



**HAL**  
open science

# From explicit estimates for primes to explicit estimates for the Möbius function

Olivier Ramaré

► **To cite this version:**

Olivier Ramaré. From explicit estimates for primes to explicit estimates for the Möbius function. *Acta Arithmetica*, 2013, 157 (4), pp.365-379. 10.4064/aa157-4-4 . hal-02572812

**HAL Id: hal-02572812**

**<https://hal.science/hal-02572812v1>**

Submitted on 13 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Olivier Ramaré,  
CNRS, Laboratoire Paul Painlevé,  
Université Lille 1,  
59 655 Villeneuve d'Ascq, France  
Email: ramare@math.univ-lille1.fr

June 28, 2012

## Abstract

We prove two explicit estimates respectively slightly stronger than  $|\sum_{d \leq D} \mu(d)|/D \leq 0.013/\text{Log } D$  for every  $D \geq 1\,100\,000$  and than  $|\sum_{d \leq D} \mu(d)/d| \leq 0.026/\text{Log } D$  for every  $D \geq 61\,000$ .

## 1 Introduction

There is a long litterature concerning explicit estimates for the summatory function of the Moebius function, and we cite for instance [21], [1], [4], [3], [6], [7], [10], [11]. The paper [5] proposes a very usefull 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 Tchebychef 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 [19], [22] 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:

---

2010 *Mathematics Subject Classification*: Primary 11N37, 11Y35; Secondary 11A25.  
*Key words and phrases*: Explicit estimates, Moebius function.

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

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

$$|M(x)|/x \leq c_1 \max_{c_2x < y \leq x/c_2} |\psi(y) - y|/y + c_1x^{-1/4}.$$

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  $M_{\mathcal{P}}(x) = o(x)$  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 [21] also used in [10]. We describe it in the section 3, after a combinatorial preparation. Here is our main Theorem.

**Theorem 1.1.** *For  $D \geq 1\,078\,853$ , we have*

$$\left| \sum_{d \leq D} \mu(d) \right| \leq \frac{0.0130 \operatorname{Log} D - 0.118}{(\operatorname{Log} D)^2} D.$$

The last result of this shape is from [10] and has 0.10917 (starting from  $D = 695$ ) instead of 0.0130.

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

**Corollary 1.2.** *For  $D \geq 60\,298$ , we have*

$$\left| \sum_{d \leq D} \mu(d)/d \right| \leq \frac{0.0260 \operatorname{Log} D - 0.118}{(\operatorname{Log} D)^2}.$$

The last result of this shape is from [11] and has 0.2185 (starting from  $x = 33$ ) instead of 0.0260. Here are two results that are simpler to remember:

**Corollary 1.3.** *For  $D \geq 60\,200$ , we have*

$$\left| \sum_{d \leq D} \mu(d)/d \right| \leq \frac{\operatorname{Log} D - 4}{40(\operatorname{Log} D)^2}.$$

*If we replace the  $-4$  by  $0$ , the resulting bound is valid from  $24\,270$  onward.*

**Corollary 1.4.** *For  $D \geq 50\,000$ , we have*

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

*If we replace the  $-10$  by  $0$ , the resulting bound is valid from  $11\,815$  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  $\text{Log } x - \gamma$ . This problem is surprisingly difficult but [16] 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  $r(x) = \tilde{\psi}(x) - \text{Log } x + \gamma$ , where we recall that

$$(1.1) \quad \tilde{\psi}(x) = \sum_{n \leq x} \Lambda(n)/n.$$

We shall use square-brackets to denote the integer part and parenthesis to denote the fractionnal part, so that  $D = [D] + \{D\}$ . 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

$$(2.1) \quad P_0 = F, \quad P_1 = Z, \quad P_2 = Z' + Z^2, \quad P_3 = Z'' + 3ZZ' + Z^3.$$

In general, the following recursion formula holds

$$(2.2) \quad P_k = F(P_{k-1}/F)' = P'_{k-1} + ZP_{k-1}.$$

Here is the result this leads to:

**Theorem 2.1.** *We have*

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

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

$$-\text{Log}(F(s+X)/F(s)) = \sum_{k \geq 1} [Z(s)]^{(k-1)} \frac{X^k}{k!}$$

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

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

where the  $Y_m(x_1, x_2, \dots)$  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

$$(3.1) \quad \sum_{\ell \leq L} \mathbb{1} \star (\mu \text{Log}^k)(\ell) = \mathfrak{R}_k L + o(L).$$

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  $\tilde{\psi}(x) - \text{Log } x + \gamma$ . 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

$$(3.2) \quad \sum_{\ell \leq L} \mu(\ell) \text{Log}^k \ell = \sum_{d \leq L} \mu(d) \left( \mathfrak{R}_k \frac{L}{d} + o(L/d) \right)$$

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

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

$$(3.3) \quad \sum_{\ell \leq L} \mu(\ell) \text{Log}^2 \ell = \sum_{d\ell \leq L} \mu(\ell) (\Lambda \star \Lambda(d) - \Lambda(d) \text{Log} d).$$

As it turns out, the main term of the summatory function of  $\Lambda \text{Log}$  (namely  $L \text{Log} L$ ) cancels the one of  $\Lambda \star \Lambda$ . This requires the prime number Theorem. In deriving the prime number theorem from Selberg's formula  $\mu \star \text{Log}^2 = \Lambda \text{Log} + \Lambda \star \Lambda$ , 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 (3.3) as follows:

$$(3.4) \quad 2\gamma + \sum_{\ell \leq L} \mu(\ell) \text{Log}^2 \ell = \sum_{d\ell \leq L} \mu(\ell) (\Lambda \star \Lambda(d) - \Lambda(d) \text{Log} d + 2\gamma).$$

Case  $k = 1$  is classical, but it is interesting to note that this is the starting point of [21].

## 4 Some known estimates and straightforward consequences

**Lemma 4.1** ([18]).  $\max_{t \geq 1} \psi(t)/t = \psi(113)/113 \leq 1.04$ .

Concerning small values, we quote from [17] the following result

$$(4.1) \quad |\psi(x) - x| \leq \sqrt{x} \quad (8 \leq x \leq 10^{10}).$$

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

$$(4.2) \quad |\psi(x) - x| \leq 0.8 \sqrt{x} \quad (1\,500 \leq x \leq 10^{10}).$$

**Lemma 4.2.**

$$|\psi(x) - x| \leq 0.0065x / \text{Log} x \quad (x \geq 1\,514\,928).$$

*Proof.* By [8, Théorème 1.3] improving on [22, Theorem 7], we have

$$(4.3) \quad |\psi(x) - x| \leq 0.0065x / \text{Log} x \quad (x \geq \exp(22)).$$

We readily extend this estimate to  $x \geq 3\,430\,190$  by using (4.2). We then use the function `WalkPsi` from the script `IntR.gp` (with the proper `model` function). □

**Lemma 4.3.** *For  $x \geq 7\,105\,266$ , we have*

$$|\psi(x) - x|/x \leq 0.000\,213.$$

*Proof.* We start with the estimate from [20, (4.1)]

$$(4.4) \quad |\psi(x) - x|/x \leq 0.000\,213 \quad (x \geq 10^{10}).$$

We extend it to  $x \geq 14\,500\,000$  by using (4.2). We complete the proof by using the following Pari/Gp script (see [15]):

```
{CalculerLambdas(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 = CalculerLambdas(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 (4.2), 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  $x \geq 7\,105\,266$ . On using (4.2), 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. □

We quote from [16] the following Lemma.

**Lemma 4.6.** *When  $x \geq 23$ , we have*

$$\tilde{\psi}(x) = \text{Log } x - \gamma + \mathcal{O}^*\left(\frac{0.0067}{\text{Log } x}\right).$$

Let us turn our attention to the summatory function of the Moebius function. In [6], we find the bound

$$(4.5) \quad |M(x)| \leq 0.571\sqrt{x} \quad (33 \leq x \leq 10^{12})$$

In [7], we find

$$(4.6) \quad |M(x)| \leq x/2360 \quad (x \geq 617973)$$

(see also [4]) which [2] (published also in [3]) improves in

$$(4.7) \quad |M(x)| \leq x/4345 \quad (x \geq 2\,160\,535).$$

### **Bounds for squarefree numbers**

**Lemma 4.7.** *We have for  $D \geq 1$*

$$\sum_{d \leq D} \mu^2(d) = \frac{6}{\pi^2}D + \mathcal{O}^*(0.7\sqrt{D}).$$

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

*Proof.* [1] (see also [2]) proves that

$$\sum_{d \leq D} \mu^2(d) = \frac{6}{\pi^2}D + \mathcal{O}^*(0.1333\sqrt{D}) \quad (D \geq 1\,664)$$

and we use direct inspection using Pari/Gp to conclude.  $\square$

**Lemma 4.8.** *Let  $D/K \geq 1$ . Let  $f$  be a non-negative non-decreasing  $C^1$  function. We have*

$$\sum_{D/L < d \leq D/K} \mu^2(d)f(D/d) \leq 1.31f(L) + \frac{6D}{\pi^2} \int_K^L \frac{f(t)dt}{t^2} + 0.35\sqrt{D} \int_K^L \frac{f(t)dt}{t^{3/2}}.$$

*Proof.* We use a simple integration by parts to write

$$\begin{aligned} \sum_{D/L < d \leq D/K} \mu^2(d)f(D/d) &= \sum_{D/L < d \leq D/K} \mu^2(d) \left( f(K) + \int_K^{D/d} f'(t)dt \right) \\ &= \sum_{D/L < d \leq D/K} \mu^2(d)f(K) + \int_K^L \left( \sum_{D/L < d \leq D/t} \mu^2(d) \right) f'(t)dt. \end{aligned}$$



We then employ Lemma 4.7 to get the bound:

$$\frac{6D}{\pi^2 K} f(K) + \int_K^L \frac{6D}{\pi^2 t} f'(t) dt + 0.7 \sqrt{\frac{D}{K}} f(K) + 0.7 \int_K^L \sqrt{\frac{D}{t}} f'(t) dt$$

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

$$(5.1) \quad R_4(D) = \sum_{d_1 \leq \sqrt{D}} \Lambda(d_1) R(D/d_1).$$

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*

$$\sum_{d \leq T} \frac{\Lambda(d)}{d \operatorname{Log} \frac{D}{d}} \leq 1.04 \operatorname{Log} \frac{\operatorname{Log} D}{\operatorname{Log}(D/T)} + \frac{1.04}{\operatorname{Log} D}.$$

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

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

as required.  $\square$

**Lemma 5.2.** *We have  $|R_4(D)|/D \leq 0.0065$  when  $D \geq 10^{10}$ . When  $D \geq 1\,300\,000\,000$ , we have  $|R_4(D)|/D \leq 0.0073$ .*

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  $10^{10}$ . 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 lose 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  $x - \sum_{|\gamma| \leq G} x^{\frac{1}{2} + i\gamma} / (\frac{1}{2} + i\gamma)$  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 erratic. 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:

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

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

$$\begin{aligned} |R_4(D)|/D &\leq 0.000213 \sum_{d \leq T} \frac{\Lambda(d)}{d} + \frac{1}{D^{1/2}} \sum_{T < d \leq \sqrt{D}} \frac{\Lambda(d)}{\sqrt{d}} \\ &\leq 0.000213 \tilde{\psi}(T) \\ &\quad + \frac{1}{D^{1/2}} \left( \frac{\psi(\sqrt{D}) - \psi(T)}{D^{1/4}} + \frac{1}{2} \int_T^{\sqrt{D}} \frac{\psi(u) - \psi(T)}{u^{3/2}} du \right) \end{aligned}$$

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

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

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

$$\begin{aligned} |R_4(D)|/D &\leq 0.000213 \operatorname{Log} T \\ &\quad + \frac{1}{D^{1/2}} \left( 2D^{1/4} - 2\sqrt{T} + 2 + \operatorname{Log} \frac{\sqrt{D}}{T} \right). \end{aligned}$$

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:

$$|R_4(D)|/D \leq \frac{1}{D^{1/2}} \left( \frac{\psi(1500)}{1500^{1/2}} + \frac{1}{2} \int_1^{1500} \frac{\psi(u)}{u^{3/2}} du \right) + \frac{0.8}{D^{1/2}} \left( \frac{\psi(\sqrt{D}) - \psi(1500)}{D^{1/4}} + \frac{1}{2} \int_{1500}^{\sqrt{D}} \frac{\psi(u) - \psi(1500)}{u^{3/2}} du \right).$$

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

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

The right hand side is not more than 0.0073 when  $D \geq 1\,300\,000\,000$ .  $\square$

## 6 The relevant error term for the primes

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

$$(6.1) \quad \sum_{d \leq D} (\Lambda \star \Lambda(d) - \Lambda(d) \operatorname{Log} d) = -2[D]\gamma + R_2^*(D).$$

The object of this section is to derive explicit estimate 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  $D \geq 1\,435\,319$ , we have  $|R_2^*(D)|/D \leq 0.0213$ .*

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

**Lemma 6.2.**

$$|R_2^*(D)| \leq 2D|r(\sqrt{D})| + 2D^{1/2}R(\sqrt{D}) + R(\sqrt{D})^2 + R(D) \operatorname{Log} D + 1 + 2\gamma + 2R_4(D) + \left| \int_1^D R(t) \frac{dt}{t} \right|$$

where  $R_4$  is defined in (5.1).

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

$$\begin{aligned} \sum_{d \leq D} \Lambda(d) \operatorname{Log} d &= \psi(D) \operatorname{Log} D - \int_1^D \psi(t) dt/t \\ &= D \operatorname{Log} D - D + 1 + R(D) \operatorname{Log} D - \int_1^D R(t) dt/t. \end{aligned}$$

Concerning the other summand, Dirichlet hyperbola formula yields

$$\begin{aligned}
 \sum_{d_1 d_2 \leq D} \Lambda(d_1) \Lambda(d_2) &= 2 \sum_{d_1 \leq \sqrt{D}} \Lambda(d_1) \sum_{d_2 \leq D/d_1} \Lambda(d_2) - \psi(\sqrt{D})^2 \\
 &= 2D \sum_{d_1 \leq \sqrt{D}} \frac{\Lambda(d_1)}{d_1} - D \\
 &\quad - 2\sqrt{D}R(\sqrt{D}) - R(\sqrt{D})^2 + 2 \sum_{d_1 \leq \sqrt{D}} \Lambda(d_1) R(D/d_1) \\
 &= D \operatorname{Log} D - 2D\gamma - D \\
 &\quad + 2Dr(\sqrt{D}) - 2\sqrt{D}R(\sqrt{D}) - R(\sqrt{D})^2 + 2R_4(D).
 \end{aligned}$$

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

$$(6.2) \quad R_3(D) = 2Dr(\sqrt{D}) - 2\gamma\{D\} - 2\sqrt{D}R(\sqrt{D}) - R(\sqrt{D})^2.$$

The Lemma follows readily.  $\square$

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

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

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

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

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

**Lemma 6.4.**

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

See script `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  $R(D) \operatorname{Log} D$  by using Lemma 4.2 (this requires  $D \geq 1514928$ ). 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:

$$\begin{aligned}
 |R_2^*(D)| \leq & \frac{4 \cdot 0.0067 D}{\operatorname{Log} D} + 0.006 D + (0.003)^2 D + 0.0065 D \\
 & + 0.0073 D + 132 + 0.000213 D - 0.000213 \cdot 10^8.
 \end{aligned}$$

We reach

$$(6.3) \quad |R_2^*(D)|/D \leq 0.0213$$

when  $D \geq 1.3 \cdot 10^9$ . Thanks to Lemma 6.3, we extend this bound to  $D \geq 1\,435\,319$ .  $\square$

## 7 Estimating $M(D)$

We appeal to (3.4) and use Dirichlet hyperbola formula. We get in this manner our starting equation:

$$(7.1) \quad \sum_{d \leq D} \mu(d) \operatorname{Log}^2 d = 2\gamma + \sum_{d \leq D/K} \mu(d) R_2^*(D/d) \\ + \sum_{k \leq K} R_2^*(k) \sum_{D/(k+1) < d \leq D/k} \mu(d).$$

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

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

as in [21] for instance. However, when we bound  $M(D/k) - M(D/(k+1))$  roughly by  $D/(k(k+1))$  in (7.1), we get  $D \sum_{k \leq K} |R_2^*(k)|/(k(k+1))$  which is expected to be  $\mathcal{O}(D)$ . On bounding  $M(D/k) - M(D/K)$  by  $D/k$  in the second expression, we only get  $D \sum_{k \leq K} |\Lambda \star \Lambda - \Lambda \operatorname{Log} - 2\gamma|(k)/k$  which is of size  $D \operatorname{Log}^2 K$ . Practically, if we want to use a bound of the shape  $|M(x)| \leq x/4345$ , we will loose the differentiating aspect and will bound  $|M(D/k) - M(D/(k+1))|$  by  $2D/(4345k)$  and not by  $D/(4345k^2)$ . 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

$$(7.2) \quad \sum_{k \leq K} R_2^*(k) (M(D/k) - M(D/(k+1))) \\ = \sum_{k \leq K} (\Lambda \star \Lambda - \Lambda \operatorname{Log} + 2\gamma)(k) M(D/k) + R_2^*(K) M(D/K).$$

**Lemma 7.1.** *When  $K = 462\,848$ , we have*

$$\sum_{k \leq K} \frac{|\Lambda \star \Lambda - \Lambda \operatorname{Log} + 2\gamma|(k)}{k} + \frac{|R_2^*(K)|}{K} \leq 0.03739 \times 4345.$$

We can use the simple bound (6.3) and get, for  $D/K \geq 2\,160\,535$

$$\left| \sum_{d \leq D} \mu(d) \operatorname{Log}^2 d \right| / D \leq \frac{2\gamma}{D} + 0.0213 \left( \frac{6}{\pi^2} \operatorname{Log} \frac{D}{K} + 1.166 \right) + 0.03739$$

$$\leq 0.0130 \operatorname{Log} D - 0.144$$

with  $K = 462\,848$ . Note that this lower bound of  $K$  has been chosen to verify

$$462\,848 \times 2\,160\,535 \leq 10^{12}.$$

Concerning the smaller values, we use summation by parts:

$$\sum_{d \leq D} \mu(d) \operatorname{Log}^2 d = \sum_{d \leq D} \mu(d) \operatorname{Log}^2 D - 2 \int_1^D \sum_{d \leq t} \mu(d) \frac{\operatorname{Log} t \, dt}{t}$$

which gives, when  $33 \leq D \leq 10^{12}$ ,

$$\left| \sum_{d \leq D} \mu(d) \operatorname{Log}^2 d \right| \leq 0.571\sqrt{D} \operatorname{Log}^2 D + 2 \left| \int_1^{33} \sum_{d \leq t} \mu(d) \frac{\operatorname{Log} t \, dt}{t} \right|$$

$$+ 2 \cdot 0.571 \int_{33}^D \frac{\operatorname{Log} t \, dt}{\sqrt{t}}$$

$$\leq 0.571\sqrt{D} \operatorname{Log}^2 D + 2.284\sqrt{D} \operatorname{Log} D + 4.568\sqrt{D} - 43$$

and this is  $\leq 0.0130 \operatorname{Log} D - 0.144$  when  $D \geq 8\,613\,000$ . We extend this bound to  $D \geq 2\,161\,205$  by direct computations using Pari/Gp.

Let us state formally:

**Lemma 7.2.** *For  $D \geq 2\,161\,205$ , we have*

$$\left| \sum_{d \leq D} \mu(d) \operatorname{Log}^2 d \right| / D \leq 0.0130 \operatorname{Log} D - 0.144.$$

## 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

$$(8.1) \quad M_k(f, D) = \sum_{n \leq D} f(n) \operatorname{Log}^k n.$$

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  $D \geq D_0$ ,*

$$M_0(f, D) = \frac{M_k(f, D)}{\text{Log}^k D} + M_0(f, D_0) - \frac{M_k(f, D_0)}{\text{Log}^k D_0} - k \int_{D_0}^D \frac{M_k(f, t)}{t \text{Log}^{k+1} t} dt.$$

This formula in a special case is also used in [21] and [10].

*Proof.* Indeed, we have

$$k \int_{D_0}^D \frac{M_k(f, t)}{t \text{Log}^{k+1} t} dt = -\frac{M_k(f, D_0)}{\text{Log}^k D_0} + \sum_{n \leq D} f(n) \frac{\text{Log}^k n}{\text{Log}^k D} - \sum_{D_0 < n \leq D} f(n)$$

□

*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  $D_0 = 2\,161\,205$ :

$$\begin{aligned} |M(D)| &\leq \frac{0.0130 \text{Log} D - 0.144}{\text{Log}^2 D} D + M(D_0) - \frac{M_2(\mu, D_0)}{\text{Log}^2 D_0} \\ &\quad + 2 \int_{D_0}^D \frac{0.0130 \text{Log} t - 0.144}{\text{Log}^3 t} dt. \\ &\leq \frac{0.0130 \text{Log} D - 0.144}{\text{Log}^2 D} D - 3.48 + 2 \int_{D_0}^D \frac{0.0130 \text{Log} t - 0.144}{\text{Log}^3 t} dt. \\ &\leq \frac{0.0130 \text{Log} D - 0.118}{\text{Log}^2 D} D - 3.48 \\ &\quad - 0.0260 \frac{D_0}{\text{Log}^2 D_0} - \int_{D_0}^D \frac{0.236}{t \text{Log}^3 t} dt. \end{aligned}$$

(We used Pari/Gp to compute the quantity  $M(D_0) - M_2(\mu, D_0)/\text{Log}^2 D_0$ . We conclude by direct verification, again by relying on Pari/Gp. □

## 9 From $M$ to $m$

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

**Lemma 9.1** (El Marraki). *We have*

$$|m(D)| \leq \frac{|M(D)|}{D} + \frac{1}{D} \int_1^D \frac{|M(t)| dt}{t} + \frac{\text{Log} D}{D}.$$

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  $ct/\text{Log} t$  would hence result in the useless bound  $m(D) \ll \text{Log} \text{Log} D$ .

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

$$(9.1) \quad m(D) = \frac{M(D)}{D} + \int_1^D \frac{M(t)dt}{t^2}$$

and

$$(9.2) \quad \int_1^D \left[ \frac{D}{t} \right] \frac{M(t)dt}{t} = \text{Log } D.$$

We deduce from the above that

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

The Lemma follows readily. □

*Proof.* We have, when  $D \geq D_0 = 1\,078\,853$ ,

$$\begin{aligned} |m(D)| &\leq \frac{0.0130 \text{Log } D - 0.118}{(\text{Log } D)^2} + \frac{1}{D} \int_{D_0}^D \frac{0.0130 \text{Log } t - 0.118}{(\text{Log } t)^2} dt \\ &\quad + \frac{1}{D} \int_1^{D_0} \frac{|M(t)|dt}{t} + \frac{\text{Log } D}{D}, \\ &\leq \frac{0.0130 \text{Log } D - 0.118}{(\text{Log } D)^2} + \frac{1}{D} \int_{D_0}^D \frac{0.0130 dt}{\text{Log } t} \\ &\quad - \frac{1}{D} \int_{D_0}^D \frac{0.118 dt}{(\text{Log } t)^2} + \frac{301 + \text{Log } D}{D}. \end{aligned}$$

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

$$\begin{aligned} |m(D)| &\leq \frac{0.0260 \text{Log } D - 0.118}{(\text{Log } D)^2} - \frac{0.105}{D} \int_{D_0}^D \frac{dt}{(\text{Log } t)^2} + \frac{-9795 + \text{Log } D}{D}, \\ &\leq \frac{0.0260 \text{Log } D - 0.118}{(\text{Log } D)^2} - \frac{1}{D} \int_{D_0}^D \frac{dt}{t} + \frac{-9795 + \text{Log } D}{D} \end{aligned}$$

This proves that  $|m(D)|(\text{Log } D)^2 \leq 0.0260 \text{Log } D - 0.118$  as soon as  $D \geq 1\,078\,853$ . We extend this bound by direct inspection. □

## References

- [1] H. Cohen and F. Dress. Estimations numériques du reste de la fonction sommatoire relative aux entiers sans facteur carré. *Prépublications mathématiques d'Orsay : Colloque de théorie analytique des nombres, Marseille*, pages 73–76, 1988.



- [2] H. Cohen, F. Dress, and M. El Marraki. Explicit estimates for summatory functions linked to the Möbius  $\mu$ -function. *Univ. Bordeaux 1*, Pré-publication(96-7), 1996.
- [3] H. Cohen, F. Dress, and M. El Marraki. Explicit estimates for summatory functions linked to the Möbius  $\mu$ -function. *Funct. Approx. Comment. Math.*, 37(part 1):51–63, 2007.
- [4] N. Costa Pereira. Elementary estimates for the Chebyshev function  $\psi(X)$  and for the Möbius function  $M(X)$ . *Acta Arith.*, 52:307–337, 1989.
- [5] F. Dress. Théorèmes d’oscillations et fonction de Möbius. *Sémin. Théor. Nombres, Univ. Bordeaux I*, Exp. No 33:33pp, 1983/84. <http://resolver.sub.uni-goettingen.de/purl?GDZPPN002545454>.
- [6] F. Dress. Fonction sommatoire de la fonction de Möbius 1. Majorations expérimentales. *Exp. Math.*, 2(2), 1993.
- [7] F. Dress and M. El Marraki. Fonction sommatoire de la fonction de Möbius 2. Majorations asymptotiques élémentaires. *Exp. Math.*, 2(2), 1993.
- [8] P. Dusart. *Autour de la fonction qui compte le nombre de nombres premiers*. PhD thesis, Limoges, [http\string://www.unilim.fr/laco/theses/1998/T1998\\_01.pdf](http://www.unilim.fr/laco/theses/1998/T1998_01.pdf), 1998. 173 pp.
- [9] P. Dusart. Inégalités explicites pour  $\psi(x)$ ,  $\theta(x)$ ,  $\pi(x)$  et les nombres premiers. *C. R. Math. Acad. Sci., Soc. R. Can.*, 21(2):53–59, 1999.
- [10] M. El Marraki. Fonction sommatoire de la fonction  $\mu$  de Möbius, majorations asymptotiques effectives fortes. *J. Théor. Nombres Bordx.*, 7(2), 1995.
- [11] M. El Marraki. Majorations de la fonction sommatoire de la fonction  $\frac{\mu(n)}{n}$ . *Univ. Bordeaux 1*, Pré-publication(96-8), 1996.
- [12] A. Kienast. Über die Äquivalenz zweier Ergebnisse der analytischen Zahlentheorie. *Mathematische Annalen*, 95:427–445, 1926. 10.1007/BF01206619.
- [13] Edmund Landau. Über einige neuere Grenzwertsätze. *Rendiconti del Circolo Matematico di Palermo (1884 - 1940)*, 34:121–131, 1912. 10.1007/BF03015010.
- [14] B.V. Levin and A.S. Fainleib. Application of some integral equations to problems of number theory. *Russian Math. Surveys*, 22:119–204, 1967.

- [15] The PARI Group, Bordeaux. *PARI/GP, version 2.5.2*, 2011. <http://pari.math.u-bordeaux.fr/>.
- [16] O. Ramaré. Explicit estimates: from  $\Lambda(n)$  to  $\Lambda(n)/n$ . *Submitted to Math. Comp.*, 2012.
- [17] O. Ramaré and R. Rumely. Primes in arithmetic progressions. *Math. Comp.*, 65:397–425, 1996.
- [18] J.B. Rosser. Explicit bounds for some functions of prime numbers. *American Journal of Math.*, 63:211–232, 1941.
- [19] J.B. Rosser and L. Schoenfeld. Approximate formulas for some functions of prime numbers. *Illinois J. Math.*, 6:64–94, 1962.
- [20] J.B. Rosser and L. Schoenfeld. Sharper bounds for the Chebyshev Functions  $\vartheta(x)$  and  $\psi(x)$ . *Math. Comp.*, 29(129):243–269, 1975.
- [21] L. Schoenfeld. An improved estimate for the summatory function of the Möbius function. *Acta Arith.*, 15:223–233, 1969.
- [22] L. Schoenfeld. Sharper bounds for the Chebyshev Functions  $\vartheta(x)$  and  $\psi(x)$  ii. *Math. Comp.*, 30(134):337–360, 1976.
- [23] Wen-Bin Zhang. A generalization of Halász’s theorem to Beurling’s generalized integers and its application. *Illinois J. Math.*, 31(4):645–664, 1987.