

DUALITY BETWEEN PRIME FACTORS AND THE PRIME NUMBER THEOREM FOR ARITHMETIC PROGRESSIONS - II

Krishnaswami Alladi and Jason Johnson

Dedicated to George Andrews and Bruce Berndt for their 85th birthdays

ABSTRACT: *In the first paper under this title (1977), the first author utilized a duality identity between the largest and smallest prime factors involving the Moebius function, to establish the following result as a consequence of the Prime Number Theorem for Arithmetic Progressions: If k and ℓ are positive integers, with $1 \leq \ell \leq k$ and $(\ell, k) = 1$, then*

$$\sum_{n \geq 2, p(n) \equiv \ell \pmod{k}} \frac{\mu(n)}{n} = \frac{-1}{\phi(k)},$$

where $\mu(n)$ is the Moebius function, $p(n)$ is the smallest prime factor of n , and $\phi(k)$ is the Euler function. Here we utilize the next level Duality identity between the second largest prime factor and the smallest prime factor, involving the Moebius function and $\omega(n)$, the number of distinct prime factors of n , to establish the following result as a consequence of the Prime Number Theorem for Arithmetic Progressions: For all ℓ and k as above,

$$\sum_{n \geq 2, p(n) \equiv \ell \pmod{k}} \frac{\mu(n)\omega(n)}{n} = 0.$$

A quantitative version of this result is proved.

Keywords: Duality between prime factors, Moebius function, number of prime factors, smallest prime factor, largest prime factor, second largest prime factor, prime number theorem for arithmetic progressions, extension to number fields.

Mathematics Subject Classification: 11M06, 11M11, 11N25, 11N37, 11N60

1 Background

Two famous results of Edmund Landau are that

$$M(x) := \sum_{1 \leq n \leq x} \mu(n) = o(x), \quad \text{as } x \rightarrow \infty \tag{1.1}$$

and

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0 \tag{1.2}$$

are (elementarily) equivalent to the Prime Number Theorem (PNT), where $\mu(n)$ is the Moebius function. Similarly, there are results equivalent to the Prime Number Theorem for Arithmetic Progressions (PNTAP) in which $\mu(n)$ is replaced by $\mu(n)\chi(n)$, where $\chi(n)$ is a Dirichlet character modulo k , when the arithmetic progression under consideration has common difference k .

In [2], the first author noticed the following interesting Duality identities involving the Moebius function that connect the smallest and largest prime factors of integers:

$$\sum_{2 \leq d|n} \mu(d)f(p(d)) = -f(P(n)), \tag{1.3}$$

and

$$\sum_{2 \leq d|n} \mu(d)f(P(d)) = -f(p(n)), \tag{1.4}$$

where for $d > 1$, $p(d)$ and $P(d)$ denote the smallest and largest prime factors of d respectively, and f is ANY function on the primes. Using (1.3) and properties of the Moebius function, it was shown in [2] that if f is a bounded function on the primes such that

$$\lim_{x \rightarrow \infty} \frac{1}{x} \sum_{1 \leq n \leq x} f(P(n)) = c, \tag{1.5}$$

then

$$\sum_{n=2}^{\infty} \frac{\mu(n)f(p(n))}{n} = -c, \tag{1.6}$$

and vice-versa. This is a surprising generalization of Landau's result (1.2). To realize this is a generalization, rewrite (1.2) as

$$\sum_{n=2}^{\infty} \frac{\mu(n)}{n} = -1. \tag{1.7}$$

Then (1.7) follows from (1.5) and (1.6) by taking $f(p) = 1$ for all primes p .

Next it was shown in [2] that the PNTAP implies that the sequence $P(n)$ of largest prime factors is uniformly distributed in the reduced residue classes modulo a positive integer k . So if f is chosen to be the characteristic function of primes in an arithmetic progression $\ell(mod k)$, then for such f , (1.5) holds with $c = 1/\phi(k)$, and therefore

$$\sum_{n \geq 2, p(n) \equiv \ell(mod k)} \frac{\mu(n)}{n} = \frac{-1}{\phi(k)}, \tag{1.8}$$

for ALL positive integers k and any ℓ satisfying $(\ell, k) = 1$. This is even more surprising because it gives a way of slicing the convergent series in (1.7) into $\phi(k)$ subseries all converging to the same value! As far as we know, this is the first example of slicing convergent series into equal valued subseries. In the last

few years, (1.8) has received considerable attention and has been generalized in the setting of algebraic number theory (see [5], [8], [10], [16], [21], and [22]).

In [2], the following more general duality identities were noted: For a positive integer k let $P_k(n)$ and $p_k(n)$ denote the k -th largest and k -th smallest prime factors n respectively (defined by strict inequalities), if n has at least k distinct prime factors. Also let $\omega(n)$ denote the number of distinct prime factors of n . Then

$$\sum_{1 < d|n}^* \mu(d) f(P_k(d)) = (-1)^k \binom{\omega(n) - 1}{k - 1} f(p_1(n)), \quad (1.9)$$

and

$$\sum_{1 < d|n}^* \mu(d) f(p_k(d)) = (-1)^k \binom{\omega(n) - 1}{k - 1} f(P_1(n)), \quad (1.10)$$

where the $*$ over the summation means that if n has fewer than k distinct prime factors, then the sum is zero. (NOTE: When $k = 1$, we often write $p(n)$ and $P(n)$ in place of $p_1(n)$ and $P_1(n)$, respectively.) From (1.9) and (1.10), it follows by Moebius inversion that

$$\sum_{1 < d|n} \mu(d) \binom{\omega(d) - 1}{k - 1} f(P_1(d)) = (-1)^k f(p_k(n)), \quad (1.11)$$

and

$$\sum_{1 < d|n} \mu(d) \binom{\omega(d) - 1}{k - 1} f(p_1(d)) = (-1)^k f(P_k(n)). \quad (1.12)$$

In these identities, we adopt the convention that $f(P_k(n)) = f(p_k(n)) = 0$ if $\omega(n) < k$, that is, if n has fewer than k distinct prime factors.

In this paper we will discuss consequences of (1.12) in the case $k = 2$, that is the identity

$$\sum_{1 < d|n} \mu(d) (\omega(d) - 1) f(p(d)) = f(P_2(n)) \quad (1.13)$$

and its implications.

As a start, analogous to (1.2), we establish (see Theorem 4 of §2) that

$$\sum_{n=1}^{\infty} \frac{\mu(n) \omega(n)}{n} = 0. \quad (1.14)$$

Since $\omega(1) = 0$, (1.14) gives

$$\sum_{n=2}^{\infty} \frac{\mu(n) \omega(n)}{n} = 0, \quad (1.15)$$

unlike (1.7). Next we establish (see Theorem 7 of §5), that for each positive integer k , the sequence $P_2(n)$ of second largest prime factors, is uniformly distributed in the reduced residue classes $\ell \pmod{k}$. From

this uniform distribution result, (1.15), and by choosing f to be the characteristic function for primes $\equiv \ell \pmod{k}$, we prove (see Theorem 10 of §6) that

$$\sum_{n \geq 2, p_1(n) \equiv \ell \pmod{k}} \frac{\mu(n)\omega(n)}{n} = 0, \quad (1.16)$$

and this is the main result of the paper.

To establish Theorems 7 and 10, several auxiliary results are proved. All the theorems in this paper are proved in quantitative form.

Notations and Conventions: In what follows, c, c_1, c_2, \dots are absolute positive constants whose values will not concern us. The \ll and O notations are equivalent and will be used interchangeably as is convenient. We also adopt the convention that

$$f(x) \ll g(x) \quad \text{means} \quad |f(x)| < K|g(x)|,$$

with x ranging in some domain, and K a positive constant. Implicit constants are absolute unless otherwise indicated with a subscript. Although our results can be established with uniformity by allowing the modulus k to grow slowly as a function of x , we only consider here an arbitrary but fixed modulus k . The alphabet n whether used as the argument of a function, or in a summation, will always be a positive integer. Also, any time we have a sum over p , or have p as an argument of a function, it is to be understood that p is prime.

By $E(x, k, \ell)$ we mean the difference

$$E(x, k, \ell) = \pi(x, k, \ell) - \frac{\ell i(x)}{\phi(k)},$$

where

$$\pi(x, k, \ell) = \sum_{p \leq x, p \equiv \ell \pmod{k}} 1, \quad \text{and} \quad \ell i(x) = \int_2^x \frac{dt}{\log t}.$$

When k and ℓ are specific, we simply use $E(x)$ in place of $E(x, k, \ell)$. Finally, by $R(x)$ we mean any decreasing function of x that tends to zero as $x \rightarrow \infty$ and bounds from above the relative error in the PNTAP; that is

$$\left| \pi(x, k, \ell) - \frac{\ell i(x)}{\phi(k)} \right| < \frac{\ell i(x)}{\phi(k)} R(x).$$

In what follows, we will choose

$$R(x) = e^{-c\sqrt{\log x}},$$

where c can be any positive constant. Also $T = T(x)$ will be a function which will be chosen optimally to get suitable bounds in various estimates, but T will not necessarily be the same in different contexts. We shall use the standard notation $[x]$ for the integral part of a real number x , and $\{x\}$, where indicated, will denote the fractional part of x , namely $x - [x]$. Finally, complex numbers will be denoted either by z , or by $s = \sigma + it$ when dealing with Dirichlet series.

2 The Moebius function and the number of prime factors

Here and throughout, by $C = C(x)$ we mean the rectangular contour whose corners in the complex plane are given by

$$\left(1 + \frac{1}{\log x}, -T\right), \quad \left(1 + \frac{1}{\log x}, T\right), \quad \left(-\frac{1}{\log T}, T\right), \quad \left(-\frac{1}{\log T}, -T\right), \quad (2.1)$$

where $x \geq 3$ and $T = \exp\sqrt{\log x}$. Inside and on this contour, we know that for some absolute constant $c_1 > 0$,

$$\left|\frac{1}{\zeta(s)}\right| \ll (\log x)^{c_1}, \quad (2.2)$$

where $\zeta(s)$ is the Riemann zeta function. Since

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}, \quad \text{for } \operatorname{Re}(s) > 1, \quad (2.3)$$

the Perron integral method applied to the contour C , together with (2.2) gives

$$M(x) := \sum_{n \leq x} \mu(n) \ll x e^{-c_2 \sqrt{\log x}}. \quad (2.4)$$

Instead of summing over all positive integers n as in (2.3), if we sum only over those integers which are not multiples of a certain prime p , then we have

$$\sum_{n \geq 1, (n,p)=1} \frac{\mu(n)}{n^s} = \left(1 - \frac{1}{p^s}\right)^{-1} \frac{1}{\zeta(s)}, \quad \text{for } \operatorname{Re}(s) > 1. \quad (2.5)$$

Now on the contour C , we have the bound

$$\left|\left(1 - \frac{1}{p^\sigma}\right)^{-1}\right| \leq \left(1 - \frac{1}{p^\sigma}\right)^{-1} \leq \left(1 - \frac{1}{2^\sigma}\right)^{-1} \leq \left(1 - \frac{1}{2\sqrt{\log 3}}\right)^{-1} \quad (2.6)$$

valid uniformly for ALL primes p , where $\sigma = 1 - 1/\log T$. Thus the Perron integral method that yielded the bound for $M(x)$ in (2.4), now gives in view of (2.6), the following bound

$$M^{(p)}(x) := \sum_{n \leq x, (n,p)=1} \mu(n) \ll x e^{-c_2 \sqrt{\log x}}, \quad (2.7)$$

valid uniformly for ALL primes p .

Next we consider the sum

$$M_\omega(x) := \sum_{1 \leq n \leq x} \mu(n) \omega(n).$$

Even though our focus here is n square-free, for the purpose of employing the Moebius inversion formula, it is convenient to also consider the function $\Omega(n)$, which is the (total) number of prime factors of n counted with multiplicity, because the function Ω is a totally additive.

If χ_P denotes the characteristic function of the prime powers, then

$$\Omega(n) = \sum_{d|n} \chi_P(d).$$

Thus by Moebius inversion, we have

$$\begin{aligned} \chi_P(n) &= \sum_{d|n} \mu(d) \Omega\left(\frac{n}{d}\right) \\ &= \Omega(n) \sum_{d|n} \mu(d) - \sum_{d|n} \mu(d) \Omega(d) = - \sum_{d|n} \mu(d) \Omega(d), \end{aligned} \quad (2.8)$$

because $\Omega(n) \sum_{d|n} \mu(d)$ is identically zero. Since $\mu(d) = 0$ if d is not square-free, we may rewrite (2.8) as

$$\sum_{d|n} \mu(d) \omega(d) = -\chi_P(n), \quad (2.9)$$

and Moebius inversion applied to (2.9) yields

$$\mu(n) \omega(n) = - \sum_{d|n} \chi_P(d) \mu\left(\frac{n}{d}\right). \quad (2.10)$$

Our first result is:

Theorem 1: *For $x \geq 2$, we have*

$$M_\omega(x) := \sum_{n \leq x} \mu(n) \omega(n) \ll \frac{x}{\log x}.$$

Proof: Use

$$\begin{aligned} M_\omega(x) &= \sum_{n \leq x} \mu(n) \omega(n) = \sum_{n \leq x} \mu(n) \sum_{p|n} 1 \\ &= \sum_{p \leq x} \sum_{n \leq x, n \equiv 0 \pmod{p}} \mu(n). \end{aligned} \quad (2.11)$$

In (2.11), in the inner sum on the right, put $n = mp$, with $(m, p) = 1$, to rewrite it as

$$M_\omega(x) = - \sum_{p \leq x} \sum_{m \leq x/p, (m,p)=1} \mu(m) = - \sum_{p \leq x} M^{(p)}\left(\frac{x}{p}\right) = \Sigma_1 + \Sigma_2, \quad (2.12)$$

where

$$\Sigma_1 = \sum_{p \leq T} M^{(p)}\left(\frac{x}{p}\right) \quad \text{and} \quad \Sigma_2 = \sum_{T < p \leq x} M^{(p)}\left(\frac{x}{p}\right), \quad (2.13)$$

and T will be chosen optimally below.

To estimate Σ_1 and Σ_2 we use the well-known estimate

$$\sum_{p \leq x} \frac{1}{p} = \log \log x + c_3 + O(e^{-c_4 \sqrt{\log x}}), \quad \text{for } x \geq 3, \quad (2.14)$$

and its consequence

$$\sum_{y < p \leq x} \frac{1}{p} = \log \log x - \log \log y + O(e^{-c_4 \sqrt{\log y}}), \quad \text{for } 3 \leq y \leq x. \quad (2.15)$$

So from (2.7) and (2.14) it follows that

$$\Sigma_1 \ll \sum_{p \leq T} \frac{x}{p} e^{-c_2 \sqrt{\log(x/p)}} \ll x \log \log x e^{-c_2 \sqrt{\log(x/T)}}. \quad (2.16)$$

To estimate Σ_2 , we break it up as follows:

$$\Sigma_2 = \sum_{m < (x/T)^{-1}} \sum_{\frac{x}{m+1} < p \leq \frac{x}{m}}, \quad (2.17)$$

and denote the inner sum in (2.17) as $\Sigma^{(m)}$. We will set $T = x^{1-\varepsilon}$ and will choose $\varepsilon \rightarrow 0$ optimally as $x \rightarrow \infty$. Thus we get the following bound for $\Sigma^{(m)}$ by using (2.7) and (2.15):

$$\begin{aligned} \Sigma^{(m)} &\ll x e^{-c_2 \sqrt{\log m}} \left\{ \left(\log \log \frac{x}{m} - \log \log \frac{x}{m+1} \right) + O(e^{-c_4 \sqrt{\log(x/m)}}) \right\} \\ &\ll x e^{-c_2 \sqrt{\log m}} \frac{1}{m \log x} + x e^{-c_2 \sqrt{\log m}} e^{-c_4 \sqrt{\log(x/m)}}. \end{aligned} \quad (2.18)$$

Note that $\frac{x}{T} = x^\varepsilon$, and that

$$\sum_{m=1}^{\infty} \frac{e^{-c_2 \sqrt{\log m}}}{m} < \infty.$$

Thus by summing the expression on the right in (2.18) over $m \leq x^\varepsilon$, we get

$$\begin{aligned} \Sigma_2 &\ll \frac{x}{\log x} + x \sum_{m < x^\varepsilon} e^{-c_2 \sqrt{\log m}} e^{-c_4 \sqrt{\log(x/m)}} \\ &\ll \frac{x}{\log x} + x^{1+\varepsilon} e^{-\sqrt{(1-\varepsilon) \log x}}. \end{aligned} \quad (2.19)$$

Finally, we choose

$$\varepsilon = c_5 \frac{(\log \log x)^2}{\log x}, \quad \text{with } c_5 = 4c_2^{-2}. \quad (2.20)$$

So (2.16), (2.19) and (2.20) yield

$$\Sigma_1 \ll \frac{x \log \log x}{\log^2 x} \quad \text{and} \quad \Sigma_2 \ll \frac{x}{\log x}. \quad (2.21)$$

Theorem 1 follows from (2.12) and (2.21). □

Remark: If we choose $c_5 = N^2 c_2^{-2}$, with N arbitrarily large, then we would get

$$\Sigma_1 \ll_N \frac{x \log \log x}{\log^N x},$$

but this is of no use since Σ_2 is bounded only by $x/\log x$, and not any better by the above method. For a sharper estimate for $M_\omega(x)$ due to Tenenbaum by analytic methods see (3.15) below.

Next, using Theorem 1 and following an idea of Axer, we prove:

Theorem 2: *With $\{w\}$ denoting the fractional part of w , we have*

$$\sum_{n \leq x} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} \ll \frac{x \sqrt{\log \log x}}{\sqrt{\log x}}.$$

Proof: Begin by splitting

$$\sum_{n \leq x} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} = \sum_{n \leq T} + \sum_{T < n \leq x} = \Sigma_3 + \Sigma_4, \quad (2.22)$$

with $T = T(x)$ to be determined below.

For Σ_3 , we use the trivial bound

$$|\Sigma_3| \leq \sum_{n \leq T} \omega(n) \ll T \log \log T. \quad (2.23)$$

To estimate Σ_4 , we use partial summation:

$$\begin{aligned} \Sigma_4 &= \sum_{T < n \leq x} (M_\omega(n) - M_\omega(n-1)) \left\{ \frac{x}{n} \right\} \\ &= \sum_{T < n \leq x-1} M_\omega(n) \left(\left\{ \frac{x}{n} \right\} - \left\{ \frac{x}{n+1} \right\} \right) + O(|M_\omega(x)|) + (|M_\omega\left(\frac{x}{T}\right)|). \end{aligned} \quad (2.24)$$

From (2.24) and Theorem 1 we deduce that

$$|\Sigma_4| \ll \frac{x}{\log x} + \sum_{T < n \leq x-1} |M_\omega(n)| \left| \left\{ \frac{x}{n} \right\} - \left\{ \frac{x}{n+1} \right\} \right| \ll \frac{x}{\log x} \sum_{T < n \leq x-1} \left| \left\{ \frac{x}{n} \right\} - \left\{ \frac{x}{n+1} \right\} \right|. \quad (2.25)$$

At this stage, we note that

$$\sum_{T < n \leq x} \left| \left\{ \frac{x}{n} \right\} - \left\{ \frac{x}{n+1} \right\} \right| \leq V_\Omega[1, \frac{x}{T}] \ll \frac{x}{T}, \quad (2.26)$$

where $V_\Omega[a, b]$ is the total variation of $\{x\}$ on $[a, b]$. So by (2.25) and (2.26) we have

$$\Sigma_4 \ll \frac{x}{\log x} \frac{x}{T}. \quad (2.27)$$

Thus (2.22), (2.23) and (2.27) yield

$$\sum_{n \leq x} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} \ll T \log \log T + \frac{x}{\log x} \frac{x}{T}. \quad (2.28)$$

We need to choose T optimally to minimize the right hand side of (2.28). The choice

$$T = \frac{x}{\sqrt{\log x \log \log x}}$$

in (2.28) yields

$$\sum_{n \leq x} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} \ll \frac{x \sqrt{\log \log x}}{\sqrt{\log x}}$$

which proves Theorem 2. □

Since Theorem 2 deals with the fractional part function as the weight, we establish next the corresponding result with the weight as the integral part function:

Theorem 3:

$$\sum_{n \leq x} \mu(n) \omega(n) \left[\frac{x}{n} \right] = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right).$$

Proof: Note that (2.9) yields

$$\begin{aligned} \sum_{n \leq x} \mu(n) \omega(n) \left[\frac{x}{n} \right] &= \sum_{n \leq x} \sum_{d|n} \mu(d) \omega(d) \\ &= \sum_{n \leq x} \chi_P(n) = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right), \end{aligned}$$

which proves Theorem 3.

Theorems 2 and 3 lead to the the main result of this section:

Theorem 4:

$$m_\omega(x) := \sum_{n \leq x} \frac{\mu(n) \omega(n)}{n} = O\left(\frac{\sqrt{\log \log x}}{\sqrt{\log x}}\right). \quad (2.29)$$

In particular

$$\sum_{n=1}^{\infty} \frac{\mu(n) \omega(n)}{n} = \sum_{n=2}^{\infty} \frac{\mu(n) \omega(n)}{n} = 0.$$

Proof: From Theorems 2 and 3, we get

$$\begin{aligned} \sum_{n \leq x} \mu(n) \omega(n) \frac{x}{n} &= \sum_{n \leq x} \mu(n) \omega(n) \left[\frac{x}{n} \right] + \sum_{n \leq x} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} \\ &\ll \frac{x \sqrt{\log \log x}}{\sqrt{\log x}} + \frac{x}{\log x}. \end{aligned} \quad (2.30)$$

By cancelling x on both extremes of (2.30), we get (2.29) of Theorem 4. By letting $x \rightarrow \infty$ in (2.29), we get (1.14) and (1.15), which is the second assertion of Theorem 4. \square

Remark: When I communicated the results of this section to Tenenbaum, he responded [19] by saying that he can establish stronger quantitative versions of the Theorems 1 and 4 by the Selberg-Delange analytic method. We describe Tenenbaum's approach and state his stronger quantitative results in the next section. But we mention here that the actual results of §2 and §3 are not used to establish our main result in §6, but the elementary method used in this section is what is employed in §6. The discussion on $M_\omega(x)$ and $m_\omega(x)$ in this and the next section provides a context to understand our main result.

3 Analytic approach to sums of $\mu(n)\omega(n)$

For a complex number z , represent the function $\zeta(s)^z$ as a Dirichlet series

$$\zeta(s)^z = \sum_{n=1}^{\infty} \frac{d_z(n)}{n^s}, \quad \text{for } \operatorname{Re}(s) > 1. \quad (3.1)$$

The function $d_z(n)$ is called the generalized divisor function because $d_2(n) = d(n)$ is the standard divisor function. In a fundamental paper, Selberg [14] showed that

$$\sum_{n \leq x} d_z(n) = \frac{x(\log x)^{z-1}}{\Gamma(z)} + O_R(x(\log x)^{z-2}) \quad (3.2)$$

is valid uniformly for $|z| \leq R$.

In order to establish (3.2), Selberg used the Perron integral method, but since $\zeta(s)^z$ has a branch point singularity at $s = 1$ when z is not an integer, he modified the contour $C(x)$ in §2 by replacing the short line segment from $b - i(2 \log T)^{-1}$ to $b + i(\log T)^{-1}$ on $C(x)$ with the following lacet L around $s = 1$:

L starts at $b - i(2 \log x)^{-1}$, runs parallel to the x -axis until $1 - i(2 \log x)^{-1}$, then encircles $s = 1$ in a semi-circle of radius $(2 \log x)^{-1}$, and ends with a line segment from $1 + i(\log x)^{-1}$ to $b + i(\log x)^{-1}$.

The contribution around L leads to the Hankel contour for the Gamma function, and that explains the presence of $\Gamma(z)$ in (3.2).

Selberg's method applies more generally to sums of coefficients of Dirichlet series representing functions of the type

$$\zeta(s)^z \cdot H(s),$$

where $H(s)$ would be analytic in the half plane $\operatorname{Re}(s) > \frac{1}{2}$, and indeed Selberg considered a few important such $H(s)$ in [14]. A case of interest to us here is the sum

$$S_{-z}(x) := \sum_{n \leq x} \mu(n) z^{\omega(n)} \quad (3.3)$$

which can be viewed as the sum of the coefficients of the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{\mu(n) z^{\omega(n)}}{n^s} = \prod_p \left(1 - \frac{z}{p^s}\right) \quad \text{for } \operatorname{Re}(s) > 1. \quad (3.4)$$

We could rewrite (3.4) as

$$\sum_{n=1}^{\infty} \frac{\mu(n)z^{\omega(n)}}{n^s} = \zeta(s)^{-z}G(s, z), \quad (3.5)$$

where

$$G(s, z) = \prod_p \left(1 - \frac{z}{p^s}\right) \left(1 - \frac{1}{p^s}\right)^{-z}, \quad (3.6)$$

is analytic in $Re(s) > \frac{1}{2}$, and uniformly bounded in the half plane $Re(s) \geq \frac{1}{2} + \delta$, for each $\delta > 0$. So by the Selberg method, one gets

$$S_{-z}(x) = \frac{x(\log x)^{-z-1}G(1, z)}{\Gamma(-z)} + O_R(x(\log x)^{-z-2}), \quad (3.7)$$

is uniformly valid for $|z| \leq R$, and indeed this is implicit in [13].

Selberg's method was extended by Delange to deal with sums of coefficients of Dirichlet series convergent in $Re(s) > 1$, and can be represented as

$$\zeta(s)^z (\log \zeta(s))^k H(s), \quad (3.8)$$

where k is a non-negative integer, and $H(s)$ is analytic in $Re(s) > \frac{1}{2}$. This is a natural extension of Selberg's method (and one to be expected) because $\zeta^z(s) = e^{z \log \zeta(s)}$, but it is a useful extension. Tenenbaum [18] has a thorough account of the Selberg-Delange method in its most general form, and in doing so, has improved on the quantitative aspects as well.

We can view $M_\omega(x)$ as

$$M_\omega(x) = \frac{d}{dz} S_{-z}(x)|_{z=1}. \quad (3.9)$$

So one may heuristically get from (3.7) that

$$M_\omega(x) \sim \frac{d}{dz} \left(\frac{x(\log x)^{-z-1}G(1, z)}{\Gamma(-z)} \right) \Big|_{z=1}. \quad (3.10)$$

Note that,

$$\frac{G(1, z)}{\Gamma(-z)}$$

has a simple zero at $z = 1$, and so (3.10) would yield

$$M_\omega(x) \sim \frac{c_7 x}{(\log x)^2}, \quad (3.11)$$

with some non-zero constant. But Tenenbaum in his letter [19] established (3.11) by the Selberg-Delange method, as well as a precise series expansion as detailed below.

Observe that

$$\sum_{n=1}^{\infty} \frac{\mu(n)\omega(n)}{n^s} = \frac{d}{dz} \left(\sum_{n=1}^{\infty} \frac{\mu(n)z^{\omega(n)}}{n^s} \right) \Big|_{z=1}, \quad \text{for } Re(s) > 1, \quad (3.12)$$

because the term-by-term differentiation of the Dirichlet series on the right is valid in $Re(s) > 1$. So by (3.5) and (3.12) we have

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{\mu(n)\omega(n)}{n^s} &= \frac{d}{dz}(\zeta^{-z}(s)G(s, z))|_{z=1} \\ &= -\zeta^{-z}(s) \log \zeta(s)G(s, z)|_{z=1} + \zeta^{-z}(s)G'(s, z)|_{z=1} \\ &= -\zeta^{-1}(s) \log \zeta(s)G(s, 1) + \zeta^{-1}(s)G'(s, 1), \quad \text{for } Re(s) > 1. \end{aligned} \tag{3.13}$$

In view of the representation in (3.13) for the Dirichlet series, the Selberg-Delange method can be applied using the function on the right in (3.13), to deduce that

$$M_{\omega}(x) = \frac{c_7 x}{\log^2 x} + O\left(\frac{x}{\log^3 x}\right). \tag{3.14}$$

Tenenbaum [19] notes that Theorem 11.5.2 in his book [18] readily yields the following more precise estimate:

$$M_{\omega}(x) = x \sum_{0 \leq k \leq N} \frac{\lambda_k}{\log^{k+2} x} + O(xR_{N+2}(x)), \tag{3.15}$$

where the λ_k are constants, with $\lambda_0 = 1$ (consequently $c_7 = 1$), and

$$R_N(x) = e^{-\sqrt{\log x}} + O\left(\left(\frac{c_8 N + 1}{\log x}\right)^{N+1}\right). \tag{3.16}$$

But, as noted in the previous section, such a precise result as (3.16) for $M_{\omega}(x)$ is not needed for us here.

It follows from (3.15) by partial summation that

$$\sum_{n \leq x} \frac{\mu(n)\omega(n)}{n} = c_9 + O\left(\frac{1}{\log x}\right),$$

and so the series

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

is convergent to c_9 . If we call the expression on the right in (3.13) as $F(s, 1)$, then Tenenbaum [19] notes that

$$c_9 = \lim_{\sigma \rightarrow 1^+} F(\sigma, 1) = 0$$

and so

$$\sum_{n \leq x} \frac{\mu(n)\omega(n)}{n} \ll \frac{1}{\log x} \tag{3.17}$$

which is stronger than our Theorem 4, which we derived elementarily from the strong form of the Prime Number Theorem.

Indeed, in the spirit of (3.15), Tenenbaum [19] proved the stronger result

$$\sum_{n \leq x} \frac{\mu(n)\omega(n)}{n} = \sum_{1 \leq k \leq N} \frac{\nu_k}{\log^k x} + O(R_{N+1}(x)). \quad (3.18)$$

which follows by directly applying the Selberg-Delange method to evaluate the sum by considering the associated Dirichlet series

$$\sum_{n=1}^{\infty} \frac{\mu(n)z^{\omega(n)}}{n^{s+1}}.$$

Some key differences:

Denote by

$$m(x) := \sum_{n \leq x} \frac{\mu(n)}{n}. \quad (3.19)$$

It is known by a theorem of Landau on Dirichlet series whose coefficients are eventually of the same sign (with a similar result for Dirichlet type integrals), that $M(x)$ changes sign infinitely often, and that $m(x)$ changes sign infinitely often as it converges to 0 as $x \rightarrow \infty$. In contrast, in view of the fact that $M_\omega(x)$ can be estimated asymptotically with a leading term as in (3.14), $M_\omega(x)$ will eventually be of the same sign, and therefore will NOT change sign infinitely often. Similarly,

$$m_\omega(x) := \sum_{n \leq x} \frac{\mu(n)\omega(n)}{n} \quad (3.20)$$

will be eventually of the same sign as it tends to 0 when $x \rightarrow \infty$, and so will NOT change sign infinitely often.

The other key difference is in their sizes. Whereas the strong form of the Prime Number Theorem implies that

$$M(x) = O(xe^{-c\sqrt{\log x}}) \quad \text{and} \quad m(x) = O(e^{-c\sqrt{\log x}}),$$

we have

$$M_\omega(x) \sim c_7 \frac{x}{\log^2 x} \quad \text{and} \quad m_\omega(x) \sim c_{10} \frac{1}{\log x}.$$

It is however to be noted that by writing

$$\sum_{n=1}^{\infty} \frac{\mu(n)\omega(n)}{n} = \int_{1-}^{\infty} \frac{dM_\omega(x)}{x}, \quad (3.21)$$

and using integration-by-parts, we get

$$\sum_{n=1}^{\infty} \frac{\mu(n)\omega(n)}{n} = \int_1^{\infty} \frac{M_\omega(x)}{x^2} dx = 0. \quad (3.22)$$

Thus $M_\omega(x)$ must change sign, even though it does not change sign infinitely often!

4 The sizes of the largest and second largest prime factors

The fundamental counting function associated with the largest prime factor $P(n)$ is

$$\Psi(x, y) = \sum_{n \leq x, P(n) \leq y} 1. \quad (4.1)$$

Here and in what follows, we shall denote by α , the quantity $\frac{\log x}{\log y}$. In an important paper [7], de Bruijn showed that with some constant $c > 0$,

$$\Psi(x, y) \ll x e^{-c\alpha}, \quad \text{uniformly for } 2 \leq y \leq x. \quad (4.2)$$

Tenenbaum [19; Theorem III.5.1] has shown that (4.2) holds with $c = 1/2$. de Bruijn also proved that

$$\Psi(x, y) \ll x \log^2 y e^{-\alpha \log \alpha - \alpha \log \log \alpha + O(\alpha)}, \quad \text{for } y > \log^2 x, \quad (4.3)$$

and indeed the uniform asymptotic estimate

$$\Psi(x, y) \sim x \rho(\alpha), \quad \text{for } e^{(\log x)^{3/5}} \leq y \leq x, \quad (4.4)$$

where ρ satisfies the integro-difference equation

$$\rho(\alpha) = 1 - \int_1^\alpha \frac{\rho(u-1) du}{u}$$

and (de Bruijn [6])

$$\rho(\alpha) = e^{-\alpha \log \alpha - \alpha \log \log \alpha + O(\alpha)}.$$

Thus $\Psi(x, y)$ is quite small in comparison with x when α is large.

Remark: Note that (4.3) is of no use when $\alpha > 1$ is fixed, because trivially $\Psi(x, y) \leq x$ (!); so (4.3) is used only when $\alpha \rightarrow \infty$ with x . Much better estimates for $\Psi(x, y)$ are known including those that significantly extend the range of the asymptotic formula (4.4); for such superior results, see Hildebrand and Tenenbaum [9]. For our purpose here, these superior results on $\Psi(x, y)$ are not needed; the above bounds suffice.

Next consider $P_2(n)$, the second largest prime factor of n . Note that whereas $P(n)$ is uniquely defined, there are two ways to define the second largest prime factor. We could define $P_2(n) = P(n/P(n))$ or $P_2(n)$ as the largest prime factor of n strictly less than $P(n)$. In the former definition, we set $P_2(n) = 1$ if $\Omega(n) < 2$, and in the latter definition we set $P_2(n) = 1$, if $\omega(n) < 2$. From the point of view of asymptotic estimates, there is little difference between the two ways of defining $P_2(n)$. This is made precise by:

Theorem 5: *Let $N(x)$ denote the number of positive integers $n \leq x$ for which $P(n)$ repeats. Then*

$$N(x) \ll \frac{x}{e^{(\frac{1}{2}+o(1))\sqrt{\log x \log \log x}}}.$$

Proof: By (4.3),

$$\Psi(x, e^{\sqrt{\log x \log \log x}}) \ll \frac{x}{e^{(\frac{1}{2}+o(1))\sqrt{\log x \log \log x}}}.$$

So it suffices to consider those integers n for which $P(n) > e^{\sqrt{(\log x \log \log x)}}$. Among these integers $\leq x$, the number of those with largest prime $P(n) = p$ repeating, is trivially $O(x/p^2)$. Thus

$$\begin{aligned} N(x) &<< \Psi(x, e^{\sqrt{(\log x \log \log x)}}) + \sum_{p > \exp \sqrt{(\log x \log \log x)}} \frac{x}{p^2} \\ &<< \frac{x}{e^{(\frac{1}{2} + o(1)) \sqrt{(\log x \log \log x)}}} + \sum_{n > \exp \sqrt{(\log x \log \log x)}} \frac{x}{n^2} << \frac{x}{e^{(\frac{1}{2} + o(1)) \sqrt{(\log x \log \log x)}}}, \end{aligned}$$

which proves Theorem 5. □

Here we shall use the definition for $P_2(n)$ as the largest prime factor strictly less than the largest prime factor. From Theorem 5 we see that we can focus on those integers for which $P(n)$ occurs square-free.

Consider the counting function

$$\Psi_2(x, y) = \sum_{n \leq x, P_2(n) \leq y} 1. \tag{4.5}$$

In contrast to $\Psi(x, y)$ which is very small in comparison with x when α is large, the function $\Psi_2(x, y)$ is not that small. To realize this, observe that all integers of the form $2p \leq x$ where p is prime, will have $P_2(n) = 2$, and so

$$\Psi_2(x, y) \geq \Psi_2(x, 2) >> \frac{x}{\log x}, \quad \text{for all } y \geq 2. \tag{4.6}$$

What we need here is a quantitative version of the fact that for ‘‘almost all’’ integers, $P_2(n)$ is large. This (and much more) is provided by a result of Tenenbaum [], on the size of $P_k(n)$, when $k \geq 2$. We state Tenenbaum’s result for $k = 2$ (his eqns (1.5) and (1.6) in []) in the form of:

Theorem 6: (Tenenbaum) *There exists a function $\rho_2(\alpha)$ such that*

$$\Psi_2(x, y) = x \rho_2(\alpha) (1 + O(\frac{1}{\log y})), \quad \text{uniformly for } 2 \leq y \leq x. \tag{4.7}$$

The function $\rho_2(\alpha)$ satisfies

$$\frac{1}{\alpha} << \rho_2(\alpha) << \frac{1}{\alpha}. \tag{4.8}$$

Tenenbaum’s proof of his stronger quantitative result on the joint distribution of the $P_k(n)$ for $k \geq 2$, is quite intricate, and makes use of sharp estimates for $\Psi(x, y)$. He has a precise formula for $\rho_2(\alpha)$ which we do not need here. For our purpose, all we need is:

Theorem 6*: *Uniformly for $2 \leq T \leq x$, we have*

$$\Psi_2(x, T) << \frac{x \log T}{\log x}.$$

which follows from Theorem 6. But we point out that the bound in Theorem 6* for $T \leq \exp\{(\log x)^{1-\delta}\}$, for any $\delta > 0$, can be proved using the bounds in (4.2a) and (4.2b); the implicit constant will depend on δ . Theorem 6* will be used in what follows.

5 The uniform distribution of $P_2(n)$ modulo k

In this section, we shall prove:

Theorem 7 *For each integer $k \geq 2$, the sequence $P_2(n)$ of the second largest prime factors, is uniformly distributed in the reduced residue classes modulo k . More precisely, for each fixed $k \geq 2$, and any $1 \leq \ell < k$ with $(\ell, k) = 1$, we have*

$$N_2(x, k, \ell) := \sum_{n \leq x, P_2(n) \equiv \ell \pmod{k}} 1 = \frac{x}{\phi(k)} + O\left(\frac{x(\log \log x)^2}{\log x}\right). \quad (5.1)$$

Remark: Note that the number of positive integers up to x with $\omega(n) = 1$ or $\Omega(n) = 1$ is

$$\pi(x) + O(\sqrt{x}) = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right) = o(x) \quad (5.2)$$

Thus it does not matter in Theorem 7 whether the sum in (5.1) is taken over all integers $n \leq x$ for which $P_2(n) \equiv \ell \pmod{k}$, or restricted to integers for which $\omega(n) \geq 2$.

Proof: For a prime p , denote by $S_2(x, p)$ the set of integers $n \leq x$ for which $P_2(n) = p$. Then

$$\sum_{p \leq \sqrt{x}} |S_2(x, p)| = x - \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right). \quad (5.3)$$

Now if $N \in S_2(x, p)$, then we may write

$$N = mpq, \quad \text{where } q \geq p \text{ and } P(m) \leq p,$$

with q being prime. In particular $m \leq (x/p^2)$. Thus

$$|S_2(x, p)| = \sum_{m \leq (x/p^2), P(m) \leq p} \sum_{q > p, mpq \leq x} 1 = \sum_{p < q \leq x/p} \sum_{m \leq x/(pq), P(m) \leq p} 1. \quad (5.4)$$

Thus by summing the expression in (5.4) over $p \leq \sqrt{x}$, we get

$$\sum_{p \leq \sqrt{x}} |S_2(x, p)| = \sum_{p \leq \sqrt{x}} \sum_{p \leq q \leq x/p} \Psi\left(\frac{x}{pq}, p\right) = \sum_{q \leq (x/2)} \sum^* \Psi\left(\frac{x}{pq}, p\right), \quad (5.5)$$

where the $*$ over the inner sum on the right means that the conditions on p and q are

$$p \leq \sqrt{x}, \quad p \leq q, \quad \text{and} \quad p \leq \frac{x}{q}. \quad (5.6)$$

To deal efficiently with the double sum on the right in (5.5), we consider two cases, namely $q \leq \sqrt{x}$ and $q > \sqrt{x}$, together with the inequalities in (5.6). This gives

$$\sum_{p \leq \sqrt{x}} |S_2(x, p)| = \sum_{q \leq \sqrt{x}} \sum_{p \leq q} \Psi\left(\frac{x}{pq}, p\right) + \sum_{\sqrt{x} < q \leq (x/2)} \sum_{p \leq (x/q)} \Psi\left(\frac{x}{pq}, p\right). \quad (5.7)$$

At this point we note that we have already established an effective version of the statement that $P_2(n)$ is “almost always” large, namely, Theorem 6*. In view of Theorem 6*, we may consider in (5.7) only the integers $n \leq x$ for which $P_2(n) > y$, with y to be chosen later to satisfy $\log y = o(\log x)$. Thus we modify (5.7) to

$$\begin{aligned} \sum &:= \sum_{y < p \leq \sqrt{x}} |S_2(x, p)| = \sum_{y < q \leq \sqrt{x}} \sum_{y < p < q} \Psi\left(\frac{x}{pq}, p\right) + \sum_{\sqrt{x} < q \leq (x/y)} \sum_{y < p \leq (x/q)} \Psi\left(\frac{x}{pq}, p\right) \\ &= x + O\left(\frac{x \log y}{\log x}\right). \end{aligned} \tag{5.8}$$

Next we shall compare \sum with

$$I := \sum_{y < q \leq \sqrt{x}} \int_y^q \Psi\left(\frac{x}{tq}, t\right) \frac{dt}{\log t} + \sum_{\sqrt{x} < q \leq (x/y)} \int_y^{x/q} \Psi\left(\frac{x}{tq}, t\right) \frac{dt}{\log t} \tag{5.9}$$

and estimate the difference (error) $E = \sum - I$ by using the strong form of the Prime Number Theorem. We first consider the absolute value of the difference

$$\begin{aligned} E_1 &:= \left| \sum_{y < q \leq \sqrt{x}} \left\{ \sum_{y < p < q} \Psi\left(\frac{x}{pq}, p\right) - \int_y^q \Psi\left(\frac{x}{tq}, t\right) \frac{dt}{\log t} \right\} \right| \\ &= \left| \sum_{y < q \leq \sqrt{x}} \left\{ \sum_{y < p < q} \sum_{n \leq (x/pq), P(n) \leq p} 1 - \int_y^q \left(\sum_{n \leq (x/tq), P(n) \leq t} 1 \right) \frac{dt}{\log t} \right\} \right| \\ &\leq \sum_{y < q \leq \sqrt{x}} \sum_{n \leq (x/yq)} \left| \sum_{\max(P(n), y) \leq p \leq \min(x/nq, q)} 1 - \int_{\max(P(n), y)}^{\min(x/nq, q)} \frac{dt}{\log t} \right|. \end{aligned} \tag{5.10}$$

It is to be noted that for the final term in (5.10), we have dropped the condition on $P(n)$ in the sum over n on the right for simplicity since this will not lessen the effectiveness of the upper bound on E_1 that we will get. We now use the strong form of the Prime Number Theorem on the expression on the right in (5.10) to deduce that

$$\begin{aligned} E_1 &\ll \sum_{y < q \leq \sqrt{x}} \sum_{n \leq (x/yq)} \frac{x}{nq \exp\{\sqrt{\log(x/nq)}\}} \leq \frac{x}{\exp(\sqrt{\log y})} \sum_{y < q \leq \sqrt{x}} \frac{1}{q} \sum_{n \leq (x/yq)} \frac{1}{n} \\ &\ll \frac{x \log x}{\exp(\sqrt{\log y})} \sum_{y < q \leq \sqrt{x}} \frac{1}{q} \ll \frac{x \log x \log \log x}{\exp(\sqrt{\log y})}. \end{aligned} \tag{5.11}$$

In obtaining this upper bound for E_1 , we have used the fact that $z/\exp\sqrt{\log z}$ is an increasing function of z , and so in deriving (5.11) we only used the error term in the strong form of the Prime Number Theorem with $z = x/nq$.

Similarly, we bound the difference

$$\begin{aligned}
E_2 &:= \left| \sum_{\sqrt{x} < q \leq (x/y)} \left\{ \sum_{y < p < (x/q)} \Psi\left(\frac{x}{pq}, p\right) - \int_y^{x/q} \Psi\left(\frac{x}{tq}, t\right) \frac{dt}{\log t} \right\} \right| \\
&= \left| \sum_{\sqrt{x} < q \leq (x/y)} \left\{ \sum_{y < p < (x/q)} \sum_{n \leq (x/pq), P(n) \leq p} 1 - \int_y^{x/q} \left(\sum_{n \leq (x/tq), P(n) \leq t} 1 \right) \frac{dt}{\log t} \right\} \right| \\
&\leq \sum_{\sqrt{x} < q \leq (x/y)} \sum_{n \leq (x/yq)} \left| \sum_{\max(P(n), y) \leq p \leq (x/nq)} 1 - \int_{\max(P(n), y)}^{x/nq} \frac{dt}{\log t} \right| \\
&\ll \sum_{\sqrt{x} < q \leq (x/y)} \sum_{n \leq (x/yq)} \frac{x}{nq \exp\{\sqrt{\log(x/nq)}\}} \leq \frac{x}{\exp(\sqrt{\log y})} \sum_{\sqrt{x} < q \leq (x/y)} \frac{1}{q} \sum_{n \leq (x/yq)} \frac{1}{n} \\
&\ll \frac{x \log x}{\exp(\sqrt{\log y})} \sum_{\sqrt{x} < q \leq (x/y)} \frac{1}{q} \ll \frac{x \log x \log \log x}{\exp(\sqrt{\log y})}. \tag{5.12}
\end{aligned}$$

So from (5.11) and (5.12), we see that

$$|E| = \left| \sum -I \right| \leq E_1 + E_2 \ll \frac{x \log x \log \log x}{\exp(\sqrt{\log y})}. \tag{5.13}$$

From (5.13) and (5.8) we deduce that

$$I = x + O\left(\frac{x \log y}{\log x}\right) + O\left(\frac{x \log x \log \log x}{\exp(\sqrt{\log y})}\right). \tag{5.14}$$

At this stage, we make the choice

$$y = \exp\{(2 \log \log x)^2\}, \tag{5.15}$$

to conclude that

$$I = x + O\left(\frac{x(\log \log x)^2}{\log x}\right) \quad \text{and} \quad \sum = x + O\left(\frac{x(\log \log x)^2}{\log x}\right). \tag{5.16}$$

This will be crucial in establishing Theorem 7.

Now for an arbitrary but fixed integer $k \geq 2$, and for any $1 \leq \ell < k$ with $(\ell, k) = 1$, we consider the set $S_2^{k, \ell}(x)$ of integers $n \leq x$ such that, $\omega(n) \geq 2$, and $P_2(n) \equiv \ell \pmod{k}$. By classifying the members of this set in terms of their second largest prime factor, we see that

$$N_2(x, k, \ell) = |S_2^{k, \ell}(x)| = \sum_{p \leq \sqrt{x}, p \equiv \ell \pmod{k}} |S_2(x, p)|. \tag{5.17}$$

In view of Theorem 6*, we have for any $y \leq \exp(\log x)^{1-\delta}$

$$\sum_{p \leq y, p \equiv \ell \pmod{k}} |S_2(x, p)| \leq \sum_{p \leq y} |S_2(x, p)| \ll \frac{x \log y}{\log x}. \tag{5.18}$$

Thus

$$N_2(x, k, \ell) = |S_2^{k,\ell}(x)| = \sum_{y < p \leq \sqrt{x}, p \equiv \ell \pmod{k}} |S_2(x, p)| + O\left(\frac{x \log y}{\log x}\right) \quad (5.19)$$

Let us denote the sum on the right of (5.19) as $\sum^{k,\ell}$. Then by reasoning as above, we get

$$\sum^{k,\ell} = \sum_{y < q \leq \sqrt{x}} \sum_{y < p < q, p \equiv \ell \pmod{k}} \Psi\left(\frac{x}{pq}, p\right) + \sum_{\sqrt{x} < q \leq (x/y)} \sum_{y < p \leq (x/q), p \equiv \ell \pmod{k}} \Psi\left(\frac{x}{pq}, p\right). \quad (5.20)$$

We now want to compare the expression in (5.20) with

$$I^{k,\ell} := \sum_{y < q \leq \sqrt{x}} \frac{1}{\phi(k)} \int_y^q \Psi\left(\frac{x}{tq}, t\right) \frac{dt}{\log t} + \sum_{\sqrt{x} < q \leq (x/y)} \frac{1}{\phi(k)} \int_y^{x/q} \Psi\left(\frac{x}{tq}, t\right) \frac{dt}{\log t} \quad (5.21)$$

Clearly

$$I^{k,\ell} = \frac{I}{\phi(k)}, \quad (5.22)$$

with I as in (5.9). Just as we obtained the bound in (5.13) for the difference $\sum -I$ using the strong form of the Prime Number Theorem, we can use the same reasoning together with the strong form of the Prime Number Theorem for Arithmetic Progressions to deduce that

$$\sum^{k,\ell} - I^{k,\ell} \ll \frac{x \log x \log \log x}{\exp(\sqrt{\log y})}. \quad (5.23)$$

So from the above estimates, we deduce that

$$N_2(x, k, \ell) = |S_2^{k,\ell}(x)| = \frac{x}{\phi(k)} + O\left(\frac{x \log y}{\log x}\right) + O\left(\frac{x \log x \log \log x}{\exp(\sqrt{\log y})}\right). \quad (5.24)$$

Once again, we make the choice

$$y = \exp\{(2 \log \log x)^2\}, \quad (5.25)$$

to deduce Theorem 7 from (5.24). □

6 Proof of the Main Result

Theorem 7 paves the way to the proof of our main result (Theorem 10 below). Enroute to Theorem 10, we establish two theorems, the first of which relies on Theorem 7:

Theorem 8: *For integers ℓ, k satisfying $1 \leq \ell \leq k$ with $(\ell, k) = 1$, we have*

$$\sum_{n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n) \omega(n) \ll \frac{x (\log \log x)^4}{\log x}.$$

Proof: Let f be a function on the primes defined by

$$f(p) = 1 \quad \text{if } p \equiv \ell \pmod{k}, \quad f(p) = 0, \quad \text{otherwise.} \quad (6.1)$$

Then by taking $k = 2$ in (1.12), and by Moebius inversion, we get

$$\begin{aligned} \sum_{1 < n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n)(\omega(n) - 1) &= \sum_{1 < n \leq x} \mu(n)(\omega(n) - 1)f(p(n)) \\ &= \sum_{1 < n \leq x} \sum_{d|n} \mu\left(\frac{n}{d}\right) f(P_2(d)), \end{aligned}$$

which we rewrite as

$$\begin{aligned} \sum_{1 < n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n)\omega(n) &= \sum_{1 < n \leq x} \sum_{d|n} \mu\left(\frac{n}{d}\right) f(P_2(d)) + \sum_{1 < n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n) \\ &:= \Sigma_5 + \Sigma_6 \quad \text{respectively} \end{aligned} \quad (6.2)$$

It was already established in [2] that

$$\Sigma_6 \ll x \exp\{-(\log x)^{1/3}\}. \quad (6.3)$$

With regard to Σ_5 , we employ the hyperbola method and write it as

$$\begin{aligned} \Sigma_5 &= \sum_{m \leq T} \mu(m) \sum_{d \leq (x/m)} f(P_2(d)) + \sum_{d \leq (x/T)} f(P_2(d)) \sum_{T \leq m \leq (x/d)} \mu(m) \\ &:= \Sigma_7 + \Sigma_8, \quad \text{respectively.} \end{aligned} \quad (6.4)$$

Clearly from the strong form of the Prime Number Theorem, we get

$$\Sigma_8 \ll \sum_{d \leq (x/T)} f(P_2(d)) \frac{x}{d \exp \sqrt{\log(x/d)}} \ll \frac{x \log x}{\exp \sqrt{\log T}}. \quad (6.5)$$

Regarding Σ_7 , Theorem 7 gives

$$\Sigma_7 = \sum_{m \leq T} \mu(m) \left(\frac{x}{\phi(k)m} \right) + O\left(\frac{x(\log \log x)^2}{m \log(x/m)} \right)$$

which by the strong form of the Prime Number Theorem is

$$\ll \frac{x}{\phi(k) \exp \sqrt{(\log T)}} + \frac{x \log T (\log \log x)^2}{\log(x/T)}. \quad (6.6)$$

At this point we choose

$$T = \exp\{4(\log \log x)^2\} \quad \Leftrightarrow \quad \sqrt{\log T} = 2 \log \log x. \quad (6.7)$$

With this choice of T , we deduce from (6.2) - (6.6) that

$$\sum_{n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n)\omega(n) \ll \frac{x(\log \log x)^4}{\log x} \quad (6.8)$$

which proves Theorem 8. □

We next prove:

Theorem 9: *Let ℓ, k be integers satisfying $1 \leq \ell \leq k$ with $(\ell, k) = 1$. Then*

$$\sum_{1 < n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} \ll \frac{x (\log \log x)^{5/2}}{\sqrt{\log x}},$$

where $\{y\}$ denotes the fractional part of y .

Proof: Begin with the decomposition

$$\begin{aligned} & \sum_{1 < n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} \\ = & \sum_{1 < n \leq T, p(n) \equiv \ell \pmod{k}} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} + \sum_{T < n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} \\ & := \Sigma_9 + \Sigma_{10} \quad \text{respectively,} \end{aligned} \tag{6.9}$$

where T will be chosen optimally below.

Clearly

$$\Sigma_9 \ll T \log \log T. \tag{6.10}$$

To estimate Σ_{10} , put

$$M_\omega(x, \ell, k) = \sum_{1 < n \leq x, p(n) \equiv \ell \pmod{k}} \mu(n) \omega(n). \tag{6.11}$$

So we have

$$\begin{aligned} \Sigma_{10} &= \sum_{T \leq n \leq x} (M_\omega(n, \ell, k) - M_\omega(n-1, \ell, k)) \left\{ \frac{x}{n} \right\} \\ &= \sum_{T \leq n \leq x} M_\omega(n, \ell, k) \left(\left\{ \frac{x}{n} \right\} - \left\{ \frac{x}{n+1} \right\} \right) \\ &\ll \sum_{T \leq n \leq x} |M_\omega(n, \ell, k)| \left| \left\{ \frac{x}{n} \right\} - \left\{ \frac{x}{n+1} \right\} \right| \ll \frac{x (\log \log x)^4}{\log x} V_\Omega \left[1, \frac{x}{T} \right] \\ &\ll \frac{x (\log \log x)^4}{\log x} \frac{x}{T}, \end{aligned} \tag{6.12}$$

using Theorem 8, where in (6.11), as in the proof of Theorem 2, $V_\Omega[a, b]$ denotes the total variation of $\{y\}$ in the interval $[a, b]$.

On comparing the bounds in (6.9) and (6.11), we see that the optimal choice of T is given by setting

$$T \log \log T \sim \frac{x (\log \log x)^4}{\log x} \frac{x}{T}$$

and this leads to the choice

$$T = \frac{x(\log \log x)^{3/2}}{\sqrt{\log x}}. \quad (6.13)$$

With this value of T , we get the upper bound in Theorem 9. \square

We are now in a position to prove our main result:

Theorem 10: *For integers ℓ, k satisfying $1 \leq \ell \leq k$ and $(\ell, k) = 1$, we have*

$$m_\omega(x; \ell, k) := \sum_{n \leq x, p(n) \equiv \ell \pmod{k}} \frac{\mu(n)\omega(n)}{n} \ll \frac{(\log \log x)^{5/2}}{\sqrt{\log x}}.$$

Letting $x \rightarrow \infty$, we get

$$\sum_{n \geq 1, p(n) \equiv \ell \pmod{k}} \frac{\mu(n)\omega(n)}{n} = \sum_{n \geq 2, p(n) \equiv \ell \pmod{k}} \frac{\mu(n)\omega(n)}{n} = 0$$

Proof: With $f(p)$ defined on the primes as above, note that

$$\begin{aligned} \sum_{1 < d \leq x} \mu(d)(\omega(d) - 1)f(p(d))\left[\frac{x}{d}\right] &= \sum_{n \leq x} \sum_{1 < d|n} \mu(d)(\omega(d) - 1)f(p(d)) \\ &= \sum_{n \leq x} f(P_2(n)) = \frac{x}{\phi(k)} + O\left(\frac{x(\log \log x)^2}{\log x}\right) \end{aligned} \quad (6.14)$$

by (1.12) and Theorem 7.

It was already established in [2] that

$$\begin{aligned} \sum_{1 < n \leq x} \mu(d)f(p(d))\left[\frac{x}{d}\right] &= \sum_{n \leq x} \sum_{1 < d|n} \mu(d)f(p(d)) \\ &= -\sum_{n \leq x} f(P(n)) = \frac{-x}{\phi(k)} + O\left(\frac{x}{\exp\{(\log x)^{1/3}\}}\right). \end{aligned} \quad (6.15)$$

On comparing (6.13) and (6.14), we see that the main term $x/\phi(k)$ cancels, and this leads to

$$\sum_{1 < d \leq x} \mu(d)\omega(d)f(p(d))\left[\frac{x}{d}\right] = O\left(\frac{x(\log \log x)^2}{\log x}\right). \quad (6.16)$$

But we know by Theorem 9 that

$$\sum_{1 < n \leq x} \mu(d)\omega(d)f(p(d))\left\{\frac{x}{d}\right\} = O\left(\frac{x(\log \log x)^{5/2}}{\sqrt{\log x}}\right). \quad (6.17)$$

Finally by adding the expressions in (6.15) and (6.16), we get

$$x \sum_{1 < n \leq x} \frac{\mu(d)\omega(d)f(p(d))}{d} = O\left(\frac{x(\log \log x)^{5/2}}{\sqrt{\log x}}\right). \quad (6.18)$$

On dividing both sides of (6.17) by x , we get Theorem 10. □

Remarks: The estimates when $k = 1$

It is to be noted that when $k = 1$, the method of this section would yield quantitative estimates for $M_\omega(x)$, and $m_\omega(x)$, but these would be weaker than what we got in Theorems 1 and 4 in Section 2. First observe that when $k = 1$, the uniform distribution of $P_2(n)(\text{mod } 1)$ is trivial, but then the error term in Theorem 7 for $k = 1$ would be sharper. More precisely,

$$N_2(x, 1, 1) = x + O\left(\frac{x}{\log x}\right). \tag{6.19}$$

If we work through the proof of Theorem 8 for $k = 1$ with f being the characteristic function of the primes, and use (6.18), the following estimates would hold:

The bound in (6.3) would be

$$\Sigma_6 \ll x \exp\{-c\sqrt{\log x}\}. \tag{6.20}$$

The bound for Σ_8 in (6.5) would not change, but the bound for Σ_7 in (6.6) would be

$$\Sigma_7 \ll \frac{x}{\exp\sqrt{\log T}} + \frac{x \log T}{\log(x/T)}. \tag{6.21}$$

With T chosen as in (6.7), the final estimate that we would get is

$$M_\omega(x) \ll \frac{x(\log \log x)^2}{\log x}, \tag{6.22}$$

which is sharper than Theorem 8 when $k = 1$, but weaker than Theorem 1 which was proved by a different method.

If we now follow the proof of Theorem 9, and use (6.21) instead of Theorem 8 for $k = 1$, the optimal choice of T would be

$$T = \frac{x \sqrt{\log \log x}}{\sqrt{\log x}} \tag{6.23}$$

in place of T in (6.12). This would then yield

$$\sum_{2 \leq n \leq x} \mu(n) \omega(n) \left\{ \frac{x}{n} \right\} \ll \frac{x(\log \log x)^{3/2}}{\sqrt{\log x}}, \tag{6.24}$$

which is stronger than Theorem 9 for $k = 1$, but weaker than Theorem 2.

Finally, if we use (6.23) in the proof of Theorem 10 for $k = 1$, we would get

$$m_\omega(x) \ll \frac{(\log \log x)^{3/2}}{\sqrt{\log x}}, \tag{6.25}$$

which is sharper than Theorem 10 for $k = 1$, but weaker than Theorem 4.

7 Sums involving the exceptional primes

In §2 we proved (Theorem 4) that

$$\sum_{n=2}^{\infty} \frac{\mu(n)\omega(n)}{n} = 0, \quad (7.1)$$

by establishing some preliminary results. Then using the methods of §2, and by establishing several results, we proved in §6 that if $k \geq 2$ is an arbitrary modulus, then for every ℓ that satisfies $(\ell, k) = 1$

$$\sum_{n=2, p(n) \equiv \ell \pmod{k}}^{\infty} \frac{\mu(n)\omega(n)}{n} = 0. \quad (7.2)$$

When we sum the expression on the left in (7.2) over all $1 \leq \ell < k$ with $(\ell, k) = 1$, we do not get the full sum in (7.1) because the primes

$$p \equiv \ell \pmod{k} \quad \text{with} \quad (\ell, k) > 1, \quad (7.3)$$

have not been accounted for. But there will be primes satisfying the conditions in (7.3), which we call *exceptional primes*, if and only if ℓ is a prime divisor of k , and in this case there is just a single prime p in the residue class $\ell \pmod{k}$, namely $p = \ell$. It turns out that the sum in (7.2) is 0 when taken over n satisfying $p(n) = p$ for any fixed prime regardless of whether p divides k or not. That is we have:

Theorem 11: *Let p be an arbitrary but fixed prime. Then*

$$\sum_{n=1, p(n)=p}^{\infty} \frac{\mu(n)\omega(n)}{n} = \sum_{n=2, p(n)=p}^{\infty} \frac{\mu(n)\omega(n)}{n} = 0 \quad (7.4)$$

Proof: The square-free integers n with $p(n) = p$ are those of the form

$$n = mp, \quad \text{with} \quad (m, N_p) = 1, \quad \text{where} \quad N_p = \prod_{q \leq p, q = \text{prime}} q. \quad (7.5)$$

Thus using $\omega(mp) = \omega(m) + 1$, we get

$$\begin{aligned} \sum_{n=2, p(n)=p}^{\infty} \frac{\mu(n)\omega(n)}{n} &= \frac{-1}{p} \sum_{(m, N_p)=1} \frac{\mu(m)\omega(mp)}{m} \\ &= -\frac{1}{p} \sum_{(m, N_p)=1} \frac{\mu(m)}{m} - \frac{1}{p} \sum_{(m, N_p)=1} \frac{\mu(m)\omega(m)}{m} \\ &\quad \Sigma_{11} + \Sigma_{12}. \end{aligned} \quad (7.6)$$

It is a classical result the $\Sigma_{11} = 0$. The methods of §2 can be used to show that $\Sigma_{12} = 0$. Thus Theorem 11 follows from (7.5).

Remarks:

- (i) Since the exceptional primes, namely those that divide the modulus k , are finite in number, the sum of the expression in Theorem 11 taken over all exceptional primes is 0 since it is a sum of a finite number of zeros. Thus by Theorem 11, the exceptional primes are accounted for in the full sum in (7.1). When $(\ell, k) = 1$, there are infinitely many primes $p \equiv \ell \pmod{k}$, and for each p in the residue class $\ell \pmod{k}$, the sum as in Theorem 11 is 0. What makes Theorem 10 interesting is that we are summing “infinitely many zeros”, yet the sum is 0.
- (ii) As was the case with our earlier theorems, a quantitative version of Theorem 11 can be established.

8 The case of general f

In the penultimate section of [2], it was shown that if f is ANY bounded function on the primes, then

$$M_f(x) := \sum_{2 \leq n \leq x} \mu(n) f(p(n)) = o(x). \tag{8.1}$$

From (8.1), it follows by Axer’s theorem that

$$\sum_{2 \leq n \leq x} \mu(n) f(p(n)) \left\{ \frac{x}{n} \right\} = o(x), \tag{8.2}$$

where $\{t\}$ denotes the fractional part of t . Next, by the Duality identity (1.3) we have

$$\sum_{2 \leq n \leq x} \mu(n) f(p(n)) \left[\frac{x}{n} \right] = - \sum_{2 \leq n \leq x} f(P(n)). \tag{8.3}$$

Hence by adding the expressions in (8.2) and (8.3), we get

$$x \sum_{2 \leq n \leq x} \frac{\mu(n) f(p(n))}{n} = - \sum_{2 \leq n \leq x} f(P(n)) + o(x). \tag{8.4}$$

From (8.4), the equivalence of (1.5) and (1.6) follows, and this was how this equivalence was proved in [2].

In [2], the following simple bound

$$M_f(x) \ll \frac{x}{\log \log \log x} \tag{8.5}$$

but was established, but subsequently in [3] it was refined to

$$\max_{|f| \leq 1} |M_f(x)| \sim \frac{2x}{\log x}. \tag{8.6}$$

Of course, for specific functions f , such as f being the characteristic function of primes in an arithmetic progression $\ell \pmod{k}$, where $(\ell, k) = 1$, the bound for $M_f(x)$ is vastly superior (see [2]).

Similar in spirit to (8.1), it can be shown that:

Theorem 12: *If f is any bounded function on the primes, then*

$$M_{f,\omega}(x) := \sum_{n \leq x} \mu(n) \omega(n) f(p(n)) = o(x).$$

A proof of a quantitative version of Theorem 12 will be given in a subsequent paper []. From Theorem 12, by Axer's Theorem arguments, it will follow that

$$\sum_{n \leq x} \mu(n)\omega(n)f(p(n))\left\{\frac{x}{n}\right\} = o(x). \quad (8.7)$$

While all this seems to be similar to (8.1) and (8.2), an important difference occurs here. In order to apply the Duality identity (1.12) when $k = 2$, we have to consider the sum

$$\sum_{2 \leq n \leq x} \mu(n)(\omega(n) - 1)f(p(n))\left[\frac{x}{n}\right] = \sum_{2 \leq n \leq x} \sum_{1 < d|n} \mu(d)(\omega(d) - 1)f(p)d = \sum_{2 \leq n \leq x} f(P_2(n)). \quad (8.8)$$

We note that the first sum on the left hand side of (8.8) is

$$\begin{aligned} & \sum_{2 \leq n \leq x} \mu(n)\omega(n)f(p(n))\left[\frac{x}{n}\right] - \sum_{2 \leq n \leq x} \mu(n)f(p(n))\left[\frac{x}{n}\right] \\ &= \sum_{2 \leq n \leq x} \mu(n)\omega(n)f(p(n))\left[\frac{x}{n}\right] + \sum_{2 \leq n \leq x} f(P(n)) \end{aligned} \quad (8.9)$$

in view of (8.3). So from (8.7), (8.8), and (8.9), we get

$$x \sum_{2 \leq n \leq x} \frac{\mu(n)\omega(n)f(p(n))}{n} = \sum_{2 \leq n \leq x} f(P_2(n)) - \sum_{2 \leq n \leq x} f(P(n)). \quad (8.10)$$

Now (8.10) yields the following result:

Theorem 13: *If f is a bounded function on the primes such that*

$$\sum_{2 \leq n \leq x} f(P(n)) \sim \kappa x \quad (8.11)$$

and

$$\sum_{2 \leq n \leq x} f(P_2(n)) \sim \kappa x, \quad (8.12)$$

for some constant κ , then

$$\sum_{n=2}^{\infty} \frac{\mu(n)\omega(n)f(p(n))}{n} = 0. \quad (8.13)$$

Remarks: When $f(p)$ defined on primes p is the characteristic function of primes in the residue class $\ell \pmod k$, where $(\ell, k) = 1$, then $\kappa = 1/\phi(k)$ in (8.11) and (8.12), in which case (8.13) is Theorem 10. If instead of the same constant κ in (8.12) and (8.13), we had two different constants, κ_1 in (8.11) and κ_2 in (8.12), then the sum in (8.13) will converge to $\kappa_2 - \kappa_1$. But we wish to stress that we know of no natural example of a bounded function on the primes for which the constants κ_1 and κ_2 have different values. Thus we pose:

PROBLEM: Does there exist a bounded function f on the primes such that

$$\sum_{2 \leq n \leq x} f(P(n)) \sim \kappa_1 x, \quad \text{and} \quad \sum_{2 \leq n \leq x} f(P_2(n)) \sim \kappa_2 x, \quad \text{with} \quad \kappa_1 \neq \kappa_2.$$

Consider now the following situation: Given x arbitrarily large, define a function f on the primes as follows:

$$f(p) = 1 \text{ if } \sqrt{x} < p \leq x, \quad f(p) = 0 \text{ if } p \leq \sqrt{x}. \quad (8.14)$$

With f as in (8.14), we have

$$\begin{aligned} \sum_{2 \leq n \leq x} f(P(n)) &= \sum_{\sqrt{x} < p \leq x} \sum_{n \leq x, P(n)=p} 1 \\ &= \sum_{\sqrt{x} < p \leq x} \left[\frac{x}{p} \right] = x \log 2 + O\left(\frac{x}{\log x}\right). \end{aligned} \quad (8.15)$$

On the other hand, since $P_2(n) \leq \sqrt{x}$ if $n \leq x$, we clearly have

$$\sum_{2 \leq n \leq x} f(P_2(n)) = 0.$$

So in this example, $\kappa_1 = \log 2$, and $\kappa_2 = 0$. But note that the definition of f in (8.14) depends on x , whereas in Problem 1 we ask for a function f just defined on the primes (without dependency on x).

The importance of the consideration of general functions f in this section will be clear in the next section when we will discuss algebraic analogues to the results of Alladi [2] by various authors, and algebraic analogues of the results in this paper by Sengupta [15].

9 Algebraic and q -analogues, and higher order duality

The Duality identity (1.3), and the result (1.8) established in quantitative form in Alladi [2], have attracted a lot of attention in the last decade. It all started with the paper [5] of Dawsey who obtained the following algebraic analogue and extension of (1.8) to Galois extensions of the field of rationals \mathbb{Q} :

Let K , be a Galois extension of \mathbb{Q} , and \mathbb{O}_K the ring of integers in K . If p is a prime in the integers, then let P denote the prime ideal that is contained in \mathbb{O}_K which lies above p . If p is unramified, let $\left[\frac{K/\mathbb{Q}}{P} \right]$ denote the Artin symbol. For simplicity, let

$$\left[\frac{K/\mathbb{Q}}{p} \right] := \left[\frac{K/\mathbb{Q}}{P} \right].$$

Then:

Theorem D (Dawsey): *Let K be a finite Galois extension of \mathbb{Q} with Galois group $G = \text{Gal}(K/\mathbb{Q})$. Let C be a conjugacy class in G . Then*

$$- \sum_{n \geq 2, \left[\frac{K/\mathbb{Q}}{p(n)} \right] = C} \frac{\mu(n)}{n} = \frac{|C|}{|G|}.$$

Dawsey notes that Theorem D is a generalization of (1.8), because in the special case when K is a cyclotomic extension of \mathbb{Q} , the group $Gal(K/\mathbb{Q})$ can be identified with \mathbb{Z}_k^* , the set of reduced residues modulo k for some positive integer k ; since \mathbb{Z}_k^* is Abelian, each conjugacy class has just one element and so

$$\frac{|C|}{|G|} = \frac{1}{\phi(k)}.$$

The way Dawsey proves Theorem D is to use the Chebotarev Density Theorem to show that

$$\sum_{2 \leq n \leq x, [\frac{K/\mathbb{Q}}{P(n)}]=C} 1 \sim \frac{|C|}{|G|}. \quad (9.1)$$

With (9.1) established, then by the methods in [2] that involve Duality, Dawsey is able to get Theorem D.

Motivated by Dawsey's work, Sweeting and Woo [16] obtained a generalization of Theorem D in which the finite extensions K of \mathbb{Q} are replaced by finite extensions L of an arbitrary algebraic number field K . In discussing this more general situation, Sweeting and Woo consider a generalization of the Moebius function defined in terms of products of prime ideals instead of product of primes, and establish a duality identity that generalizes (1.3) appropriately. In this more general situation, the Chebotarev Density Theorem applies, and so an analogue of Theorem D is established in [5].

While it is true that Theorem D generalizes (1.8), it is to be noted that the more general equivalence of (1.5) and (1.6) is established as Theorem 6 in [2]. So what Dawsey confirmed is that if f is chosen to be the characteristic function of primes p for which the Artin symbol $[\frac{K/\mathbb{Q}}{P}] = C$, then the average of $f(P(n))$ exists. That is, in this case c in (1.5) is $|C|/|G|$. So the deduction of Theorem D from (9.1) is a special case of the equivalence of (1.5) and (1.6). Since, the equivalence of (1.5) and (1.6) is established in [2] for arbitrary bounded functions f , the bounds for the quantitative version of (1.6) is weak. For the Chebotarev Density Theorem, Lagarias and Odlyzko [11] have established a strong form, with the error term comparable to the error term in the strong form of the Prime Number Theorem. Thus utilizing the Lagarias-Odlyzko theorem, Dawsey is able to get a superior quantitative version of Theorem D where the bound is just as sharp as the quantitative version of (1.8) that is proved in [2] using the strong form of the Prime Number Theorem.

The results of Sweeting and Woo have been extended by Kural, McDonald and Sah [10]. A generalization in a different direction, namely replacing the Moebius function by the more general Ramanujan sum

$$c_m(n) = \sum_{k=1, (k,n)=1}^n e^{2imk\pi/n}, \quad (9.2)$$

is considered by Wang [21] ($\mu(n) = c_1(n)$). Also, Wang in collaboration with Duan and Yi [22] has discussed analogues of Alladi's duality in global function fields.

A fruitful way to generalize arithmetic results is to obtain suitable q -analogues. In two papers [12] and [13], Ono-Schneider-Wagner have discussed a variety of q -analogues of arithmetic density results and their partition implications.

With regard to the arithmetic consequences of the second order duality (namely consequences of (1.12) in the case $k = 2$), recently Sengupta [15], motivated by the work of Dawsey, has obtained the extension of Theorem 10 to the situation when K is a finite Galois extension of \mathbb{Q} . Like Dawsey, Sengupta uses the strong form of the Chebotarev Density Theorem due to Lagarias and Odlyzko [11].

We mention that Alladi and Sengupta [4] have very recently considered arithmetic consequences of higher order dualities, namely (1.12) for $k \geq 3$, and established analogues of all the results in this paper for $k \geq 3$. In this discussion of higher order dualities, it turns out that when $k \geq 3$, the bounds for certain terms have extra factors which are powers of $\log \alpha$, where $\alpha = \log x / \log y$; these factors are not present in the case $k = 2$ treated here.

Finally, we point out that all the quantitative results in [2] were established with uniformity for the moduli k of arithmetic progressions satisfying $k \leq \log^\beta x$, with implicit constants depending on β . This is because in [2], we utilized the Siegel-Walfisz theorem for primes in arithmetic progressions. If we had used the Siegel-Walfisz theorem here, then Theorem 10 would hold with uniformity for $k \leq \log^\beta x$.

10 Concluding Remarks

(i) *Arithmetic density versions:* The generalizations of (1.8) to algebraic number fields by various authors starting with Dawsey [] was motivated by rewriting (1.8) as

$$- \sum_{n \geq 2, p(n) \equiv \ell \pmod{k}} \frac{\mu(n)}{n} = \frac{1}{\phi(k)}, \quad (10.1)$$

and interpreting this as an arithmetic density result. Similarly, our Theorem 10 can be rewritten as

$$\sum_{n \geq 2, p(n) \equiv \ell \pmod{k}} \frac{\mu(n)(\omega(n) - 1)}{n} = \frac{1}{\phi(k)}, \quad (10.2)$$

and interpreted as an arithmetic density result, thereby lending itself to an arithmetic density generalization to algebraic number fields using the Chebotarev density theorem (see Sengupta [15]). The consequence of the general identity (1.12) for $k \geq 3$ discussed in Alladi-Sengupta [4] also has an arithmetic density formulation, namely

$$(-1)^k \sum_{n \geq 2, p(n) \equiv \ell \pmod{j}} \frac{\mu(n)}{n} \binom{\omega(n) - 1}{k - 1} = \frac{1}{\phi(j)}. \quad (10.3)$$

This can be generalized to algebraic number fields using the Chebotarev density theorem.

Tenenbaum's generalization of Theorem 10:

Very recently, Tenenbaum [20] has generalized Theorem 10 as follows:

Theorem T: Let \mathcal{P} be a set of primes satisfying

$$\varepsilon(t) = \frac{1}{t} \left\{ \sum_{p \leq t, p \in \mathcal{P}} \log p \right\} - \kappa = o(1), \quad \text{as } t \rightarrow \infty, \quad (10.4)$$

with some $\kappa \in [0, 1]$. Then

$$\sum_{n \geq 2, p(n) \in \mathcal{P}} \frac{\mu(n)\omega(n)}{n} = 0. \quad (10.5)$$

Tenenbaum’s proof of a quantitative form of (9.7) is analytic and quite intricate. But the main thing is that he is able to get (9.7) directly from (9.6) without relying on estimates like (8.11) and (8.12). But then, our approach using Duality connecting sums involving $\mu(n)\omega(n)f(p(n))$ with $f(P_1(n))$ and $f(P_2(n))$ is of intrinsic interest, and that is the motivation of the present paper.

Acknowledgements: KA would like to thank Gerald Tenenbaum for several helpful suggestions and critical comments.

11 References

- 1) Y. ALAMOUDI and K. ALLADI, “Asymptotic estimates for sums involving the Moebius function, and the number of prime factors, with conditions on the smallest prime factor” (in preparation)
- 2) K. ALLADI, “Duality between prime factors and an application to the Prime Number Theorem for Arithmetic Progressions”, *J. Num. Th.*, **9** (1977), 436-451.
- 3) K. ALLADI, “Asymptotic estimates of sums involving the Moebius function”, *J. Num. Th.*, **14** (1982), 86-98.
- 4) K. ALLADI and S. SENGUPTA, “Higher order duality between prime factors and primes in arithmetic progressions”, (in preparation).
- 5) M. L. DAWSEY, “A new formula for Chebotarev densities”, *Res. Num. Th.*, **3** (2017).
<https://doi.org/10.1007/s40993-017-0093-7>
- 6) N. G. DE BRUIJN, “On the asymptotic behavior of a function occurring in the theory of primes”, *J. Indian Math. Soc. (N.S.)*, **15** (1951), 25-32.
- 7) N. G. DE BRUIJN, “On the number of positive integers $\leq x$ and free of prime factors $> y$ ”, *Indag. Math.*, **13** (1951), 50-60.
- 8) L. DUAN, B. WANG, and S. YI, “Analogues of Alladi’s formula over global function fields”, *Finite Fields and their Applications*, **74** (2021), <https://doi.org/10.1016/j.ffa.2021.101874>
- 9) A. J. HILDEBRAND and G. TENENBAUM, “On integers free of large prime factors”, *Trans. Amer. Math. Soc.*, **296** (1986), 265-290.
- 10) M. KURAL, V. MCDONALD, and A. SAH, “Moebius formulas for densities of sets of prime ideals”, *Arch. Math.* **135** (2020), 53-60.
- 11) J. C. LAGARIAS and A. M. ODLYZKO, “Effective versions of the Chebotarev density theorem”, in *Algebraic Number Fields, L-functions, and Galois properties*, A. Fröhlich (Ed.), Acad. Press. London (1977), 409-464.

- 12) K. ONO, R. SCHNEIDER, and I. WAGNER, “Partition theoretic formulas for arithmetic densities”, in *Analytic Number Theory, Modular Forms, and q -Hypergeometric Series* (G. E. Andrews and F. Garvan, Eds.) - Conf. in honor of Krishna Alladi’s 60th birthday, Springer Proceedings in Math. and Stat., **221** (2017), 611-624.
- 13) K. ONO, R. SCHNEIDER, and I. WAGNER, “Partition-theoretic formulas for arithmetic densities - II”, *Hardy-Ramanujan J.*, **43** (2020), 1-16.
- 14) A. SELBERG, “Note on a paper by L. G. Sathe”, *J. Indian Math. Soc.*, **18** (1954), 83-87.
- 15) S. SENGUPTA, “Algebraic analogues of theorems of Alladi-Johnson relating to second order duality among prime factors”
- 16) N. SWEETING and K. WOO, “Formulas for Chebotarev densities of Galois extensions of number fields”, *Research Num. Th.* **5** (2019) - <https://doi.org/10.1007/s40993-018-0142-x>
- 17) G. TENENBAUM, “A rate estimate in Billingsley’s theorem for the size distribution of the number of prime factors”, *Quart. J. Math.*, **51** (2000), 385-403.
- 18) G. TENENBAUM, *Probabilistic and Analytic Number Theory*, Grad. Studies in Math., **163** (2015), Amer. Math. Soc, Providence RI, 630 pp.
- 19) G. TENENBAUM, “Private Communication to Alladi” (2019)
- 20) G. TENENBAUM, “On a family of arithmetic series related to the Moebius function” (2024 preprint).
- 21) B. WANG, “The Ramanujan sum and Chebotarev densities”, *Ramanujan J.*, **55** (2021), 1105-1111.
- 22) B. WANG, “Analogues of Alladi’s formula”, *J. Num. Th.* **221** (2021), 232-246.