiy Klurman, Jiajun Mai, Sacha Mangerel, Kaisa Matomäki, Andrei Shubin, Joni Teräväinen, and Peter Zenz, for many helpful discussions and comments. Finally, some of this work was conducted while the author was visiting CalTech, he is grateful for their hospitality.
References

[FI10] John Friedlander and Henryk Iwaniec. *Opera de cribro*, volume 57 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2010.

[HT93] Adolf Hildebrand and Gérard Tenenbaum. Integers without large prime factors. *Journal de Théorie des Nombres de Bordeaux*, 5(2):197–211, 1993.

[IK04] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[MM09] Helmut Maier and Hugh L. Montgomery. The sum of the Mbius function. *Bulletin of the London Mathematical Society*, 41(2):213–226, 2009.

[Mot76] Yoichi Motohashi. On the sum of the Mbius function in a short segment. *Proc. Japan Acad.*, 52(9):477–479, 1976.

[MR15] Kaisa Matomäki and Maksym Radziwiłł. A note on the Liouville function in short intervals. *arXiv e-prints*, page arXiv:1502.02374, 2015.

[MR16] Kaisa Matomäki and Maksym Radziwiłł. Multiplicative functions in short intervals. *Annals of Mathematics*, 183(3):1015–1056, 2016.

[MT19] Kaisa Matomäki and Joni Teräväinen. On the Möbius function in all short intervals. *arXiv e-prints*, page arXiv:1911.09076, 2019.

[Ram76] Kanakanahalli Ramachandra. Some problems of analytic number theory. *Acta Arithmetica*, 31(4):313–324, 1976.

[Sel43] Atle Selberg. On the normal density of primes in small intervals, and the difference between consecutive primes. *Arch. Math. Naturvid.*, 47(6):87–105, 1943.

[Sou09] Kannan Soundararajan. Partial sums of the Mbius function. *Journal Fur Die Reine Und Angewandte Mathematik - J REINE ANGEW MATH*, 2009:141–152, 06 2009.

[Sou16] Kannan Soundararajan. The Liouville function in short intervals [after Matomäki and Radziwiłł]. *arXiv e-prints*, page arXiv:1606.08021, 2016.

[Ste70] Elias M. Stein. *Singular Integrals and Differentiability Properties of Functions (PMS-30)*. Princeton University Press, 1970.

[Ten15] Gérard Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 163 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.

[THB86] Edward C. Titchmarsh and David R. Heath-Brown. *The theory of the Riemann zeta-function*. Clarendon Press, Oxford, 1986.