

*EXPLICIT AVERAGES OF NON-NEGATIVE MULTIPLICATIVE
FUNCTIONS: GOING BEYOND THE MAIN TERM*

BY

O. RAMARÉ (Marseille) and P. AKHILESH (Allahabad)

Abstract. We produce an explicit formula for averages of the type $\sum_{d \leq D} (g \star \mathbb{1})(d)/d$, where \star is the Dirichlet convolution and g a function that vanishes at infinity (more precise conditions are needed, a typical example of an acceptable function is $g(m) = \mu(m)/m$). This formula enables one to exploit the changes of sign of $g(m)$. We use this formula for the classical family of sieve-related functions $G_q(D) = \sum_{d \leq D, (d,q)=1} \mu^2(d)/\varphi(d)$ for an integer parameter q , improving noticeably on earlier results. The remainder of the paper deals with the special case $q = 1$ to show how to practically exploit the changes of sign of the Möbius function. It is proven in particular that $|G_1(D) - \log D - c_0| \leq 4/\sqrt{D}$ and $|G_1(D) - \log D - c_0| \leq 18.4/(\sqrt{D} \log D)$ when $D > 1$, for a suitable constant c_0 .

1. Introduction. Evaluating the averages of multiplicative functions is a classical and important problem. It has been approached in several manners: in [41], [20], [12] from the “elementary” side; in [16], [40], [4], [15], [18] from the Tauberian side; and in [11], [9], [23] following the Halász method. See also [1], [2] for better analytical estimates valid for a narrower class of arithmetical functions.

The theory went also into special evaluations that were then systematized, as in [38] or [17]. This first series of works essentially handles the “main term” and treats what remains as an error term. Subsequent investigations, as in [42] and [19], went on to study the error term more precisely, in some very special cases, and a main term of size \sqrt{x} was indeed extracted. The two typical cases are the characteristic function of the square-free integers whose Dirichlet series is $\zeta(s)/\zeta(2s)$, and the multiplicative function whose Dirichlet series is $\zeta(s)\zeta(2s)$. The articles [38] and [17] present these series in the form $\zeta(s)H(s)$ where the only hypothesis on H is that it is absolutely convergent for $\Re s > 1/2$. The articles [42] and [19] use the fact that $\zeta(2s)$ and $1/\zeta(2s)$ can be controlled beyond $\Re s = 1/2$. In the first case, a main term of size \sqrt{x} is indeed extracted from the error term, while in the

2010 *Mathematics Subject Classification*: Primary 11N37, Secondary 11N36, 11N56.

Key words and phrases: multiplicative functions, explicit evaluations.

Received 21 November 2013; revised 10 September 2015.

Published online 31 January 2017.

second one, the error term is simply improved, corresponding to a pole of residue 0.

The present work can be seen as an elementary and explicit counterpart of such studies. The task of getting explicit estimates on particular questions has been undertaken by several authors as in [5], and systematically in [26, Lemma 3.2]. Concerning the analytical approach, the amount of information available is sparse; one can find a very explicit truncated Perron formula in [27, Theorem 7.1], but the bounds for the usual Dirichlet series (like $1/\zeta(s)$) are still too weak or simply missing. We develop here a strategy that achieves two things: Theorem 1.4 gets the most of the convolution method, and Theorem 1.3 goes *beyond* in some problems and incorporates cancellations from the Möbius function.

We use the following family of functions as our primary challenge:

$$(1.1) \quad G_q(D) = \sum_{\substack{d \leq D \\ (d,q)=1}} \frac{\mu^2(d)}{\varphi(d)}.$$

This family has been quite extensively studied (see for instance [38], [33], [26, Lemma 3.4]) and used (see for instance [36, Lemma 4.6] and [14, (7.12), (B.5)]). It occurs in sieving, when bounding from above the number of primes in a given arithmetic progression (the so-called Brun–Titchmarsh Theorem) via the Selberg sieve [13] or via the Montgomery sieve [22], or, more generally, when considering the resulting sequence as an enveloping sieve, as initially done in [26], [32] (see also [10]).

The question we address here is to find as good an explicit error term as possible. We first give a fairly simple proof of the following theorem. As usual we write $g = \mathcal{O}^*(f)$ to mean that $|f| \leq g$.

THEOREM 1.1. *We have*

$$G_q(D) = \frac{\varphi(q)}{q}(\log D + c(q)) + \mathcal{O}^*(5.9j(q)/\sqrt{D})$$

where

$$j(q) = \prod_{\substack{p|q \\ p \neq 2}} \frac{p^{3/2} + p - \sqrt{p} - 1}{p^{3/2} - \sqrt{p} + 1} \prod_{2|q} \frac{21}{25}$$

and (cf. [34, (2.11)])

$$c(q) = c_0 + \sum_{p|q} \frac{\log p}{p}, \quad c_0 = \gamma + \sum_{p \geq 2} \frac{\log p}{p(p-1)} = 1.332\,582\,275\dots$$

For the moduli in the array below, the value $5.9j(q)$ may be replaced by $j^*(q)$ given by

q	6	30	210	2310	30 030	510 510
$j^*(q)$	5.2	6.0	7.1	8.4	9.9	11.6

The above error term is for instance important in [14, (7.29)]. The term “ $\prod_{2|q} \frac{21}{25}$ ” means that the factor $\frac{21}{25}$ is present only when q is even. Asymptotic comparison with [26, Lemma 3.4] is not easy: the D -dependence in the error term is $7.284/D^{1/3}$ there while it is now $5.9/\sqrt{D}$, but the q -dependence is a multiplicative function (f_1 there and j here) that behaves at primes there like $1 + p^{-2/3} + \mathcal{O}(1/p)$ and here like $1 + p^{-1/2} + \mathcal{O}(1/p)$. This implies that $j(q)/f_1(q)$ can be arbitrarily large. A quick computation using the fact that $j(p)/f_1(p)$ is non-increasing shows however that the present error term is better when q has no more than one million prime factors and D is at least 250. Hence, for all practical purposes, the approximation given by Theorem 1.1 is always better.

The error term in Theorem 1.1 is limited to $\mathcal{O}(1/\sqrt{D})$. Theorem 1.3 below enables one to get $\mathcal{O}(1/\sqrt{D \log D})$ in a fully *explicit* manner. The details are however difficult when considering a general modulus q , so we restrict our attention to $q = 1$.

THEOREM 1.2. *When $D \geq 1$, we have $G_1(D) = \log D + c_0 + \mathcal{O}^*(3.95/\sqrt{D})$. When $D > 1$, we have $G_1(D) = \log D + c_0 + \mathcal{O}^*(18.4/\sqrt{D \log D})$.*

If one has access to bounds for the summatory function of the Möbius function, better than the ones given by Lemma 7.14, or if one can rely on more extensive computations concerning small values and further uses Lemma 8.4, then the saving will be automatically transferred on the error term under examination. For instance, if the computations in Lemma 8.4 were pushed till 10^{15} , and expecting a similar output, we could get an error term $\mathcal{O}^*(2.94/\sqrt{D})$. Surely these parts will be improved upon. Furthermore, other sections of this proof can also be enhanced.

Our main tool is the following general and fully explicit theorem.

THEOREM 1.3. *Let $(g(m))_{m \geq 1}$ be a sequence of complex numbers such that both series $\sum_{m \geq 1} g(m)/m$ and $\sum_{m \geq 1} g(m)(\log m)/m$ converge. Define $G^\sharp(x) = \sum_{m > x} g(m)/m$ and assume that $\int_1^\infty |G^\sharp(t)| dt/t$ converges. Let $A_0 \geq 1$ be a real parameter. Then*

$$\sum_{n \leq D} \frac{(g \star \mathbb{1})(n)}{n} = \sum_{m \geq 1} \frac{g(m)}{m} \left(\log \frac{D}{m} + \gamma \right) + \int_{D/A_0}^\infty G^\sharp(t) \frac{dt}{t} + \mathcal{O}^*(\mathfrak{R})$$

where

$$\begin{aligned} \mathfrak{R} = & \left| \sum_{1 \leq a \leq A_0} \frac{1}{a} G^\sharp\left(\frac{D}{a}\right) + G^\sharp\left(\frac{D}{A_0}\right) \left(\log \frac{A_0}{[A_0]} - R([A_0]) \right) \right| \\ & + \frac{6/11}{D} \sum_{m \leq D/A_0} |g(m)| \end{aligned}$$

where $[A_0]$ is the integer part of A_0 , while the remainder R is defined in Lemma 2.1.

By (2.1), when $[A_0]$ is large, $R([A_0])$ is equivalent to $1/(2[A_0])$, and thus, on taking $A_0 = [A_0] + 1/2$, we conclude that $\log(A_0/[A_0]) - R([A_0])$ decreases at least like $1/A_0^2$. In our case study, we have $G^\sharp(D/A_0)(\log(A_0/[A_0]) - R([A_0])) = \mathcal{O}(D^{-1/2}A_0^{-3/2})$ while the bound for $\sum_{1 \leq a \leq A_0} G^\sharp(D/a)/a$ grows with A_0 and the last term of \mathfrak{R} is bounded by $\mathcal{O}(D^{-1/2}A_0^{-1/2})$.

Theorem 1.4 below corresponds to $A_0 \leq 1$. Both theorems are refinements of [26, Lemma 3.2] obtained via the classical *convolution method* (see [3] for a pedagogical presentation) to which we add a finite version of the Dirichlet hyperbola principle as in [28, (19)]. In particular, only by following the convolution method, we were for a very long time unable to reach an error term of size $\mathcal{O}_q(1/\sqrt{D})$ in Theorem 1.1 (see the beginning of Subsection 8.2 for an explanation regarding this bound); we could only get $\mathcal{O}_q((\log D)/\sqrt{D})$, which is why we settled for the asymptotically weaker but numerically better $\mathcal{O}_q(1/D^{1/3})$.

Theorem 1.3 enables us to go even beyond $\mathcal{O}_q(1/\sqrt{D})$. The additional ingredient comes from a good upper bound for $|G^\sharp(x)|$ that is often accessible and is introduced via the Covering Remainder Lemma (Lemma 3.1). In the case study of Theorem 1.1, Theorem 1.3 leads to

$$G_q(D) = \frac{\varphi(q)}{q}(\log D + c(q)) + \mathcal{O}_q(\exp(-c_1\sqrt{\log D})/\sqrt{D})$$

for some constant $c_1 > 0$.

There are several occasions when selecting $A_0 = 1$ is in fact better: when the information on G^\sharp is scarce, when g is rather large (see the example $g(d) = \mu(d)/d$ below), or when g is non-negative. One can then modify the proof so as to get a somewhat better result.

THEOREM 1.4. *Let $(g(m))_{m \geq 1}$ be a sequence of complex numbers such that both $\sum_{m \geq 1} g(m)/m$ and $\sum_{m \geq 1} g(m)(\log m)/m$ converge. Define $G^\sharp(x) = \sum_{m > x} g(m)/m$ and assume that $\int_1^\infty |G^\sharp(t)| dt/t$ converges. Then, for any real $\eta \geq 1$,*

$$\begin{aligned} \sum_{n \leq D} \frac{(g \star \mathbb{1})(n)}{n} &= \sum_{m \geq 1} \frac{g(m)}{m} \left(\log \frac{D}{m} + \gamma \right) + \int_{\eta D}^\infty G^\sharp(t) \frac{dt}{t} \\ &\quad - (\gamma - \log \eta)G^\sharp(\eta D) + \mathcal{O}^* \left(\frac{\gamma}{D} \sum_{m \leq \eta D} |g(m)| \right). \end{aligned}$$

When $\eta = 1$, the constant γ in the error term can be reduced to $6/11$.

This theorem is tailored for cases like $g(m) = \mu(m)/m$. The above yields an error term of size $\mathcal{O}((\log D)/D)$. Following [39] (see also [25]), one can

improve slightly on this bound, but [21] (see also the intriguing [17]) implies that one cannot even replace the $\log D$ in the error term by a constant. The introduction of the A_0 parameter in Theorem 1.1 may be inefficient, while the η parameter of Theorem 1.4 may give some improvements.

The major difficulty one faces when putting this program into practice, for instance when proving Theorem 1.1 or 1.2, are the coprimality conditions that come naturally into play, even for Theorem 1.2, though no coprimality condition was introduced at the beginning. This is apparent for instance in (7.7) which we repeat here:

$$\frac{\mu(k)}{\varphi(k)} = \sum_{dab^2=k} \frac{\mu(d)}{d} \frac{\mu^2(abd)}{a\varphi(a)b^2\varphi(b)}.$$

The most inner and main variable is d , but the factor $\mu^2(abd)$ forces d to be coprime to ab . This question has already been addressed in [30], [29] and [31] in some special instances, and these results are used here. In general, though, we start from a non-negative function, and we aim at studying the second term of the asymptotic expansion. This implies that we may have to deal with an oscillating function, as is the case here. If this second term corresponds to a pole, like for the coefficients of $\zeta(s)\zeta(2s)$, cancellations do not occur; however, our series is closer to $\zeta(s)/\zeta(2s)$. Most of our effort concerns the function $\mu(k)/(k\varphi(k))$, and for instance two of our crucial (and novel) bounds are given in Lemmas 5.6 and 7.18. We state here a consequence of them for the reader to appreciate their content:

THEOREM 1.5. *For any real $K \geq 14$ and any modulus q^* ,*

$$\left| \sum_{\substack{k>K \\ (k,q^*)=1}} \frac{\mu(k)}{\varphi(k)k} \right| \leq \frac{0.37}{K}.$$

When further $K > q^$, we have*

$$\left| \sum_{\substack{k>K \\ (k,q^*)=1}} \frac{\mu(k)}{k\varphi(k)} \right| \leq \frac{2.33 \prod_{p|q^*} (1 + p^{-1})}{K \log(K/q^*)} + \frac{5}{K} \sqrt{\frac{q^*}{K}} + \frac{6.6}{K} \sqrt{\frac{q^*}{K}} \mathbb{1}_{K \geq 1970q^*}.$$

We devised some special analysis for low values of the parameters, but we face a major hurdle: the low values of K/q^* do not correspond to a bounded set of parameters (the fact that K/q^* is bounded does not delimit a finite set of parameters in q^* nor a bounded one for K). Overcoming this hurdle costs us quite a lot, hence the constant 2.33 above, while for comparison, [28] shows for instance that when $D \geq 50\,000$, we have

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

We also rely on precise local estimates for some particular functions. These estimates have an independent interest, but apart from some local improvements, they are essentially routine.

About the computations. We have to perform several computations by computer. We have used commonly Pari/GP [24] with the GMP [7] multiprecision library and checked the results with Sage [35] with interval arithmetic, thus using GMP, MPFR [6] and MPFI. Both kinds of scripts have been run on independent machines; the GP-scripts have been written and run by one author and the Sage-scripts have been written and run by the other. Whenever required, we have speeded up the GP script by using GP2C, a software that converts the GP script into a C program that can then be compiled. Sage being much slower than GP and this last process being absent, the last script used for Lemma 8.4 has been run with Perl and the MPFI library for interval arithmetic. It ran for three months (while the GP counterpart ran for ten days).

Let us describe in more detail how to compile a GP script, say contained in the file `Check.gp`. We run the command `gp2c -g Check.gp > Check.gp.c` with the “-g” flag to enable garbage collection. At the top of `Check.gp.c`, the reader will find the compilation command starting with `/*-*- compile-command:`. This command has to be run to get the compiled form. From then onward, a simple way to proceed is to start GP with a large enough table of primes, say with `gp -p10000000000`. At the top of `Check.gp.c`, the reader will find several lines starting with `GP; install`. The `install` commands have to be entered in GP, the first one being

```
install("init_Check","v","init_Check","./Check.gp.so").
```

Then one has to run the command `allocatemem(7500000000)` to extend the stack to a bit more than 7 Giga, `init_Check()` to take care of the global variables, and one is finally ready to use the functions in the script `Check.gp`.

Concerning the Perl scripts, the interesting part of the header reads:

```
use Math::MPFI qw(:mpfi);
use Math::Prime::Util qw(primes);
use Math::Factor::XS qw(factors prime_factors matches);
[...]
Rmpfi_set_default_prec(100); # Set default precision to 100 bits
```

2. Proof of Theorem 1.4 and beginning of proof of Theorem 1.3

LEMMA 2.1. *For every real $X \geq 1$,*

$$\sum_{n \leq X} \frac{1}{n} = \log X + \gamma + \mathcal{O}^*\left(\frac{6}{11X}\right).$$

For every real $X > 0$,

$$\sum_{n \leq X} \frac{1}{n} = \log X + \gamma + \mathcal{O}^*\left(\frac{\gamma}{X}\right).$$

In general, we define $R(X) = \sum_{n \leq X} 1/n - \log X - \gamma$.

The first inequality with $7/12 = 0.583\dots$ appears in [5, (3.1)]. The constant $6/11 = 0.545\dots$ is nearly optimal: the optimum is $2(\log 2 + \gamma - 1) = 0.540\dots$

Proof of Lemma 2.1.

- We inspect $x \mapsto |\log x + \gamma - 1|x$ for $x \in [1, 2)$ and find its maximum to be $2(\log 2 + \gamma - 1) = 0.540725\dots$
- We inspect $x \mapsto |\log x + \gamma - 3/2|x$ for $x \in [2, 3)$ and find its maximum to be $3(\log 3 + \gamma - 3/2) = 0.527\dots$
- We inspect $x \mapsto |\log x + \gamma - 11/6|x$ for $x \in [3, 4)$ and find its maximum to be $3(\log 3 + \gamma - 11/6) = 0.520\dots$

We proceed in a similar fashion for all intervals up to 11 and obtain the maxima below, all reached (almost!) at the endpoint of the interval:

[4, 5)	[5, 6)	[6, 7)	[7, 8)	[8, 9)	[9, 10)	[10, 11)
0.516...	0.513...	0.511...	0.510...	0.508...	0.508...	0.507...

For larger values, the Euler–MacLaurin formula gives (see for instance [37, Théorème 5])

$$(2.1) \quad \sum_{n \leq N} \frac{1}{n} = \log N + \gamma + \frac{1}{2N} - \frac{1}{12N^2} + \mathcal{O}^*\left(\frac{1}{60N^4}\right)$$

for any positive integer N . We have $\frac{N+1}{N}(\frac{1}{2} - \frac{1}{12N} + \frac{1}{60N^3}) \leq 0.538$ when $N \geq 11$. The lemma follows readily. ■

Proof of Theorem 1.4. We write

$$\sum_{n \leq D} \frac{(g \star \mathbb{1})(n)}{n} = \sum_{m \geq 1} \frac{g(m)}{m} \sum_{n \leq D/m} \frac{1}{n}.$$

The summation in m can be restricted to $m \leq D$, and more generally to $m \leq \eta D$ for some $\eta \geq 1$. This together with the definition of $R(x)$ after Lemma 2.1 gives

$$\sum_{n \leq D} \frac{(g \star \mathbb{1})(n)}{n} = \sum_{m \leq \eta D} \frac{g(m)}{m} \left(\log \frac{D}{m} + \gamma\right) + \sum_{m \leq \eta D} \frac{g(m)}{m} R\left(\frac{D}{m}\right).$$

The next task is to complete the summation over m , for which we add and subtract the quantity $\sum_{m > \eta D} \frac{g(m)}{m} (\log \frac{D}{m} + \gamma)$ and we rewrite the subtracted

part by using G^\sharp :

$$\begin{aligned}
 - \sum_{m > \eta D} \frac{g(m)}{m} \left(\log \frac{D}{m} + \gamma \right) &= \sum_{m > \eta D} \frac{g(m)}{m} \left(\log \frac{m}{\eta D} - \gamma + \log \eta \right) \\
 &= \sum_{m > \eta D} \frac{g(m)}{m} \int_{\eta D}^m \frac{dt}{t} - (\gamma - \log \eta) G^\sharp(\eta D) \\
 &= \int_{\eta D}^\infty \sum_{m > t} \frac{g(m)}{m} \frac{dt}{t} - (\gamma - \log \eta) G^\sharp(\eta D) \\
 &= \int_{\eta D}^\infty G^\sharp(t) \frac{dt}{t} - (\gamma - \log \eta) G^\sharp(\eta D).
 \end{aligned}$$

This leads to the following key expression:

$$\begin{aligned}
 (2.2) \quad \sum_{n \leq D} \frac{(g \star \mathbb{1})(n)}{n} &= \sum_{m \geq 1} \frac{g(m)}{m} \left(\log \frac{D}{m} + \gamma \right) + \int_{\eta D}^\infty G^\sharp(t) \frac{dt}{t} \\
 &\quad - (\gamma - \log \eta) G^\sharp(\eta D) + \sum_{m \leq \eta D} \frac{g(m)}{m} R\left(\frac{D}{m}\right).
 \end{aligned}$$

On appealing to Lemma 2.1, one gets Theorem 1.4. ■

The main point above has been to write (when $\eta = 1$)

$$\sum_{d > D} \frac{g(d)}{d} \log \frac{d}{D} = \int_D^\infty \sum_{d > t} \frac{g(d)}{d} \frac{dt}{t},$$

while the more common treatment is comparable to using $\log \frac{d}{D} = \log d - \log D$ and treating the resulting two sums independently.

Second proof of (2.2). Equation (2.2) being a linear identity in g , we can use algebraic means to prove it. The following decomposition holds:

$$g = \sum_{k \geq 1} g(k) \delta_{.=k}.$$

It is thus enough to prove our identity for $\delta_{.=k}$. In this case, we readily check that

$$G^\sharp(t) = \begin{cases} 1/k & \text{when } t \leq k, \\ 0 & \text{otherwise.} \end{cases}$$

The function $g \star \mathbb{1}$ is the characteristic function of those integers that are divisible by k . The proof of (2.2) splits into two cases:

- (1) When $k \leq \eta D$, the LHS of (2.2) is $(1/k) \sum_{n \leq D/k} 1/n$, while the RHS is $\frac{1}{k} (\log \frac{D}{k} + \gamma) + 0 - (\gamma - \log \eta) \times 0 + \frac{1}{k} R(D/k)$. Both sides agree by the definition of $R(D/k)$.
- (2) When $\eta D < k$, the LHS of (2.2) vanishes, while the RHS is

$$\frac{1}{k} \left(\log \frac{D}{k} + \gamma \right) + \frac{1}{k} \int_{\eta D}^k \frac{dt}{t} - (\gamma - \log \eta) \frac{1}{k},$$

which is also verified to vanish.

The proof is complete. ■

3. The Covering Remainder Lemma and proof of Theorem 1.3

LEMMA 3.1 (Covering Remainder Lemma). *Let $(g(m))_{m \geq 1}$ be a sequence of complex numbers such that $\sum_{m \geq 1} g(m)/m$ converges. Define $G^\#(x) = \sum_{m > x} g(m)/m$. The function R is given after Lemma 2.1. Let $A_0 \geq 1$ be a real number and let $[A_0]$ denote its integer part. We have*

$$\begin{aligned} \sum_{X/A_0 < m \leq X} \frac{g(m)}{m} R\left(\frac{X}{m}\right) &= \gamma G^\#(X) + \int_1^{A_0} G^\#\left(\frac{X}{t}\right) \frac{dt}{t} - \sum_{1 \leq a \leq A_0} \frac{1}{a} G^\#\left(\frac{X}{a}\right) \\ &\quad - G^\#\left(\frac{X}{A_0}\right) \left(\log \frac{A_0}{[A_0]} - R([A_0]) \right), \end{aligned}$$

When g is non-negative, this lemma does not lead to any further saving. But otherwise, the saving is interesting. In case $g(m)/m = \mu(m)/m^2$, the estimate $R(x) \ll 1$ leads to a bound of at least $\mathcal{O}(1/X)$ for the LHS, while this lemma gives the bound $\mathcal{O}(\exp(-c\sqrt{\log X})/X)$ for some positive constant c . This has the same effect as [21, Lemma 1]: shortening the sum that was responsible for the size of the remainder term.

Proof of Lemma 3.1. We set $B = [A_0]$ to ease notation. When a is a positive integer, $b \in [a, a + 1]$, and $m \in (X/b, X/a]$, we have

$$\log \frac{X}{m} + \gamma + R(X/m) = \sum_{n \leq X/m} \frac{1}{n} = \sum_{n \leq a} \frac{1}{n} = \log a + \gamma + R(a),$$

from which we infer that $R(X/m) = R(a) - \int_a^{X/m} dt/t$. This implies that

$$\sum_{X/b < m \leq X/a} \frac{g(m)}{m} R(X/m) = \sum_{X/b < m \leq X/a} \frac{g(m)}{m} R(a) - \int_a^b \sum_{X/b < m \leq X/t} \frac{g(m)}{m} \frac{dt}{t}.$$

We sum this over $a \in \{1, \dots, B\}$ with the choice $b = \min(a + 1, A_0)$. On

using the notation G^\sharp , we get

$$\sum_{X/A_0 < m \leq X} \frac{g(m)}{m} R(X/m) = \sum_{a \leq B} \left(G^\sharp \left(\frac{X}{\min(a+1, A_0)} \right) - G^\sharp \left(\frac{X}{a} \right) \right) R(a) + \sum_{a \leq B} \int_a^{\min(a+1, A_0)} \left(G^\sharp \left(\frac{X}{t} \right) - G^\sharp \left(\frac{X}{\min(a+1, A_0)} \right) \right) \frac{dt}{t}.$$

Some shuffling is called for. Here is the first step:

$$\sum_{X/A_0 < m \leq X} \frac{g(m)}{m} R(X/m) = \sum_{2 \leq a \leq B+1} G^\sharp \left(\frac{X}{\min(a, A_0)} \right) R(a-1) - \sum_{a \leq B} G^\sharp \left(\frac{X}{a} \right) R(a) + \int_1^{A_0} G^\sharp \left(\frac{X}{t} \right) \frac{dt}{t} - \sum_{a \leq B} G^\sharp \left(\frac{X}{\min(A_0, a+1)} \right) \log \frac{\min(A_0, a+1)}{a},$$

which we rewrite in the form (we set $R(0) = 0$)

$$\sum_{X/A_0 < m \leq X} \frac{g(m)}{m} R(X/m) = \sum_{1 \leq a \leq B} G^\sharp \left(\frac{X}{a} \right) (R(a-1) - R(a)) + G^\sharp \left(\frac{X}{A_0} \right) R(B) + \int_1^{A_0} G^\sharp \left(\frac{X}{t} \right) \frac{dt}{t} - \sum_{a \leq B} G^\sharp \left(\frac{X}{\min(A_0, a+1)} \right) \log \frac{\min(A_0, a+1)}{a}.$$

Here is the second step:

$$\sum_{X/A_0 < m \leq X} \frac{g(m)}{m} R(X/m) = -G^\sharp(X)R(1) + \int_1^{A_0} G^\sharp \left(\frac{X}{t} \right) \frac{dt}{t} + \sum_{2 \leq a \leq B} G^\sharp \left(\frac{X}{a} \right) \left(R(a-1) - R(a) - \log \frac{a}{a-1} \right) + G^\sharp \left(\frac{X}{A_0} \right) \left(R(B) - \log \frac{A_0}{B} \right).$$

It is then obvious how to establish the claimed formula since $R(1) = 1 - \gamma$ and $-1/a = R(a-1) + \log(a-1) - (R(a) + \log(a))$ when $a \geq 2$. ■

Proof of Theorem 1.3. Theorem 1.3 is a simple application of the lemma. We start from the key expression (2.2) of the previous section, select $\eta = 1$, and treat the last term via the Covering Remainder Lemma. ■

4. A convolution identity. The convolution method consists in comparing an unknown arithmetical function to a known one; here $\mathbb{1}_{(d,q)=1} \frac{\mu^2(d)d}{\varphi(d)}$ is compared to $\mathbb{1}$.

LEMMA 4.1.

$$\sum_{\substack{k^2 \ell r | d \\ r|q \\ (k,\ell)=1 \\ (k,q)=1}} \frac{\mu(rk)\mu^2(\ell)k}{\varphi(k)\varphi(\ell)} = \mathbb{1}_{(d,q)=1} \frac{\mu^2(d)d}{\varphi(d)}.$$

Proof. Let $g(d)$ be the multiplicative function on the right-hand side. When $h \geq 2$, we check that $g(p^h) = g(p^2)$. It is then easy to verify the claimed identity. ■

For the rest of this paper, we define

$$(4.1) \quad r_2(X; q) = \sum_{\substack{k^2 \ell r > X \\ r|q \\ (k,\ell)=1 \\ (k,q)=1}} \frac{\mu(rk)\mu^2(\ell)}{rk\varphi(k)\ell\varphi(\ell)},$$

$$(4.2) \quad r_1(X) = \sum_{\substack{k^2 \ell r \leq X \\ r|q \\ (k,\ell)=1 \\ (k,q)=1}} \frac{\mu^2(rk\ell)k}{\varphi(k)\varphi(\ell)}.$$

We now proceed to evaluate these quantities.

5. Cancellation in $\sum_{k > K} \mu(k)/(\varphi(k)k)$. Elementary methods. Evaluating $\sum_{k > K, (k,q^*)=1} \frac{\mu(k)}{\varphi(k)k}$ is a crucial step in bounding $r_2(X; q)$. This section is devoted to rather elementary/algorithmical results, the high point being Lemma 5.7. The proof of Theorem 1.1 requires in fact less work, but we prepare for the proof of Theorem 1.2 and we have gathered in a single place results of the same flavor.

LEMMA 5.1 ([26, Lemma 3.4]). *For any real $K > 0$,*

$$\sum_{k \leq K} \frac{\mu^2(k)}{\varphi(k)} = \log K + c_0 + \mathcal{O}^*(7.3/K^{1/3}).$$

Note that the left-hand side is also $G_1(K)$.

LEMMA 5.2. *For $K \geq 50$,*

$$0.946 \leq (1/K) \sum_{k \leq K} \frac{\mu^2(k)k}{\varphi(k)} \leq 1.066.$$

After we prove Theorem 1.1, we will be able to replace 0.946 by 0.989. Further assuming $K \geq 10^6$, we can replace 0.946 by 0.996 and 1.066 by 1.004.

Proof. We simply write

$$\begin{aligned} \sum_{k \leq K} \frac{\mu^2(k)k}{\varphi(k)} &= K \sum_{k \leq K} \frac{\mu^2(k)}{\varphi(k)} - \int_1^K \sum_{k \leq t} \frac{\mu^2(k)}{\varphi(k)} dt \\ &= K(\log K + c_0 + \mathcal{O}^*(7.3/K^{1/3})) - \int_1^K (\log t + c_0 + \mathcal{O}^*(7.3/t^{1/3})) dt \\ &= K + c_0 + \mathcal{O}^*\left(\frac{5}{2}7.3K^{2/3}\right). \end{aligned}$$

We ran a GP-script [24] and a Sage script that checked that

$$(5.1) \quad 0.98999 \leq \frac{1}{K} \sum_{k \leq K} \frac{\mu^2(k)k}{\varphi(k)} \leq 1.066 \quad (100 \leq K \leq 10^8).$$

Once we prove Theorem 1.1, we will be able to replace the error term $\frac{5}{2}7.3K^{2/3}$ by $3 \times 10.8\sqrt{K}$. ■

We will later need the following estimate, but it is clearer to state it here.

LEMMA 5.3. *Assuming Theorem 1.1 has been proved, for $K \geq 50$ we have*

$$0.972 \leq \frac{2}{K} \sum_{\substack{k \leq K \\ (k,2)=1}} \frac{\mu^2(k)k}{\varphi(k)} \leq 1.066.$$

When $K \geq 10^6$, we can replace the couple (0.972, 1.066) by (0.995, 1.005).

LEMMA 5.4. *Assuming Theorem 1.1 has been proved, for $K \geq 10^6$ and $b \in \{210, 2310, 30\,030, 510\,510\}$ we have*

$$\sum_{\substack{k \leq K \\ (k,b)=1}} \frac{\mu^2(k)k}{\varphi(k)} = \frac{\varphi(b)K}{b} (1 + \mathcal{O}^*(0.0001)).$$

Of course, the following factorisations hold: $210 = 2 \cdot 3 \cdot 5 \cdot 7$, $2310 = 11 \cdot 210$, $30\,030 = 13 \cdot 2310$ and $510\,510 = 17 \cdot 30\,030$.

Proof of Lemmas 5.3 and 5.4. We proceed as in the proof of Lemma 5.2, though we now have Theorem 1.1 at our disposal. With $c_2(b) = j^*(b)b/\varphi(b)$ we write

$$\begin{aligned} \sum_{\substack{k \leq K \\ (k,b)=1}} \frac{\mu^2(k)k}{\varphi(k)} &= K \sum_{\substack{k \leq K \\ (k,b)=1}} \frac{\mu^2(k)}{\varphi(k)} - \int_1^K \sum_{\substack{k \leq t \\ (k,b)=1}} \frac{\mu^2(k)}{\varphi(k)} dt \\ &= K \frac{\varphi(b)}{b} (\log K + c(b) + \mathcal{O}^*(c_2(b)/K^{1/2})) \\ &\quad - \int_1^K \frac{\varphi(b)}{b} (\log t + c(b) + \mathcal{O}^*(c_2(b)/t^{1/2})) dt \\ &= \frac{\varphi(b)}{b} (K + c(b)) + \mathcal{O}^*(3 \cdot j^*(b)K^{1/2}). \end{aligned}$$

We ran a GP-script [24] that checked the claimed inequality for $K \leq 10^{10}$ (as a matter of fact, it is enough to check it until $K = 6 \cdot 10^9$ when $q = 210$, and for all the moduli considered until $K = 9.6 \cdot 10^9$). We have

$$\sum_{\substack{k \leq K \\ (k,b)=1}} \frac{\mu^2(k)k}{\varphi(k)} = \frac{\varphi(b)K}{b} (1 + \mathcal{O}^*(0.0001)) \quad (10^6 \leq K \leq 10^{10}).$$

When $b = 2$, we extend the result to $K \geq 50$ by direct computations. ■

LEMMA 5.5. *Assuming Theorem 1.1 has been proved, for $K \geq K(b)$ and $b \in \{210, 2310, 30030\}$ we have*

$$1 - \epsilon_b^- \leq \frac{Kb}{\varphi(b)} \sum_{\substack{k > K \\ (k,b)=1}} \frac{\mu^2(k)}{k\varphi(k)} \leq 1 + \epsilon_b^+$$

where one can choose:

$K(210)$	$K(2310)$	$K(30030)$	$K(2310)$	$K(30030)$
9523	865	66	3500	5
ϵ_{210}^+	ϵ_{2310}^+	ϵ_{30030}^+	ϵ_{2310}^+	ϵ_{30030}^+
0.00169	0.0156	0.0911	0.0025	0.5344
ϵ_{210}^-	ϵ_{2310}^-	ϵ_{30030}^-	ϵ_{2310}^-	ϵ_{30030}^-
0.0265	0.0290	0.1073	0.0044	0.5488

Moreover

$$\sum_{\substack{k > K \\ (k,210)=1}} \frac{\mu^2(k)}{k\varphi(k)} \leq \begin{cases} 0.3/K & \text{when } K \geq 14, \\ 0.24/K & \text{when } K \geq 190, \end{cases}$$

and

$$\sum_{\substack{k > K \\ (k,30)=1}} \frac{\mu^2(k)}{k\varphi(k)} \leq \begin{cases} 0.35/K & \text{when } K \geq 14, \\ 0.276/K & \text{when } K \geq 190. \end{cases}$$

It is difficult to understand the choice of $K(b)$ without reading the end of the proof of Lemma 5.6. Let us say that these values are below $2 \cdot 10^6/b$ and are chosen as small as possible without worsening the value of ϵ_b^\pm .

Proof of Lemma 5.5. Summation by parts reduces the problem to using Lemma 5.4:

$$\begin{aligned} \sum_{\substack{k>K \\ (k,b)=1}} \frac{\mu^2(k)}{k\varphi(k)} &= 2 \int_K^\infty \sum_{\substack{k\leq t \\ (k,b)=1}} \frac{k\mu^2(k)}{k\varphi(k)} \frac{dt}{t^3} - \frac{1}{K^2} \sum_{\substack{k\leq K \\ (k,b)=1}} \frac{k\mu^2(k)}{k\varphi(k)} \\ &= \frac{\varphi(b)}{b} \frac{2 + \mathcal{O}^*(0.0002)}{K} - 1 + \mathcal{O}^*(0.0001) \end{aligned}$$

when $K \geq 10^6$. The lower bound is obtained similarly. It is easy to complete the proof by a direct computation. ■

Here is an intriguing lemma.

LEMMA 5.6. *Assuming Theorem 1.1 has been proved, for $K \geq 200$ and q^* arbitrary we have*

$$\left| \sum_{\substack{k>K \\ (k,q^*)=1}} \frac{\mu(k)}{\varphi(k)k} \right| \leq \frac{0.3616}{K}.$$

If q^* is assumed to be odd, we can replace 0.3616 by 0.3464.

The bound $K \geq 200$ will be reduced to $K \geq 14$ in the next lemma. It is not clear from the proof whether one can reach $o(1/K)$ when $K \rightarrow \infty$, due to the presence of the coprimality condition. The proof goes by discussing the gcd of k with the product of the prime factors below 11 or 13 in the odd case. One can think that by selecting a larger bound, say B , the resulting constant would go to 0, i.e. that the RHS of (5.4) below (replacing 0.026 by 0) is $o(\log B)$ uniformly for $q_0 | q^* = \prod_{p \leq B} p$. This seems difficult to prove. From a practical viewpoint, this lemma will have to be completed by Lemma 5.7 below.

Proof of Lemma 5.6. Let $S(q^*, K)$ be the sum to be estimated. When $30 | q^*$, the lemma follows from the last estimate of Lemma 5.5. Now assume that $30 \nmid q^*$. We choose a base $b \in \{210, 2\,310, 30\,030, 510\,510\}$, and set $\gcd(q^*, b) = b/q_0$. Notice that if q^* is odd, then q_0 is even. We classify the summation variable k in $S(q^*, x)$ according to its gcd δ with q_0 . We get

$$S(q^*, K) = \sum_{\delta|q_0} \frac{\mu(\delta)}{\varphi(\delta)\delta} \sum_{\substack{\ell>K/\delta \\ (\ell,q_0q^*)=1}} \frac{\mu(\ell)}{\varphi(\ell)\ell} = \sum_{\substack{\ell>K/q_0 \\ (\ell,q_0q^*)=1}} \frac{\mu(\ell)}{\varphi(\ell)\ell} \sum_{\substack{\delta|q_0 \\ \delta>K/\ell}} \frac{\mu(\delta)}{\varphi(\delta)\delta}.$$

We order the divisors δ of q_0 as $1 = \delta_1 < \delta_2 < \dots < \delta_I = q_0$, and thus

$$(5.2) \quad S(q^*, K) = \prod_{p|q_0} \frac{p^2 - p - 1}{p^2 - p} \sum_{\substack{\ell > K \\ (\ell, q_0 q^*) = 1}} \frac{\mu(\ell)}{\varphi(\ell)\ell} \\ + \sum_{1 \leq i \leq I-1} \sum_{\substack{K/\delta_i \geq \ell > K/\delta_{i+1} \\ (\ell, q_0 q^*) = 1}} \frac{\mu(\ell)}{\varphi(\ell)\ell} \sum_{i < j \leq I} \frac{\mu(\delta_j)}{\varphi(\delta_j)\delta_j}.$$

Since $b | q_0 q^*$, we have reached our main expression:

$$(5.3) \quad |S(q^*, K)| \leq \prod_{p|q_0} \frac{p^2 - p - 1}{p^2 - p} \sum_{\substack{\ell > K \\ (\ell, b) = 1}} \frac{|\mu(\ell)|}{\varphi(\ell)\ell} \\ + \sum_{2 \leq i \leq I} \sum_{\substack{K/\delta_{i-1} \geq \ell > K/\delta_i \\ (\ell, b) = 1}} \left| \frac{|\mu(\ell)|}{\varphi(\ell)\ell} \right| \left| \sum_{i \leq j \leq I} \frac{\mu(\delta_j)}{\varphi(\delta_j)\delta_j} \right|.$$

When $K/b \geq K(b)$, this gives, by Lemma 5.5,

$$(5.4) \quad \frac{b}{\varphi(b)} K |S(q^*, K)| \leq \prod_{p|q_0} \frac{p^2 - p - 1}{p^2 - p} (1 + \epsilon_b^+) \\ + \sum_{1 \leq i \leq I-1} ((1 + \epsilon_b^+) \delta_{i+1} - (1 - \epsilon_b^-) \delta_i) \left| \sum_{i < j \leq I} \frac{\mu(\delta_j)}{\varphi(\delta_j)\delta_j} \right|.$$

It is easy to run a program to check all possibilities of q_0 .

```
{getlocbound(q0, epsplus, epsmoins, base, fixeddivisor = 1) =
my(res = 0.0, aux = 1.0, divisorlist, di);
forprime(p = 2, q0, if (q0%p==0, aux *= (1-1/p/(p-1)),));
res += aux*(1 + epsplus);
```

```
divisorlist = divisors(q0/fixeddivisor);
aux = moebius(q0)/eulerphi(q0)/q0;
forstep(i = length(divisorlist)-1, 1, -1,
di = divisorlist[i]*fixeddivisor;
res += ((1+epsplus)*divisorlist[i+1]
-(1-epsmoins)*di)* abs(aux);
aux += moebius(di)/eulerphi(di)/di);
return(res * eulerphi(base)/base);}
```

```
{getbound(epsplus=0.00214, epsmoins = 0.0265, base = 210, fixeddivisor = 1)
= my(res = 0.0);
fordiv(base/fixeddivisor, q0p,
res = max(res, getlocbound(q0p*fixeddivisor, epsplus,
epsmoins, base, fixeddivisor)));
return(res);}
```

Anticipating on Lemma 5.7 below, we have to cover the range $K \geq 200$. We first used the above script with the ideal choice `epsplus=epsmoins=0` and found the following values:

210	2310	30030	510510
0.3827...	0.3560...	0.3340...	0.3167...

Now we have to get true values with effective `epsplus` and `epsmoins`, but it is better to be somewhat general to understand our choices. Say some values for $(\epsilon_b^\pm, K(b))$ are available in Lemma 5.5; then the value given by our script will be valid provided that $K/b \geq K(b)$. The resulting bound for K will be too large: Lemma 5.7 covers only the range $K \leq 200$, so we would need to have $bK(b) \leq 200$. This is not realistic. Instead in the range $K \in [200, bK(b)]$, we compute the upper bound provided by (5.3). This enables us to extend the upper bound to $K \geq 200$. We are now blocked by a decent bound for $bK(b)$ in these computations; we found that going beyond $2 \cdot 10^6$ in general and beyond 10^7 for a specific choice would require too much work, so we used these values. Here are the maxima obtained by this additional script:

210	2310	30030
0.3967...	0.3686...	0.3405...

The worst values are obtained from $q_0 = b$ in each of the three cases. Now that this is fixed, we understand the values chosen in Lemma 5.5.

Now, *when q^* is even*, we select $b = 2310$, and run the same computations but only with this modulus and for all $K \leq 10^7$. Note that we thus only need $2310K(2310) \geq 10^7$, i.e. $K(2310) \geq 4329$ is enough. We ran `getbound(0.0025, 0.0044, 2310)` (where the values 0.0025 and 0.0044 are given by Lemma 5.5) and reached the bound 0.3616.

We have 32 divisors q_0 of $b = 2310$ to consider, but we may discard those (four) for which $30 \mid 2310/q_0$. This still took between three and four hours per divisor with GP. For $K \in [1000, 10^7]$ (resp. $K \in [200, 10^7]$), the worst constant was $0.3643 + \mathcal{O}^*(10^{-4})$ (resp. $0.3687 + \mathcal{O}^*(10^{-4})$), reached when $q_0 = 2310$. We can exclude this divisor since our q_0 is even; the next worst constant is $0.3530 + \mathcal{O}^*(10^{-4})$ (resp. $0.3608 + \mathcal{O}^*(10^{-4})$), reached when $q_0 = 210$. We can again exclude this divisor; the next worst constant is $0.3369 + \mathcal{O}^*(10^{-4})$ (resp. $0.3454 + \mathcal{O}^*(10^{-4})$), reached when $q_0 = 1155$.

When q^ is odd*, q_0 is even, we need only cover the range $K \geq 378$. On this assumption, `getbound(0.5344, 0.5487, 30030, 2)=0.3369...` is small, so we only have to cover the range $[378, 30030 \times 5]$ on selecting $b = 30\,030$. We ran `beprimal(378, 160000, 13)` and saw that the worst case is when $q_0 = 30030$ and $K = 817$ with value $0.3442 + \mathcal{O}^*(10^{-4})$. ■

LEMMA 5.7. For any positive real K and any modulus q^* ,

$$\left| \sum_{\substack{k > K \\ (k, q^*) = 1}} \frac{\mu(k)}{\varphi(k)k} \right| \leq \frac{1.26}{K}.$$

Once Theorem 1.1 has been proved, and assuming $K \geq 14$, we can further reduce 1.26 to 0.3616. Assuming $K \in [14, 202]$, we can reduce 1.26 to 0.3185.

Once Theorem 1.1 has been proved, and assuming that $K \geq 14$ and q^* is odd, we can reduce 1.26 to 0.3442.

See Lemma 7.18 for a much better result when q^* is small. When K is large enough, say larger than K_0 , Lemma 5.6 does the job. The present lemma completes the work for K below K_0 ; it relies on an algorithm of large complexity with respect to K_0 .

Proof of Lemma 5.7. For $K < 1$, this is trivially checked. Let $h(q) = \prod_{p|q} p(p-1)/(p^2 - p - 1)$ and $C = \prod_{p \geq 2} h(p)^{-1} = 0.373959 \dots$. We have

$$\sum_{\substack{k > K \\ (k, q^*) = 1}} \frac{\mu(k)}{\varphi(k)k} = Ch(q^*) - \sum_{\substack{k \leq K \\ (k, q^*) = 1}} \frac{\mu(k)}{\varphi(k)k}.$$

- When $1 \leq K < 2$, this quantity is $Ch(q^*) - 1$ which is negative, $\geq C - 1$ and this last quantity is $\geq -1.26/2$.
- When $2 \leq K < 3$ and q^* is odd, this is $Ch(q^*) - 1/2$ and $Ch(q^*) \in [h(2)^{-1}, C]$. We check that $|h(2)^{-1} - 1/2| \leq 1.26/3$ and $|C - 1/2| \leq 1.26/3$.
- When $2 \leq K < 3$ and q^* is even, this is $Ch(q^*) - 1$ and $Ch(q^*) \in [2C, 1]$. We check that $|2C - 1| \leq 1.26/3$ and $|1 - 1| \leq 1.26/3$.

It is straightforward to write an algorithm. Assume we have a bound K_0 and let q^b be the product of all the primes below K_0 . Given q_1 dividing q^b , we consider all the q^* with $(q^b, q^*) = q_1$. We have

$$(5.5) \quad Ch(q_1) \leq Ch(q^*) \leq h(q_1)/h(q^b).$$

Note that when q^b is large enough, the difference between C and $1/h(q^b)$ is small. Let us now look at the sum $S(K, q^*) = \sum_{\substack{k \leq K \\ (k, q^*) = 1}} \frac{\mu(k)}{\varphi(k)k}$ which we compare to $S(K, q_1)$. We have

$$S(K, q^*) - S(K, q_1) = - \sum_{\substack{k \leq K \\ (k, q_1) = 1 \\ (k, q^*) > 1}} \frac{\mu(k)}{\varphi(k)k}.$$

When $K_0 \geq K$, this new sum is empty. But if $(k, q_1) = 1$ and $(k, q^*) > 1$ then k has at least one prime factor $> K_0$. Assuming that $K_0^2 > K$, only

one such prime p divides k . We write $k = sp$, where $s \leq K/K_0$ is prime to q_1 . Set $q^\sharp = \prod_{p \leq c} p$ and $q_0 = (q^\sharp, q_1)$. Let

$$(5.6) \quad \Sigma^+(K, q_0) = \sum_{K_0 < p \leq K} \frac{1}{p(p-1)} \max \left(\sum_{\substack{s \leq K/p \\ (s, q_0)=1}} \frac{\mu(s)}{s\varphi(s)}, 0 \right),$$

and let $\Sigma^-(K, q_0)$ be the corresponding sum with max instead of min. We have

$$(5.7) \quad S(K, q_1) - \Sigma^-(K, q) \leq S(K) \leq S(K, q_1) + \Sigma^+(K, q_0).$$

Combining (5.5) and (5.7), we reach

$$\begin{aligned} Ch(q_1) - S(K, q_1) - \Sigma^+(K, q_0) &\leq Ch(q^*) - S(K) \\ &\leq h(q_1)/h(q^b) - S(K, q_1) + \Sigma^-(K, q_0). \end{aligned}$$

Let us establish an algorithm from this inequality.

- (1) We want to cover the range $K_{\min} \leq K \leq cK_0$, where $c \in [1, K_0)$ is a (small) real number.
- (2) We define $q^\sharp = \prod_{p \leq c} p$ and $q^b = \prod_{p \leq K_0} p$. It is easy to (pre)compute all the $\Sigma^-(K, q_0)$ and $\Sigma^+(K, q_0)$ for $K \leq cK_0$ and $q_0 \leq c$.
- (3) We loop over q^b and K .

We ran such a GP-script with up to 25 primes with $c = 1$, and this takes care of $K < 101$. The worst constant was $2(1 - C) = 1.252\dots$, which is indeed not more than 1.26. When $K \geq 5$, the worst constant is 439/1008. When $K \geq 10$, the worst constant is 2/5.

On taking $c = 1.99$ and using the first 26 primes for q^b to go to $K \leq 202$, we prove that the worst constant is 0.3199... provided that $K \geq 14$ and $K \leq 202$. This constant is reached at $K = 202$ with $q_1 = 6$. Selecting c larger worsens the bound severely.

Assuming q^* to be odd, on taking $c = 3.5$ and using the first 28 primes for q^b to go to $K \leq 378$, we prove that the worst constant is 0.3252... provided that $K \geq 14$ and $K \leq 378$. This constant is reached at $K = 378$ with $q_1 = 3 \cdot 5 \cdot 7 \cdot 17 \cdot 19 \cdot 23 \cdot 29$.

For $K \geq 50$, we proceed by integration by parts as follows:

$$\begin{aligned} \sum_{\substack{k > K \\ (k, q^*)=1}} \frac{|\mu(k)|}{\varphi(k)k} &\leq 2 \int_K^\infty \sum_{k \leq t} \frac{|\mu(k)|k}{\varphi(k)} \frac{dt}{t^3} - \frac{1}{K^2} \sum_{k \leq K} \frac{|\mu(k)|k}{\varphi(k)} \\ &\leq \frac{2 \times 1.066 - 0.946}{K} \leq \frac{1.186}{K}. \end{aligned}$$

Theorem 1.1 is not required for this estimate, and when we assume Theorem 1.1, we can rely directly on Lemma 5.6. ■

6. Proof of Theorem 1.1. Recall that r_1 and r_2 are defined by (4.2) and (4.1).

LEMMA 6.1. *We have*

$$r_1(X) \leq 4.95 \sqrt{X} j_1(q), \quad j_1(q) = \prod_{p|q} \frac{\sqrt{p}(p^{3/2} + p - \sqrt{p} - 1)}{\sqrt{p}(p^{3/2} - \sqrt{p} + 1) + 1}.$$

Proof. We use $\mu^2(k)k/\phi(k) = \sum_{ab=k} \mu^2(ab)/\varphi(a)$ in (4.2) to infer that

$$r_1(X) = \sum_{\substack{a^2b^2\ell r \leq X \\ r|q \\ (ab\ell, q)=1 \\ (ab, \ell)=1}} \frac{\mu^2(r\ell ab)}{\varphi(a)\varphi(\ell)}.$$

At this level, we relax the coprimality conditions on b and reach

$$r_1(X) \leq \sum_{\substack{a^2\ell r \leq X \\ r|q \\ (a\ell, q)=1 \\ (a, \ell)=1}} \frac{\mu^2(r\ell a)}{\varphi(a)\varphi(\ell)} \sqrt{\frac{X}{a^2\ell r}}.$$

We then simply extend the summations over a, ℓ and r :

$$\begin{aligned} r_1(X) &\leq \sqrt{X} \prod_{p|q} \frac{1 + \frac{1}{\sqrt{p}}}{1 + \frac{1}{\sqrt{p}(p-1)} + \frac{1}{p(p-1)}} \prod_{p \geq 2} \left(1 + \frac{1}{\sqrt{p}(p-1)} \frac{1}{1 + \frac{1}{p(p-1)}} \right) \\ &\quad \times \prod_{p \geq 2} \left(1 + \frac{1}{p(p-1)} \right) \\ &\leq 4.95 \sqrt{X} \prod_{p|q} \frac{p(p-1) + \sqrt{p}(p-1)}{p(p-1) + \sqrt{p} + 1} = 4.95 \sqrt{X} j_1(q). \quad \blacksquare \end{aligned}$$

LEMMA 6.2. *We have*

$$|r_2(X; q)| \leq \frac{3.9}{\sqrt{X}} j_2(q), \quad j_2(q) = \prod_{p|q} \frac{p^{3/2} + p - \sqrt{p} - 1}{p^{3/2} - \sqrt{p} + 1}.$$

The expressions given for j_1 and j_2 make it clear that $j_1 \leq j_2$ so we can simplify the error term. Note however that $j_1(2) = 0.773\dots$, which is rather smaller than $j_2(2) = 2$.

Proof of Lemma 6.2. We treat the summation over k by Lemma 5.7 and obtain

$$\begin{aligned}
 |r_2(X; q)| &\leq \frac{1.26}{\sqrt{X}} \sum_{\substack{\ell \geq 1 \\ r|q \\ (\ell, q)=1}} \frac{\mu^2(r)\mu^2(\ell)}{\sqrt{r}\sqrt{\ell}} \varphi(\ell) \\
 &\leq \frac{1.26}{\sqrt{X}} \prod_{p|q} \frac{\sqrt{p}(p-1)}{1 + \sqrt{p}(p-1)} \frac{1 + \sqrt{p}}{\sqrt{p}} \prod_{p \geq 2} \left(1 + \frac{1}{\sqrt{p}(p-1)}\right) \\
 &\leq \frac{3.87}{\sqrt{X}} \prod_{p|q} \frac{(p-1)(1 + \sqrt{p})}{1 + \sqrt{p}(p-1)} = \frac{3.87}{\sqrt{X}} j_2(q). \blacksquare
 \end{aligned}$$

Proof of Theorem 1.1. We use Theorem 1.4 together with Lemmas 6.1 and 6.2. Notice that $G^\sharp = r_2$. We thus derive that

$$\begin{aligned}
 G_q(D) &= \sum_{\substack{k, \ell \geq 1 \\ r|q \\ (k\ell, q)=1}} \frac{\mu^2(k\ell)\mu(rk)}{\varphi(k)\varphi(\ell)k\ell r} \left(\log \frac{D}{k^2\ell r} + \gamma \right) \\
 &\quad + \mathcal{O}^* \left((2 + |\gamma - \log \eta|) \frac{3.9}{\sqrt{\eta D}} j_2(q) + \frac{\gamma}{D} 4.95 \sqrt{\eta D} j_1(q) \right).
 \end{aligned}$$

The expression given in the theorem for the main term is easily computed or identified via [26, Lemma 3.4]. We select $\eta = e^\gamma$ in the error term. In the cases $q = 6, 30, 210, 2310$ and $q = 30\,030$, we compute the expression for the error term directly. In general, we prefer to simplify this error term further. We now discuss according to whether $2|q$ or not. When q is odd, we use $j_1(q) \leq j_2(q)$; this leads to the error term $\mathcal{O}^*(5.82j_2(q)/\sqrt{D})$. When $q = 2q'$ with q' odd, we obtain the error term $\mathcal{O}^*(4.95j_2(q')/\sqrt{D})$.

Our formula follows readily. \blacksquare

7. Cancellation in $\sum_{k>K} \mu(k)/(\varphi(k)k)$. Analytical estimates. We continue with the problem of bounding $\sum_{k>K, (k, q^*)=1} \frac{\mu(k)}{\varphi(k)k}$ but we use the results on the Möbius function recalled in Lemma 7.14. The main point is Lemma 7.18.

7.1. Some asymptotic estimates

LEMMA 7.1. *For $L > 0$,*

$$\sum_{\ell \leq L} \frac{\mu^2(\ell)}{\varphi(\ell)} \prod_{p|\ell} (1 + p^{-1}) = \frac{15}{\pi^2} (\log L + c_2) + \mathcal{O}^*(16.4/L^{1/3})$$

where

$$c_2 = \gamma - \sum_{p \geq 2} \frac{2 \log p}{p^3 - p^2 + p} = 0.187\,529\dots$$

Proof. We use [26, Lemma 3.4]. We define

$$f(\ell) = \frac{\mu^2(\ell)}{\varphi(\ell)} \prod_{p|\ell} (1 + p^{-1}).$$

First note that

$$\begin{aligned} \sum_{\ell \geq 1} \frac{f(\ell)}{\ell^s} &= \prod_{p \geq 2} \left(1 + \frac{p+1}{(p-1)p^{s+1}} \right) \\ &= \zeta(s+1) \prod_{p \geq 2} \left(1 + \frac{2}{(p-1)p^{s+1}} - \frac{p+1}{(p-1)p^{2s+2}} \right) = \zeta(s+1)H(s) \end{aligned}$$

say, which gives $\ell f(\ell) = \sum_{d|\ell} g(d)$ where g is the multiplicative function defined on prime powers by

$$g(p) = \frac{2}{p-1}, \quad g(p^2) = -\frac{p+1}{p-1}, \quad g(p^k) = 0 \quad (k \geq 3).$$

We note that

$$H(0) = \prod_{p \geq 2} \left(1 + \frac{2}{(p-1)p} - \frac{p+1}{(p-1)p^2} \right) = \prod_{p \geq 2} \left(1 + \frac{1}{p^2} \right) = \frac{\zeta(2)}{\zeta(4)} = \frac{15}{\pi^2}.$$

Furthermore

$$\frac{-1}{H(0)} \sum_{m \geq 1} \frac{g(m) \log m}{m} = \frac{H'(0)}{H(0)} = - \sum_{p \geq 2} \frac{2 \log p}{p^3 - p^2 + p},$$

while

$$\bar{H}(-1/3) = \prod_{p \geq 2} \left(1 + \frac{2}{(p-1)p^{2/3}} + \frac{p+1}{(p-1)p^{4/3}} \right) \leq 18. \blacksquare$$

LEMMA 7.2. For $L > 0$,

$$\sum_{\ell \leq L} \mu^2(\ell) \prod_{p|\ell} \frac{p^2 + p}{p^3 - p^2 + p + 1} = d_0(\log L + c_3) + \mathcal{O}^*(9.7/L^{1/3})$$

where $d_0 = 1.026\dots$ and

$$c_3 = \gamma - \sum_{p \geq 2} \frac{(3p+1) \log p}{p^4 - p^3 + 3p^2 - p - 2} = 0.048757\dots$$

Proof. We again use [26, Lemma 3.4]. We define

$$f(\ell) = \mu^2(\ell) \prod_{p|\ell} \frac{p^2 + p}{p^3 - p^2 + p + 1}.$$

First note that

$$\begin{aligned} \sum_{\ell \geq 1} \frac{f(\ell)}{\ell^s} &= \prod_{p \geq 2} \left(1 + \frac{p^2 + p}{(p^3 - p^2 + p + 1)p^s} \right) \\ &= \zeta(s + 1) \prod_{p \geq 2} \left(1 + \frac{2p^2 - p - 1}{(p^3 - p^2 + p + 1)p^{s+1}} - \frac{p + 1}{(p^3 - p^2 + p + 1)p^{2s}} \right) \\ &= \zeta(s + 1)H(s) \end{aligned}$$

say, which gives $\ell f(\ell) = \sum_{d|\ell} g(d)$ where g is the multiplicative function defined on prime powers by

$$g(p) = \frac{2p^2 - p - 1}{p^3 - p^2 + p + 1}, \quad g(p^2) = -\frac{p^3 + p^2}{p^3 - p^2 + p + 1}, \quad g(p^k) = 0 \quad (k \geq 3).$$

We note that

$$\begin{aligned} d_0 = H(0) &= \prod_{p \geq 2} \left(1 + \frac{2p^2 - p - 1}{(p^3 - p^2 + p + 1)p} - \frac{p + 1}{p^3 - p^2 + p + 1} \right) \\ &= \prod_{p \geq 2} \left(1 + \frac{p^2 - 2p - 1}{(p^3 - p^2 + p + 1)p} \right) = 1.026\dots \end{aligned}$$

Furthermore

$$\frac{-1}{H(0)} \sum_{m \geq 1} \frac{g(m) \log m}{m} = \frac{H'(0)}{H(0)} = -\sum_{p \geq 2} \frac{(3p + 1) \log p}{p^4 - p^3 + 3p^2 - p - 2},$$

while

$$\overline{H}(-1/3) = \prod_{p \geq 2} \left(1 + \frac{2p^2 - p - 1}{(p^3 - p^2 + p + 1)p^{2/3}} + \frac{p + 1}{(p^3 - p^2 + p + 1)p^{-2/3}} \right) \leq 9.7. \blacksquare$$

7.2. Majorising tails of averages

LEMMA 7.3. *For every $L > 0$, we have $\sum_{\ell > L} \frac{\mu^2(\ell)}{\ell \varphi(\ell)} \leq 1.96/L$. When $L \geq 1$, we can replace 1.96 by 1.14.*

In [26, Lemma 3.10], a similar estimate is proved, but with 4 instead of 1.14.

Proof of Lemma 7.3. Denote by $S(L)$ the sum we want to estimate. Summation by parts gives

$$S(L) = 2 \int_L^\infty \sum_{L < \ell \leq t} \frac{\mu^2(\ell)\ell}{\varphi(\ell)} \frac{dt}{t^3} = -\sum_{\ell \leq L} \frac{\mu^2(\ell)\ell}{\varphi(\ell)L^2} + 2 \int_L^\infty \sum_{L \leq \ell \leq t} \frac{\mu^2(\ell)\ell}{\varphi(\ell)} \frac{dt}{t^3},$$

and this is not more than $(2 \times 1.004 - 0.996)/L \leq 1.012/L$ when $L \geq 10^6$ (as in the proof of Lemma 5.7) by Lemma 5.2. We complete by a direct verification up to 10^6 . \blacksquare

LEMMA 7.4. For every $L \geq 1$, we have $\sum_{\ell > L, (\ell, 2)=1} \frac{\mu^2(\ell)}{\ell\varphi(\ell)} \leq 0.592/L$.

Proof. Denote the sum by $S(L)$. Summation by parts gives

$$S(L) = 2 \int_L^\infty \sum_{L < \ell \leq t} \frac{\mu^2(\ell)\ell}{\varphi(\ell)} \frac{dt}{t^3} = - \sum_{\substack{\ell \leq L \\ (\ell, 2)=1}} \frac{\mu^2(\ell)\ell}{\varphi(\ell)L^2} + 2 \int_L^\infty \sum_{\substack{\ell \leq t \\ (\ell, 2)=1}} \frac{\mu^2(\ell)\ell}{\varphi(\ell)} \frac{dt}{t^3},$$

and this is not more than $(2 \times 1.005 - 0.995)/2$ by Lemma 5.3. We complete by a direct verification up to 10^6 . ■

LEMMA 7.5. For every $L \geq 1$, we have $\sum_{\ell > L, (\ell, 2)=1} \frac{\mu^2(\ell)}{\sqrt{\ell}\varphi(\ell)} \leq 1.19/\sqrt{L}$.

Proof. We have

$$\sum_{\substack{\ell > L \\ (\ell, 2)=1}} \frac{\mu^2(\ell)}{\sqrt{\ell}\varphi(\ell)} = \sum_{\substack{\ell > L \\ (\ell, 2)=1}} \frac{\mu^2(\ell)\sqrt{Y}}{\ell\varphi(\ell)} + \int_Y^\infty \sum_{\substack{\ell > t \\ (\ell, 2)=1}} \frac{\mu^2(\ell)}{\ell\varphi(\ell)} \frac{dt}{2\sqrt{t}}$$

and an appeal to Lemma 7.4 concludes the proof when $L \geq 1$. ■

LEMMA 7.6. For every $L > 0$, we have $\sum_{\ell > L} \frac{\mu^2(\ell) \prod_{p|\ell} (1+p^{-1})}{\ell\varphi(\ell)} \leq 2.83/L$. When $L \geq 6$, we can decrease 2.83 to 1.85, and when $L \geq 30$, we can decrease it further to 1.63.

Proof. We proceed as in the proof of Lemma 7.3. Let $S(L)$ be the sum we want to bound, and $f(\ell) = \mu^2(\ell) \prod_{p|\ell} (1+p^{-1})/\varphi(\ell)$. Summation by parts gives, via Lemma 7.1,

$$\begin{aligned} S(L) &= \int_L^\infty \sum_{L < \ell \leq t} f(\ell) \frac{dt}{t^2} = - \sum_{L < \ell \leq t} \frac{f(\ell)}{L} + \int_L^\infty \sum_{\ell \leq t} f(\ell) \frac{dt}{t^2} \\ &= -\frac{15}{\pi^2 L} (\log L + c_2) + \int_L^\infty \frac{15}{\pi^2} (\log t + c_2) \frac{dt}{t^2} + \mathcal{O}^* \left(\frac{16.4(1+3/4)}{L^{4/3}} \right) \\ &= \frac{15}{\pi^2 L} + \mathcal{O}^*(29/L^{4/3}), \end{aligned}$$

and this is not more than 2.5 when $L \geq 10^5$. We complete by a direct verification up to 10^6 . ■

LEMMA 7.7. For every $L > 0$, we have $\sum_{\ell > L} \frac{\mu^2(\ell) \prod_{p|\ell} (1+p^{-1})}{\ell^2\varphi(\ell)} \leq 2.02/L^2$.

Proof. We denote by $S(L)$ the sum to be bounded and write $f(\ell) = \mu^2(\ell) \prod_{p|\ell} (1+p^{-1})/\varphi(\ell)$. Summation by parts gives, via Lemma 7.1,

$$\begin{aligned}
 S(L) &= 2 \int_L^\infty \sum_{L < \ell \leq t} f(\ell) \frac{dt}{t^3} = - \sum_{L < \ell \leq t} \frac{f(\ell)}{L^2} + 3 \int_L^\infty \sum_{\ell \leq t} f(\ell) \frac{dt}{t^3} \\
 &= - \frac{15}{\pi^2 L^2} (\log L + c_2) + 3 \int_L^\infty \frac{15}{\pi^2} (\log t + c_2) \frac{dt}{t^3} + \mathcal{O}^* \left(\frac{16.4(1 + 3/7)}{L^{7/3}} \right) \\
 &= \frac{15}{2\pi^2 L^2} + \mathcal{O}^*(24/L^{7/3}),
 \end{aligned}$$

and this is not more than 2 when $L \geq 10^4$. We complete by a direct verification up to 10^6 . ■

LEMMA 7.8. *When $L > 0$, we have $\sum_{\ell > L} \frac{\mu^2(\ell) \prod_{p|\ell} (1+p^{-1})}{\sqrt{\ell} \varphi(\ell)} \leq 4.43/\sqrt{L}$. This is ≤ 3.80 when $L \geq 2$, and ≤ 3.12 when $L \geq 100$.*

Proof. We denote by $S(L)$ the sum we want to bound and set $f(\ell) = \mu^2(\ell) \prod_{p|\ell} (1 + p^{-1})/\varphi(\ell)$. Summation by parts gives, via Lemma 7.1,

$$\begin{aligned}
 S(L) &= \frac{3}{2} \int_L^\infty \sum_{L < \ell \leq t} f(\ell) \frac{dt}{t^{3/2}} = - \sum_{L < \ell \leq t} \frac{f(\ell)}{\sqrt{L}} + \frac{3}{2} \int_L^\infty \sum_{\ell \leq t} f(\ell) \frac{dt}{t^{3/2}} \\
 &= - \frac{15}{\pi^2 \sqrt{L}} (\log L + c_2) + \frac{3}{2} \int_L^\infty \frac{15}{\pi^2} (\log t + c_2) \frac{dt}{t^{3/2}} + \mathcal{O}^* \left(\frac{16.4(1 + 3/2)}{L^{2/3}} \right) \\
 &= \frac{30}{\pi^2 \sqrt{L}} + \mathcal{O}^*(41/L^{2/3}),
 \end{aligned}$$

and this is not more than 2 when $L \geq 10^4$. We complete by a direct verification up to 10^6 . ■

LEMMA 7.9. *For every $L > 0$, we have $\sum_{\ell > L} \frac{\mu^2(\ell)}{\ell^2} \prod_{p|\ell} \frac{p^2+p}{p^3-p^2+p+1} \leq 1.25/L^2$. When $L \geq 20$, we can decrease 1.25 to 0.66.*

Proof. We denote the sum by $S(L)$, and set $f(\ell) = \mu^2(\ell) \prod_{p|\ell} \frac{p^2+p}{p^3-p^2+p+1}$. Summation by parts gives, via Lemma 7.2,

$$\begin{aligned}
 S(L) &= 2 \int_L^\infty \sum_{L < \ell \leq t} f(\ell) \frac{dt}{t^3} = - \sum_{L < \ell \leq t} \frac{f(\ell)}{L^2} + 3 \int_L^\infty \sum_{\ell \leq t} f(\ell) \frac{dt}{t^3} \\
 &= - \frac{d_0}{L^2} (\log L + c_3) + 2 \int_L^\infty d_0 (\log t + c_3) \frac{dt}{t^3} + \mathcal{O}^* \left(\frac{9.7(1 + 3/7)}{L^{7/3}} \right) \\
 &= \frac{d_0}{2L^2} + \mathcal{O}^*(13.9/L^{7/3}),
 \end{aligned}$$

and this is not more than 2 when $L \geq 10^4$. We complete by a direct verification up to 10^6 . ■

LEMMA 7.10. *For every $K > 0$, we have $\sum_{k > K} \frac{\mu^2(k)}{k^2 \varphi(k)} \leq 0.606/K^2$.*

Proof. Summation by parts gives

$$\begin{aligned} \sum_{k>K} \frac{\mu^2(k)}{k^2\varphi(k)} &= 3 \int_K^\infty \sum_{K < \ell \leq t} \frac{\mu^2(k)k}{\varphi(k)} \frac{dt}{t^4} \\ &= - \sum_{k \leq K} \frac{\mu^2(k)k}{\varphi(k)K^3} + 3 \int_K^\infty \sum_{k \leq t} \frac{\mu^2(k)k}{\varphi(k)} \frac{dt}{t^4}, \end{aligned}$$

and this is not more than $(\frac{3}{2} \times 1.004 - 0.996)/K \leq 0.51/K$ when $K \geq 10^6$ (as in the proof of Lemma 5.7) by Lemma 5.2. We complete by a direct verification up to 10^6 . ■

LEMMA 7.11. *For every $X > 0$,*

$$\sum_{a^2b^3 > X} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)} \leq \frac{2.45}{\sqrt{X}}.$$

Proof. We split the sum in two:

$$\sum_{a^2b^3 > X} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)} \leq \sum_{b^3 > X, a \geq 1} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)} + \sum_{b^3 \leq X} \sum_{\substack{a^2 > X/b^3 \\ (a,b)=1}} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)}.$$

In the second sum, we forget the condition $(a, b) = 1$ when b is odd, and degrade it to $(a, 2) = 1$ when b is even. Via Lemmas 7.3 and 7.4 we get

$$\begin{aligned} \sum_{a^2b^3 > X} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)} &\leq 1.96 \sum_{b^3 > X} \frac{\mu^2(b)}{b^2\varphi(b)} + 1.14 \sum_{\substack{b^3 \leq X \\ (b,2)=1}} \frac{\mu^2(b)}{b^2\varphi(b)} \sqrt{\frac{b^3}{X}} \\ &\quad + 0.592 \sum_{\substack{b^3 \leq X/8 \\ (b,2)=1}} \frac{\mu^2(b)}{4b^2\varphi(b)} \sqrt{\frac{b^3}{X}} \end{aligned}$$

and, via Lemma 7.10,

$$\sum_{\substack{a^2b^3 > X \\ (ab, q^*)=1}} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)} \leq 1.96 \times \frac{0.606}{X^{2/3}} + \frac{1.14 + 0.592/4}{\sqrt{X}} \sum_{\substack{b^3 \leq X \\ (b,2)=1}} \frac{\mu^2(b)}{b^{1/2}\varphi(b)},$$

yielding the upper bound $1.19/X^{2/3} + 2.32/\sqrt{X}$. We check directly that this quantity is below $2.45/\sqrt{X}$ when $1 \leq X \leq 10^6$. Assuming that $X \geq 5000$, we can reduce 2.45 to 2.2, but we would have to be more careful in the above analysis. ■

LEMMA 7.12. *For every $X > 0$,*

$$\sum_{a^2b^3 > X} \frac{\mu^2(ab) \prod_{p|ab} (1 + p^{-1})}{a\varphi(a)b^2\varphi(b)} \leq \frac{6.21}{\sqrt{X}}.$$

Proof. We proceed as in Lemma 7.11 and split the sum in two:

$$\begin{aligned} \sum_{a^2b^3 > X} \frac{\mu^2(ab) \prod_{p|ab} (1 + p^{-1})}{a\varphi(a)b^2\varphi(b)} &\leq \sum_{b^3 > X/6^2, a \geq 1} \frac{\mu^2(ab) \prod_{p|ab} (1 + p^{-1})}{a\varphi(a)b^2\varphi(b)} \\ &+ \sum_{b^3 \leq X/6^2} \sum_{a^2 > X/b^3} \frac{\mu^2(ab) \prod_{p|ab} (1 + p^{-1})}{a\varphi(a)b^2\varphi(b)}, \end{aligned}$$

and thus via Lemma 7.6 and on using $\prod_{p \geq 2} (1 + \frac{p+1}{p^2(p-1)}) \leq 2.42$, we get

$$\begin{aligned} \sum_{a^2b^3 > X} \frac{\mu^2(ab)\mu^2(ab) \prod_{p|ab} (1 + p^{-1})}{a\varphi(a)b^2\varphi(b)} &\leq 2.42 \sum_{b^3 > X/6^2} \mu^2(b) \prod_{p|b} \frac{p+1}{p(p^3 - p^2 + p + 1)} \\ &+ 1.85 \sum_{b^3 \leq X} \frac{\mu^2(b) \prod_{p|b} (1 + p^{-1})}{b^2\varphi(b)} \sqrt{\frac{b^3}{X}}. \end{aligned}$$

Lemma 7.9 gives

$$\begin{aligned} \sum_{\substack{a^2b^3 > X \\ (ab, q^*)=1}} \frac{\mu^2(ab)\mu^2(ab) \prod_{p|ab} (1 + p^{-1})}{a\varphi(a)b^2\varphi(b)} &\leq 2.42 \times \frac{0.66 \cdot 6^{4/3}}{X^{2/3}} + \frac{1.85}{\sqrt{X}} \prod_{p \geq 2} \left(1 + \frac{1}{\sqrt{p}(p-1)} \right), \end{aligned}$$

yielding the upper bound $5.69/\sqrt{X} + 17.5/X^{2/3}$. We check directly that this quantity is below $4.4/\sqrt{X}$ when $1 \leq X \leq 1.5 \cdot 10^9$. ■

7.3. External lemmas concerning the Möbius function. Here is part of [8, Lemma 10.2]. See also [29, Theorem 1.1].

LEMMA 7.13. *When $q^* \geq 1$ and $x > 0$, we have*

$$\left| \sum_{\substack{d \leq x \\ (d, q^*)=1}} \frac{\mu(d)}{d} \right| \leq 1.$$

Here is part of [31, Theorems 1.12 and 1.13].

LEMMA 7.14. *When $x > q^* \geq 1$, we have*

$$\left| \sum_{\substack{d \leq x \\ (d, q^*)=1}} \frac{\mu(d)}{d} \right| \leq \kappa_1(x/q^*) \prod_{p|q^*} (1 + p^{-1}) / \log(x/q^*).$$

where

$$\kappa_1(t) = \begin{cases} 4/5 & \text{when } 1 < t < 296, \\ 1/2 & \text{when } 296 \leq t < 687, \\ 5/16 & \text{when } 687 \leq t < 882, \\ 5/38 & \text{when } 882 \leq t < 11\,811, \\ 1/7 & \text{when } 11\,811 \leq t. \end{cases}$$

7.4. Taking care of the $1/\log(X/\ell)$ factor. In several instances, the use of Lemma 7.14 leads to expressions of the shape

$$S_\kappa(f; U, c) = \sum_{u \leq U/c} \frac{\kappa(U/u)f(u)}{\log(U/u)}$$

where c is uniformly taken as 50. This value is however irrelevant in the general discussion. The function κ is either the function κ_1 defined in Lemma 7.14 or the constant function 1; and f is some non-negative arithmetical function such that $\sum_{u \geq 1} f(u)$ converges. Since κ is a decreasing step function with jumps at $c = a_1 < a_2 < \dots < a_I$, we have

$$S_\kappa(f; U, c) = \kappa(a_I^+)S(f; U, a_I) + \sum_{1 \leq i \leq I-1} \kappa(a_i^+)(S(f; U, a_i) - S(f; U, a_{i+1}))$$

where

$$S(f; U, a) = \sum_{u \leq U/a} \frac{f(u)}{\log(U/u)}.$$

This gives

$$(7.1) \quad S_\kappa(f; U, c) = \kappa(c^+)S(f; U, c) - \sum_{2 \leq i \leq I} S(f; U, a_i)(\kappa(a_{i-1}^+) - \kappa(a_i^+)).$$

We first handle $S(f; U, a)$ by integration by parts:

$$(7.2) \quad S(f; U, a) = \sum_{u \leq U/a} \frac{f(u)}{\log a} - \int_a^U \sum_{u \leq U/t} f(u) \frac{dt}{t(\log t)^2}.$$

This expression is convenient to compute $S(f; U, a)$ for bounded values of U . Set $C(f) = \sum_{u \geq 1} f(u)$ and $G(f; t) = \sum_{u \leq t} f(t)$ as well as

$$(7.3) \quad S_0(f; U, a) = \int_a^U \sum_{U/t < u} f(u) \frac{dt}{t(\log t)^2}.$$

We have

$$S(f; U, a) = \sum_{u \leq U/a} \frac{f(u)}{\log a} - C(f) \left(\frac{1}{\log a} - \frac{1}{\log U} \right) + S_0(f; U, a),$$

i.e.

$$(7.4) \quad S(f; U, a) = - \sum_{u > U/a} \frac{f(u)}{\log U} + \frac{C(f)}{\log U} + S_0(f; U, a).$$

In order to evaluate $S(f; U, a)$ for large U , we use some upper bound for $\sum_{U/t < u} f(u)$. In the process, the numerical difficulty comes from the fact that we need such an evaluation when U/t can be as small as 1, and this required uniformity is a drag on the constants. For instance, when we have a bound of the shape $\sum_{x < u} f(u) \ll 1/\sqrt{x}$, the term $S_0(f; U, a)$ is $\ll 1/(\log x)^2$, but this happens only when x is rather large. It is much better to handle the values when x (i.e. initially U/t) is small by direct computations, when f is indeed directly computable (this is not the case for instance when $f(u)$ vanishes as soon as u is not coprime to some parameter q). We have, for an integer parameter $H \geq 1$,

$$\begin{aligned} 0 \leq S_0(f; U, a) - \sum_{1 \leq h \leq \min(H, U/c)} \int_{\max(a, U/(h+1))}^{U/h} (C(f) - G(f; U/t)) \frac{dt}{t(\log t)^2} \\ \leq \int_{a \leq t \leq U/(H+1)} \sum_{U/t < u} f(u) \frac{dt}{t(\log t)^2}. \end{aligned}$$

We have used the notation $\int_{a \leq t \leq b}$ and not \int_a^b : in the first case the integral has value 0 when $b < a$, while the second one is usually understood as $-\int_b^a$. This upper bound develops into

$$(7.5) \quad \begin{aligned} 0 \leq S_0(f; U, a) \\ - \sum_{1 \leq h \leq \min(H, U/a)} \left(C(f) - \sum_{u \leq h} f(u) \right) \left(\frac{1}{\log \frac{U}{h}} - \frac{1}{\log \max(a, \frac{U}{h+1})} \right) \\ \leq \int_{a \leq t \leq U/(H+1)} \sum_{U/t < u} f(u) \frac{dt}{t(\log t)^2}. \end{aligned}$$

The very last problem consists in estimating the first summand in (7.4). We use

$$(7.6) \quad \sum_{u > U/a} f(u) = \begin{cases} C(f) - \sum_{u \leq U/a} f(u) & \text{when } U/a \leq H + 1, \\ \mathcal{O}^*(\sum_{u > H+1} f(u)) & \text{when } U/a \geq H + 1. \end{cases}$$

The cutting point $a \cdot (H + 1)$ in the conditions $U \leq a \cdot (H + 1)$ and $U > a \cdot (H + 1)$ is somewhat arbitrary. It turns out that the small values of U are going to have a predominant role; for such values, the partial sum $\sum_{u \leq U/a} f(u)$ may be noticeably smaller than $C(f)$. In case the jump around $c(H + 1)$ is too large, it can be reduced by increasing the value of H .

If we gather the above, here is how we proceed:

- We split $S_\kappa(f; U, c)$ according to (7.1).
- We compute a lower bound $S^{(1)}(f; U, a)$ and an upper bound $S^{(2)}(f; U, a)$ for $S(f; U, a)$ following (7.4)–(7.6) and a given bound for $\sum_{U/t < u} f(u)$.
- From these bounds we produce an upper bound for $S_\kappa(f; U, c)$. This bound is a “simple” function that we study easily. Practically, it is enough to compute and plot it on a large enough range; it vanishes at infinity, and this infinity is close enough that the remaining range can be covered by computations.

This is what we do in the next two lemmas.

LEMMA 7.15. *For all $x \geq 1800$,*

$$(\log x) \sum_{a^2 b^3 \leq x/50} \frac{\mu^2(ab) \kappa_1(x/(a^2 b^3)) \prod_{p|ab} (1 + 1/p)}{a \varphi(a) b^2 \varphi(b) \log(x/a^2 b^3)} \leq 1.89.$$

Proof. We use (7.5) with

$$f(\ell) = \begin{cases} \frac{\mu^2(ab) \prod_{p|ab} (1 + 1/p)}{a \varphi(a) b^2 \varphi(b)} & \text{when } \ell = a^2 b^3 \text{ with } \mu^2(ab) = 1 \\ 0 & \text{else} \end{cases}$$

and $c = 50$. We have

$$C(f) = \prod_{p \geq 2} \left(1 + \frac{1 + 1/p}{p(p-1)} + \frac{1 + 1/p}{p^2(p-1)} \right).$$

We select $H = 10^6$ in the method above. The bound for the tail is provided by Lemma 7.12. The lengthy discussion above results in the following simple script:

```
{getval30(n)=
my(res=1.0, p, fac = factor(n)~);
if(n==1, return(1),);
for(k = 1, length(fac),
  p = fac[1,k];
  if(fac[2, k] == 2, res *= (p+1)/p^2/(p-1),
    if(fac[2, k] == 3, res *= (p+1)/p^3/(p-1),
      return(0))););
return(res);}

cc8 = prodeuler(p = 2, 1000000, 1.0+(p+1)^2/p^3/(p-1));

{upper8inner(x, c, H, notail=0)=
my(res = 0.0, aux = 0.0);
for(h = 1, min(x/c, H),
  aux += getval30(h);
  res += (cc8 - aux)*(-1/log(x/h)+1/log(max(c,x/(h+1)))));
if(x <= c*(H+1), res += aux/log(x), res += cc8/log(x));
if(notail==1, return(res - 6.21/sqrt(H+1)/log(x)),);
```

```

if((x <= c*(H+1)), return(res),
    res += 6.21/sqrt(x)*intnum( t = c, x/(H+1), 1/sqrt(t)/log(t)^2);
return(res);}

{upper8lower(x, c=50, H=300) = upper8inner(x, c, H, 1);}
{upper8upper(x, c=50, H=300) = upper8inner(x, c, H, 0);}

{upper8(x, H=300)=
  4/5*upper8upper(x, 50, H)
  -(4/5-1/2)*upper8lower(x, 296, H)
  -(1/2-5/16)*upper8lower(x, 687, H)
  -(5/16-5/38)*upper8lower(x, 882, H)
  -(5/38-1/7)*upper8lower(x, 11811, H);}

{bupper8(x, H=400)= upper8(x, H)*log(x);} ■
    
```

LEMMA 7.16. *For all $x \geq 50$,*

$$(\log x) \sum_{\ell \leq x/50} \frac{\mu^2(\ell) \prod_{p|\ell} (1 + p^{-1})}{\sqrt{\ell} \varphi(\ell) \log(x/\ell)} \leq 5.02.$$

Proof. We use (7.5) with $f(\ell) = \mu^2(\ell) \prod_{p|\ell} (1 + p^{-1}) / (\sqrt{\ell} \varphi(\ell))$ and $c = 50$. We have

$$C(f) = \prod_{p \geq 2} \left(1 + \frac{p + 1}{p^{3/2}(p - 1)} \right).$$

We use Lemma 7.8 and select $H = 10^6$ in the method above. The bound for the tail is provided by Lemma 7.8. ■

7.5. Applications

LEMMA 7.17. *When $K/q^* \geq 1800$, we have*

$$\left| \sum_{\substack{k > K \\ (k, q^*) = 1}} \frac{\mu(k)}{\varphi(k)} \right| \leq \frac{1.89}{\log(K/q^*)} \prod_{p|q^*} \left(1 + \frac{1}{p} \right) + 17.4 \left(\frac{q^*}{K} \right)^{1/2}.$$

Proof. Denote the sum by $S(K)$. We check by multiplicativity that

$$(7.7) \quad \frac{\mu(k)}{\varphi(k)} = \sum_{dab^2=k} \frac{\mu(d)}{d} \frac{\mu^2(abd)}{a\varphi(a)b^2\varphi(b)}.$$

This identity leads to

$$S(K) = \sum_{\substack{a, b \geq 1 \\ (ab, q^*) = 1}} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)} \sum_{\substack{d > K/(ab^2) \\ (d, q^*ab) = 1}} \frac{\mu(d)}{d}.$$

Since $\sum_{d \geq 1, (d, q^* ab)=1} \mu(d)/d = 0$, we also have

$$S(K) = - \sum_{\substack{a, b \geq 1 \\ (ab, q^*)=1}} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)} \sum_{\substack{d \leq K/(ab^2) \\ (\bar{d}, q^* ab)=1}} \frac{\mu(d)}{d}.$$

When $q^* a^2 b^3 \leq K/50$, we can use Lemma 7.14, while otherwise we rely on Lemma 7.13. We get

$$\begin{aligned} |S(K)| \leq & \sum_{\substack{q^* a^2 b^3 \leq K/50 \\ (ab, q^*)=1}} \frac{\kappa_1(K/(q^* a^2 b^3)) \mu^2(ab)}{a\varphi(a)b^2\varphi(b) \log(K/(q^* a^2 b^3))} \prod_{p|q^* ab} (1 + p^{-1}) \\ & + \sum_{\substack{q^* a^2 b^3 > K/50 \\ (ab, q^*)=1}} \frac{\mu^2(ab)}{a\varphi(a)b^2\varphi(b)}. \end{aligned}$$

Lemma 7.15 with $x = K/q^*$ implies that the first sum is not more than $1.89 \prod_{p|q^*} (1 + p^{-1}) / \log(K/q^*)$. This together with Lemma 7.11 (and noting that $2.45\sqrt{50} \leq 17.4$) yields

$$|S(K)| \leq \prod_{p|q^*} \frac{p+1}{p} \frac{1.89}{\log(K/q^*)} + 17.4 \left(\frac{q^*}{K}\right)^{1/2}. \blacksquare$$

LEMMA 7.18. *When $K > q^*$ and $K \geq 14$, we have*

$$\begin{aligned} & \left| \sum_{\substack{k > K \\ (k, q^*)=1}} \frac{\mu(k)}{k\varphi(k)} \right| \\ & \leq \frac{1.89 \prod_{p|q^*} (1 + p^{-1})}{K \log(K/q^*)} + \frac{5}{K} \left(\frac{q^*}{K}\right)^{1/2} + \frac{6.6}{K} \left(\frac{q^*}{K}\right)^{1/2} \mathbb{1}_{K \geq 1970q^*}. \end{aligned}$$

The last statement means that the third summand is to be added only when K/q^* is larger than 1970.

This lemma is qualitatively much better than Lemma 5.7 when q^* is small: it indeed uses the oscillations of the Möbius function.

Proof of Lemma 7.18. When $K \geq 14$ and $K/q^* \leq 1970$, this is a consequence of Lemma 5.7. Otherwise, we use integration by parts to get

$$\sum_{\substack{k > K \\ (k, q^*)=1}} \frac{\mu(k)}{k\varphi(k)} = \int_K^\infty \sum_{\substack{k > t \\ (k, q^*)=1}} \frac{\mu(k)}{\varphi(k)} \frac{dt}{t^2}.$$

The lemma follows readily after invoking Lemma 7.17. \blacksquare

8. Bounding $r_2(X; 1)$

8.1. Some asymptotic estimates

LEMMA 8.1. *For all $L > 0$, we have $\sum_{\ell > L} \frac{\mu^2(\ell)}{\sqrt{\ell}\varphi(\ell)} \leq 3.08/\sqrt{L}$. When $L \geq 1$, we can replace 3.08 by 2.28.*

Proof. We have

$$\sum_{\ell > L} \frac{\mu^2(\ell)}{\sqrt{\ell}\varphi(\ell)} = \sum_{\ell > L} \frac{\mu^2(\ell)\sqrt{Y}}{\ell\varphi(\ell)} + \int_Y \sum_{\ell > t} \frac{\mu^2(\ell)}{\ell\varphi(\ell)} \frac{dt}{2\sqrt{t}},$$

and an appeal to Lemma 7.3 concludes the proof when $L \geq 1$. Else a direct computation of $\prod_{p \geq 2} (1 + \frac{1}{\sqrt{p}(p-1)})$ is enough. ■

LEMMA 8.2. *For all $A > 0$, we have $\sum_{a \leq A} \mu^2(a)a^{1/4}/\varphi(a) \leq 4A^{1/4}$.*

Proof. We use Theorem 1.1 and integration by parts:

$$\begin{aligned} \sum_{a \leq A} \mu^2(a)a^{1/4}/\varphi(a) &= A^{1/4} \left(\log A + c_0 + \mathcal{O}^* \left(\frac{5.9}{\sqrt{A}} \right) \right) - \frac{1}{4} \int_1^A (\log t + c_0) \frac{dt}{t^{3/4}} \\ &\quad - \frac{1}{4} \int_1^\infty (G(t) - \log t - c_0) \frac{dt}{t^{3/4}} + \mathcal{O}^* \left(5.9 \int_A^\infty \frac{dt}{t^{5/4}} \right), \end{aligned}$$

i.e.

$$\sum_{a \leq A} \mu^2(a)a^{1/4}/\varphi(a) = 4A^{1/4} + c_1 + \mathcal{O}^* \left(\frac{30}{A^{1/4}} \right)$$

where

$$c_1 = -4 + c_0 - \frac{1}{4} \int_1^\infty (G(t) - \log t - c_0) \frac{dt}{t^{3/4}} = -1.424\dots$$

We compute this constant directly. ■

8.2. Main engine. Most of the remainder term is controlled by $r_2(X; q)$ from (4.1). The p -factor of the Dirichlet series associated with its summand reads, with the notation $u = 1/p$ and $v = 1/p^s$,

$$\begin{aligned} D_p(s) &= 1 - \frac{v^2}{1-u} + \frac{uv}{1-u} = 1 - v^2 + \frac{(1-v)uv}{1-u} \\ &= (1-v^2) \left(1 + \frac{uv}{(1-u)(1+v)} \right), \end{aligned}$$

and thus $D(s) = \prod_{p \geq 2} D_p(s)$ is given by

$$(8.1) \quad D(s) = \frac{1}{\zeta(2s)} \prod_{p \geq 2} \left(1 + \frac{1}{(p-1)(p^s+1)} \right).$$

The second series is absolutely convergent (in the sense of Godement, i.e. $\sum_{p \geq 2} |1/[(p-1)(p^s+1)]| < \infty$) when $\Re s > 0$. Finding an extension of it beyond this line, or showing that this line is a natural boundary, is an open problem. The main part of $D(s)$ comes from $1/\zeta(2s)$, which implies in particular that $r_2(X; q) \ll_{q,\varepsilon} 1/X^{3/4+\varepsilon}$ under the Riemann Hypothesis. Numerical computations (see Lemma 8.4) concerning $|r_2(X; q)/X^{3/4}|$ when $q = 1$ exhibit an oscillating behavior that is as mysterious as that of $|\sqrt{X} \sum_{n > X} \mu(n)/n|$.

LEMMA 8.3. *When $X \geq (700q)^3$, we have*

$$|r_2(X; q)| \leq \frac{\prod_{p|q}(1+p^{-1/2})(1+p^{-1})}{\sqrt{X} \log(X/q^3)} \left(6.4 + \frac{18.3 \log(X/q^3)}{X^{1/6}} \sqrt{q} \right).$$

Proof. Recall (4.1):

$$r_2(X; q) = \sum_{\substack{k^2 \ell r > X \\ r|q \\ (k\ell, q) = (k, \ell) = 1}} \frac{\mu(rk)\mu^2(\ell)}{rk\varphi(k)\ell\varphi(\ell)}.$$

We split the summation over ℓ according to whether $\ell q \leq Y = X^{1/3}/50 \geq 14q$ or not, say into $\Sigma(\ell q \leq Y) + \Sigma(\ell q > Y)$. For the second sum, we use Lemmas 5.7 and 8.1 to get

$$\begin{aligned} \Sigma(\ell q > Y) &\leq \sum_{\substack{\ell > Y/q \\ (\ell, q) = 1}} \frac{\mu^2(\ell)}{\ell\varphi(\ell)} \sum_{r|q} \frac{\mu^2(r)}{r} 0.3616 \sqrt{\frac{r\ell}{X}} \\ &\leq 0.3616 \times 2.28 \frac{\prod_{p|q}(\sqrt{p} + 1)}{\sqrt{XY}} \leq \frac{0.825 \prod_{p|q}(\sqrt{p} + 1)}{\sqrt{XY}}. \end{aligned}$$

For the first sum, we use Lemma 7.18 with $q^* = \ell q$ and $K = \sqrt{X/(\ell r)}$, as well as $\theta = 50/(1970)^{2/3}$, to get

$$\begin{aligned} |\Sigma(\ell q \leq Y)| &\leq \sum_{\substack{\ell \leq Y/q \\ (\ell, q) = 1}} \frac{\mu^2(\ell)}{\ell\varphi(\ell)} \sum_{r|q} \frac{\mu^2(r)}{r} \prod_{p|q\ell} \left(1 + \frac{1}{p} \right) \frac{1.89\sqrt{r\ell}}{\sqrt{X}^{\frac{3}{2}} \log(X^{1/3}/(q\ell))} \\ &\quad + \frac{5q^{1/2}}{X^{3/4}} \sum_{r|q} \mu^2(r)r^{-1/4} \sum_{\substack{\ell \leq 50Y/(rq^2)^{1/3} \\ (\ell, q) = 1}} \frac{\mu^2(\ell)\ell^{5/4}}{\ell\varphi(\ell)} \\ &\quad + \frac{6.6q^{1/2}}{X^{3/4}} \sum_{r|q} \mu^2(r)r^{-1/4} \sum_{\substack{\ell \leq \theta Y/(rq^2)^{1/3} \\ (\ell, q) = 1}} \frac{\mu^2(\ell)\ell^{5/4}}{\ell\varphi(\ell)} \end{aligned}$$

i.e., with Lemmas 7.16 and 8.2,

$$\begin{aligned}
 |\Sigma(\ell q \leq Y)| &\leq \frac{1.89 \times 5.02 \times \frac{2}{3}}{\sqrt{X} \log(X^{1/3}/q)} \prod_{p|q} (1 + p^{-1/2})(1 + p^{-1}) \\
 &\quad + \frac{5 \cdot 50^{1/4} + 6.6\theta^{1/4}}{X^{3/4}} \prod_{p|q} (p^{1/3} + 1)4Y^{1/4} \\
 &\leq \frac{6.33}{\sqrt{X} \log(X^{1/3}/q)} \prod_{p|q} (1 + p^{-1/2})(1 + p^{-1}) \\
 &\quad + \frac{12.46}{X^{2/3}} \prod_{p|q} (1 + p^{1/4}).
 \end{aligned}$$

On putting the pieces together, we reach

$$\begin{aligned}
 |r_2(X; q)| &\leq \frac{6.33}{\sqrt{X} \log(X/q^3)} \prod_{p|q} (1 + p^{-1/2})(1 + p^{-1}) \\
 &\quad + \frac{0.825 \times 50^{1/2} + 12.46}{X^{2/3}} \prod_{p|q} (\sqrt{p} + 1),
 \end{aligned}$$

which simplifies into

$$|r_2(X; q)| \leq \frac{\prod_{p|q} (1 + p^{-1/2})(1 + p^{-1})}{\sqrt{X} \log(X/q^3)} \left(6.33 + \frac{18.3 \log(X/q^3)}{X^{1/6}} \prod_{p|q} \sqrt{p} \right). \blacksquare$$

LEMMA 8.4. *When $10^6 \leq X \leq 1.3 \cdot 10^{11}$, we have*

$$|r_2(X; 1)| \leq 1.03/X^{3/4}.$$

The quantity $|r_2(X; 1)|X^{3/4}$ is larger than 1.02 for some X inside $[8 \cdot 10^7, 9 \cdot 10^7]$ and some X inside $[358 \cdot 10^7, 359 \cdot 10^7]$.

Proof. First of all, a simple computation gives $r_2(0; 1) = 1$. Thus

$$r_2(X; 1) = 1 - \sum_{\substack{k^2 \ell \leq X \\ (k, \ell) = 1}} \frac{\mu(k)\mu^2(\ell)}{k\phi(k)\ell\phi(\ell)}.$$

In other words, we have to go through the n 's from 1 to $1.3 \cdot 10^{11}$, and for each n , we compute $a_n = \sum_{k^2 \ell = n} \mu^2(\ell)\mu(k)/(k\phi(k)\ell\phi(\ell))$, and then subtract a_n from our sum.

We wrote the script `Cr2.gp`, which we converted into a C-program with garbage collecting via `gp2c -g Cr2.gp > Cr2.gp.c`. We then started `gp -s600000000 -p1000000000`, installed the relevant function and let the program run. A very similar script was developed in Perl and used interval arithmetic and much more RAM memory. The C-program stopped at $1.1 \cdot 10^{11}$,

while the Perl script went up to $1.3 \cdot 10^{11}$ in about three months. We describe the algorithm used; the times indicated are relevant to the C-program and are of course much larger for its Perl counterpart.

We did part of these computations twice. The first batch went up to $2 \cdot 10^{10}$ by intervals of length 10^7 , while the second batch went up to $1.1 \cdot 10^{11}$ by intervals of $2 \cdot 10^7$. The computation took about sixty-five hours in the first case and about ten days in the second one. Let us give some details concerning the general process. We treated every interval $I = [(h-1) \cdot 10^7, h \cdot 10^7]$ as a whole. A loop over primes $\leq \sqrt{h \cdot 10^7}$ detected integers from I divisible by such a prime, a second loop over primes $\leq \sqrt{h \cdot 10^7}$ detected integers from I divisible by the square of such a prime, a third loop over primes $\leq (h \cdot 10^7)^{1/3}$ discarded integers from I divisible by the cube of such a prime. We add the corresponding contributions to a_n . The case when $n = p \cdot \ell^l \cdot k^2$ where p is a prime $p > \sqrt{h \cdot 10^7}$ remains, but we detect it by storing the product $P(n)$ of all the prime divisors of the previous shape for a given n , so that we recover p in the form $n/P(n)$. In this manner we avoided decomposing every integer into prime factors. Each interval of 10^7 integers took about 120 seconds to proceed: closer to 115 at the beginning of the range and closer to 125 at the end of it. The limitation of this method is the live-memory size; with blocks of size 10^7 we used a bit less than 6 Gigaoctets. (For the second batch: about 145 seconds at the beginning of the range and about 160 at the end, and 12 Gigaoctets of RAM.)

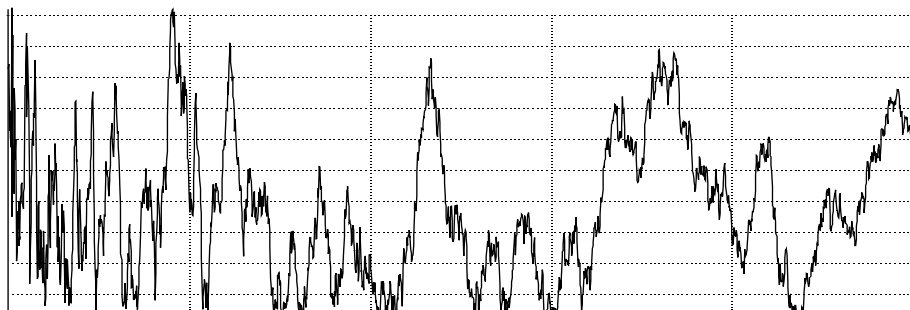


Fig. 1

Figure 1 plots an approximation of

$$\rho(t) = \max_{t < u/10^7 \leq t+1} |r_2(u; 1) \cdot u^{3/4}|$$

where t ranges over the integers from $[1, 1999]$. The points are simply joined by segments. The bottom horizontal line is $x = 0$, the top one is $x = 1$. One can see that the behavior seems to vary less when the variable becomes larger. However a plot in logarithmic scale would be somewhat less regular. Figure 2 is the plot of the value from 10^7 to $1.3 \cdot 10^{11}$.

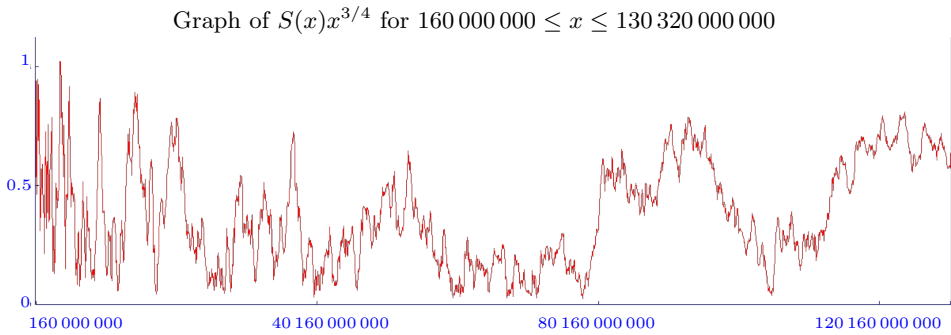


Fig. 2

Here is a sample of approximate values obtained, to enable checking.

t	$\rho(t)$	t	$\rho(t)$	t	$\rho(t)$	t	$\rho(t)$	t	$\rho(t)$	t	$\rho(t)$
1	0.8272	16	0.6010	31	0.3915	46	0.4391	61	0.4649	76	0.2475
2	0.8435	17	0.2590	32	0.3759	47	0.1331	62	0.3243	77	0.3051
3	0.6276	18	0.2065	33	0.4291	48	0.2502	63	0.1487	78	0.1468
4	0.6920	19	0.2516	34	0.4437	49	0.3323	64	0.1577	79	0.1169
5	0.6525	20	0.3682	35	0.4742	50	0.4734	65	0.3828	80	0.0629
6	0.6044	21	0.3110	36	0.7747	51	0.4998	66	0.4008	81	0.0898
7	0.5746	22	0.4012	37	0.7761	52	0.5777	67	0.2505	82	0.2136
8	1.0249	23	0.2566	38	0.8418	53	0.6271	68	0.1798	83	0.2601
9	0.9370	24	0.2861	39	0.9435	54	0.5811	69	0.1893	84	0.2455
10	0.5984	25	0.2757	40	0.6719	55	0.6851	70	0.3084	85	0.1508
11	0.4799	26	0.4099	41	0.7976	56	0.6705	71	0.3009	86	0.2774
12	0.4245	27	0.4210	42	0.9013	57	0.7157	72	0.1862	87	0.4153
13	0.7654	28	0.4534	43	0.7998	58	0.8566	73	0.2029	88	0.4321
14	0.6972	29	0.4599	44	0.6974	59	0.7082	74	0.1363	89	0.5492
15	0.5337	30	0.3937	45	0.6805	60	0.6620	75	0.2491	90	0.3114

LEMMA 8.5. For all $X \geq 49$, we have $|r_2(X; 1)| \leq 0.508/\sqrt{X}$.

Proof. We use Lemma 8.3 for $X \geq 1.1 \cdot 10^{11}$, Lemma 8.4 for $10^6 \leq X \leq 1.1 \cdot 10^{11}$, and direct computation otherwise. ■

LEMMA 8.6. For all $X \geq 2$, we have $|r_2(X; 1)| \leq \frac{13}{\sqrt{X} \log X}$.

Proof. We use Lemma 8.3 for $X \geq 1.3 \cdot 10^{11}$, Lemma 8.4 for $10^6 \leq X \leq 1.1 \cdot 10^{11}$, and direct computation otherwise. ■

9. Proof of Theorem 1.2. We prove the first estimate. We use Theorem 1.3 with \mathfrak{R}' . The function G^\sharp is $r_2(X; 1)$ defined in (4.1) and bounded

in Lemma 8.5, while $\sum_{m \leq D} |g(m)|$ is $r_1(X)$ defined in (4.2) and bounded in Lemma 6.1 with $q = 1$. When $A_0 \in [1, 2)$, we see that it is best to use A_0 around 5.526, with value 3.9800 We readily check that

$$G_1(D) = \log D + c_0 + \mathcal{O}^*(1/\sqrt{D}) \quad (2 \leq D \leq 100)$$

and

$$G_1(D) = \log D + c_0 + \mathcal{O}^*(1.5/\sqrt{D}) \quad (1 \leq D \leq 100).$$

The proof of the estimate is complete.

The second estimate is slightly more difficult to prove, though most of the work has been done. We select $A_0 = \log D$ in Theorem 1.3 together with Lemma 6.1 (with $q = 1$) and Lemma 8.6. We have (in the notation of Theorem 1.3)

(9.1)

$$\begin{aligned} \mathfrak{R} \leq & \left| \sum_{1 \leq a \leq A_0} \frac{13\sqrt{a}}{a\sqrt{D} \log(D/a)} + \frac{13\sqrt{A_0}}{\sqrt{D} \log(D/A_0)} \left(\log \frac{A_0}{A_0 - 1} - \frac{6/11}{A_0 - 1} \right) \right| \\ & + \frac{\frac{6}{11}4.95}{\sqrt{DA_0}}. \end{aligned}$$

We use $0 \leq \log \frac{A_0}{A_0 - 1} \leq 1/(A_0 - 1)$ and (we single out the term with $a = 1$) the inequality $\sum_{2 \leq a \leq A_0} 1/\sqrt{a} \leq \int_1^{A_0} dt/\sqrt{t} \leq \frac{1}{2}(\sqrt{A_0} - 1)$. This yields

$$\mathfrak{R}\sqrt{D} \leq \frac{13}{\log D} + \frac{13.2}{2} \frac{\sqrt{A_0} - 1}{\log(D/A_0)} + \frac{13\sqrt{A_0}}{\log(D/A_0)} \frac{5/11}{A_0 - 1} + \frac{\frac{6}{11}4.95}{\sqrt{A_0}}.$$

We have to add $13\sqrt{D} \int_{D/A_0}^{\infty} dt/[t^{3/2} \log t]$. This quantity is not more than $26/\log(D/A_0)$. Some numerical analysis shows that, when $D \geq \exp(21.25)$,

$$\begin{aligned} \frac{13}{\log D} + \frac{13}{2} \frac{\sqrt{A_0} - 1}{\log(D/A_0)} + \frac{13\sqrt{A_0}}{\log(D/A_0)} \frac{5/11}{A_0 - 1} + \frac{\frac{6}{11}4.95}{\sqrt{A_0}} + \frac{26}{\log(D/A_0)} \\ \leq \frac{18.4}{\sqrt{A_0}} = \frac{18.4}{\sqrt{\log D}}. \end{aligned}$$

When $D \leq \exp(21.25)$, the desired estimate follows from the first part of Theorem 1.2.

REFERENCES

[1] U. Balakrishnan and Y.-F. S. Pétermann, *Asymptotic estimates for a class of summatory functions*, J. Number Theory 70 (1998), 1–36.
 [2] U. Balakrishnan and Y.-F. S. Pétermann, *Asymptotic estimates for a class of summatory functions. II*, Comment. Math. Univ. St. Pauli 50 (2001), 89–109.
 [3] P. Berment et O. Ramaré, *Ordre moyen d’une fonction arithmétique par la méthode de convolution*, Revue de la filière mathématiques (RMS) 122 (2012), 1–15.

- [4] H. Delange, *A theorem on multiplicative arithmetic functions*, Proc. Amer. Math. Soc. 18 (1967), 743–749.
- [5] J.-M. Deshouillers et F. Dress, *Sommes de diviseurs et structure multiplicative des entiers*, Acta Arith. 49 (1988), 341–375.
- [6] L. Fousse, G. Hanrot, V. Lefèvre, P. Pélissier, and P. Zimmermann, *MPFR: A multiple-precision binary floating-point library with correct rounding*, ACM Trans. Math. Software 33 (2007), no. 2, art. 13, 15 pp.
- [7] T. Granlund and the GMP development team, *GNU MP: The GNU Multiple Precision Arithmetic Library*, 5.0.5 edition, 2012; <http://gmplib.org/>.
- [8] A. Granville and O. Ramaré, *Explicit bounds on exponential sums and the scarcity of squarefree binomial coefficients*, Mathematika 43 (1996), 73–107.
- [9] A. Granville and K. Soundararajan, *The spectrum of multiplicative functions*, Ann. of Math. (2) 153 (2001), 407–470.
- [10] B. J. Green, *Roth’s theorem in the primes*, Ann. of Math. 161 (2005), 1609–1636.
- [11] G. Halász, *On the distribution of additive and the mean values of multiplicative arithmetic functions*, Studia Sci. Math. Hungar. 6 (1971/72), 211–233.
- [12] H. Halberstam and H. E. Richert, *Mean value theorems for a class of arithmetic functions*, Acta Arith. 43 (1971), 243–256.
- [13] H. Halberstam and H. E. Richert, *Sieve Methods*, Academic Press, London, 1974.
- [14] H. A. Helfgott, *Major arcs for Goldbach’s theorem*, arXiv:1305.2897 (2013).
- [15] H. Helson, *Convergence of Dirichlet series*, in: L’analyse harmonique dans le domaine complexe (Montpellier, 1972), Lecture Notes in Math. 336, Springer, Berlin, 1973, 153–160.
- [16] S. Ikehara, *An extension of Landau’s theorem in the analytic theory of numbers*, J. Math. Phys. 10 (1931), 1–12.
- [17] J. Kaczorowski and K. Wiertelak, *Oscillations of the remainder term related to the Euler totient function*, J. Number Theory 130 (2010), 2683–2700.
- [18] J. Korevaar, *The Wiener–Ikehara theorem by complex analysis*, Proc. Amer. Math. Soc. 134 (2006), 1107–1116.
- [19] Y.-K. Lau, *Summatory formula of the convolution of two arithmetical functions*, Monatsh. Math. 136 (2002), 35–45.
- [20] B. V. Levin and A. S. Fainleib, *Application of some integral equations to problems of number theory*, Russian Math. Surveys 22 (1967), 119–204.
- [21] H. L. Montgomery, *Fluctuations in the mean of Euler’s phi function*, Proc. Indian Acad. Sci. Math. Sci. 97 (1987), 239–245.
- [22] H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika 20 (1973), 119–133.
- [23] H. L. Montgomery and R. C. Vaughan, *Mean values of multiplicative functions*, Period. Math. Hungar. 43 (2001), 199–214.
- [24] The PARI Group, Bordeaux, PARI/GP version 2.5.2, 2011; <http://pari.math.u-bordeaux.fr/>.
- [25] Y.-F. S. Pétermann, *On an estimate of Walfisz and Saltykov for an error term related to the Euler function*, J. Théor. Nombres Bordeaux 10 (1998), 203–236.
- [26] O. Ramaré, *On Snirel’man’s constant*, Ann. Scuola Norm. Sup. Pisa 21 (1995), 645–706.
- [27] O. Ramaré, *Eigenvalues in the large sieve inequality*, Funct. Approx. Comment. Math. 37 (2007), 7–35.
- [28] O. Ramaré, *From explicit estimates for the primes to explicit estimates for the Möbius function*, Acta Arith. 157 (2013), 365–379.

- [29] O. Ramaré, *Some elementary explicit bounds for two mollifications of the Moebius function*, *Funct. Approx. Comment. Math.* 49 (2013), 229–240.
- [30] O. Ramaré, *Explicit estimates on the summatory functions of the Möbius function with coprimality restrictions*, *Acta Arith.* 165 (2014), 1–10.
- [31] O. Ramaré, *Explicit estimates on several summatory functions involving the Moebius function*, *Math. Comp.* 84 (2015), 1359–1387.
- [32] O. Ramaré and I. M. Ruzsa, *Additive properties of dense subsets of sifted sequences*, *J. Théor. Nombres Bordeaux* 13 (2001), 559–581.
- [33] H. Riesel and R. C. Vaughan, *On sums of primes*, *Ark. Mat.* 21 (1983), 45–74.
- [34] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, *Illinois J. Math.* 6 (1962), 64–94.
- [35] W. A. Stein et al., *Sage Mathematics Software (Version 5.9)*, 2013; <http://www.sagemath.org>.
- [36] T. Tao, *Every odd number greater than 1 is the sum of at most five primes*, *Math. Comp.* 83 (2014), 997–1038.
- [37] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, 2nd ed., *Cours Spéc. 1*, Soc. Math. France, Paris, 1995.
- [38] J. E. van Lint and H. E. Richert, *On primes in arithmetic progressions*, *Acta Arith.* 11 (1965), 209–216.
- [39] A. Walfisz, *Weylsche Exponentialsummen in der neueren Zahlentheorie*, *Deutsch. Verlag Wiss.*, 1963.
- [40] N. Wiener, *A note on Tauberian theorems*, *Ann. of Math. (2)* 33 (1932), 787.
- [41] E. Wirsing, *Das asymptotische Verhalten von Summen über multiplikative Funktionen*, *Math. Ann.* 143 (1961), 75–102.
- [42] J. Wu, *Sur un problème de Rényi*, *Monatsh. Math.* 117 (1994), 303–322.

O. Ramaré
CNRS / Institut de Mathématiques de Marseille
Aix Marseille Université, U.M.R. 7373
Site Sud, Campus de Luminy, Case 907
13288 Marseille Cedex 9, France
E-mail: olivier.ramare@univ-amu.fr

P. Akhilesh
Harish-Chandra Research Institute
Chhatnag Road
Jhusi, Allahabad 211 019, India
E-mail: akhileshp.clt@gmail.com

