From explicit estimates for the primes to explicit estimates for the Moebius function

O. Ramaré

March 8, 2012

Abstract

We prove an estimate slightly stronger than $|\sum_{d \leq D} \mu(d)/d| \leq 0.03/\log D$ for every $D \geq 11815$.

1 Introduction

There is a long literature concerning explicit estimates for the summatory function of the Moebius function, and we cite for instance [20], [1], [4], [3], [6], [7], [10], [11]. The paper [5] proposes a very useful annotated bibliography covering relevant items up to 1983. It has been known since the beginning of the 20th century at least (see for instance [13]) that showing that $M(x) = \sum_{n \leq x} \mu(n)$ is $o(x)$ is equivalent to showing that the Tchebychev function $\psi(x) = \sum_{n \leq x} \Lambda(n)$ is asymptotic to $x$. We have good explicit estimates for $\psi(x) - x$, see for instance [18], [21] and [9]. This is due to the fact that we can use analytic tools in this problem since the residues at the poles of the Dirichlet generating series (namely here $-\zeta'(s)/\zeta(s)$) are known. However this situation has no counterpart in the Moebius function case. It would thus be highly valuable to deduce estimates for $M(x)$ from estimates for $\psi(x) - x$, but a precise quantitative link is missing. I proposed some years back the following conjecture:

Conjecture (Strong form of Landau’s equivalence Theorem, II).

There exist positive constants $c_1$ and $c_2$ such that

$$|M(x)|/x \leq c_1 \max_{c_2x < y \leq x/t} |\psi(y) - y|/y + c_1 x^{-1/4}.$$
Such a conjecture is trivially true under the Riemann Hypothesis. In this respect, we note that [23] proves that in case of the Beurling’s generalized integers, one can have $M_P(x) = o(x)$ without having $\psi(x) \sim x$. This reference has been kindly shown to me by Harold Diamond whom I warmly thank here.

We are not able to prove such a strong estimate, but we are still able to derive estimate for $M(x)$ from estimates for $\psi(x) - x$. Our process can be seen as a generalization of the initial idea of [20] also used in [10]. We describe it in the section 3, after a combinatorial preparation. Here is our main Theorem.

**Theorem 1.1.** For $D \geq 464402$, we have

$$\left| \sum_{d \leq D} \mu(d) \right| \leq \frac{0.0146 \log D - 0.1098}{(\log D)^2} \cdot D.$$ 

The last result of this shape is from [10] and has 0.10917 (starting from $D = 695$) instead of 0.0160.

On following an idea of [11] which we recall in the last section, we deduce from the above the following estimate.

**Corollary 1.1.** For $D \geq 59839$, we have

$$\left| \sum_{d \leq D} \mu(d)/d \right| \leq \frac{0.0292 \log D - 0.1098}{(\log D)^2}.$$ 

The last result of this shape is from [11] and has 0.2185 (starting from $x = 33$) instead of 0.0320. Here is result which is simpler to remember:

**Corollary 1.2.** For $D \geq 50000$, we have

$$\left| \sum_{d \leq D} \mu(d)/d \right| \leq \frac{3 \log D - 10}{100(\log D)^2}.$$ 

If we replace the $-10$ by 0, the resulting bound is valid from 11815 onward.

We will meet another problem in between, which is to relate quantitatively the error term $\psi(x) - x$ with the error term concerning the approximation of $\tilde{\psi}(x) = \sum_{n \leq x} \Lambda(n)/n$ by $\log x - \gamma$. This problem is surprisingly difficult but [15] offers a good enough solution.

I thank Harald Helfgott for interesting discussions that pushed me into pulling this note out of its drawer and François Dress for giving me the
preprint [11]. This paper was done in majority when I was enjoying the hospitality of the Mathematical Sciences Institute in Chennai, and I thank this institution and my hosts Ramachandran Balasubramanian, Anirban Mukhopadhyay and Sanoli Gun for this opportunity to work in peace and comfort.

Notation

We define the shortcuts $R(x) = \psi(x) - x$ and $r(x) = \tilde{\psi}(x) - \log x + \gamma$, where we recall that

$$\tilde{\psi}(x) = \sum_{n \leq x} \Lambda(n)/n. \quad (1)$$

We shall use square-brackets to denote the integer part and parenthesis to denote the fractional part, so that $D = \lfloor D \rfloor + \{D\}$. But since this notation is used seldomly we shall also use square brackets in their usual function.

2 A combinatorial tool

We prove a formal identity in this section. Let $F$ be a function and $Z = -F'/F$ the opposite of its logarithmic derivative. We look at

$$F[1/F]^{(k)} = P_k.$$

It is immediate to compute the first values and we find that

$$P_0 = F, \quad P_1 = Z, \quad P_2 = Z' + Z^2, \quad P_3 = Z'' + 3ZZ' + Z^3. \quad (2)$$

In general, the following recursion formula holds

$$P_k = F(P_{k-1}/F)' = P'_{k-1} + ZP_{k-1}. \quad (3)$$

Here is the result this leads to:

**Theorem 2.1.** We have

$$F[1/F]^{(k)} = \sum_{\sum i \geq 1, k_i = k} \frac{k!}{k_1!k_2! \cdots (1!)^{k_1}(2!)^{k_2} \cdots \prod_{k_i} Z^{(i-1)k_i}}.$$

We can prove it by using the recursion formula given above. We present now a different line. Let us expand $1/F(s + X)$ in Taylor series around $X = 0$.

$$\frac{1}{F(s + X)} = \sum_{k \geq 0} [1/F(s)]^{(k)} \frac{X^k}{k!}.$$
We do the same for $-F'(s + X)/F(s + X)$ getting:

$$
\frac{-F'(s + X)}{F(s + X)} = \sum_{k \geq 0} [Z(s)]^{(k)} \frac{X^k}{k!}.
$$

Integrating formally this expression, we get

$$
-\log \left( \frac{F(s + X)}{F(s)} \right) = \sum_{k \geq 1} [Z(s)]^{(k-1)} \frac{X^k}{k!}
$$

where the constant term is chosen so that the constant term is indeed 0. We then apply the exponential formula

$$
\exp \left( \sum_{k \geq 1} x_k \frac{X^k}{k!} \right) = \sum_{m \geq 0} Y_m(x_1, x_2, \ldots) \frac{x^m}{m!}
$$

where the $Y_m(x_1, x_2, \ldots)$ are the complete exponential Bell polynomials whose expression yields the Theorem above.

### 3 The general argument

Let us specialize $F = \zeta$ in Theorem 2.1. The left hand side therein has a simple pole in $s = 1$ with a residu being the $k$-th Taylor coefficient of $1/\zeta(s)$ around $s = 1$, coefficient that we are to multiply by $k!$. Let us call $\mathfrak{R}_k$ this residue. By a routine argument, we get

$$
\sum_{\ell \leq L} 1 \ast (\mu \log^k \ell) = \mathfrak{R}_k L + o(L).
$$

Note that, thanks to Theorem 2.1, the error term is quantified in terms of the error term in the approximations of both $\psi(x) - x$ and $\tilde{\psi}(x) - \log x + \gamma$. Getting to this error term in fact requires using a good enough error term for both these quantities (see for instance [12]). We then continue

$$
\sum_{\ell \leq L} \mu(\ell) \log^k \ell = \sum_{d \leq L} \mu(d) \left( \mathfrak{R}_k \frac{L}{d} + o(L/d) \right)
$$

which ensures us that $\sum_{\ell \leq L} \mu(\ell) \log^k \ell$ is $o(L \log L)$.

Case $k = 2$ is most enlightening. In this case, our method consist in writing

$$
\sum_{\ell \leq L} \mu(\ell) \log^2 \ell = \sum_{d \leq L} \mu(\ell) \left( \Lambda \ast \Lambda(d) - \Lambda(d) \log d \right).
$$
As it turns out, the main term of the summand function of $\Lambda \log$ (namely $L \log L$) cancels the one of $\Lambda \ast \Lambda$. This requires the prime number Theorem. In deriving the prime number theorem from Selberg’s formula $\mu \ast \log^2 = \Lambda \log + \Lambda \ast \Lambda$, it is a well known difficulty to show that both summands indeed contribute and this is another show-up of the parity principle. We modify (6) as follows:

$$2\gamma + \sum_{\ell \leq L} \mu(\ell) \log^2 \ell = \sum_{d \leq L} \mu(\ell)(\Lambda \ast \Lambda(d) - \Lambda(d) \log d + 2\gamma).$$

Case $k = 1$ is classical, but it is interesting to note that this is the starting point of [20].

### 4 Some known estimates and straightforward consequences

**Lemma 4.1** ([17]), $\max_{t \geq 1} \psi(t)/t = \psi(113)/113 \leq 1.04$.

Concerning small values, we quote from [16] the following result

$$|\psi(x) - x| \leq \sqrt{x} \quad (8 \leq x \leq 10^{10}).$$

If we change this $\sqrt{x}$ by $\sqrt{2x}$, this is valid from $x = 1$ onwards. Furthermore

$$|\psi(x) - x| \leq 0.8 \sqrt{x} \quad (1500 \leq x \leq 10^{10}).$$

**Lemma 4.2.**

$$|\psi(x) - x| \leq 0.0065x/\log x \quad (x \geq 1514928).$$

**Proof.** By [8, Théorème 1.3] improving on [21, Theorem 7], we have

$$|\psi(x) - x| \leq 0.0065x/\log x \quad (x \geq \exp(22)).$$

We readily extend this estimate to $x \geq 3430190$ by using (9). We then use the function $\text{WalkPsi}$ from the script $\text{IntR.gp}$ (with the proper $\text{model}$ function).

**Lemma 4.3.** For $x \geq 7105266$, we have

$$|\psi(x) - x|/x \leq 0.000213.$$
Proof. We start with the estimate from [19, (4.1)]
\[ |\psi(x) - x|/x \leq 0.000213 \quad (x \geq 10^{10}). \] (11)
We extend it to \( x \geq 14,500,000 \) by using (9). We complete the proof by using
the following Pari/GP script (see [22]):

```gp
{CalculeLambdas(Taille)=
    my(pk, Lambdas);
    Lambdas = vector(Taille);
    forprime(p = 2,Taille,
        pk = p;
        while(pk <= Taille, Lambdas[pk] = p; pk*=p));
    return(Lambdas);}
{model(n)=n}
{WalkPsi(zmin, zmax)=
    my(res = 0.0, mo, maxi, psiaux = 0.0, Lambdas);
    Lambdas = CalculeLambdas(zmax);
    for(y = 2, zmin,
        if(Lambdas[y]!=0, psiaux += log(Lambdas[y]),));
    maxi = abs(psiaux-zmin)/model(zmin);
    for(y = zmin+1, zmax,
        mo = 1/model(y);
        maxi = max(maxi, abs(psiaux-y)*mo);
        if(Lambdas[y]!=0, psiaux += log(Lambdas[y]),);
        maxi = max(maxi, abs(psiaux-y)*mo));
    print("|psi(x)-x|/model(x) <= ", maxi, " pour ",
        zmin, " <= x <= ", zmax);
    return(maxi);}
```

Lemma 4.4. For \( x \geq 59,843 \), we have
\[ |\psi(x) - x|/x \leq 0.0025. \]

Proof. The preceding Lemma proves it for \( x \geq 7,105,266 \). On using (9), we
extend it to \( x \geq 102,500 \). We complete the proof by using the same script as
in the proof of Lemma 4.3.

Lemma 4.5. For \( x \geq 32,054 \), we have
\[ |\psi(x) - x|/x \leq 0.003. \]
Proof. The preceding Lemma proves it for \( x \geq 7105266 \). On using (9), we extend it to \( x \geq 102500 \). We complete the proof by using the same script as in the proof of Lemma 4.3.

We quote from [15] the following Lemma.

**Lemma 4.6.** When \( x \geq 23 \), we have
\[
\hat{\psi}(x) = \log x - \gamma + O^*(\frac{0.0067}{\log x}).
\]

Let us turn our attention to the summatory function of the Moebius function. In [6], we find the bound
\[
|M(x)| \leq 0.571\sqrt{x} \quad (33 \leq x \leq 10^{12})
\] (12)
In [7], we find
\[
|M(x)| \leq x/2360 \quad (x \geq 617973)
\] (13)
(see also [4]) which [2] improves in
\[
|M(x)| \leq x/4345 \quad (x \geq 2160535).
\] (14)

**Bounds for squarefree numbers**

**Lemma 4.7.** We have for \( D \geq 1 \)
\[
\sum_{d \leq D} \mu^2(d) = \frac{6}{\pi^2} D + O^*(0.7\sqrt{D}).
\]

For \( D \geq 10 \), we can replace 0.7 by 0.5.

**Proof.** [1] (see also [2]) proves that
\[
\sum_{d \leq D} \mu^2(d) = \frac{6}{\pi^2} D + O^*(0.1333\sqrt{D}) \quad (D \geq 1664)
\]
and we use direct inspection using Pari/Gp to conclude.

**Lemma 4.8.** Let \( D/K \geq 1 \). Let \( f \) be a non-negative non-decreasing \( C^4 \) function. We have
\[
\sum_{D/L < d \leq D/K} \mu^2(d)f(D/d) \leq 1.31f(L) + \frac{6D}{\pi^2} \int_K^L \frac{f(t)dt}{t^2} + 0.35\sqrt{D} \int_K^L \frac{f(t)dt}{t^{3/2}}.
\]

7
Proof. We use a simple integration by parts to write
\[
\sum_{D/L<d\leq D/K} \mu^2(d) f(D/d) = \sum_{D/L<d\leq D/K} \mu^2(d) \left( f(K) + \int_{K}^{D/d} f'(t) dt \right).
\]

We then employ Lemma 4.7 to get the bound:
\[
\frac{6D}{\pi^2} f(K) + \int_{K}^{L} \frac{6D}{\pi^2 t} f'(t) dt + 0.7 \frac{D}{K} f(K) + 0.7 \int_{K}^{L} \frac{D}{t} f'(t) dt.
\]

Two integrations by parts gives the expression
\[
\frac{6}{\pi^2} f(L) + \int_{K}^{L} \frac{6D}{\pi^2 t^2} f(t) dt + 0.7 f(L) + 0.35 \sqrt{D} \int_{K}^{L} \frac{f(t) dt}{t^{3/2}}.
\]

The Lemma follows readily.

\[\square\]

5 A preliminary estimate on primes

Our aim here is to evaluate
\[
R_4(D) = \sum_{d_1 \leq \sqrt{D}} \Lambda(d_1) R(D/d_1).
\]
(15)

This remainder term is crucial in the final analysis and will be numerically one of the dominant terms.

Lemma 5.1. When \( D \geq 1 \), and \( \sqrt{D} \geq T \geq 1 \), we have
\[
\sum_{d \leq T} \frac{\Lambda(d)}{d \log \frac{D}{d}} \leq 1.04 \log \frac{\log D}{\log(D/T)} + 1.04 \log D.
\]

Proof. Let us define \( f(t) = 1/(t \log \frac{D}{t}) \). We have by a classical summation
by parts:

$$
\sum_{d \leq T} \Lambda(d) f(d) = \sum_{d \leq T} \Lambda(d) f(T) - \sum_{d \leq T} \Lambda(d) \int_d^T f'(t) dt
\leq \frac{1.04}{\log(D/T)} - 1.04 \int_1^T t f'(t) dt
\leq \frac{1.04}{\log(D/T)} - 1.04 [t f(t)]_1^T + 1.04 \int_1^T f(t) dt
\leq \frac{1.04}{\log D} + 1.04 \int_{D/T}^D \frac{dt}{t \log t} \leq \frac{1.04}{\log D} + 1.04 \log \frac{\log D}{\log(D/T)}
$$

as required.

**Lemma 5.2.** We have $|R_4(D)|/D \leq 0.0065$ when $D \geq 10^{10}$. When $D \geq 1300000000$, we have $|R_4(D)|/D \leq 0.01$.

The proof that follows is somewhat clumsy due to the fact that we have not been able to compute $R_4(D)$ for $D$ up to $10^{12}$. By inspecting the expression defining $R_4$ and the proof below, the reader will see one could try to get a better bound for

$$
\sum_{D^{1/4} < d \leq \sqrt{D}} \Lambda(d) R(D/d).
$$

Indeed one can compute the exact values of $R(D/d)$ and try to approximate them properly so as not to loose the sign changes in the expression. A proper model is even given by the explicit formula for $\psi(x)$. We have however tried to use the resulting polynomial, namely $x - \sum_{|\gamma| \leq G} x^{1/2 + i\gamma}/(1/2 + i\gamma)$ with $G = 20$, $G = 30$ and $G = 200$, but the approximation was very weak. It may be better to find directly a numerical fit for $R(x)$ on this limited range. It should be noted that the function $R(x)$ is highly erratical. Such a process would be important since the value 0.0065 that we get here decides for a large part of the final value in Theorem 1.1.

**Proof.** When $D \geq 1514928^2$, we have by Lemma 4.2 and Lemma 5.1:

$$
|R_4(D)|/D \leq 0.0065 \sum_{d \leq \sqrt{D}} \frac{\Lambda(d)}{d \log(D/d)} \leq 0.0065 \cdot \left(0.73 + \frac{1.04}{\log D}\right).
$$

This implies that $|R_4(D)|/D \leq 0.00499$ in the given range. When $10^{10} \leq$
$D \leq 1514928^2$, we set $T = D/10^9$, we write

$$|R_4(D)|/D \leq 0.000213 \sum_{d \leq T} \frac{\Lambda(d)}{d} + \frac{1}{D^{1/2}} \sum_{T < d \leq \sqrt{D}} \frac{\Lambda(d)}{\sqrt{d}}$$

$$\leq 0.000213 \psi(T) + \frac{1}{D^{1/2}} \left( \frac{\psi(\sqrt{D}) - \psi(T)}{D^{1/4}} + \frac{1}{2} \int_T^{\sqrt{D}} \frac{\psi(u) - \psi(T)}{u^{3/2}} du \right)$$

i.e. on using $\psi(u) \leq u + \sqrt{u}$,

$$|R_4(D)|/D \leq 0.000213 \tilde{\psi}(T) + \frac{1}{D^{1/2}} \left( \frac{\psi(\sqrt{D})}{D^{1/4}} - \frac{\psi(T)}{T^{1/2}} + \frac{1}{2} \int_T^{\sqrt{D}} \frac{\psi(u)}{u^{3/2}} du \right)$$

$$\leq 0.000213 \tilde{\psi}(T) + \frac{1}{D^{1/2}} \left( \frac{\sqrt{D} + D^{1/4}}{D^{1/4}} - \frac{T - \sqrt{T}}{T^{1/2}} + D^{1/4} - \sqrt{T} + \log \frac{\sqrt{D}}{T} \right)$$

i.e. since $\tilde{\psi}(x) \leq \log x$ when $x \geq 1$

$$|R_4(D)|/D \leq 0.000213 \log T + \frac{1}{D^{1/2}} \left( 2D^{1/4} - 2\sqrt{T} + 2 + \log \frac{\sqrt{D}}{T} \right).$$

We deduce that $|R_4(D)|/D \leq 0.0065$ when $D \geq 10^{10}$. When now $10^9 \leq D \leq 10^{10}$, we proceed as follows:

$$|R_4(D)|/D \leq \frac{1}{D^{1/2}} \left( \frac{\psi(1500)}{1500^{1/4}} + \frac{1}{2} \int_1^{1500} \frac{\psi(u)}{u^{3/2}} du \right)$$

$$+ \frac{0.8}{D^{1/2}} \left( \frac{\psi(\sqrt{D}) - \psi(1500)}{D^{1/4}} + \frac{1}{2} \int_{1500}^{\sqrt{D}} \frac{\psi(u) - \psi(1500)}{u^{3/2}} du \right).$$

$\psi(1500) = 1509.27 + O^*(0.01)$

$$|R_4(D)|/D^{1/2} \leq 0.2 \frac{1509.3}{1500^{1/4}} + 0.642 + 0.8 \cdot 1.04 \left( 2D^{1/4} - 1500^{1/4} \right).$$

The right hand side is not more than $0.00999$ when $D \geq 1300\,000\,000.$  \qed
6 The relevant error term for the primes

The main actor of this section is the remainder term $R_2^*$ defined by

$$\sum_{d \leq D} (\Lambda \ast \Lambda(d) - \Lambda(d) \log d) = -2[D] \gamma + R_2^*(D). \quad (16)$$

The object of this section is to derive explicit estimates for $R_2^*$ from explicit estimates for the $\psi$. Most of the original work has been achieved already in the previous section, and we essentially put things in shape. Here is our result.

**Lemma 6.1.** When $D \geq 1086579$, we have $|R_2^*(D)|/D \leq 0.0240$.

We start by an expression for $R_2^*$.

**Lemma 6.2.**

$$|R_2^*(D)| \leq 2D |r(\sqrt{D})| + 2D^{1/2} R(\sqrt{D}) + R(\sqrt{D})^2 + R(D) \log D$$

$$+ 1 + 2\gamma + 2R_4(D) + \left| \int_1^D R(t) \frac{dt}{t} \right|$$

where $R_4$ is defined in (15).

**Proof.** The proof is fully pedestrian. We have

$$\sum_{d \leq D} \Lambda(d) \log d = \psi(D) \log D - \int_1^D \psi(t) dt/t$$

$$= D \log D - D + 1 + R(D) \log D - \int_1^D R(t) dt/t.$$  

Concerning the other summand, Dirichlet hyperbola formula yields

$$\sum_{d_1d_2 \leq D} \Lambda(d_1)\Lambda(d_2) = 2 \sum_{d_1 \leq \sqrt{D}} \Lambda(d_1) \sum_{d_2 \leq D/d_1} \Lambda(d_2) - \psi(\sqrt{D})^2$$

$$= 2D \sum_{d_1 \leq \sqrt{D}} \frac{\Lambda(d_1)}{d_1} - D$$

$$- 2\sqrt{D} R(\sqrt{D}) - R(\sqrt{D})^2 + 2 \sum_{d_1 \leq \sqrt{D}} \Lambda(d_1) R(D/d_1)$$

$$= D \log D - 2D\gamma - D$$

$$+ 2Dr(\sqrt{D}) - 2\sqrt{D} R(\sqrt{D}) - R(\sqrt{D})^2 + 2R_4(D).$$
We reach $R_2^*(D) = R_3(D) - 1 + 2R_4(D) - R(D) \log D + \int_1^D R(t)dt/t$, where

$$R_3(D) = 2Dr(\sqrt{D}) - 2\gamma\{D\} - 2\sqrt{D}R(\sqrt{D}) - R(\sqrt{D})^2.$$  \hspace{1cm} (17)

The Lemma follows readily. \qed

**Lemma 6.3.** For the real number $D$ verifying $3 \leq D \leq 110\,000\,000$, we have

$$|R_2^*(D)| \leq 1.80\sqrt{D} \log D.$$  

When $110\,000\,000 \leq D \leq 1\,800\,000\,000$, we have

$$|R_2^*(D)| \leq 1.93\sqrt{D} \log D.$$  

We used a Pari/Gp script. The only non-obvious point is that we have precomputed the values of $\Lambda \ast \Lambda - \Lambda \ast \log$ on intervals of length $2 \cdot 10^6$. On letting this script run longer (about twenty days), I would most probably able to show that the bound $|R_2^*(D)| \leq 2\sqrt{D} \log D$ holds when $D \leq 10^{10}$. This would improve a bit on the final result.

**Lemma 6.4.**

$$\int_1^{10^8} R(t)dt/t = -129.559 + O^*(0.01).$$

See script $\text{IntR.gp}$. 

**Proof.** We prove Lemma 6.1 here. Let us assume that $D \geq 1.3 \cdot 10^9$. We start with Lemma 6.2. We bound $r(\sqrt{D})$ via Lemma 4.6 (this requires $D \geq 23^2$), then $R(\sqrt{D})$ by Lemma 4.4 (this requires $D \geq 32054^2$), and $R(D) \log D$ by using Lemma 4.2 (this requires $D \geq 1\,514\,928$). We bound $R_4$ by appealing to Lemma 5.2. We conclude by appealing to Lemma 4.3. All of that amounts to the bound:

$$|R_2^*(D)| \leq \frac{4 \cdot 0.0067 D}{\log D} + 0.006 D + (0.003)^2 D + 0.0065 D + 0.01 D + 132 + 0.000213 D - 0.000213 \cdot 10^8.$$  

We reach

$$|R_2^*(D)|/D \leq 0.0240$$  \hspace{1cm} (18)

when $D \geq 1.3 \cdot 10^9$. Thanks to Lemma 6.3, we extend this bound to $D \geq 1\,086\,579$. \qed


7 Estimating $M(D)$

We appeal to (7) and use Dirichlet hyperbola formula. We get in this manner our starting equation:

$$
\sum_{d \leq D} \mu(d) \log^2 d = 2\gamma + \sum_{d \leq D/K} \mu(d) R_2^*(D/d) \\
+ \sum_{k \leq K} R_2^*(k) \sum_{D/(k+1) < d \leq D/k} \mu(d). \quad (19)
$$

This equation is much more important than it looks since a bound for $R_2^*(k)$ that is $\ll k/(\log k)^2$ shows that the second sum converges. A more usual treatment would consist in writing

$$
\sum_{d \leq D} \mu(d) \log^2 d = 2\gamma + \sum_{d \leq D/K} \mu(d) R_2^*(D/d) \\
+ \sum_{k \leq K} (\Lambda \ast \Lambda - \Lambda \log + 2\gamma)(k) \sum_{D/K < d \leq D/k} \mu(d).
$$

as in [20] for instance. However, when we bound $M(D/k) - M(D/(k + 1))$ roughly by $D/(k(k + 1))$ in (19), we get $D \sum_{k \leq K} |R_2^*(k)|/(k(k + 1))$ which is expected to be $O(D)$. On bounding $M(D/k) - M(D/K)$ by $D/k$ in the second expression, we only get $D \sum_{k \leq K} |\Lambda \ast \Lambda - \Lambda \log - 2\gamma|(k)/k$ which is of size $D \log^2 K$. Practically, if we want to use a bound of the shape $|M(x)| \leq x/2360$, we will loose the differentiating aspect and will bound $|M(D/k) - M(D/(k + 1))|$ by $2D/(2360 k)$ and not by $D/(2360 k^2)$. It is thus better to use differentiation with respect to $R_2^*$ when $k$ is fairly small.

It turns out that small is large enough! We write

$$
\sum_{k \leq K} R_2^*(k)(M(D/k) - M(D/(k + 1))) \\
= \sum_{k \leq K} (\Lambda \ast \Lambda - \Lambda \log + 2\gamma)(k)M(D/k) + R_2^*(K)M(D/K). \quad (20)
$$

Lemma 7.1. When $K = 100\,000$, we have

$$
\sum_{k \leq K} \frac{|\Lambda \ast \Lambda - \Lambda \log + 2\gamma|(k)}{k} + \frac{|R_2^*(K)|}{K} \leq 0.02503 \times 2360.
$$
We can use the simple bound (18) and get, for \( D/K \geq 1086579 \)
\[
\left| \sum_{d \leq D} \mu(d) \log^2 d \right| / D \leq \frac{2\gamma}{D} + 0.0240 \left( \frac{6}{\pi^2} \log \frac{D}{K} + 1.166 \right) + 0.03660
\leq 0.0146 \log D - 0.139
\]
with \( K = 10^5 \).

Concerning the smaller values, we use summation by parts:
\[
\sum_{d \leq D} \mu(d) \log^2 d = \sum_{d \leq D} \mu(d) \log^2 D - 2 \int_1^D \sum_{d \leq t} \mu(d) \frac{\log t \, dt}{t}
\]
which gives, when \( 33 \leq D \leq 10^{12} \),
\[
\left| \sum_{d \leq D} \mu(d) \log^2 d \right| \leq 0.571 \sqrt{D} \log^2 D + 2 \left| \int_1^{33} \sum_{d \leq t} \mu(d) \frac{\log t \, dt}{t} \right|
+ 2 \cdot 0.571 \int_33^D \frac{\log t \, dt}{\sqrt{t}}
\leq 0.571 \sqrt{D} \log^2 D + 2.284 \sqrt{D} \log D + 4.568 \sqrt{D} - 43
\]
and this is \( \leq 0.0146 \log D - 0.139 \) when \( D \geq 4225000 \). We extend this bound to \( D \geq 464405 \) by direct computations using Pari/Gp.

Let us state formally:

**Lemma 7.2.** For \( D \geq 1078806 \), we have
\[
\left| \sum_{d \leq D} \mu(d) \log^2 d \right| / D \leq 0.0146 \log D - 0.139.
\]

### 8 A general formula and proof of Theorem 1.1

Let \((f(n))\) be a sequence of complex numbers. We consider, for integer \( k \geq 0 \), the weighted summatory function
\[
M_k(f, D) = \sum_{n \leq D} f(n) \log^k n.
\] (21)

We want to derive information on \( M_0(f, D) \) from information on \( M_k(f, D) \). The traditional way to do that is in essence due to [14] and goes via a differential equation. It turns out that it is clearer and somewhat more precise to use the identity that follows.
Lemma 8.1. We have, when \( k \geq 0 \), and for \( D \geq D_0 \),

\[
M_0(f, D) = \frac{M_k(f, D)}{\log^k D} + M_0(f, D_0) - \frac{M_k(f, D_0)}{\log^k D_0} - k \int_{D_0}^D \frac{M_k(f, t)}{t \log^{k+1} t} dt.
\]

This formula in a special case is also used in [20] and [10].

Proof. Indeed, we have

\[
k \int_{D_0}^D \frac{M_k(f, t)}{t \log^{k+1} t} dt = -\frac{M_k(f, D_0)}{\log^k D_0} + \sum_{n \leq D} f(n) \frac{\log^k n}{\log^k D} - \sum_{D_0 < n \leq D} f(n).
\]

\( \square \)

Proof. We proceed to the proof of Theorem 1.1. In the notation of Lemma 8.1, we have \( M(D) = M_0(\mu, D) \). We have by Lemma 7.2 and with \( D_0 = 1078806 \):

\[
|M(D)| \leq \frac{0.0146 \log D - 0.139}{\log^2 D} D + M(D_0) - \frac{M_2(\mu, D_0)}{\log^2 D_0} + 2 \int_{D_0}^D \frac{0.0146 \log t - 0.139}{\log^3 t} dt.
\]

\[
\leq \frac{0.0146 \log D - 0.139}{\log^2 D} D - 1.25 + 2 \int_{D_0}^D \frac{0.0146 \log t - 0.139}{\log^3 t} dt.
\]

\[
\leq \frac{0.0146 \log D - 0.1098}{\log^2 D} D - 1.25 - 0.0292 \frac{D_0}{\log^2 D_0} - \int_{D_0}^D \frac{0.2196}{t \log^3 t} dt.
\]

(We use Pari/Gp to compute the quantity \( M(D_0) - M_2(\mu, D_0)/\log^2 D_0 \)). We conclude by direct verification, again by relying on Pari/Gp. \( \square \)

9 From \( M \) to \( m \)

We take the following Lemma from [11, (1.1)].

Lemma 9.1 (El Marraki). We have

\[
|m(D)| \leq \frac{|M(D)|}{D} + \frac{1}{D} \int_1^D \frac{|M(t)| dt}{t} + \frac{\log D}{D}.
\]
This Lemma may look trivial enough, but its teeth are hidden. Indeed, a usual summation by parts would bound $|m(D)|$ by an expression containing the integral of $|M(t)|/t^2$. An upper bound for $|M(t)|$ of the shape $ct/\log t$ would hence result in the useless bound $m(D) \ll \log \log D$.

**Proof.** We reproduce the proof, as it is short and the preprint we refer to is difficult to find. We have two equations, namely:

$$m(D) = \frac{M(D)}{D} + \int_1^D \frac{M(t)dt}{t}$$  \quad (22)

and

$$\int_1^D \left[ \frac{D}{t} \right] \frac{M(t)dt}{t^2} = \log D.$$  \quad (23)

We deduce from the above that

$$m(D) = \frac{M(D)}{D} + \frac{1}{D} \int_1^D \left( \frac{D}{t} - \left[ \frac{D}{t} \right] \right) \frac{M(t)dt}{t} + \frac{\log D}{D}.$$  

The Lemma follows readily.

**Proof.** We have, when $D \geq D_0 = 464402$,

$$|m(D)| \leq \frac{0.0146 \log D - 0.1098}{(\log D)^2} + \frac{1}{D} \int_{D_0}^D \frac{0.0146 \log t - 0.1098}{(\log t)^2} dt$$

$$+ \frac{1}{D} \int_1^{D_0} \frac{|M(t)|dt}{t} + \frac{\log D}{D},$$

$$\leq \frac{0.0146 \log D - 0.1098}{(\log D)^2} + \frac{1}{D} \int_{D_0}^D \frac{0.0146 dt}{\log t}$$

$$- \frac{1}{D} \int_{D_0}^D \frac{0.1098 dt}{(\log t)^2} + \frac{196 + \log D}{D}.$$  

We continue by an integration by parts and some numerical computations:

$$|m(D)| \leq \frac{0.0292 \log D - 0.1098}{(\log D)^2} - \frac{0.0952}{D} \int_{D_0}^D \frac{dt}{(\log t)^2} + \frac{-323 + \log D}{D},$$

$$\leq \frac{0.0292 \log D - 0.1098}{(\log D)^2} - \frac{1}{D} \int_{D_0}^D \frac{dt}{t} + \frac{-271 + \log D}{D}.$$  

This proves that $|m(D)| (\log D)^2 \leq 0.0292 \log D - 0.1098$ as soon as $D \geq 464402$. We extend this bound by direct inspection. \qed
References


