Lower bounds on the Mertens function $M(x)$ for $x \gg 2.3315 \times 10^{1656520}$

Maxie Dion Schmidt

Georgia Institute of Technology

School of Mathematics

Abstract

The Mertens function, $M(x) = \sum_{n \leq x} \mu(n)$, is classically defined to be the summatory function of the Möbius function $\mu(n)$. The Mertens conjecture which stated that $|M(x)| < C \cdot \sqrt{x}$ for all $x \geq 1$ has a well-known disproof due to Odlyzko et. al. given in the early 1980's by computation of non-trivial zeta function zeros in conjunction with integral formulas expressing $M(x)$. It is conjectured and widely believed that $M(x)/\sqrt{x}$ changes sign infinitely often and grows unbounded in the direction of both $\pm\infty$ along subsequences of integers $x \geq 1$. Our proof of this property of $q(x) \equiv M(x)/\sqrt{x}$ is not based on standard estimates of $M(x)$ by Mellin inversion, which are intimately tied to the distribution of the non-trivial zeros of the Riemann zeta function. There is a distinct stylistic flavor and element of combinatorial analysis peppered in with the standard methods from analytic number theory which distinguishes our methods from other proofs of established upper, rather than lower, bounds on $M(x)$.

Keywords and Phrases: *Möbius function sums; Mertens function; summatory function; arithmetic functions; Dirichlet inverse; Liouville lambda function; prime omega functions; prime counting functions; Dirichlet series and DGFs; asymptotic lower bounds; Mertens conjecture.*
Primary Math Subject Classifications (2010): *11N37; 11A25; 11N60; 11N64; and 11-04.*

Reference on special notation and other conventions

Symbol	Definition
$\mathbb{E}[f(x)]$	We break with the notation $\sim$ used to denote the average order of a function from Hardy and Wright. In its place, we use the clearer notation $\mathbb{E}[f(x)] = h(x)$ to denote that f has a so-called average order growth rate of $h(x)$. What this means is that $\frac{1}{x}\sum_{n\leq x} f(n) \sim h(x)$, or equivalently that $$\lim_{x\to\infty} \frac{\frac{1}{x}\sum_{n\leq x} f(n)}{h(x)} = 1.$$

$o(f), O_\alpha(g), \Omega(h)$ Using standard notation, we write that $f = o(g)$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

We adapt the stock big-O notation, writing $f = O_{\alpha_1,\ldots,\alpha_k}(g)$ for some parameters $\alpha_1, \ldots, \alpha_k$ if $f = O(g)$ subject only to some potentially fluctuating parameters that depend on the fixed α_i. In contrast to the notion of $O(g)$ as a means for stating a bound for a function from above, we borrow the Hardy-Littlewood definition of Big-Ω notation under which we may write that $f(x) = \Omega(g(x))$ if and only if

$$\limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| > 0.$$

The signed function notation $f = \Omega_\pm(g)$ means that $f = \Omega_+(g)$ and $f = \Omega_-(g)$ where

$$f(x) = \Omega_+(g(x)) \iff \limsup_{x\to\infty} \frac{f(x)}{g(x)} > 0,$$

and

$$f(x) = \Omega_-(g(x)) \iff \liminf_{x\to\infty} \frac{f(x)}{g(x)} < 0.$$

$C_k(n)$ Auxillary component functions in obtaining asymptotic bounds on $g^{-1}(n)$. These functions are defined for $k \geq 0$ and $n \geq 1$ by the formula

$$C_k(n) := \begin{cases} \varepsilon(n), & \text{if } k = 0; \\ \sum_{d\mid n} \omega(d) C_{k-1}(n/d), & \text{if } k \geq 1. \end{cases}$$

$[q^n]F(q)$ The coefficient of q^n in the power series expansion of $F(q)$ about zero.

Symbol	Definition
DGF	*Dirichlet generating function (or DGF).* Given a sequence $\{f(n)\}_{n \geq 0}$, its DGF enumerates the sequence in a different way than formal generating functions in an auxiliary variable. Namely, for $|s| < \sigma_a$, the abcissa of absolute convergence of the series, the DGF $D_f(s)$ constitutes an analytic function of s given by: $D_f(s) := \sum_{n \geq 1} f(n)/n^s$. type
$\sigma_0(n), d(n)$	The ordinary divisor function, $d(n) := \sum_{d|n} 1$.
$\varepsilon(n)$	The multiplicative identity with respect to Dirichlet convolution, $\varepsilon(n) = \delta_{n,1}$, defined such that for any arithmetic f we have that $f * \varepsilon = \varepsilon * f = f$.
$f * g$	The Dirichlet convolution of f and g, $f * g(n) := \sum_{d|n} f(d)g(n/d)$, for $n \geq 1$. This symbol for the discrete convolution of two arithmetic functions is the only notion of convolution of functions we employ within the article that is not explicitly expanded by the definition of another relvant convolution operation, e.g., integral formula or summation with exactly specified indices as input to the functions at hand.
$f^{-1}(n)$	The Dirichlet inverse of f with respect to convolution defined recursively by $f^{-1}(n) = -\frac{1}{f(1)} \sum_{\substack{d|n \\ d>1}} f(d)f^{-1}(n/d)$ provided that $f(1) \neq 0$. The inverse function, when it exists, is unique and satisfies the relations that $f^{-1} * f = f * f^{-1} = \varepsilon$.
$\lfloor x \rfloor, \lceil x \rceil$	The floor function is defined as $\lfloor x \rfloor := x - \{x\}$ where $0 \leq \{x\} < 1$ denotes the fractional part of $x \in \mathbb{R}$. The corresponding ceiling (or greatest integer) function $\lceil x \rceil := x + 1 - \{x\}$. The ceiling function is sometimes also written as $[x] \equiv \lceil x \rceil$.
$\Gamma(z), \Gamma(a,z)$	Euler's (complete) gamma function is defined for $\Re(z) > -1$ by the integral representation $$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt,$$ where $\Gamma(n+1) = n!$ for non-negative integers n. The gamma function $\Gamma(z+1) \sim \sqrt{2\pi z} \left(\frac{z}{e}\right)^z$ satisfies a generalized form of Stirling's approximation of the single factorial function. It also satisfies a functional equation of the form $\Gamma(z+1) = z \cdot \Gamma(z)$ for $\Re(z) > 0$. The corresponding notion of the (upper) incomplete gamma function is given for real $a \geq 0$ by $$\Gamma(a,z) = \int_a^\infty t^{s-1} e^{-t} dt, \Re(z) > -1.$$
$g^{-1}(n), G^{-1}(x)$	The Dirichlet inverse function, $g^{-1}(n) = (\omega + 1)^{-1}(n)$ with corresponding summatory function $G^{-1}(x) := \sum_{n \leq x} g^{-1}(n)$.
$\text{Id}_k(n)$	The power-scaled identity function, $\text{Id}_k(n) := n^k$ for $n \geq 1$.

3

Symbol	Definition
$\mathbb{1}_{\mathbb{S}}, \chi_{\text{cond}(x)}$	We use the notation $\mathbb{1}, \chi : \mathbb{N} \rightarrow \{0,1\}$ to denote indicator, or characteristic functions. In paticular, $\mathbb{1}_{\mathbb{S}}(n) = 1$ if and only if $n \in \mathbb{S}$, and $\chi_{\text{cond}}(n) = 1$ if and only if n satisfies the condition cond.
$\sum_{n \leq x}, \sum_{n \geq m}$	We use the notation $\sum_{n \leq x} f(n) \equiv \sum_{n=1}^{x} f(n) = f(1) + f(2) + \cdots + f(\lfloor x \rfloor)$, though we typically assume that x is integer valued even when these summatory functions grow as smooth functions for large x. Similarly, we often abbreviate the upper bounds on infinite sums by writing $\sum_{n \geq m} f(n) = f(m) + f(m+1) + f(m+2) + \cdots$ to mean that we are taking the infinite sum over $f(n)$ for $n \in [m, \infty)$: $\sum_{n \geq m} f(n) \equiv \sum_{n=m}^{\infty} f(n)$.
$\log_*^m(x)$	The iterated logarithm function defined recursively for integers $m \geq 0$ and any $x > 0$ taken so that the function is non-negative (e.g., with $x \geq e^e$ if $m = 2$, $x \geq e^{e^e}$ if $m = 3$, and so on) by $$\log_*^m(x) := \begin{cases} x, & \text{if } m = 0; \\ \log x, & \text{if } m = 1; \\ \log \left[\log_*^{m-1}(x) \right], & \text{if } m \geq 2. \end{cases}$$
$[n = k]_\delta$	Synonym for $\delta_{n,k}$ which is one if and only if $n = k$, and zero otherwise.
$[\text{cond}]_\delta$	For a boolean-valued cond, $[\text{cond}]_\delta$ evaluates to one precisely when cond is true, and zero otherwise.
$\lambda(n)$	The Liouville lambda function, $\lambda(n) := (-1)^{\Omega(n)}$, denotes the parity of $\Omega(n)$, the number of distinct prime factors of n counting multiplicity. That is, $\lambda(n) \in \{\pm 1\}$ with $\lambda(n) = +1$ if and only if $\Omega(n) \equiv 0 \bmod 2$.
$\gcd(m,n); (m,n)$	The greatest common divisor of m and n. Both notations for the GCD are used interchangably within the article.
$\mu(n)$	The Möbius function defined such that $\mu^2(n)$ is the indicator function of the squarefree integers, and so that $\mu(n) = (-1)^{\omega(n)}$ whenever n is squarefree, i.e., has no prime power divisors with exponent greater than one.
$M(x)$	The Mertens function which is the summatory function over $\mu(n)$, $M(x) := \sum_{n \leq x} \mu(n)$.
$\nu_p(n)$	The valuation function that extracts the maximal exponent of p in the prime factorization of n, e.g., $\nu_p(n) = 0$ if $p \nmid n$ and $\nu_p(n) = \alpha$ if $p^\alpha \| n$ (p^α exactly divides n) for p prime and $n \geq 2$.
$\omega(n), \Omega(n)$	We define the distinct prime factor counting functions as the strongly additive function $\omega(n) := \sum_{p \mid n} 1$ and the completely additive function $\Omega(n)$ by $\Omega(n) := \sum_{p^\alpha \| n} \alpha$. Equivalently, if the factorization of $n \geq 2$ is given by $n := p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, then $\omega(n) = r$ and $\Omega(n) = \alpha_1 + \cdots + \alpha_r$. By convention, we define that $\omega(1) = \Omega(1) = 0$.

Symbol	Definition
$\pi_k(x), \widehat{\pi}_k(x)$	The prime counting function variant $\pi_k(x)$ denotes the number of integers $1 \leq n \leq x$ for $x > 1$ with exactly k distinct prime factors: $\pi_k(x) := \#\{n \leq x : \omega(n) = k\}$. Similarly, the function $\widehat{\pi}_k(x) := \#\{n \leq x : \Omega(n) = k\}$ for $x \geq 2$. Montgomery and Vaughan use the alternate notation of $\sigma_k(x)$, which we intentionally avoid due to conflicting notation with other special arithmetic functions used in this article, in place of $\widehat{\pi}_k(x)$.
$\sum_{p \leq x}, \prod_{p \leq x}$	Unless otherwise specified by context, we use the index variable p to denote that the summation (product) is to be taken only over prime values within the summation bounds.
$P(s)$	For complex s with $\Re(s) > 1$, we define the prime zeta function to be the DGF $P(s) = \sum_{p \text{ prime}} p^{-s}$. This function has an analytic continuation to $\Re(s) \in (0,1)$ with a logarithmic singularity near $s := 1$: $P(1 + \varepsilon) = -\log \varepsilon + C + O(\varepsilon)$.
$\sigma_\alpha(n)$	The generalized sum-of-divisors function, $\sigma_\alpha(n) := \sum_{d|n} d^\alpha$, defined for any $n \geq 1$ and $\alpha \in \mathbb{C}$.
$\begin{bmatrix} n \\ k \end{bmatrix}$	The unsigned Stirling numbers of the first kind, $\begin{bmatrix} n \\ k \end{bmatrix} = (-1)^{n-k} \cdot s(n,k)$.
$\sim, \approx, \gtrsim, \lesssim, \gg, \ll$	See the first section of the introduction to the article for clarification of the asymptotic notation we employ in the article.
$\sum'_{n \leq x}$	We denote by $\sum'_{n \leq x} f(n)$ the summatory function of f at x minus $\frac{f(x)}{2}$ if $x \in \mathbb{Z}$.
$\tau_m(n)$	Let $\tau_m(n) \equiv \mathbb{1}_{*_m}(n)$ denote the m-fold Dirichlet convolution of one with itself at n, e.g., the arithmetic function with DGF given by $\zeta(s)^m$. Note that $\tau_2(n)$ yields the divisor function, $d(n) \equiv \sigma_0(n)$, sometimes also denoted $\tau(n)$ – a distinction we avoid to remove confusion with other standard notation for Ramanujan's tau function. Many standard references also use the notation $d_k(n)$ in place of $\tau_k(n)$ for clarity.
$\zeta(s)$	The Riemann zeta function, defined by $\zeta(s) := \sum_{n \geq 1} n^{-s}$ when $\Re(s) > 1$, and by analytic continuation to the entire complex plane with the exception of a simple pole at $s = 1$.

1 Preface: Explanations of unconventional notions and pre-conceptions of asymptotics and traditional notation for asymptotic relation symbols

We note that the next careful explanation in the subtle distinctions in our usage of what we consider to be traditional notation for asymptotic relations are key to understanding our choices of upper and lower bound expressions given throughout the article. Thus, to avoid any confusion that may linger as we begin to state our new results and bounds on the functions we work with in this article, we preface the article starting with this section detailing our precise definitions, meanings and assumptions on the uses of certain symbols, operators, and relations that we use to convey the growth rates of arithmetic functions on their domain of x when x is taken to be very large, and tending to infinity [13, *cf.* §2] [3].

1.1 Average order, similarity and approximation of asymptotic growth rates of quantities

1.1.1 Similarity and average order (expectation)

First, we say that two functions $A(x), B(x)$ satisfy the relation $A \sim B$ if

$$\lim_{x \to \infty} \frac{A(x)}{B(x)} = 1.$$

It is sometimes standard to express the *average order* of an arithmetic function $f \sim h$ that may actually oscillate, or say have value of one infinitely often, in the cases that $\frac{1}{x} \cdot \sum_{n \leq x} f(n) \sim h(x)$.

For example, in the acceptably classic language of [7] we would normally write that $\Omega(n) \sim \log \log n$, even though technically, $1 \leq \Omega(n) \leq \frac{\log n}{\log 2}$. To be absolutely clear about notation, we choose to be explicit and not re-use the $\sim$ relation by instead writing $\mathbb{E}[f(x)] = h(x)$ to denote that f has a limiting average order growing at the rate of h. A related conception of f having *normal order* of g holds whenever

$$f(n) = (1 + o(1))g(n), \text{a.e.}$$

1.1.2 Approximation

We choose the convention to write that $f(x) \approx g(x)$ if $|f(x) - g(x)| = O(1)$. That is, we write $f(x) \approx g(x)$ to denote that f is approximately equal to g at x modulo at most a small constant difference between the functions. For example, for a non-decreasing arithmetic function $f \geq 0$ and some (finite) upper bound $M > 1$, we can express that

$$\sum_{n \leq M} f(n) \approx \int_1^M f(x)dx,$$

provided that f is integrable on $[1, M]$. The previous approximation generalizes (on shorter intervals of integration) the notion of the so-called familiar *integral test* from introductory calculus.

This convention also happens to be useful in applying standard analytic number theoretic constructs of approximating the growth rates of arithmetic functions, and finite bounded summations of them, by smooth functions and in using Abel summation. The formula we

prefer for the Abel summation variant of summation by parts of finite sums of a product of two functions is stated as follows [1, *cf.* §4.3] *:

Proposition 1.1 (Abel Summation Integral Formula). *Suppose that $t > 0$ is real-valued, and that $A(t) \sim \sum_{n \leq t} a(n)$ for some weighting arithmetic function $a(n)$ with $A(t)$ continuously differentiable on $(0, \infty)$. Furthermore, suppose that $b(n) \sim f(n)$ with f a differentiable function of $n \geq 0$ – that is, $f'(t)$ exists and is smooth for all $t \in (0, \infty)$. Then for $0 \leq y < x$, where we typcially assume that the bounds of summation satisfy $x, y \in \mathbb{Z}^+$, we have that*

$$\sum_{y < n \leq x} a(n)b(n) \sim A(x)b(x) - A(y)b(y) - \int_y^x A(t)f'(t)dt.$$

Remark 1.2. The classical proof of the Abel summation formula given in Apostol's book has an alternate proof method noted in Section 4.3 of this reference. In particular, since $A(x)$ is a step function with jump of $a(n)$ at each integer-valued $n \geq 1$, the integral formula stated in Proposition 1.1 can be expressed in the following Riemann-Stieltjes integral notation:

$$\sum_{y < n \leq x} a(n)b(n) = \int_y^x f(t)dA(t).$$

A notable special case yields the integral approximation to summations we stated above where $[t]$ is the *nearest integer function*:

$$\sum_{y < n \leq x} f(n) = f(x)[x] - f(y)[y] - \int_y^x [t]f'(t)dt.$$

1.1.3 Vinogradov's notation for asymptotics

We use the conventional relations $f(x) \gg g(x)$ and $h(x) \ll r(x)$ to symbolically express that we expect f to be "substantially" larger than g, and h to be "significantly" smaller, in asymptotic order (e.g., rate of growth when x is large). In practice, we adopt a somewhat looser definition of these symbols which allows $f \gg g$ and $h \ll r$ provided that there are constants $C, D > 0$ such that whenever x is sufficiently large we have that $f(x) \geq C \cdot g(x)$ and $h(x) \leq D \cdot r(x)$. This notation is sometimes called *Vinogradov's asymptotic notation*. Another way of expressing our meaning of these relations is by writing

$$f \gg g \iff g = O(f),$$

and

$$h \ll r \iff r = \Omega(h),$$

using Knuth's well-trodden style of big-O (and Landau notation) and big-Ω (Hardy-Littlewood notation) notation from theoretical computer science and the analysis of algorithms. However, we prefer the standard notation and conventions from mathematical analysis in the form of $\gg, \ll$ be used to express our bounds within this article.

---

*Compare to the exact formula for *summation by parts* of any arithmetic functions, u_n, v_n, stated as in [13, §2.10(ii)] for $U_j := u_1 + u_2 + \cdots + u_j$ when $j \geq 1$:

$$\sum_{j=1}^{n-1} u_j \cdot v_j = U_{n-1}v_n + \sum_{j=1}^{n-1} U_j \left(v_j - v_{j+1}\right), n \geq 2.$$

1.2 An unconventional pair of asymptotic relations employed to drop lower-order terms in upper and lower bounds on arithmetic functions

We say that $h(x) \overset{\blacktriangle}{\gtrsim} r(x)$ if $h \gg r$ as $x \to \infty$, and define the relation $\overset{\blacktriangle}{\precsim}$ similarly as $h(x) \overset{\blacktriangle}{\precsim} r(x)$ if $h \ll r$ as $x \to \infty$. This usage of the notation of $\overset{\blacktriangle}{\gtrsim}, \overset{\blacktriangle}{\precsim}$ intentionally breaks with the usual conventions for the use of the standard ralations $\gtrsim, \precsim$ where these relations are employed elsewhere. Our distinct, intentional usage of these relations in our different context is intended to simplify the ways we express otherwise tricky and complicated expressions for upper and lower bounds that hold only exactly in limiting cases where x is large as $x \to \infty$.

That is to say that our convention is particularly convenient for expressing upper and lower bounds on functions given by asymptotically dominant main terms in the expansion of more complicated symbolic expansions. Where possible, we aim to carefully distinguish where these operators are applied to signed versus unsigned function variants.

An example motivating this usage of these relations clarifies the point of making this distinction.

Example 1.3. Suppose that exactly

$$f(x) \geq -(\log \log \log x)^2 + 3 \times 10^{1000000} \cdot (\log \log \log x)^{1.999999999} + E(x),$$

where $E(x) = o\left((\log \log \log x)^2\right)$ and the unusually complicated expression for $E(x)$ requires more than 100000 ascii characters to typeset accurately. Then naturally, we prefer to work with only the expression for the asymptotically dominant main term in the lower bounds stated above. Note that since this main term contribution does not dominate the bound until x is very large, so that replacing the right-hand-side expression with just this term yields an invalid inequality except for in limiting cases. In this instance, we prefer to write

$$f(x) \overset{\blacktriangle}{\gtrsim} -(\log \log \log x)^2, \text{ as } x \to \infty,$$

or more conventionally that

$$|f(x)| \overset{\blacktriangle}{\gtrsim} (\log \log \log x)^2, \text{ as } x \to \infty,$$

which indicates that this substantially simplifed form of the lower bound on f holds as $x \to \infty$. Thus it is problematic to only write that

$$f(x) \geq -(\log \log \log x)^2,$$

since there is a substantial (however, asymptotically negligible) initial range of $x \geq 1$ where this lower bound is invalid as stated in the previous equation. The use of the new (modified) notation for $\overset{\blacktriangle}{\gtrsim}$ is intended to capture both that we are conveying a lower bound for the function, and crucially that this lower bound is valid only when x is very large, i.e., in some sense that the lower bound holds in the same sense as the relation $\sim$: for example, entending a notion similar to $|f(x)| \geq g(x)$ with $g(x) \sim h(x)$. This is a subtle distinction that comes into play when we later use it to state lower bounds in our new results.

Remark 1.4 (Emphasizing the rationale of the use of the new notation). Hence, we emphasize that our new uses of these traditional symbols, $\overset{\blacktriangle}{\gtrsim}, \overset{\blacktriangle}{\precsim}$, are as asymptotic relations defined to simplify our results by dropping expressions involving more precise, exact terms that are nonetheless asymptotically insignificant, to obtain accurate statements in limiting

cases of large x that hold as $x \to \infty$. In principle, this convention allows us to write out simplified bounds that still capture the most simple essence of the upper or lower bound as we choose to view it when x is very large. This take on the new meanings denoted by $\overset{\blacktriangle}{\gtrsim}, \overset{\blacktriangle}{\lesssim}$ is particularly powerful and is utilized in this article when we express many lower bound estimates for functions that would otherwise require literally pages of typeset symbols to state exactly, but which have simple enough formulae when considered as bounds that hold in this type of limiting asymptotic context.

1.3 Asymptotic expansions and uniformity

Because a subset of the results we cite that are proved in the references (e.g., borrowed from Chapter 7 of [11]) provide statements of asymptotic bounds that hold *uniformly* for x large, though in a bounded range depending on parameters, we need to briefly make precise what our preconceptions are of this terminology. We introduce the notation for asymptotic expansions of a function $f : \mathbb{R} \to \mathbb{R}$ from [13, §2.1(iii)].

1.3.1 Ordinary asymptotic expansions of a function

Let $\sum_n a_n x^{-n}$ denote a formal power series expansion in x where we ignore any necessary conditions on convergence of the series. For each integer $n \geq 1$, suppose that

$$f(x) = \sum_{s=0}^{n-1} a_s x^{-s} + O(x^{-n}),$$

as $|x| \to \infty$ where this limiting bound holds for $x \in \mathbb{X}$ in some unbounded set $\mathbb{X} \subseteq \mathbb{R}, \mathbb{C}$. When such a bound holds, we say that $\sum_s a_s x^{-s}$ is a *Poincaré asymptotic expansion*, or just *asymptotic series expansion*, of $f(x)$ as $x \to \infty$ in the fixed set $\mathbb{X}$. The condition in the previous equation is equivalent to writing

$$f(x) \sim a_0 + a_1 x^{-1} + a_2 x^{-2} + \cdots ; x \in \mathbb{X}, \text{ for } |x| \to \infty.$$

The prior two characterizations of an asymptotic expansion for f are also equivalent to the statement that

$$x^n \left(f(x) - \sum_{s=0}^{n-1} a_s x^{-s} \right) \xrightarrow{x \to \infty} a_n.$$

1.3.2 Uniform asymptotic expansions of a function

Let the set $\mathbb{X}$ from the definition in the last subsection correspond to a closed sector of the form

$$\mathbb{X} := \{x \in \mathbb{C} : \alpha \leq \arg(x) \leq \beta\}.$$

Then we say that the asymptotic property

$$f(x) = \sum_{s=0}^{n-1} a_s x^{-s} + O(x^{-n}),$$

from before holds *uniformly* with respect to $\arg(x) \in [\alpha, \beta]$ as $|x| \to \infty$.

Another useful, important notion of uniform asymptotic bounds is taken with respect to some parameter u (or set of parameters, respectively) that ranges over the point set (point

sets, respectively) $u \in \mathbb{U}$. In this case, if we have that the u-parameterized expressions

$$\left| x^n \left(f(u, x) - \sum_{s=0}^{n-1} a_s(u) x^{-s} \right) \right|,$$

are bounded for all integers $n \geq 1$ for $x \in \mathbb{X}$ as $|x| \to \infty$, then we say that the asymptotic expansion of f holds uniformly for $u \in \mathbb{U}$. Note that the function $f \equiv f(u, x)$ and the asymptotic series coefficients $a_s(u)$ now may have an implicit dependence on the parameter u. If the previous boundedness condition holds for all positive integers n, we write that

$$f(u, x) \sim \sum_{s=0}^{\infty} a_s(u) x^{-s}; x \in \mathbb{X}, \text{ as } |x| \to \infty,$$

and say that this asymptotic expansion, or bound, holds *uniformly with respect to $u \in \mathbb{U}$*.

2   An introduction to the Mertens function – definition, properties, known results and conjectures

Suppose that $n \geq 1$ is a natural number with factorization into distinct primes given by $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}$. We define the *Möebius function* to be the signed indicator function of the squarefree integers:

$$\mu(n) = \begin{cases} 1, & \text{if } n = 1; \\ (-1)^k, & \text{if } \alpha_i = 1, \forall 1 \leq i \leq k; \\ 0, & \text{otherwise.} \end{cases}$$

There are many known variants and special properties of the Möebius function and its generalizations [15, *cf.* §2], however, for our purposes we seek to explore the properties and asymptotics of weighted summatory functions over $\mu(n)$. The Mertens summatory function, or *Mertens function*, is defined as [17, A002321]

$$M(x) = \sum_{n \leq x} \mu(n), \ x \geq 1,$$

$$\longmapsto \{1, 0, -1, -1, -2, -1, -2, -2, -2, -1, -2, -2, -3, -2, -1, -1, -2, -2, -3, -3, -2, -1, -2, -2\}$$

A related function which counts the number of *squarefree* integers than x sums the average order of the Möbius function as [17, A013928]

$$Q(n) = \sum_{n \leq x} |\mu(n)| \sim \frac{6x}{\pi^2} + O\left(\sqrt{x}\right).$$

It is known that the asymptotic density of the positively versus negatively weighted sets of squarefree numbers are in fact equal as $x \to \infty$:

$$\mu_+(x) = \frac{\#\{1 \leq n \leq x : \mu(n) = +1\}}{Q(x)} = \mu_-(x) = \frac{\#\{1 \leq n \leq x : \mu(n) = -1\}}{Q(x)} \xrightarrow[n \to \infty]{} \frac{3}{\pi^2}.$$

While this limiting law suggests an even bias for the Mertens function, in practice $M(x)$ has a noted negative bias in its values, and the actual local oscillations between the approximate densities of the sets $\mu_\pm(x)$ lend an unpredictable nature to the function and its characteistic oscillatory sawtooth shaped plot.

2.1   Properties

The well-known approach to evaluating the behavior of $M(x)$ for large $x \to \infty$ results from a formulation of this summatory function as a predictable exact sum involving x and the non-trivial zeros of the Riemann zeta function for all real $x > 0$. This formula is easily expressed via an inverse Mellin transformation over the reciprocal zeta function. In particular, we notice that since by Perron's formula we have

$$\frac{1}{\zeta(s)} = \int_1^\infty \frac{s \cdot M(x)}{x^{s+1}} dx,$$

we then obtain that

$$M(x) = \lim_{T \to \infty} \frac{1}{2\pi i} \int_{T-i\infty}^{T+i\infty} \frac{x^s}{s \cdot \zeta(s)} ds.$$

This representation along with the standard Euler product representation for the reciprocal zeta function leads us to the exact expression for $M(x)$ when $x > 0$ given by the next theorem.

Theorem 2.1 (Analytic Formula for $M(x)$). *Assuming the RH, we can show that there exists an infinite sequence $\{T_k\}_{k\geq 1}$ satisfying $k \leq T_k \leq k+1$ for each k such that for any $x \in \mathbb{R}_{>0}$*

$$M(x) = \lim_{k \to \infty} \sum_{\substack{\rho:\zeta(\rho)=0 \\ |\Im(\rho)|<T_k}} \frac{x^\rho}{\rho \cdot \zeta'(\rho)} - 2 + \sum_{n \geq 1} \frac{(-1)^{n-1}}{n \cdot (2n)!\zeta(2n+1)} \left(\frac{2\pi}{x}\right)^{2n} + \frac{\mu(x)}{2} \left[x \in \mathbb{Z}^+\right]_\delta .$$

An unconditional bound on the Mertens function due to Walfisz (1963) states that there is an absolute constant $C > 0$ such that

$$M(x) \ll x \cdot \exp\left(-C \cdot \log^{3/5}(x)(\log\log x)^{-3/5}\right).$$

Under the assumption of the RH, Soundararajan in 2009 proved new updated estimates bounding $M(x)$ for large x of the following forms [18]:

$$M(x) \ll \sqrt{x} \cdot \exp\left(\log^{1/2}(x)(\log\log x)^{14}\right),$$
$$M(x) = O\left(\sqrt{x} \cdot \exp\left(\log^{1/2}(x)(\log\log x)^{5/2+\epsilon}\right)\right), \quad \forall \epsilon > 0.$$

Other explicit bounds due to the article by Kotnik include the following simpler estimates for the Mertens function when x is sufficiently large:

$$|M(x)| < \frac{x}{4345}, \quad \forall x > 2160535,$$
$$|M(x)| < \frac{0.58782 \cdot x}{\log^{11/9}(x)}, \quad \forall x > 685.$$

2.2 Conjectures

The *Riemann Hypothesis* (RH) is equivalent to showing that $M(x) = O\left(x^{1/2+\varepsilon}\right)$ for any $0 < \varepsilon < \frac{1}{2}$. It is still unresolved whether

$$\limsup_{x \to \infty} |M(x)|/\sqrt{x} = \infty,$$

although computational evidence suggests that this is a likely conjecture [10, 8]. There is a rich history to the original statement of the *Mertens conjecture* which states that

$$|M(x)| < c \cdot x^{1/2}, \quad \text{some constant } c > 0,$$

which was first verified by Mertens for $c = 1$ and $x < 10000$, although since its beginnings in 1897 has since been disproved by computation by Odlyzko and té Riele in the early 1980's.

There are a number of other interesting unsolved and at least somewhat accessible open problems related to the asymptotic behavior of $M(x)$ at large x. It is believed that the sign of $M(x)$ changes infinitely often. That is to say that it is widely believed that $M(x)$ is ocsillatory and exhibits a negative bias insomuch as $M(x) < 0$ more frequently than $M(x) > 0$ over all $x \in \mathbb{N}$. One of the most famous still unanswered questions about the Mertens function concerns whether $|M(x)|/\sqrt{x}$ is unbounded on the natural numbers. In particular, the precise statement of this problem is to produce an affirmative answer whether $\limsup_{x \to \infty} |M(x)|/\sqrt{x} = +\infty$, or equivalently whether there is an infinite sequence of natural numbers $\{x_1, x_2, x_3, \ldots\}$ such that $M(x_i)x_i^{-1/2}$ grows without bound along this subsequence.

Extensive computational evidence has produced a conjecture due to Gonek that in fact the limiting behavior of $M(x)$ satisfies that

$$\limsup_{x \to \infty} \frac{|M(x)|}{\sqrt{x}(\log\log x)^{5/4}},$$

corresponds to some bounded constant. To date an exact rigorous proof that $M(x)/\sqrt{x}$ is unbounded still remains elusive, though there is suggestive probabilistic evidence of this property established by Ng in 2008. We cite that prior to this point it is known that [14, *cf.* §4.1]

$$\limsup_{x \to \infty} \frac{M(x)}{\sqrt{x}} > 1.060 \qquad (\text{now } 1.826054),$$

and

$$\liminf_{x \to \infty} \frac{M(x)}{\sqrt{x}} < -1.009 \qquad (\text{now } -1.837625),$$

although based on work by Odlyzko and te Riele it seems probable that each of these limits should be $\pm\infty$, respectively [12, 9, 10, 8]. It is also known that $M(x) = \Omega_\pm(\sqrt{x})$ and $M(x)/\sqrt{x} = \Omega_\pm(1)$.

3 Introduction to our new methodology: An concrete approach to bounding $M(x)$ from below

3.1 Summing series over Dirichlet convolutions

Theorem 3.1 (Summatory functions of Dirichlet convolutions). *Let $f, g : \mathbb{Z}^+ \to \mathbb{C}$ be any arithmetic functions such that $f(1) \neq 0$. Suppose that $F(x) := \sum_{n \leq x} f(n)$ and $H(x) := \sum_{n \leq x} h(n)$ denote the summatory functions of f, g, respectively, and that $F^{-1}(x)$ denotes the summatory function of the Dirichlet inverse $f^{-1}(n)$ of f, i.e., the unique arithmetic function such that $f * f^{-1} = \varepsilon$ where $\varepsilon(n) = \delta_{n,1}$ is the multiplicative identity with respect to Dirichlet convolution. Then, letting the counting function $\pi_{f*h}(x)$ be defined as in the first equation below, we have the following equivalent expressions for the summatory function of $f * h$ for integers $x \geq 1$:*

$$\pi_{f*h}(x) = \sum_{n \leq x} \sum_{d \mid n} f(d) h(n/d)$$

$$= \sum_{d \leq x} f(d) H\left(\left\lfloor \frac{x}{d} \right\rfloor\right)$$

$$= \sum_{k=1}^{x} H(k) \left[F\left(\left\lfloor \frac{x}{k} \right\rfloor\right) - F\left(\left\lfloor \frac{x}{k+1} \right\rfloor\right) \right].$$

Moreover, we can invert the linear system determining the coefficients of $H(k)$ for $1 \leq k \leq x$ naturally to express $H(x)$ as a linear combination of the original left-hand-side summatory function as:

$$H(x) = \sum_{j=1}^{x} \pi_{f*h}(j) \left[F^{-1}\left(\left\lfloor \frac{x}{j} \right\rfloor\right) - F^{-1}\left(\left\lfloor \frac{x}{j+1} \right\rfloor\right) \right]$$

$$= \sum_{n=1}^{x} f^{-1}(n) \pi_{f*h}\left(\left\lfloor \frac{x}{n} \right\rfloor\right).$$

Corollary 3.2 (Convolutions Arising From Möbius Inversion). *Suppose that g is an arithmetic function with $g(1) \neq 0$. Define the summatory function of the convolution of g with μ by $\widetilde{G}(x) := \sum_{n \leq x} (g * \mu)(n)$. Then the Mertens function equals*

$$M(x) = \sum_{k=1}^{x} \left(\sum_{j = \left\lfloor \frac{x}{k+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{k} \right\rfloor} g^{-1}(j) \right) \widetilde{G}(k), \forall x \geq 1.$$

Corollary 3.3 (A motivating special case). *We have exactly that for all $x \geq 1$*

$$M(x) = \sum_{k=1}^{x} (\omega + 1)^{-1}(k) \left[\pi\left(\left\lfloor \frac{x}{k} \right\rfloor\right) + 1 \right]. \tag{1}$$

3.2 Elaborating on construction behind the motivating special case

We can compute the first few terms for the Dirichlet inverse sequence of $g(n) := \omega(n) + 1$ from Corollary 3.3 numerically for the first few sequence values as

$$\{g^{-1}(n)\}_{n \geq 1} = \{1, -2, -2, 2, -2, 5, -2, -2, 2, 5, -2, -7, -2, 5, 5, 2, -2, -7, -2, -7, 5, 5, -2, 9, \ldots\}.$$

The sign of these terms is given by $\operatorname{sgn}(g^{-1}(n)) = \frac{g^{-1}(n)}{|g^{-1}(n)|} = \lambda(n)$ (see Proposition 4.2) – a useful property inherited by the distinctly additive nature of the component function $\omega(n)$. We will still require substantially simpler asymptotic formulae for $g^{-1}(n)$ than what complications are suggested by inspection of the initial numerical calculations of this sequence. It does happen that we can find fruitful and combinatorially meaningful ways to express asymptotics for this special inverse sequence. Consider first the following motivating conjecture:

A01-2020-04-26

Conjecture 3.4. *Suppose that $n \geq 1$ is a squarefree integer. We have the following properties characterizing the Dirichlet inverse function $g^{-1}(n) = (\omega + 1)^{-1}(n)$ over these integers:*

(A) $g^{-1}(1) = 1$;

(B) $\operatorname{sgn}(g^{-1}(n)) = \mu(n) \equiv \lambda(n)$;

(C) We can write the inverse function at squarefree n as

$$g^{-1}(n) = \sum_{m=0}^{\omega(n)} \binom{\omega(n)}{m} \cdot m!.$$

We illustrate parts (B)–(C) of this conjecture clearly using Table T.1 given on page 51 of the appendix section.

The realization that the beautiful and remarkably simple form of property (C) in Conjecture 3.4 holds for all squarefree $n \geq 1$ motivates our pursuit of formulas for the inverse functions $g^{-1}(n)$ based on the configuration of the exponents in the prime factorization of any $n \geq 2$. The summation methods we employ in Section 6 to weight sums of our arithmetic functions according to the sign of $\lambda(n)$ (or parity of $\Omega(n)$) is reminiscent of the combinatorially motivated sieve methods in [4, §17].

Remark 3.5 (Comparison to formative methods for bounding $M(x)$). Note that since the DGF of $\omega(n)$ is given by $D_\omega(s) = P(s)\zeta(s)$ where $P(s)$ is the *prime zeta function*, we do have a Dirichlet series for the inverse functions to invert coefficient-wise using more classical contour integral methods*. However, the uniqueness to our new methods is that our new approach does not rely on typical constructions for bounding $M(x)$ based on estimates of the non-trivial zeros of the Riemann zeta function that have so far to date been employed to bound the Mertens function from above. That is, we will instead take a more combinatorial tack to investigating bounds on this inverse function sequence in the coming sections. By Corollary 3.3, once we have established bounds on this $g^{-1}(n)$ and its summatory function, we will be able to formulate new lower bounds (in the limit supremum sense) on $M(x)$.

3.3 Fixing an exact expression for $M(x)$ through special sequences of arithmetic functions

From this point on, we fix the Dirichlet invertible function $g(n) := \omega(n) + 1$ and denote its inverse with respect to Dirichlet convolution by $g^{-1}(n) = (\omega + 1)^{-1}(n)$. For natural

---

*E.g., using [1, §11]

$$f(n) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \frac{n^{\sigma + \imath t}}{\zeta(\sigma + \imath t)(P(\sigma + \imath t) + 1)}, \sigma > 1.$$

Fröberg has also previously done some preliminary investigation as to the properties of the inversion to find the coefficients of $(1 + P(s))^{-1}$ in [5].

numbers $n \geq 1, k \geq 0$, let

$$C_k(n) := \begin{cases} \varepsilon(n) = \delta_{n,1}, & \text{if } k = 0; \\ \sum_{d|n} \omega(d) C_{k-1}(n/d), & \text{if } k \geq 1. \end{cases}$$

By Möbius inversion (see Lemma 6.1), we have that

$$(g^{-1} * 1)(n) = \lambda(n) \cdot C_{\Omega(n)}(n), \forall n \geq 1. \tag{2}$$

We have limiting asymptotics on these functions given by the following theorem:

Theorem 3.6 (Asymptotics for the functions $C_k(n)$). *The function $\sigma_0 * \tau_m$ is multiplicative with values at prime powers given by*

$$(\sigma_0 * \tau_m)(p^\alpha) = \binom{\alpha + m + 1}{m + 1}. \tag{3}$$

We have the following asymptotic base cases for the functions $C_k(n)$:

$$C_1(n) \sim \log \log n$$

$$C_2(n) \sim \frac{\sigma_0(n)n}{\log n} + O(\log \log n)$$

$$C_3(n) \sim -\frac{(\sigma_0 * 1)(n)n^2}{\log n} + O\left(n \cdot \log \log n\right).$$

For all $k \geq 4$, we obtain that the dominant asymptotic term and the error bound terms for $C_k(n)$ are given by

$$C_k(n) \sim (\sigma_0 * \tau_{k-2})(n) \times \frac{(-1)^k n^{k-1}}{(\log n)^{k-1}(k-1)!} + O_k\left(\frac{n^{k-2}}{(k-2)!} \cdot \frac{(\log \log n)^{k-2}}{(\log n)^{k-2}}\right), \quad \text{as } n \to \infty.$$

Then we can prove (see Corollary 6.8) that

$$g^{-1}(n) \sim \lambda(n) \times \sum_{d|n} C_{\Omega(d)}(d).$$

Notice that this formula is substantially easier to evaluate than the corresponding sums in (2) given directly by Möbius inversion – and hence, we prefer to work with bounds on it we prove as new results rather than the more complicated exact formula from the cited equation above. The last result in turn implies that

$$G^{-1}(x) \stackrel{\blacktriangle}{\sim} \sum_{n \leq x} \lambda(n) \cdot C_{\Omega(n)}(n) \times \sum_{d=1}^{\lfloor \frac{x}{n} \rfloor} \lambda(d). \tag{4}$$

In light of the fact that (by an integral-based interpretation of integer convolution using summation by parts, see Proposition 7.1)

$$M(x) \sim G^{-1}(x) + \sum_{k=1}^{x/2} G^{-1}(k) \cdot \frac{x}{k^2 \log(x/k)},$$

the formula in (4) implies that we can establish new *lower bounds* on $M(x)$ by appropriate estimates of the summatory function $G^{-1}(x)$ where trivially we have the bounded inner sums $L_0(x) := \sum_{n \leq x} \lambda(n) \in [-x, x]$ for all $x \geq 2$.

As explicit lower bounds for $M(x)$ along particular subsequences are not obvious, and are historically ellusive non-trivial features of the function to obtain as we expect sign changes of this function infinitely often, we find this approach to be an effective one. Now, having motivated why we must carefully estimate the $G^{-1}(x)$ bounds using our new methods, we will require the bounds suggested in the next section to work at bounding the summatory functions, $G^{-1}(x)$, for large x as $x \to \infty$.

3.4 Enumerative (or counting based) DGFs from Mongomery and Vaughan

Our inspiration for the new bounds found in the last sections of this article allows us to sum non-negative arithmetic functions weighted by the Liouville lambda function, $\lambda(n) = (-1)^{\Omega(n)}$. In particular, it uses a hybrid generating function and DGF method under which we are able to recover "good enough" asymptotics about the summatory functions that encapsulate the parity of $\lambda(n)$, e.g., through the summatory functions, $\widehat{\pi}_k(x)$. The precise statement of the theorem that we transform for these new bounds is re-stated as Theorem 3.7 below.

Theorem 3.7 (Montgomery and Vaughan, §7.4). *Recall that we have defined*

$$\widehat{\pi}_k(x) := \#\{n \leq x : \Omega(n) = k\}.$$

For $R < 2$ we have that

$$\widehat{\pi}_k(x) = \mathcal{G}\left(\frac{k-1}{\log\log x}\right) \frac{x}{\log x} \frac{(\log\log x)^{k-1}}{(k-1)!}\left(1 + O_R\left(\frac{k}{(\log\log x)^2}\right)\right),$$

uniformly for $1 \leq k \leq R\log\log x$ where

$$\mathcal{G}(z) := \frac{F(1,z)}{\Gamma(z+1)} = \frac{1}{\Gamma(z+1)} \times \prod_p \left(1 - \frac{z}{p}\right)^{-1}\left(1 - \frac{1}{p}\right)^z.$$

The precise formulations of the inverse function asymptotics proved in Section 6 depend on being able to form significant lower bounds on the summatory functions of an always positive arithmetic function weighted by $\lambda(n)$. The next theorem, proved carefully in Section 5, is the primary starting point for our new asymptotic lower bounds.

Theorem 3.8 (Generating functions of symmetric functions). *We obtain lower bounds of the following form on the function $\mathcal{G}(z)$ from Theorem 3.7 for $A_0 > 0$ an absolute constant, for $C_0(z)$ a strictly linear function only in z, and where we must take $0 \leq z \geq 1$:*

$$\mathcal{G}(z) \geq A_0 \cdot (1-z)^3 \cdot C_0(z)^z.$$

It suffices to take the components to the bound in the previous equation as

$$A_0 = \frac{2^{9/16}\exp\left(-\frac{55}{4}\log^2(2)\right)}{(3e\log 2)^3 \cdot \Gamma\left(\frac{5}{2}\right)} \approx 3.81296 \times 10^{-6}$$

$$C_0(z) = \frac{4(1-z)}{3e\log 2}.$$

In particular, with $0 \leq z \leq 1$ and $z \equiv z(k,x) = \frac{k-1}{\log\log x}$, we have that

$$\widehat{\pi}_k(x) \geq \frac{A_0 \cdot x}{\log x \cdot (\log\log x)^4 \cdot (k-1)!} \cdot \left(\frac{4}{3e\log 2}\right)^k.$$

17

3.5 Cracking the classical unboundedness result

In Section 7, we provide the culmination of what we build up to in the proofs established in prior sections of the article. Namely, we prove the form of an explicit limiting lower bound for the summatory function, $G^{-1}(x) := \sum_{n \leq x} g^{-1}(n)$, along a specific subsequence. What we obtain is the following important summary corollary resolving the classical question of the unboundedness of the scaled function Mertens function $|M(x)|/\sqrt{x}$ in the limit supremum sense:

Corollary 3.9 (Bounds for the classically scaled Mertens function). *Let* $u_0 := e^{e^{e^e}}$ *and define the infinite increasing subsequence,* $\{x_n\}_{n \geq 1}$, *by* $x_n := e^{e^{e^{e^{e^{4n}}}}}$. *We have that along the increasing subsequence* x_y *for large* $y \geq \max\left(\left\lceil e^{e^{e^e}} \right\rceil, u_0 + 1 \right)$:

$$\frac{|M(x_y)|}{\sqrt{x_y}} \overset{\blacktriangle}{\sim} 2C_{\ell,1} \cdot \frac{(\log \log \log \sqrt{x_y})^{5 + 2\log 2 + \frac{1}{3\log 2}}}{\log \log \log \log \sqrt{x_y}} + o(1),$$

as $y \to \infty$. *In the previous equation, we adopt the notation for the absolute constant* $C_{\ell,1} > 0$ *defined more precisely by*

$$C_{\ell,1} := \frac{2^{\frac{33}{16}}}{81 \cdot \pi^2 e^3 \log^4(2)} \exp\left(-\frac{55}{4} \log^2(2)\right) \approx 1.52355 \times 10^{-6}.$$

This is all to say that in establishing the rigorous proof of Corollary 3.9 based on our new methods, we not only show that

$$\limsup_{x \to \infty} \frac{|M(x)|}{\sqrt{x}} = +\infty,$$

but also set a minimal rate (along a large subsequence) at which the scaled Mertens function grows without bound.

3.6 A summary outline: Listing the core logical steps and critical components to the proof in order of exposition

3.6.1 Step-by-step overview

We offer another brief step-by-step summary overview of the critical components to our proof outlined in the introduction above, and then which are proved piece-by-piece in the next sections of the article below. This outline is provided to help the reader see our logic and proof methodology as easily and quickly as possible.

(1) We prove a matrix inversion formula relating the summatory functions of an arithmetic function f and its Dirichlet inverse f^{-1} (for $f(1) \neq 0$). See Theorem 3.1 in Section 4.

(2) This crucial step provides us with an exact formula for $M(x)$ in terms of $\pi(x)$, the seemingly unconnected prime counting function, and the Dirichlet inverse of the shifted additive function $g(n) := \omega(n) + 1$. This formula is already stated in (1) expanded above.

(3) We tighten a result from [11, §7] providing summatory functions that indicate the parity of $\lambda(n)$ using elementary arguments and more combinatorially flavored

expansions of Dirichlet series in our proof of Theorem 3.8. We use this result to sum $\sum_{n \leq x} \lambda(n) f(n)$ for general non-negative arithmetic functions f by Abel summation when x is large.

(4) We then turn to the asymptotics if the quasi-periodic $g^{-1}(n)$, estimating this inverse function's limiting asymptotics for large n (or $n \leq x$ when x is very large) in Section 6. We eventually use these estimates to prove a substantially unique new lower bound formula for the summatory function $G^{-1}(x) := \sum_{n \leq x} g^{-1}(n)$ along prescribed asymptotically large infinite subsequences that tend to $+\infty$ (see Theorem 7.4).

(5) When we return to (2) with our new lower bounds, and bootstrap, we find "magic" in the form of showing the unboundedness of $\frac{|M(x)|}{\sqrt{x}}$ along a very large increasing infinite subsequence of positive natural numbers. What we recover is a quick, and rigorous, proof of Corollary 3.9 given in Section 7.2.

3.6.2 Diagramatic flowchart of the proof logic with references to results
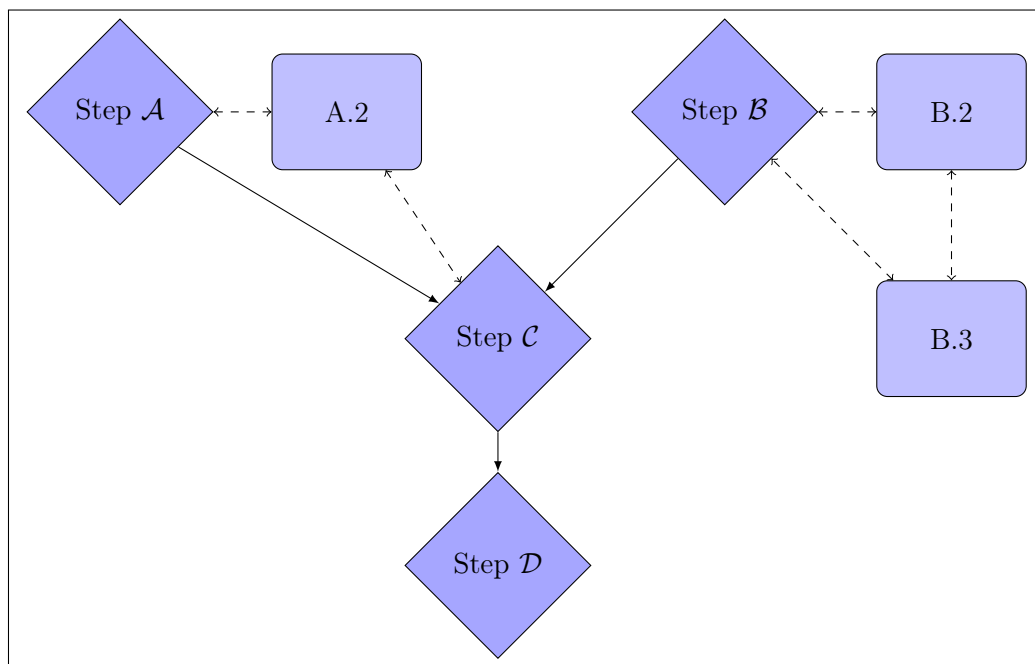
Key to the diagram stages:

Step A: *Citations and re-statements of existing theorems proved elsewhere*: E.g., statements of non-trivial theorems and key results we need that are proved in the references.

    A.A Key results and constructions:
- Theorem 3.7
- Theorem 5.1
- Corollary 5.2
- The results, lemmas, and facts cited in Section 4.3

    A.2 Lower bounds on the Abel summation based formula for $G^{-1}(x)$:
- Theorem 3.8
- Corollary 5.3
- Theorem 7.4

Step B: *Constructions of an exact formula for $M(x)$*: The exact formula we prove uses special arithmetic functions with particularly "nice" properties and bounds. This choice of the expression from Theorem 3.1 is key to how far we have traveled along the new approaches in this article. In particular, the additivity of $\omega(n)$ and the easily integrable logarithmically weighted bound on $\pi(x)$ for large x are indispensible components to why this proof works well.

Step C: *Re-writing the exact formula for $M(x)$*: Key interpretations used in formulating the lower bounds based on the re-phrased integral formula.

Step D: *The Holy Grail*: Proving that $\frac{|M(x)|}{\sqrt{x}}$ is unbounded in the limit supremum sense.

4 Preliminary proofs of lemmas and new results

4.1 Establishing the summatory function inversion identities

There are a vast number of Dirichlet convolution identities for special number theoretic functions over which we can form summatory functions and perform inversion via Theorem 3.1 [6, 16]. We will go ahead and prove this useful theorem, a crucial component to our new more combinatorial formulations used to bound $M(x)$ in later sections, before moving on. Related results for summations of Dirichlet convolutions appear in [1, §2.14; §3.10; §3.12; *cf.* §4.9, p. 95].

Proof of Theorem 3.1. Let h, g be arithmetic functions where $g(1) \neq 0$ necessarily has a Dirichlet inverse. Denote the summatory functions of h and g, respectively, by $H(x) = \sum_{n \leq x} h(n)$ and $G(x) = \sum_{n \leq x} g(n)$. We define $\pi_{g*h}(x)$ to be the summatory function of the Dirichlet convolution of g with h: $g * h$. Then we can easily see that the following expansions hold:

$$\pi_{g*h}(x) := \sum_{n=1}^{x} \sum_{d|n} g(n)h(n/d) = \sum_{d=1}^{x} g(d)H\left(\left\lfloor \frac{x}{d} \right\rfloor\right)$$

$$= \sum_{i=1}^{x} \left[G\left(\left\lfloor \frac{x}{i} \right\rfloor\right) - G\left(\left\lfloor \frac{x}{i+1} \right\rfloor\right) \right] H(i).$$

We form the matrix of coefficients associated with this system for $H(x)$, and proceed to invert it to express an exact solution for this function over all $x \geq 1$. Let the ordinary (initial, non-inverse) matrix entries be denoted by

$$g_{x,j} := G\left(\left\lfloor \frac{x}{j} \right\rfloor\right) - G\left(\left\lfloor \frac{x}{j+1} \right\rfloor\right) \equiv G_{x,j} - G_{x,j+1}.$$

The matrix we must invert in this problem is lower triangular, with ones on its diagonals – and hence is invertible. Moreover, if we let $\hat{G} := (G_{x,j})$, then this matrix is expressable by an invertible shift operation as

$$(g_{x,j}) = \hat{G}(I - U^T); \qquad U = (\delta_{i,j+1}), (I - U^T)^{-1} = ([j \leq i]_\delta).$$

Here, U is the $N \times N$ matrix whose $(i, j)^{th}$ entries are defined by $(U)_{i,j} = \delta_{i+1,j}$.

It is a useful fact that if we take successive differences of floor functions, we get non-zero behavior at divisors:

$$G\left(\left\lfloor \frac{x}{j} \right\rfloor\right) - G\left(\left\lfloor \frac{x-1}{j} \right\rfloor\right) = \begin{cases} g\left(\frac{x}{j}\right), & \text{if } j|x; \\ 0, & \text{otherwise.} \end{cases}$$

We use this property to shift the matrix $\hat{G}$, and then invert the result, to obtain a matrix involving the Dirichlet inverse of g:

$$\left[(I - U^T)\hat{G}\right]^{-1} = \left(g\left(\frac{x}{j}\right)[j|x]_\delta\right)^{-1} = \left(g^{-1}\left(\frac{x}{j}\right)[j|x]_\delta\right).$$

Now we can express the inverse of the target matrix $(g_{x,j})$ in terms of these Dirichlet inverse functions as follows:

$$(g_{x,j}) = (I - U^T)^{-1}\left(g\left(\frac{x}{j}\right)[j|x]_\delta\right)(I - U^T)$$

$$(g_{x,j})^{-1} = (I - U^T)^{-1} \left(g^{-1}\left(\frac{x}{j}\right) [j|x]_\delta \right) (I - U^T)$$

$$= \left(\sum_{k=1}^{\left\lfloor \frac{x}{j} \right\rfloor} g^{-1}(k) \right) (I - U^T)$$

$$= \left(\sum_{k=1}^{\left\lfloor \frac{x}{j} \right\rfloor} g^{-1}(k) - \sum_{k=1}^{\left\lfloor \frac{x}{j+1} \right\rfloor} g^{-1}(k) \right).$$

Thus the summatory function H is exactly expressed by the inverse vector product of the form

$$H(x) = \sum_{k=1}^{x} g_{x,k}^{-1} \cdot \pi_{g*h}(k)$$

$$= \sum_{k=1}^{x} \left(\sum_{j=\left\lfloor \frac{x}{k+1} \right\rfloor+1}^{\left\lfloor \frac{x}{k} \right\rfloor} g^{-1}(j) \right) \cdot \pi_{g*h}(k). \quad \square$$

4.2 Proving the crucial signedness property from the conjecture

Proposition 4.1 (The characteristic function of the primes). *Let $\chi_{\mathbb{P}}$ denote the characteristic function of the primes, $\varepsilon(n) = \delta_{n,1}$ be the multiplicative identity with respect to Dirichlet convolution, and denote by $\omega(n)$ the incompletely additive function that counts the number of distinct prime factors of n. Then we have the convolution identity given by*

$$\chi_{\mathbb{P}} + \varepsilon = (\omega + 1) * \mu.$$

The summatory function of the left-hand-side of the previous equation is clearly $\widetilde{G}(x) = \pi(x) + 1$ for all $x \geq 1$.

Proof. The core is to prove that for all $n \geq 1$, $\chi_{\mathbb{P}}(n) = (\mu * \omega)(n)$ – our essential claim. We notice that the Mellin transform of $\pi(x)$ – the summatory function of $\chi_{\mathbb{P}}(n)$ – at $-s$ is given by

$$s \cdot \int_1^\infty \frac{\pi(x)}{x^{s+1}} dx = \sum_{n \geq 1} \frac{\nabla[\pi](n-1)}{n^s}$$

$$= \sum_{n \geq 1} \frac{\chi_{\mathbb{P}}(n)}{n^s} = P(s),$$

where $\nabla[f](n) := f(n+1) - f(n)$ denotes the standard *forward difference operator* used to express a discrete derivative type operation on arithmetic functions. This is typical construction which more generally relates the Mellin transform $s \cdot \mathcal{M}[S_f](-s)$ to the DGF of the sequence $f(n)$ as cited, for example, in [1, §11]. Now we consider the DGF of the right-hand-side function, $f(n) := (\mu * \omega)(n)$, as

$$D_f(s) = \frac{1}{\zeta(s)} \times \sum_{n \geq 1} \frac{\omega(n)}{n^s} = P(s).$$

Thus for any $\Re(s) > 1$, the DGFs of each side of the claimed equation coincide. So by uniqueness of Dirichlet series, we see that in fact the claim holds. To obtain the full result, we add to each side of this equation a term of $\varepsilon(n) \equiv (\mu * 1)(n)$, and then factor the resulting convolution identity. $\square$

When combined with Corollary 3.2, the proof of Proposition 4.1 yields the crucial starting point providing an exact formula for $M(x)$ stated in (1) of Corollary 3.3.

Proposition 4.2 (The key signedness property of $g^{-1}(n)$). *For the Dirichlet invertible function, $g(n) := \omega(n) + 1$ defined such that $g(1) = 1$, at any $n \geq 1$, we have that $\operatorname{sgn}(g^{-1}(n)) = \lambda(n)$. Here, the notation for the operation given by $\operatorname{sgn}(h(n)) = \frac{h(n)}{|h(n)| + [h(n)=0]_\delta} \in \{0, \pm 1\}$ denotes the sign, or signed parity, of the arithmetic function h at n.*

A02-2020-04-26

Proof. Let $D_f(s) := \sum_{n \geq 1} f(n) n^{-s}$ denote the Dirichlet generating function (DGF) of $f(n)$. Then we have that

$$D_{(\omega+1)^{-1}}(s) = \frac{D_\lambda(s)}{(P(s) + 1)\zeta(2s)}.$$

Let $h^{-1}(n) := (\omega * \mu + \varepsilon)^{-1}(n) = [n^{-s}](P(s) + 1)^{-1}$. Then we have that

$$(h^{-1} * 1)(n) = -\sum_{p_1 \mid n} h^{-1}\left(\frac{n}{p_1}\right) = \lambda(n) \times \sum_{p_1 \mid n} \sum_{p_2 \mid \frac{n}{p_1}} \cdots \sum_{p_{\Omega(n)} \mid \frac{n}{p_1 p_2 \cdots p_{\Omega(n)-1}}} 1$$

$$= \begin{cases} \lambda(n) \times (\Omega(n) - 1)!, & n \geq 2; \\ \lambda(n), & n = 1. \end{cases}$$

So by Möbius inversion

$$h^{-1}(n) = \lambda(n) \left[\sum_{\substack{d \mid n \\ d < n}} \lambda(d)\mu(d)(\Omega(n/d) - 1)! + 1\right] = \lambda(n) \left[\sum_{\substack{d \mid n \\ d < n}} \mu^2(d)(\Omega(n/d) - 1)! + 1\right].$$

Then we finally have that

$$(\omega + 1)^{-1}(n) = \lambda(n) \times \sum_{d \mid n} \lambda(d) \left[\sum_{\substack{r \mid \frac{n}{d} \\ r < \frac{n}{d}}} \mu^2(r)(\Omega\left(\frac{n}{dr}\right) - 1)! + 1\right] \chi_{\mathrm{sq}}(d)\mu(\sqrt{d}),$$

where χ_{sq} is the characteristic function of the squares. In either case of $\lambda(n) = \pm 1$, there are positive constants $C_{1,n}, C_{2,n} > 0$ such that

$$\lambda(n)C_{1,n} \times \sum_{d^2 \mid n} \lambda(d^2)\mu(d) \leq g^{-1}(n) \leq \lambda(n)C_{1,n} \times \sum_{d^2 \mid n} \lambda(d^2)\mu(d),$$

where $\sum_{d^2 \mid n} \lambda(d^2)\mu(d) = \sum_{d^2 \mid n} \mu^2(n) > 0$. This proves the result. $\square$

4.3   Other facts and listings of results we will need in our proofs

Theorem 4.3 (Mertens theorem)**.**

$$P_1(x) := \sum_{p \leq x} \frac{1}{p} = \log \log x + B + o(1),$$

where $B \approx 0.2614972128476427837554$ is an absolute constant.

Corollary 4.4. *We have that for sufficiently large $x \gg 1$*

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) = \frac{e^{-B}}{\log x} \left(1 + o(1)\right).$$

Hence, for $1 < |z| < R < 2$ we obtain that

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^z = \frac{e^{-Bz}}{(\log x)^z} \left(1 + o(1)\right)^z.$$

Facts 4.5 (Exponential Integrals and Incomplete Gamma Functions)**.** The following two variants of the *exponential integral function* are defined by [13, §8.19]

$$\mathrm{Ei}(x) := \int_{-x}^{\infty} \frac{e^{-t}}{t} dt,$$

$$E_1(z) := \int_1^{\infty} \frac{e^{-tz}}{t} dt, \Re(z) \geq 0,$$

where $\mathrm{Ei}(-kz) = -E_1(kz)$. We have the following inequalities providing quasi-polynomial upper and lower bounds on $E_1(z)$:

$$1 - \frac{3}{4}z \leq E_1(z) - \gamma - \log z \leq 1 - \frac{3}{4}z + \frac{11}{36}z^2. \tag{5a}$$

A related function is the (upper) *incomplete gamma function* defined by [13, §8.4]

$$\Gamma(s, x) = \int_x^{\infty} t^{s-1} e^{-t} dt, \Re(s) > 0.$$

We have the following properties of $\Gamma(s, x)$:

$$\Gamma(s, x) = (s - 1)! \cdot e^{-x} \times \sum_{k=0}^{s-1} \frac{x^k}{k!}, s \in \mathbb{Z}^+, \tag{5b}$$

$$\Gamma(s + 1, 1) = e^{-1} \left\lfloor \frac{s!}{e} \right\rceil, s \in \mathbb{Z}^+, \tag{5c}$$

$$\Gamma(s, x) \sim x^{s-1} \cdot e^{-x}, |x| \to +\infty. \tag{5d}$$

5 Summing arithmetic functions weighted by $\lambda(n)$

5.1 Discussion: The enumerative DGF result in Theorem 3.7 from Montgomery and Vaughan

What this enumeratively-flavored result of Montgomery and Vaughan allows us to do is get a "good enough" lower bound on sums of positive and asymptotically bounded arithmetic functions weighted by the Liouville lambda function, $\lambda(n) = (-1)^{\Omega(n)}$. For comparison, we already have known, more classical bounds due to Erdös (and earlier) that we have tightly the bound [2, 11]

$$\pi_k(x) = (1 + o(1)) \cdot \frac{x}{\log x} \frac{(\log \log x)^{k-1}}{(k-1)!}.$$

We seek to approximate the right-hand-side of $\mathcal{G}(z)$ by only taking the products of the primes $p \leq u$, e.g., $p \in \{2, 3, 5, \ldots, u\}$, of which the last element in this set has average order of $\log \left[\frac{u}{\log u}\right]$ for some minimal u as $x \to \infty$.

We also state the following theorem reproduced from [11, Thm. 7.20] that handles the relative scarcity of the distribution of the $\Omega(n)$ for $n \leq x$ such that $\Omega(n) > \log \log x$. This allows us later to show that taking just $k \in [1, \log \log x]$ and summing over such k in Theorem 3.7 captures the asymptotically relevant, dominant behavior of the values of summatory functions over the $\pi_k(x)$ for all $k \leq \frac{\log x}{\log 2}$.

Theorem 5.1 (Bounds on exceptional values of $\Omega(n)$ for large n, MV 7.20). *Let*

$$B(x, r) := \# \{n \leq x : \Omega(n) \leq r \cdot \log \log x\}.$$

If $1 \leq r \leq R < 2$ and $x \geq 2$, then

$$B(x, r) \ll_R x \cdot (\log x)^{r-1-r \log r}, \quad as \ x \to \infty.$$

The proof of Theorem 5.2 is found in the cited reference as Chapter 7 of Montgomery and Vaughan. The key interpretation we need is the result stated in the next corollary.

Corollary 5.2. *Using the notation for $B(x, r)$ from Theorem 5.1, we have in particular that for $r \in (1, 2)$,*

$$\left|1 - \frac{B(x, r)}{B(x, 1)}\right| \xrightarrow{x \to \infty} 1.$$

Once again, we emphasize that Corollary 5.2 implies that for sums involving $\widehat{\pi}_k(x)$ indexed by k, we can capture the dominant asymptotic behavior of these sums by taking k in the truncated range $1 \leq k \leq \log \log x$, e.g., $0 \leq z \leq 1$ in Theorem 3.7. This fact will be important when we prove Theorem 7.4 in Section 7 using a sign-weighted summatory function in Abel summation that depends on these functions (see Lemma 7.2).

5.2 The key new results utilizing Theorem 3.7

We will require a handle on partial sums of integer powers of the reciprocal primes as functions of the integral exponent and the upper summation index x. The next corollary is not a triviality as it comes in handy when we take to the task of proving Theorem 3.8 below. The next statement of Corollary 5.3 effectively generalizes Mertens theorem stated previously as Theorem 4.3 by providing a coarse rate in x below which the reciprocal prime sums tend to absolute constants given by the prime zeta function, $P(s)$.

Corollary 5.3. *For real $s \geq 1$, let*

$$P_s(x) := \sum_{p \leq x} p^{-s}, x \gg 2.$$

When $s := 1$, we have the known bound in Mertens theorem (see Theorem 4.3). For $s > 1$, we obtain that

$$P_s(x) \approx E_1((s-1)\log 2) - E_1((s-1)\log x) + o(1).$$

It follows that for $s \geq 2$ we have that

$$P_s(x) \leq \gamma_1(s, x) + o(1),$$

where it suffices to take the bounding function as

$$\gamma_1(s, x) = -s \log\left(\frac{\log x}{\log 2}\right) + \frac{3}{4} s(s-1) \log(x/2) + \frac{11}{36} s(s-1)^2 \log^2(2).$$

A05-2020-04-26

Proof. Let $s > 1$ be real-valued. By Abel summation where our summatory function is given by $A(x) = \pi(x) \sim \frac{x}{\log x}$ and our function $f(t) = t^{-s}$ so that $f'(t) = -s \cdot t^{-(s+1)}$, we obtain that

$$P_s(x) = \frac{1}{x^s \cdot \log x} + s \cdot \int_2^x \frac{dt}{t^s \log t}$$
$$= E_1((s-1)\log 2) - E_1((s-1)\log x) + o(1), |x| \to \infty.$$

Now using the inequalities in Facts 4.5, we obtain that the difference of the exponential integral functions is bounded above and below by

$$\frac{P_s(x)}{s} \geq -\log\left(\frac{\log x}{\log 2}\right) + \frac{3}{4}(s-1)\log(x/2) - \frac{11}{36}(s-1)^2 \log^2(x)$$
$$\frac{P_s(x)}{s} \leq -\log\left(\frac{\log x}{\log 2}\right) + \frac{3}{4}(s-1)\log(x/2) + \frac{11}{36}(s-1)^2 \log^2(2).$$

This completes the proof of the bounds cited above in the statement of this lemma. $\square$

A06-2020-04-26

Proof of Theorem 3.8. We have that for all integers $0 \leq k \leq m$

$$[z^k] \prod_{1 \leq i \leq m} (1 - f(i)z)^{-1} = [z^k] \exp\left(\sum_{j \geq 1} \left(\sum_{i=1}^m f(i)^j\right) \frac{z^j}{j}\right). \tag{6}$$

In our case we have that $f(i)$ denotes the ith prime. Hence, summing over all $p \leq ux$ in place of $0 \leq k \leq m$ in the previous formula applied in tandem with Corollary 5.3, we obtain that the logarithm of the generating function series we are after when we sum over all $p \leq ux$ for some $u \geq \frac{2}{x}$ corresponds to

$$\log\left[\prod_{p \leq ux}\left(1 - \frac{z}{p}\right)^{-1}\right] \geq (B + \log\log(ux))z + \sum_{j \geq 2}\left[a(ux) + b(ux)(j-1) + c(ux)(j-1)^2\right]z^j$$
$$= (B + \log\log(ux))z - a(ux)\left(1 + \frac{1}{z-1} + z\right)$$
$$+ b(ux)\left(1 + \frac{2}{z-1} + \frac{1}{(z-1)^2}\right)$$

26

$$- c(ux) \left(1 + \frac{4}{z-1} + \frac{5}{(z-1)^2} + \frac{2}{(z-1)^3} \right)$$

$$=: \widehat{\mathcal{B}}(u, x; z).$$

In the previous equations, the lower bounds formed by the functions (a, b, c) in ux are given by the corresponding upper bounds from Corollary 5.3 due to the leading sign on the previous expansions as

$$(a_\ell, b_\ell, c_\ell) := \left(-\log \left(\frac{\log(ux)}{\log 2} \right), \frac{3}{4} \log \left(\frac{ux}{2} \right), \frac{11}{36} \log^2 2 \right).$$

Now we make a practical decision to set the uniform bound parameter to a middle ground value of $1 < R < 2$ at $R := \frac{3}{2}$ so that

$$z \equiv z(k, x) = \frac{k}{\log \log x} \in (0, R),$$

for $x \gg 1$ very large. Thus $(z-1)^{-m} \in [(-1)^m, 2^m]$ for integers $m \geq 1$, and we can thus form the lower bound stated below. Namely, these bounds on the signed reciprocals of $z - 1$ lead to an effective bound of the following form:

$$\widehat{\mathcal{B}}(u, x; z) \geq (B + \log \log(ux))z - a(ux) \left(1 + \frac{1}{z-1} + z \right)$$

$$+ b(ux) \left(1 + \frac{2}{z-1} + \frac{1}{(z-1)^2} \right) - 45 \cdot c(ux).$$

Since the function $c(ux)$ is constant, we then obtain a bound of the next form.

$$\frac{e^{-Bz}}{(\log(ux))^z} \times \exp\left(\widehat{\mathcal{B}}(u, x; z) \right) \geq \exp\left(-\frac{55}{4} \log^2(2) \right) \times \left(\frac{\log(ux)}{\log 2} \right)^{1 + \frac{1}{z-1} + z}$$

$$\times \left(\frac{ux}{2} \right)^{\frac{3}{4} \left(1 + \frac{2}{z-1} + \frac{1}{(z-1)^2} \right)}$$

$$=: \widehat{\mathcal{C}}(u, x; z). \tag{7}$$

Now we need to determine which values of $u > \frac{2}{x}$ minimize the expression in (7). For this we will use introdutory calculus in the form of the second derivative test with respect to u. In particular, we can symbolically invoke the equation solver functionality in *Mathematica* to see that

$$\frac{\partial}{\partial u} \left[\widehat{\mathcal{C}}(u, x; z) \right] \Bigg|_{u \mapsto u_0} = 0 \implies u_0 \in \left\{ \frac{1}{x}, \frac{1}{x} e^{-\frac{4}{3}(z-1)} \right\}.$$

When we substitute this outstanding parameter value of $u_0 =: \hat{u}_0 \mapsto \frac{1}{x} e^{-\frac{4}{3}(z-1)}$ into the next expression for the second derivative of the same function $\widehat{\mathcal{C}}(u, x; z)$ we obtain

$$\frac{\partial^2}{\partial u^2} \left[\widehat{\mathcal{C}}(u, x; z) \right] \Bigg|_{u = \hat{u}_0} = \exp\left(-\frac{55}{4} \log^2(2) \right) x^2 2^{\frac{8z^3 - 27z^2 + 32z - 16}{4(z-1)^2}} 3^{-z + \frac{1}{1-z} + 1} e^{\frac{5z^2 - 16z + 8}{3(z-1)}} \times$$

$$\times (1 - z)^{z + \frac{1}{z-1} - 2} z^2 \log^{\frac{z^2}{1-z}}(2) > 0,$$

provided that $z < 1$, e.g., so that $k \leq \log \log x$ in Theorem 3.7. This restriction on k to note leads to a minimum value on the partial product, or lower bound, at this $u = \hat{u}_0$ since the second derivative is positive at the zero of the first derivative whenever $z < 1$.

After a substitution of $u = \frac{1}{x}e^{-\frac{4}{3}(z-1)}$ into the expression for $\widehat{\mathcal{C}}(u, x; z)$ defined above, we have that

$$\widehat{\mathcal{C}}(u, x; z) \geq \exp\left(-\frac{55}{4}\log^2(2)\right) \cdot 2^{\frac{9}{16}}\left(\frac{1-z}{3e\log 2}\right)^3 \times \left(\frac{4(1-z)}{3e\log 2}\right)^z.$$

Finally, since $z \equiv z(k, x) = \frac{k}{\log\log x}$ and $k \in [0, R\log\log x)$, we obtain that for small k and $x \gg 1$ large $\Gamma(z+1) \approx 1$, and for k towards the upper range of its interval that $\Gamma(z+1) \approx \Gamma(5/2) = \frac{3}{4}\sqrt{\pi}$. In total, what we get out of these formulas is stated as in the theorem bounds. $\qquad\qquad\square$

6 Precisely bounding the Dirichlet inverse functions, $g^{-1}(n)$

Conjecture 3.4 is not the most accurate way to express the limiting behavior of the Dirichlet inverse functions $g^{-1}(n)$, though it does capture an important characteristic – namely, that these functions can be expressed via more simple formulas than inspection of the initial sequence properties might otherwise suggest.

With all of this in mind, we define the following sequence for integers $n \geq 1, k \geq 0$:

$$C_k(n) := \begin{cases} \varepsilon(n), & \text{if } k = 0; \\ \sum_{d|n} \omega(d) C_{k-1}(n/d), & \text{if } k \geq 1. \end{cases} \tag{8}$$

We will illustrate by example the first few cases of these functions for small k after we prove the next lemma. The sequence of important semi-diagonals of these functions begins as [17, A008480]

$$\{\lambda(n) \cdot C_{\Omega(n)}(n)\}_{n \geq 1} \mapsto \{1, -1, -1, 1, -1, 2, -1, -1, 1, 2, -1, -3, -1, 2, 2, 1, -1, -3, -1, -3, 2, 2, -1, 4, 1, 2, \ldots\}.$$

Lemma 6.1 (An exact formula for $g^{-1}(n)$)**.** *For all $n \geq 1$, we have that*

$$g^{-1}(n) = \sum_{d|n} \mu(n/d) \lambda(d) C_{\Omega(d)}(d).$$

Proof. We first write out the standard recurrence relation for the Dirichlet inverse of $\omega + 1$ as

$$g^{-1}(n) = -\sum_{\substack{d|n \\ d>1}} (\omega(d) + 1) f^{-1}(n/d) \qquad \Longrightarrow$$

$$(g^{-1} * 1)(n) = -(\omega * g^{-1})(n).$$

Now by repeatedly expanding the right-hand-side, and removing corner cases in the nested sums since $\omega(1) = 0$ by convention, we find that

$$(g^{-1} * 1)(n) = (-1)^{\Omega(n)} C_{\Omega(n)}(n) = \lambda(n) C_{\Omega(n)}(n).$$

The statement follows by Möbius inversion applied to each side of the last equation. $\qquad \square$

Example 6.2 (Special cases of the functions $C_k(n)$ for small k)**.** We cite the following special cases which should be easy enough to see on paper:

$$C_0(n) = \delta_{n,1}$$
$$C_1(n) = \omega(n)$$
$$C_2(n) = \sigma_0(n) \times \sum_{p|n} \frac{\nu_p(n)}{\nu_p(n) + 1} - \gcd\left(\Omega(n), \omega(n)\right).$$

We also can see a recurrence relation between successive $C_k(n)$ values over k of the form

$$C_k(n) = \sum_{p|n} \sum_{d | \frac{n}{p^{\nu_p(n)}}} \sum_{i=1}^{\nu_p(n)} C_{k-1}\left(d \cdot p^i\right). \tag{9}$$

Thus we can work out further cases of the $C_k(n)$ for a while until we are able to understand the general trends of its asymptotic behaviors.

29

Remark 6.3 (Preface on the growth of the functions $C_k(n)$ in the average case). We note that the average order of $\mathbb{E}[\Omega(n)] = \log \log n$, so that for large $x \gg 1$ tending to infinity, we can expect (on average) that for $p|n$, $1 \leq \nu_p(n)$ (for large $p|x, p \sim \frac{x}{\log x}$) and $\nu_p(n) \approx \log \log n$.

However, if x is primorial, we can have $\mathbb{E}[\Omega(x)] = \frac{\log x}{\log \log x}$. There is, however, a duality with the size of $\Omega(x)$ and the rate of growth of the $\nu_p(x)$ exponents. That is to say that on average, even though $\mathbb{E}[\nu_p(x)] = \log \log n$ for most $p|x$, if $\Omega(x) = m \approx O(1)$ is small, then

$$\nu_p(x) \approx \log_{\sqrt[m]{\frac{x}{\log x}}}(x) = \frac{m \log x}{\log\left(\frac{x}{\log x}\right)}.$$

Since we will be essentially averaging the inverse functions, $g^{-1}(n)$, via their summatory functions over the range $n \leq x$ for x large, we tend not to worry about bounding anything but by the average order case, which wins out when we sum (i.e., average) and tend to infinity.

Proposition 6.4. *The dominant order terms in approximating the function $C_3(n)$ for large n are given by*

$$C_3(n) \sim \frac{(\sigma_0 * 1)(n)n^2}{\log^2 n} - \frac{(\sigma_0 * 1)(n)n^2}{\log n} + O\left(n \cdot \log \log n\right).$$

We leave the terms involving the divisor function $d(n) \equiv \sigma_0(n)$ and the leading convolution terms involving it unevaluated for now.

Proof. TODO ....

We can compute the main term of $C_3(n)$ as follows where we use the notation that p, q are prime indices:

$$C_3(n) \sim \sum_{p|n} \sum_{d|\frac{n}{p^{\nu_p(n)}}} \sum_{i=1}^{\nu_p(n)} \sum_{q|dp^i} \frac{\nu_q(dp^i)}{\nu_q(dp^i)+1} \sigma_0(d)(i+1)$$

$$= \sum_{p|n} \sum_{d|\frac{n}{p^{\nu_p(n)}}} \sum_{i=1}^{\nu_p(n)} \left[\sum_{q|d} \frac{\nu_q(d)}{\nu_q(d)+1} \sigma_0(d)(i+1) + \sum_{j=1}^{i} \frac{j}{(j+1)} \sigma_0(d)(i+1) \right]$$

$$= \sum_{p|n} \sum_{d|\frac{n}{p^{\nu_p(n)}}} \sum_{q|d} \sigma_0(d) \left[\frac{\nu_p(n)(\nu_p(n)+3)}{2} \frac{\nu_q(d)}{\nu_q(d)+1} + \frac{1}{12}(\nu_p(n)+1)(\nu_p(n)+2)\left(4\nu_p(n)+9-6H^{(1)}_{\nu_p(n)+2}\right) \right].$$

We will break the two key component sums into separate calculations. First, we compute that[*]

$$C_{3,1}(n) = \sum_{p|n} \sum_{d|\frac{n}{p^{\nu_p(n)}}} \frac{\nu_p(n)(\nu_p(n)+3)}{2} \times \sum_{q|d} \frac{\nu_q(d)}{\nu_q(d)+1} \sigma_0(d)$$

---

[*]Here, the arithmetic function $\sigma_0 * 1$ is multiplicative. Its value at prime powers can be computed as

$$(\sigma_0 * 1)(p^\alpha) = \sum_{i=0}^{\alpha}(i+1) = \frac{(\alpha+1)(\alpha+2)}{2},$$

where $\sigma_0(p^\beta) = \beta + 1$.

$$= \sum_{\substack{p,q|n \\ p\neq q}} \sum_{d|\frac{n}{p^{\nu_p(n)}q^{\nu_q(n)}}} \frac{\nu_p(n)(\nu_p(n)+3)}{2} \times \sum_{i=1}^{\nu_q(n)} \frac{\nu_q(dq^i)}{\nu_q(dq^i)+1}\sigma_0(dq^i)$$

$$= \sum_{\substack{p,q|n \\ p\neq q}} \sum_{d|\frac{n}{p^{\nu_p(n)}q^{\nu_q(n)}}} \frac{\nu_p(n)(\nu_p(n)+3)\nu_q(n)(\nu_q(n)+3)}{4}\sigma_0(d)$$

$$= (\sigma_0 * 1)(n) \times \sum_{\substack{p,q|n \\ p\neq q}} \frac{\nu_p(n)(\nu_p(n)+3)\nu_q(n)(\nu_q(n)+3)}{(\nu_p(n)+1)(\nu_p(n)+2)(\nu_q(n)+1)(\nu_q(n)+2)}.$$

Next, we have that

$$C_{3,2}(n) = \sum_{p|n} \sum_{d|\frac{n}{p^{\nu_p(n)}}} \sum_{q|d} \frac{1}{12}(\nu_p(n)+1)(\nu_p(n)+2)\left(4\nu_p(n)+9-6H^{(1)}_{\nu_p(n)+2}\right)\sigma_0(d)$$

$$= \sum_{\substack{p,q|n \\ p\neq q}} \sum_{d|\frac{n}{p^{\nu_p(n)}q^{\nu_q(n)}}} \sum_{i=1}^{\nu_q(n)} \frac{1}{12}(\nu_p(n)+1)(\nu_p(n)+2)\left(4\nu_p(n)+9-6H^{(1)}_{\nu_p(n)+2}\right)\sigma_0(d)(i+1)$$

$$= (\sigma_0 * 1)(n) \times \sum_{\substack{p,q|n \\ p\neq q}} \frac{1}{6}\frac{\nu_q(n)(\nu_q(n)+3)}{(\nu_q(n)+1)(\nu_q(n)+2)}\left(4\nu_p(n)+9-6H^{(1)}_{\nu_p(n)+2}\right).$$

Now to roughly bound the error term, e.g., the GCD of prime omega functions from the exact formula for $C_3(n)$, we observe that the divisor function has average order of the form:

$$\mathbb{E}[d(n)] = \log n + (2\gamma - 1) + O\left(\frac{1}{\sqrt{n}}\right).$$

Then using that $\mathbb{E}[\omega(n)], \mathbb{E}[\Omega(n)] = \log\log n$ (except in rare cases when n is primorial, a power of 2, etc. [†]), as discussed in the next remarks, we bound the error as

$$C_{3,3}(n) = -\sum_{p|n} \sum_{d|\frac{n}{p^{\nu_p(n)}}} \sum_{i=1}^{\nu_p(n)} \gcd\left(\Omega(d)+i, \omega(d)+1\right)$$

$$= \sum_{p|n} \frac{\nu_p(n)}{\nu_p(n)+1}O\left(\sigma_0(n)\cdot\log\log n\right)$$

$$= O\left(\pi(n)\cdot\log n\cdot\log\log n\right)$$

$$= O\left(n\cdot\log\log n\right).$$

In total, we obtain that

$$C_3(n) = (\sigma_0 * 1)(n) \times \sum_{\substack{p,q|n \\ p\neq q}} \frac{1}{6}\frac{\nu_q(n)(\nu_q(n)+3)}{(\nu_q(n)+1)(\nu_q(n)+2)}\left(4\nu_p(n)+9-6H^{(1)}_{\nu_p(n)+2}\right) \qquad (10)$$

[†]In this context, we write $\mathbb{E}[\Omega(n)] = \log\log n$ to denote the *average order* of this arithmetic function – even though its actual values may fluctuate non-uniformly infinitely often, e.g., $\Omega(2^m) = m$ and for primes p we have that $\Omega(p) = 1$, but most of the time the asymptotic relation holds when we sum, or average over all possible $n \leq x$. What this notation corresponds to, or means in practice, is that the average of the summatory function satisfies: $\frac{1}{x} \cdot \sum_{n\leq x} \Omega(n) \sim \log\log x$.

$$+ \sigma_0(n) \times \sum_{\substack{p,q|n \\ q \neq p}} \frac{2^{\nu_q(n)} \nu_p(n)(\nu_p(n) + 3)}{4(\nu_p(n) + 1)(\nu_q(n) + 1)}$$

$$+ O\left(n \cdot \log \log n\right).$$

Finally, we use the law for approximating the average order cases of the exponents $\nu_p(x)$ from Remark 6.3 to approximate the exact formulas above as

$\square$

Summary 6.5 (Asymptotics of the $C_k(n)$)**.** We have the following asymptotic relations relations for the growth of small cases of the functions $C_k(n)$:

$$C_1(n) \sim \log \log n$$

$$C_2(n) \sim \frac{\sigma_0(n)n}{\log n} + O(\log \log n)$$

$$C_3(n) \sim \frac{(\sigma_0 * 1)(n)n^2}{\log^2 n} - \frac{(\sigma_0 * 1)(n)n^2}{\log n} + O\left(n \cdot \log \log n\right).(TODO)$$

The previous limiting asymptotics are computed from the explicit formulas for small k in Example 6.2 using the average order arguments outlined in Remark 6.3. Theorem 3.6 is proved next. The theorem makes precise what these formulas already suggest about the main terms of the growth rates of $C_k(n)$ as functions of k, n for limiting cases of n large for fixed k.

Proof of Theorem 3.6 – TODO. We showed how to compute the formulas for the base cases in the preceeding examples discussed above in Example 6.2. We can also see that $C_3(n)$ satsfies the formula we must establish when $k := 3$ according to our proof of Proposition 6.4. Let's proceed by using induction to prove that our asymptotics hold for all $k \geq 3$ with the recurrence formula from (9) relating $C_k(n)$ to $C_{k-1}(n)$ whenever $k \geq 2$. We will compute the main term formula first, then complete the proof by bounding the easier claimed big-O error term calculations to wrap up our induction.

A08-2020-04-26

Main term formula inductive proof. Suppose that $k \geq 4$. By the inductive assumption and the recurrence formula for $C_k(n)$ given in (9), we have that

$$C_k(n) \sim \sum_{p|n} \sum_{d|np^{-\nu_p(n)}} \sum_{i=1}^{\nu_p(n)} \frac{(-dp^i)^{k-1}}{(\log(dp^i))^{k-1} \cdot (k-2)!} \binom{i+k-1}{k-1} (\sigma_0 * \tau_{k-2})(d),$$

where the binomial coefficient term follows from the multiplicativity of the convolution $\sigma_0 * \tau_m$ at any $m \geq 1$ using the prime power formula given in (3). Now to handle the inner sum, we bound by setting $\alpha \equiv \nu_p(n)$ and invoking *Mathematica* in the form of

$$\text{IC}_k(d, \alpha) = \sum_{i=1}^{\alpha} \frac{(-dp^i)^{k-1}}{(\log(dp^i))^{k-1} \cdot (k-2)!} \binom{i+k-1}{k-1}$$

$$\approx \int -\frac{(-dp^\alpha)^{k-1}}{\log(dp^\alpha)^{k-1} \cdot (k-2)!} \binom{\alpha+k-1}{k-1} d\alpha.$$

A very careful, exact justification of the indefinite integral formula we cite below involves expanding the binomial coefficient terms in powers of α with coefficients given by the Stirling numbers of the first kind. Then we can integrate each power of α against the remaining multiplier terms. However, in general, we need not be so precise. Since we will

let $\alpha \to +\infty$ monotonely as $n \to \infty$ (in the outer sum), we should have that the main relevant term in the integrand corresponds to the monomial term in α with the largest exponent. In this case, by expanding the binomial coefficient, this term is clearly the coefficient of α^{k-1} with leading coefficient of one. Also, since

$$(\log d + \alpha \cdot \log p)^{k-1} \sim \alpha^{k-1} \cdot (\log p)^{k-1},$$

the indefinite integral we need concern ourselves with is given by

$$\begin{aligned}
\text{IC}_k(d, \alpha) &\sim \int \frac{-(-dp^{\alpha})^{k-1}}{(\log p)^{k-1} \cdot (k-1)!^2} d\alpha \\
&= \frac{(-1)^k d^{k-1} p^{(k-1)\alpha}}{(\log p)^k \cdot (k-1)!^2}.
\end{aligned}$$

As intended, we again simplify somewhat by setting average order estimates on the parameters

$$p \mapsto \left(\frac{n}{e}\right)^{\frac{1}{\log \log n}} \xrightarrow{n \to \infty} n^{\frac{1}{\log \log n}}, \alpha \mapsto \log \log n, \log p \mapsto \frac{\log n}{\log \log n}.$$

Also, since $p \gg_n d$, we obtain the dominant asymptotic growth terms of the function we denoted by $\text{IC}_k(d, \alpha) \equiv \text{IC}_k(d, n)$ as

$$\text{IC}_k(d, n) \sim \frac{-(-d)^{k-1} n^{k-1}}{(k-1)!^2 \cdot (\log n)^k}.$$

Then from the recurrence relation sums we started with from (3), we are left to sum

$$\begin{aligned}
C_k(n) &\sim \sum_{p|n} (\sigma_0 * \tau_{k-1})(n) \binom{p^{\nu_p(n)} + k - 1}{k - 1}^{-1} \times \sum_{d \mid \frac{n}{p^{\nu_p(n)}}} \text{IC}_k(d, n) \\
&\stackrel{\blacktriangle}{\sim} \sum_{p|n} (\sigma_0 * \tau_{k-1})(n) \binom{p^{\nu_p(n)} + k - 1}{k - 1}^{-1} \times \sum_{d \mid \frac{n}{p^{\nu_p(n)}}} \text{IC}_k(1, n) \\
&\stackrel{\blacktriangle}{\sim} (\sigma_0 * \tau_{k-1})(n) \cdot \pi(n) \times (\log n) \cdot \frac{(-d)^{k-1} n^{k-1}}{(k-1)! \cdot (\log n)^k} \\
&\stackrel{\blacktriangle}{\sim}
\end{aligned}$$

As we did in the previous work to prove Proposition 6.4, we handle the sums by pulling out a factor of the inner divisor sum depending only on n (and k):

$$C_k(n) = \sum_{p|n} (\sigma_0 * \tau_{k-1})(n) \binom{p^{\nu_p(n)} + k}{k}^{-1} \times \text{IC}_k(n)$$

33

$$= (\sigma_0 * \tau_{k-1})(n) \binom{p^{\nu_p(n)} + k}{k}^{-1} \cdot \pi(n) \times \text{IC}_k(n)$$

Combining with the remaining terms we get by induction a proof of our target bounds on the main term for $C_k(n)$.

Establishing the error term bound inductively. To bound the error terms, again suppose inductively that $k \geq 4$. We compute the big-O bounds as follows letting $\alpha \equiv \nu_p(n)$:

$$\text{ET}_k(n) = \sum_{i=1}^{\nu_p(n)} n^{k-2} \cdot \frac{(\log \log n)^{k-2}}{(\log n)^{k-2}}$$

$$\approx \int (dp^\alpha)^{k-2} \log \log(dp^\alpha) d\alpha$$

$$= -\frac{\text{Ei}((k-2) \log(dp^\alpha))}{(k-2) \log p} + \frac{d^{k-2} p^{(k-2)\alpha}}{(k-2) \log p} \log(dp^\alpha)$$

$$\sim \frac{d^{k-2} p^{(k-2)\alpha}}{(k-2) \log p} \log(dp^\alpha).$$

In the last expansion, we have dropped the exponential integral terms since they provide at most polynomial powers of the logarithm of their inputs.

To evaluate the outer divisor sum that results from applying the recurrence relation for $C_k(n)$, we will require the following bound providing an average order on the *generalized sum-of-divisors functions*, $\sigma_\alpha(n) := \sum_{d|n} d^\alpha$, in this case where we must take the real-valued parameter $\alpha > 0$. In particular, we have that for integers $\alpha \geq 2$ [13, §27.11]:

$$\mathbb{E}[\sigma_\alpha(n)] = \frac{\zeta(\alpha+1)}{\alpha+1} x^\alpha + O(x^{\alpha-1}).$$

Approximating the number of terms in the prime divisor sum by $\pi(x) = \frac{x}{\log x}$, we thus obtain

$$\text{ET}_k(n) \approx \frac{(\log \log n)^{k-1} e^{k-2}}{(k-1)(k-2)} x^{(k-2)\left(1 - \frac{1}{\log \log x}\right) + 1 + \log \log x} \zeta(k-1).$$

So up to what is effectively constant in k, and dropping lower order terms for a slightly suboptimal, but still sufficient for our purposes, error bound formula, we have completed the proof by induction. □

Corollary 6.6 (Asymptotics for a very special case of the functions $C_k(n)$). *For $k \gg 1$ sufficiently large, we have that*

$$C_{\Omega(n)}(n) \sim (\sigma_0 * \tau_{\log \log n-2})(n) \times \lambda(n) \frac{n^{\log \log n-1}}{(\log n)^{\log \log -1} \Gamma(\log \log n)}.$$

Moreover, by considering the average orders of the function $\nu_p(n)$ for p large and tending to infinity, we have bounds on the aysmptotic behavior of these functions of the form

$$\lambda(n) \widehat{\tau_0}(n) \overset{\blacktriangle}{\underset{\sim}{}} C_{\Omega(n)}(n) \overset{\blacktriangle}{\underset{\sim}{}} \lambda(n) \widehat{\tau_1}(n).$$

It suffices to take the functions

$$\widehat{\tau_0}(n) := \frac{1}{\log 2} \cdot \frac{\log n}{(\log n)^{\log \log n}} \cdot \frac{n^{\log \log n-1}}{\Gamma(\log \log n)}$$

$$\widehat{\tau_1}(n) := \frac{1}{2e \log 2} \cdot \frac{(\log n)^2}{(\log n)^{\log \log n}} \cdot \frac{n^{\log \log n}}{\Gamma(\log \log n)}.$$

Proof. The first stated formula follows from Theorem 3.6 by setting $k := \Omega(n) \sim \log \log n$ and simplifying. We evaluate the Dirichlet convolution functions and approximate as follows:

$$(\sigma_0 * \tau_{\log \log n - 2})(n) = \sum_{p|n} \binom{\nu_p(n) + \log \log n - 1}{\log \log n - 1}$$

$$\geq \sum_{p|n} \frac{(\nu_p(n) + \log \log n - 1)^{\log \log n - 1}}{(\log \log n)^{\log \log n - 1}}$$

$$\sim \frac{n}{\log 2}$$

$$(\sigma_0 * \tau_{\log \log n - 2})(n) \leq \left(\frac{(\nu_p(n) + \log \log n - 1)e}{\log \log n - 1} \right)^{\log \log n - 1}$$

$$\sim (2e)^{\log \log n - 1}$$

$$= \frac{n \cdot \log n}{2e \log 2}.$$

The upper and lower bounds are obtained from the next well known binomial coefficient approximations using Stirling's formula.

$$\frac{n^k}{k^k} \leq \binom{n}{k} \leq \frac{n^k}{k!} < \left(\frac{ne}{k} \right)^k \qquad \qquad \square$$

Using Lemma 6.1 directly is problematic since forming the summatory function of the exact $g^{-1}(n)$ that obey this formula leads to a nested recurrence relation involving $M(x)$ – e.g., more in-order sums of consecutive Möbius function terms yet again. Some suggestive numerical experiments suggest that this implicit recursive dependence of our new formulas for $M(x)$ can be more simply avoided by using an inexact, but still asymptotically sufficient in form expression for $g^{-1}(n)$. The next corollary provides the specific inexact, but asymptotically accurate formula for these inverse functions we have in mind.

What Corollary 6.7, and its somewhat involved (long) proof construction, allows us to do is provide a substantially simpler formula and limiting bound on the summatory functions $G^{-1}(x)$ of $g^{-1}(n)$. The form of this new formula for $G^{-1}(x)$ is established in Corollary 6.8, which is subsequently stated and easily given a short proof immediately after the next result in this section below. This is an important leap in expressing a manageable formula that we can use to bound these summatory functions from below when x is large as rigorously justified in Theorem 7.4.

Corollary 6.7 (Computing the inverse functions). *For $n \geq 2$ as $n \to \infty$ we have that*

$$g^{-1}(n) \sim \lambda(n) \times \sum_{d|n} C_{\Omega(d)}(d).$$

In particular, we can bound the error terms in the approximation of Lemma 6.1 by the previous formula to ensure that

$$\left| \frac{\lambda(n) \times \sum_{d|n} C_{\Omega(d)}(d)}{g^{-1}(n)} \right| \xrightarrow{n \to \infty} 1.$$

Proof. Using Lemma 6.1, it suffices to show that the squarefree divisors $d|n$ such that $\text{sgn}(\mu(d)\lambda(n/d)) = -1$ have an order of magnitude smaller magnitude (in the little-o

35

notation sense) than the corresponding cases of positive sign on the terms in the divisor sum from the lemma. This is because the sign of the terms in the Möbius inversion sum from the lemma will have already matched exactly that of the terms in $\lambda(n) \times \sum_{d|n} C_{\Omega(d)}(d)$ except possibly in these comparitively rare cases when $\operatorname{sgn}(\mu(d)\lambda(n/d)) = -1$.

Thus, we need only compute a reasonable bound of when such innacuracies between the exact formula from Lemma 6.1 differs from the approximation we claim works when we take the divisor sum over the unsigned $C_{\Omega(d)}(d)$ terms and weight by an overall factor of $\lambda(n)$. It is obvious that if we can show that the difference (ratio) in these two formulas is asymptotically negligible as we let $n \to \infty$, then we can use the substantially easier to evaluate approximation to calculate accurate new bounds on the summatory function of $g^{-1}(n)$ later on in the results found below (following the conclusion of this proof).

Initial strategy diverging from Möbius inversion. Let n have m_1 prime factors p_1 such that $v_{p_1}(n) = 1$, m_2 such that $v_{p_2}(n) = 2$, and the remaining $m_3 := \Omega(n) - m_1 - 2m_2$ prime factors of higher-order exponentation. We have a few cases to consider after re-writing the sum from the lemma in the following form:

$$
g^{-1}(n) = \lambda(n)C_{\Omega(n)}(n) + \sum_{i=1}^{\omega(n)} \left\{ \sum_{\substack{d|n \\ \omega(d)=\Omega(d)=i \\ \#\{p|d:\nu_p(d)=1\}=k_1 \\ \#\{p|d:\nu_p(d)=2\}=k_2 \\ \#\{p|d:\nu_p(d)\geq 3\}=k_3}} \mu(d)\lambda(n/d)C_{\Omega(n/d)}(n/d) \right\}.
$$

We obtain the following cases of the squarefree divisors contributing to the signage on the terms in the above sum:

- The sign of $\mu(d)$ is $(-1)^i = (-1)^{k_1+k_2+k_3}$;

- If $m_3 < \#\{p|n : \nu_p(n) \geq 3\}$, then $\lambda(n/d) = 1$ (since $\mu(n/d) = 0$);

- Given (k_1, k_2, k_3) as above, since $\lambda(n) = (-1)^{\Omega(n)}$, we have that $\mu(d) \cdot \lambda(n/d) = (-1)^{i-k_1-k_2}\lambda(n)$.

Thus we define the following sums, parameterized in the $(m_1, m_2, m_3; n)$, which corresponds to a change in expected parity transitioning from the Möbius inversion sum from Lemma 6.1 to the sum approximating $g^{-1}(n)$ defined at the start of this result:

$$
\widetilde{S}_{\text{odd}}(m_1, m_2, m_3; n) := \sum_{i=1}^{\omega(n)/2} \sum_{k_1=0}^{\lfloor \frac{i}{2} \rfloor} \sum_{k_2=0}^{\lfloor \frac{i}{2} \rfloor - k_1} \left[\binom{m_1}{2k_1+1}\binom{2m_2}{2k_2+1} + \binom{m_1}{2k_1}\binom{2m_2}{2k_2} \right] [i - k_1 - k_2 = k_3 \equiv m_3]_\delta
$$

$$
\widetilde{S}_{\text{even}}(m_1, m_2, m_3; n) := \sum_{i=1}^{\omega(n)/2} \sum_{k_1=0}^{\lfloor \frac{i}{2} \rfloor} \sum_{k_2=0}^{\lfloor \frac{i}{2} \rfloor - k_1} \left[\binom{m_1}{2k_1}\binom{2m_2}{2k_2+1} + \binom{m_1}{2k_1+1}\binom{2m_2}{2k_2} \right] [i - k_1 - k_2 = k_3 \equiv m_3]_\delta.
$$

Part I (Lower bounds on the inner sums of the count functions). We claim that

$$
\widetilde{S}_{\text{odd}}(m_1, m_2, m_3; n) \gg \binom{m_1}{i+1} + \binom{m_1}{\frac{i}{2}}\binom{2m_2-1}{\frac{i}{2}+1} \tag{11}
$$

$$
\widetilde{S}_{\text{even}}(m_1, m_2, m_3; n) \gg \binom{m_1}{i+1} + \binom{m_1}{\frac{i}{2}-1}\binom{2m_2}{\frac{i}{2}+1}.
$$

To prove (11) we have to provide a straightforward bound that represents the maximums of the terms in m_1, m_2. In particular, observe that for

$$\widetilde{S}_{\text{odd}}(m_1, m_2; u) = \sum_{k_1=0}^{u} \sum_{k_2=0}^{u-k_1} \left[\binom{m_1}{2k_1}\binom{2m_2}{2k_2+1} + \binom{m_1}{2k_1+1}\binom{2m_2}{2k_2} \right]$$

$$\widetilde{S}_{\text{even}}(m_1, m_2; u) = \sum_{k_1=0}^{u} \sum_{k_2=0}^{u-k_1} \left[\binom{m_1}{2k_1+1}\binom{2m_2}{2k_2+1} + \binom{m_1}{2k_1}\binom{2m_2}{2k_2} \right],$$

we have that

$$
\begin{aligned}
\widetilde{S}_{\text{odd}}(m_1, m_2; u) \overset{\blacktriangle}{\gtrsim} & \binom{m_1}{2u+1} + \max_{1 \le k_1 \le u} \binom{m_1}{2k_1+1}\binom{2m_2}{2u+1-2k_1} \\
= & \binom{m_1}{2u+1} + \binom{m_1}{2k_1+1}\binom{2m_2}{2u+1-2k_1}\bigg|_{k_1=\frac{u}{2}} \\
= & \binom{m_1}{2u+1} + \binom{m_1}{u+1}\binom{2m_2}{u+1} \\
\widetilde{S}_{\text{even}}(m_1, m_2; u) \overset{\blacktriangle}{\gtrsim} & \binom{m_1}{2u+1} + \max_{1 \le k_1 \le u} \binom{m_1}{2k_1}\binom{2m_2}{2u+1-2k_1} \\
= & \binom{m_1}{2u+1} + \binom{m_1}{u-1}\binom{2m_2}{u+1}.
\end{aligned}
$$

The lower bounds in (11) then follow by setting $u \equiv \lfloor \frac{i}{2} \rfloor$.

Part II (Bounding m_1, m_2, m_3 and effective (i, k_1, k_2) contributing to the count). We thus have to determine the asymptotic growth rate of $\widetilde{S}_{\text{odd}}(m_1, m_2, m_3; n) + \widetilde{S}_{\text{even}}(m_1, m_2, m_3; n)$, and show that it is of comparatively small order. First, we bound the count of non-zero m_3 for $n \le x$ from below. For the cases where we expect differences in signage, it's the last Iverson convention term that kills the order of growth, e.g., we expect differences when the parameter m_3 is larger than the usual configuration. We know that

$$\pi_k(x) \sim \frac{x}{\log x} \frac{(\log \log x)^{k-1}}{(k-1)!}.$$

Using the formula for $\pi_k(x)$, we can count the average orders of m_1, m_2 as

$$N_{m_1}(x) \approx \frac{1}{x}\#\{n \le x : \omega(n) = 1\} \sim \frac{\log \log x}{\log x}$$

$$N_{m_2}(x) \approx \frac{1}{x}\#\{n \le x : \omega(n) = 2\} \sim \frac{(\log \log x)^2}{\log x}.$$

Additionally, in Corollary 3.8 on page 17 we will prove a lower bound on $\widehat{\pi}_k(x)$. We use this result immediately below without proof.

When we have parameters with respect to some $n \ge 1$ such that $m_3 > 0$, it must be the case that

$$\Omega(n) - \omega(n) > \begin{cases} 0, & \text{if } \omega(n) \ge 2; \\ 1, & \text{if } \omega(n) = 1. \end{cases}$$

To count the number of cases $n \le x$ where this happens, we form the sums

$$N_{m_3}(x) \gg \pi_1(x) \times \sum_{k=3}^{\log \log x} \widehat{\pi}_k(x) + \sum_{k=2}^{\log \log x} \sum_{j=k+1}^{\log \log x} \pi_k(x)\widehat{\pi}_j(x)$$

$$\overset{\blacktriangle}{\underset{\sim}{\gtrsim}} \frac{A_0 x^2}{\sqrt{2\pi}} \cdot \frac{(\log x)^{2\log 2 + \frac{1}{3\log 2} - 2}}{(\log\log x)^{\frac{7}{2} + \log\log x}} + \frac{A_0 x^2}{2\pi} \cdot \frac{(\log x)^{2\log 2 + \frac{1}{3\log 2} - 1}}{(\log\log x)^{4 + \log\log x}}$$

$$\overset{\blacktriangle}{\underset{\sim}{\gtrsim}} \frac{A_0 x^2}{2\pi} \cdot \frac{(\log x)^{2\log 2 + \frac{1}{3\log 2} - 1 - \log\log\log x}}{(\log\log x)^4}.$$

Now in practice, we are not summing up $n \leq x$, but rather $n \leq \log\log x$. So the above function evaluates to

$$N_{m_3}(\log\log x) \gg \frac{A_0 (\log\log x)^2}{2\pi} \cdot \frac{(\log\log\log x)^{2\log 2 + \frac{1}{3\log 2} - 1 - \log\log\log\log\log x}}{(\log\log\log\log x)^4}.$$

Next, we go about solving the subproblem of finding when $i - k_1 - k_2 = m_3$. Clearly, since $k_1, k_2 \geq 0$, we have that $i \leq N_{m_3}(\log\log x)$. Moreover, since $2 \leq k_1 + k_2 \leq i/2$, when x is large, we actually obtain a number of soultions less than the order of

$$\mathcal{S}_0(x) := \frac{N_{m_3}(\log\log x)^2}{2}.$$

Part III (Putting it all together). Using the binomial coefficient inequality

$$\binom{n}{k} \geq \frac{n^k}{k^k},$$

we can work out carefully on paper using (11) that

$$\widetilde{S}_{\text{odd}}(m_1, m_2, m_3; x) \overset{\blacktriangle}{\underset{\sim}{\gtrsim}} \mathcal{S}_0(x) \left(\frac{\log\log\log x}{2\log x} \right)^{\frac{2\log\log x}{\log\log\log x} + 1} \left[1 + \frac{(\log\log\log x)^2}{\log^2 x} (4\log\log x \cdot \log\log\log x)^{\frac{\log\log x}{2\log\log\log x} + 1} \right]$$

$$\widetilde{S}_{\text{odd}}(m_1, m_2, m_3; x) \overset{\blacktriangle}{\underset{\sim}{\gtrsim}} \mathcal{S}_0(x) \left(\frac{\log\log\log x}{2\log x} \right)^{\frac{2\log\log x}{\log\log\log x} + 1} \left[1 + \left(\frac{\log x}{2\log\log\log x} \right) (8\log\log x)^{\frac{\log\log x}{\log\log\log x} + 1} \right].$$

Thus, after citing Corollary 6.6, we easily see the conclusion of this result that we stated above since

$$\left| \frac{\lambda(n) \times \sum_{d|n} C_{\Omega(d)}(d)}{g^{-1}(n)} \right| = \left| \frac{g^{-1}(n) + \widetilde{S}_{\text{even}}(m_1, m_2, m_3; n) + \widetilde{S}_{\text{odd}}(m_1, m_2, m_3; n)}{g^{-1}(n)} \right|$$

$$\leq \left| 1 + \frac{\widetilde{S}_{\text{even}}(m_1, m_2, m_3; n) + \widetilde{S}_{\text{odd}}(m_1, m_2, m_3; n)}{\widehat{\tau}_1(n)} \right| \xrightarrow{n\to\infty} 1,$$

and since

$$\left| \frac{\lambda(n) \times \sum_{d|n} C_{\Omega(d)}(d)}{g^{-1}(n)} \right| = \left| \frac{g^{-1}(n) + \widetilde{S}_{\text{even}}(m_1, m_2, m_3; n) + \widetilde{S}_{\text{odd}}(m_1, m_2, m_3; n)}{g^{-1}(n)} \right|$$

$$\geq \left| 1 - \frac{\widetilde{S}_{\text{even}}(m_1, m_2, m_3; n) + \widetilde{S}_{\text{odd}}(m_1, m_2, m_3; n)}{\widehat{\tau}_0(n)} \right| \xrightarrow{n\to\infty} 1.$$

In other words, the divisor sum in the corollary statement accuarately approximates the main term and sign of $g^{-1}(n)$ as $n \to \infty$. $\qquad\square$

Corollary 6.8. *We have that for sufficiently large x, as $x \to \infty$ that*

$$G^{-1}(x) \overset{\blacktriangle}{\gtrsim} \widehat{L}_0(\log \log x) \times \sum_{n \leq \log \log x} \lambda(n) \cdot C_{\Omega(n)}(n),$$

where the function

$$\widehat{L}_0(\log \log x) := (-1)^{\lfloor \log \log \log \log x \rfloor} \sqrt{\frac{2}{\pi}} A_0 \cdot (\log \log x) \frac{(\log \log \log x)^{2 \log 2 + \frac{1}{3 \log 2} - 1}}{(\log \log \log \log x)^{\frac{5}{2} + \log \log \log \log x}},$$

and with the exponent $2 \log 2 + \frac{1}{3 \log 2} - 1 \approx 0.867193$.

A10-2020.04-26

Proof. Using Corollary 6.7, we have that

$$G^{-1}(x) \approx \sum_{n \leq x} \lambda(n) \cdot (g^{-1} * 1)(n)$$

$$= \sum_{d \leq \log \log x} C_{\Omega(d)}(d) \times \sum_{n=1}^{\lfloor \frac{x}{d} \rfloor} \lambda(dn).$$

Now we see that by complete additivity (multiplicativity) of $\Omega(n)$ (as indicated by the sign of $\lambda(n)$) that

$$\sum_{n=1}^{\lfloor \frac{x}{d} \rfloor} \lambda(dn) = \sum_{n=1}^{\lfloor \frac{x}{d} \rfloor} \lambda(d) \lambda(n) = \lambda(d) \sum_{n \leq \lfloor \frac{x}{d} \rfloor} \lambda(n).$$

Using the result proved in Section 5 (see Theorem 3.8) we can establish that

$$\sum_{n \leq x} \lambda(n) \gg \sum_{k \leq \log \log x} (-1)^k \cdot \widehat{\pi}_k(x) =: \widehat{L}_0(x).$$

Then since for large enough x and $d \leq x$,

$$\log(x/d) \sim \log x, \log \log(x/d) \sim \log \log x,$$

we can obtain the stated result, e.g., so that $\widehat{L}_0(x) \sim \widehat{L}_0(x/d)$ for large $x \to \infty$. $\qquad \square$

7 Key applications: Establishing lower bounds for $M(x)$ by cases along infinite subsequences

7.1 The culmination of what we have done so far

As noted before in the previous subsections, we cannot hope to evaluate functions weighted by $\lambda(n)$ except for on average using Abel summation. For this task, we need to know the bounds on $\widehat{\pi}_k(x)$ we developed in the proof of Corollary 3.8. A summation by parts argument proves the crux of the next proposition providing an integral formula based expression that approximates $M(x)$ closely.

Proposition 7.1. *For all sufficiently large x, we have that*

$$M(x) \approx G^{-1}(x) + x \cdot \int_1^{x/2} \frac{G^{-1}(t)}{t^2 \cdot \log(x/t)} dt, \tag{12}$$

where $G^{-1}(x) := \sum_{n \leq x} g^{-1}(n)$ is the summatory function of $g^{-1}(n)$.

Proof. We know by applying Corollary 3.3 that

$$M(x) = \sum_{k=1}^{x} g^{-1}(k)(\pi(x/k) + 1)$$

$$= G^{-1}(x) + \sum_{k=1}^{x} g^{-1}(k)\pi(x/k),$$

where we can drop the asymptotically unnecessary floored integer-valued arguments to $\pi(x)$ in place of its approximation by $\pi(x) \sim \frac{x}{\log x}$. In fact, since we can always bound

$$\frac{Ax}{\log x} \leq \pi(x) \leq \frac{Bx}{\log x},$$

for suitably defined absolute constants, $A, B > 0$, we are not losing any precision asymptotically by making this small leap in approximation from exact summation (in the first formula) to the integral formula representing convolution (in the second formula below).

Now what we require to sum and simplify the right-hand-side summation from the last equation is nothing short of an ordinary summation by parts argument. Namely, we obtain that for sufficiently large $x \geq 2$ [*]

$$\sum_{k=1}^{x} g^{-1}(k)\pi(x/k) = G^{-1}(x)\pi(1) + \sum_{k=1}^{x-1} G^{-1}(k)\left[\pi\left(\frac{x}{k}\right) - \pi\left(\frac{x}{k+1}\right)\right]$$

$$= \sum_{k=1}^{x/2} G^{-1}(k)\left[\pi\left(\frac{x}{k}\right) - \pi\left(\frac{x}{k+1}\right)\right]$$

$$\approx \sum_{k=1}^{x/2} G^{-1}(k)\left[\frac{x}{k \cdot \log(x/k)} - \frac{x}{(k+1) \cdot \log(x/k)}\right]$$

$$= \sum_{k=1}^{x/2} G^{-1}(k)\frac{x}{k^2 \cdot \log(x/k)}.$$

---

[*]Since $\pi(1) = 0$, the actual range of summation corresponds to $k \in \left[1, \frac{x}{2}\right]$.

40

Now since for x large enough the summand is monotonic as k ranges in order over $k \in [1, x/2]$, and since the summands in the last equation are smooth functions of k (and x), and also since $G^{-1}(x)$ is a summatory function with jumps at the positive integers (signed or not), we can now write that

$$M(x) \approx G^{-1}(x) + x \cdot \int_1^{x/2} \frac{G^{-1}(t)}{t^2 \cdot \log(x/t)} dt.$$

The last equation follows by transforming a summation whose summand, say $S(k)$, is constant on each interval $k \in [n, n+1)$ for integers $n \geq 1$ into an easier to parse integral-based format. Since the bounds of integration are finite, we do not need to dwell on the oscillatory nature of the factor of $G^{-1}(t)$ in the translation from summation to integral representation. $\hfill\square$

7.1.1 From the routine: Proofs of a few cut-and-dry lemmas

The results proved next in Lemma 7.2 and Lemma 7.3 are key to completely and carefully justifying the asymptotic bounds obtained below in Theorem 7.4. We require proofs of these two lemmas to be rigorous, dotting and crossing every tick mark (so to speak) in the checklist of conditions we require in the key proof of the theorem below. Thus these two routine results are actually necessary to prove before we can return to the truly interesting matter of bounding $M(x)/\sqrt{x}$ in the next subsection.

Lemma 7.2. *Suppose that $f_k(n)$ is a sequence of arithmetic functions such that $f_k(n) > 0$ for all $n \geq 1$, $f_0(n) = \delta_{n,1}$, and $f_{\Omega(n)}(n) \overset{\blacktriangle}{\sim} \widehat{\tau}_\ell(n)$ as $n \to \infty$ where $\widehat{\tau}_\ell(t)$ is a continuously differentiable function of t for all large enough $t \gg 1$ [†]. We define the λ-sign-scaled summatory function of f as follows:*

$$F_\lambda(x) := \sum_{\substack{n \leq x \\ \Omega(n) \leq x}} \lambda(n) \cdot f_{\Omega(n)}(n).$$

Let

$$A_\Omega^{(\ell)}(t) := \sum_{k=1}^{\lfloor \log\log t \rfloor} (-1)^k \widehat{\pi}_k^{(\ell)}(t),$$

where $\widehat{\pi}_k(x) \geq \widehat{\pi}_k^{(\ell)}(x) \geq 0$ and $\widehat{\pi}_k^{(\ell)}(x)$ is a smooth monotone non-decreasing function of x when x is sufficiently large. Then we have that

$$F_\lambda(\log\log x) \overset{\blacktriangle}{\sim} A_\Omega^{(\ell)}(\log\log x)\widehat{\tau}_\ell(\log\log x) - \int_1^{\log\log x} A_\Omega^{(\ell)}(t)\widehat{\tau}_\ell'(t)dt.$$

Proof. We first note that we can form an accurate $C^1(\mathbb{R})$ approximation by the smoothness of $\widehat{\pi}_k^{(\ell)}(x)$ that allows us to apply the Abel summation formula using the summatory function $A_\Omega^{(\ell)}(t)$ for t on any connected subinterval of $[1, \infty)$. The formula for $F_\lambda(x)$ is valid

---

[†]We will require that $\widehat{\tau}_\ell(t) \in C^1(\mathbb{R})$ when we apply the Abel summation formula in the proof of Theorem 7.4. At this point, it is technically an unnecessary condition that is vacously satisfied by assumption (by requirement) and will importantly need to hold when we specialize to the actual functions employed to form our new bounds in the theorem.

by Abel summation provided that

$$\left| \frac{\displaystyle\sum_{\log\log t < k \le \frac{\log t}{\log 2}} (-1)^k \widehat{\pi}_k(t)}{A_\Omega^{(\ell)}(t)} \right| = o(1), \text{ as } t \to \infty,$$

e.g., the asymptotically dominant terms indicating the parity of $\lambda(n)$ are encompassed by the terms summed by $A_\Omega^{(\ell)}(t)$ for sufficiently large t as $t \to \infty$. Using the arguments in Montgomery and Vaughan [11, §7; Thm. 7.20] (see Theorem 5.2), we can see that uniformly in x and for any $r \in (1, 2)$

$$\left| \frac{\displaystyle\sum_{\log\log x < k \le x} (-1)^k \cdot \widehat{\pi}_k(x)}{A_\Omega(x)} \right| \le \left| \frac{B(x,r)}{\log\log x + B(x,1)} \right| \approx \left| \frac{B(x,r)}{B(x,1)} \right| = o(1), \tag{13}$$

as $x \to \infty$. Recall that the function $B(x, r)$ is defined and bounded as in the cited theorem from [11] re-stated on page 25 as it appears in the reference.

Thus we conclude that we have captured the asymptotically dominant main order terms in our formula as $x \to \infty$ using the definition of $A_\Omega(x)$. In other words, taking the sum over the summands that defines $A_\Omega(x)$ only over the truncated range of $k \in [1, \log\log x]$ does not affect the limiting asymptotically dominant terms obtained from using this formulation of the summatory function – even when we should technically index over all $k \in [1, \log_2(x)]$ to obtain a precise formula for this function. $\qquad\square$

Observe that we use the superscript and subscript of (ℓ) not to denote a formal parameter to the functions we define below, but instead to denote that these functions form *lower bound* approximations to other forms of the functions without the scripted (ℓ).

Lemma 7.3. *Suppose that $\widehat{\pi}_k(x) \ge \widehat{\pi}_k^{(\ell)}(x) \ge 0$ for sufficiently large x with $\widehat{\pi}_k^{(\ell)}(x)$ a monotone non-decreasing real-valued function. Let*

$$A_\Omega^{(\ell)}(x) := \sum_{k \le \log\log x} (-1)^k \widehat{\pi}_k^{(\ell)}(x)$$

$$A_\Omega(x) := \sum_{k \le \log\log x} (-1)^k \widehat{\pi}_k(x).$$

Then for all sufficiently large x, we have that

$$|A_\Omega(x)| \gg |A_\Omega^{(\ell)}(x)|.$$

Proof. Given an explicit smooth lower bounding function, $\widehat{\pi}_k^{(\ell)}(x)$, we define the similarly smooth and monotone residual terms in approximating $\widehat{\pi}_k(x)$ using the following notation:

$$\widehat{\pi}_k(x) = \widehat{\pi}_k^{(\ell)}(x) + \widehat{E}_k(x).$$

Then we can form the ordinary form (i.e., the exact, non-lower-bound) of the summatory functions as

$$|A_\Omega(x)| = \left| \sum_{k \le \frac{\log\log x}{2}} [\widehat{\pi}_{2k}(x) - \widehat{\pi}_{2k-1}(x)] \right|$$

42

$$\geq \left| A_\Omega^{(\ell)}(x) - \sum_{k \leq \frac{\log \log x}{2}} \widehat{E}_{2k-1}(x) \right|.$$

If the latter sum,

$$\mathrm{ES}(x) := \sum_{k \leq \frac{\log \log x}{2}} \widehat{E}_{2k-1}(x) \to \infty,$$

as $x \to \infty$, then we can always find some absolute (by monotonicity) $C_0 > 0$ such that $\mathrm{ES}(x) \leq C_0 \cdot A_\Omega(x)$. If, on the other hand, this sum becomes constant as $x \to +\infty$, then we also clearly have some absolute $C_1 > 0$ such that $|A_\Omega(x)| \geq C_1 \cdot |A_\Omega^{(\ell)}(x)|$. In either case, the claimed result holds for all large enough x. $\qquad\square$

7.1.2 A proof of the key bound on $G^{-1}(x)$

We use the result of Corollary 6.6 and Corollary 3.8 to prove the following central theorem:

Theorem 7.4 (Asymptotics and bounds for the summatory functions $G^{-1}(x)$). *We define the lower summatory function, $G_\ell^{-1}(x)$, to provide bounds on the magnitude of $G^{-1}(x)$:*

$$|G_\ell^{-1}(x)| \ll |G^{-1}(x)|,$$

for all sufficiently large $x \gg 1$. We have the next asymptotic approximations for the lower summatory function where $C_{\ell,1}$ is the absolute constant defined by

$$C_{\ell,1} = \frac{A_0}{\sqrt{2}\pi^{3/2}\log 2} = \frac{2^{\frac{33}{16}}}{81 \cdot \pi^2 e^3 \log^4(2)} \exp\left(-\frac{55}{4}\log^2(2)\right) \approx 1.52355 \times 10^{-6}.$$

That is, we have

$$\left| G_\ell^{-1}(x) \right| \overset{\blacktriangle}{\sim} C_{\ell,1} \cdot (\log x) \cdot \frac{(\log\log\log x)^{5+2\log 2 + \frac{1}{3\log 2}}}{\log\log\log\log x}.$$

The exponent in the previous equation is numerically approximated as $5 + 2\log 2 + \frac{1}{3\log 2} \approx 6.86719$.

Proof Initial Sketch: Logarithmic scaling of parameters to the accurate order of the inverse functions. For the sums given by

$$S_{g^{-1}}(x) := \sum_{n \leq x} \lambda(n) \cdot C_{\Omega(n)}(n),$$

we notice that using the asymptotic bounds (rather than the exact formulas) for the functions $C_{\Omega(n)}(n)$, we have over-summed by quite a bit. In particular, following from the intent behind the constructions in the last sections, we are really summing only over all $n \leq x$ with $\Omega(n) \leq x$. Since $\Omega(n) \leq \lfloor \log_2 n \rfloor$, many of the terms in the previous equation are actually zero (recall that $C_0(n) = \delta_{n,1}$). So we are actually only summing to the average order of $\mathbb{E}[\Omega(n)] = \log\log n$ in practice, or to the slightly larger bound if the leading sign term on $G_\ell^{-1}(x)$ is negative. Hence, the sum (in general) that we are really interested in bounding is bounded below in magnitude by $S_{g^{-1}}(\log\log x)$ as bounded in Corollary 6.8. After noting this adjustment, we can then safely apply the asymptotic formulas for the functions $C_k(n)$ from Corollary 6.6 that hold once we have verified these constraints on k, n, x. $\qquad\square$

Proof. Recall from our proof of Corollary 3.8 that a lower bound is given by

$$\widehat{\pi}_k(x) \overset{\blacktriangle}{\gtrsim} \frac{A_0 \cdot x}{\log x \cdot (\log\log x)^4 \cdot (k-1)!} \cdot \left(\frac{4}{3e \log 2}\right)^k.$$

Thus we can form a lower summatory function indicating the parity of all $\Omega(n)$ for $n \leq x$ as

$$A_\Omega^{(\ell)}(t) = \sum_{k \leq \log\log t} (-1)^k \widehat{\pi}_k(x) \tag{14}$$

$$\overset{\blacktriangle}{\gtrsim} (-1)^{\lfloor \log\log\log\log x \rfloor} \sqrt{\frac{2}{\pi}} A_0 \cdot (\log\log x) \frac{(\log\log\log x)^{2\log 2 + \frac{1}{3\log 2} - 1}}{(\log\log\log\log x)^{\frac{5}{2} + \log\log\log\log x}}.$$

Moreover, due to Lemma 7.3, we have that the right-hand-side bound in the last equation actually holds when we remove the signedness present in the formula.

Next, by Corollary 6.6 we recover from the approximation to the *polygamma function*, $\psi^{(0)}(x) \sim \log x$, that

$$\widehat{\tau}_0'(t) = \frac{d}{dx}\left[\frac{1}{\log 2 \cdot \Gamma(\log\log t)} \frac{t^{\log\log t - 1}}{(\log t)^{\log\log t - 1}}\right]$$

$$\overset{\blacktriangle}{\gtrsim} \frac{\log\log t}{\sqrt{2\pi} \log 2 \cdot \Gamma(\log\log t)} \cdot \frac{t^{\log\log t - 2}}{(\log t)^{\log\log t - 1}}.$$

As in Lemma 7.2, we apply Abel summation to obtain that

$$G_\ell^{-1}(x) = \widehat{\tau}_0(\log\log x) A_\Omega^{(\ell)}(\log\log x) - \widehat{\tau}_0(u_0) A_\Omega^{(\ell)}(u_0) - \int_{u_0}^{\log\log x} \widehat{\tau}_0'(t) A_\Omega^{(\ell)}(t) dt. \tag{15}$$

The integral term in (15) is summed approximately as

$$\int_{u_0 - 1}^{\log\log x} \widehat{\tau}_0'(t) A_\Omega^{(\ell)}(t) dt \sim \sum_{k=u_0+1}^{\frac{1}{2}\log\log\log\log x} \left(I_\ell\left(e^{e^{2k+1}}\right) - I_\ell\left(e^{e^{2k}}\right)\right) \frac{e^{e^{2k}}}{\Gamma(2k)} \tag{16}$$

$$\approx C_0(u_0) + (-1)^{\lfloor \frac{\log\log\log\log x}{2} \rfloor} \times \int_{\frac{\log\log\log\log x}{2} - 1}^{\frac{\log\log\log\log x}{2}} \frac{I_\ell\left(e^{e^{2k}}\right)}{\Gamma(2k)} e^{e^{2k}} dk.$$

We define the integrand function, $I_\ell(t) := \widehat{\tau}_0'(t) A_\Omega^{(\ell)}(t)$, from the previous equations with some limiting simplifications for $k \in \left[\frac{\log\log\log\log x}{2} - 1, \frac{\log\log\log\log x}{2}\right]$ as

$$I_\ell\left(e^{e^{2k}}\right) \frac{e^{e^{2k}}}{\Gamma(2k)} \overset{\blacktriangle}{\gtrsim} \frac{A_0}{2^{3/2}\pi \log 2 \cdot 4^k \cdot k^{2k+\frac{1}{2}}} \cdot \frac{\exp\left(2ke^{2k} - 4k^2 + 2k\left(\frac{1}{3\log 2} + 2\log 2\right)\right)}{\Gamma(2k+1)}. \tag{17}$$

So using the lower bound on the integrand in (17), and applying Stirling's formula, we find

that [‡]

$$\int_{\frac{\log\log\log\log x}{2}-1}^{\frac{\log\log\log\log x}{2}} I_\ell\left(e^{e^{2k}}\right)\frac{e^{e^{2k}}}{\Gamma\left(2k+1\right)}dk \tag{18}$$

$$\overset{\blacktriangle}{\gtrsim}\frac{A_0}{\sqrt{2}\pi^{3/2}\log 2}\cdot\frac{(\log\log\log x)^{1+2\log 2+\frac{1}{3\log 2}}}{\log\log\log\log x}\times$$

$$\times\left(\frac{\log\log x}{\log\log\log x\cdot(\log\log\log\log x)^2}\right)^{\log\log\log\log x}$$

$$\overset{\blacktriangle}{\gtrsim}\frac{A_0}{\sqrt{2}\pi^{3/2}\log 2}\cdot(\log x)\cdot\frac{(\log\log\log x)^{5+2\log 2+\frac{1}{3\log 2}}}{\log\log\log\log x}.$$

Finally, we need to expand the core components of the leading terms in (15) as

$$\widehat{\tau}_0(\log\log x)\overset{\blacktriangle}{\gtrsim}\frac{\sqrt{\log\log\log\log x}}{2\pi\cdot\log 2}\left(\frac{\log\log\log x}{\log\log x}\right)^2\times$$

$$\times\left(\frac{\log\log x}{\log\log\log x\cdot\log\log\log\log x}\right)^{\log\log\log\log x}$$

$$A_\Omega^{(\ell)}(\log\log x)\overset{\blacktriangle}{\gtrsim}(-1)^{\lfloor\log\log\log\log\log\log x\rfloor}\sqrt{\frac{2}{\pi}}A_0\cdot(\log\log\log\log x)\times$$

$$\times\frac{(\log\log\log\log x)^{2\log 2+\frac{1}{3\log 2}-1}}{(\log\log\log\log\log x)^{\frac{5}{2}+\log\log\log\log\log\log x}}.$$

These last formulas imply the forms of the stated bounds when we drop the lower-order constant term and multiply through by the bounds for the function $\widehat{L}_0(\log\log x)$ as proved in Corollary 6.8.

It is clear that the asymptotically dominant behavior of the lower bound for $|G_\ell^{-1}(x)|$ comes from the integral term calculated in the last equation from (18). Thus our otherwise large-order looking bounds obtained above actually just constribute a masked logarithmically growing factor to the main term of our bound. □

7.2 Lower bounds on the scaled Mertens function along an infinite sub-sequence

Proof of Corollary 3.9. It suffices to take $u_0 = e^{e^{e^e}}$. Now, we break up the integral over $t \in [u_0, x/2]$ into two pieces: one that is easily bounded from $u_0 \leq t \leq \sqrt{x}$, and then another that will conveniently give us our logarithmically slow-growing tendency towards infinity along the subsequence.

First, since $\pi(j) = \pi(\sqrt{x})$ for all $\sqrt{x} \leq j < x$, we can take the first chunk of the interval of integration and bound it using (12) as

$$-\int_{u_0}^{\sqrt{x}}\frac{2\sqrt{x}}{t^2\log(x)}G_\ell^{-1}(t)dt\overset{\blacktriangle}{\gtrsim}B_{\ell,2}\times\frac{2\sqrt{x}}{\log(x)}\cdot\left(\max_{u_0\leq t\leq\sqrt{x}}G_\ell^{-1}(t)\right)$$

[‡]Here, we have used that for sufficiently large x,

$$(\log\log\log x)^2\overset{\blacktriangle}{\gtrsim}\exp\left(-(\log\log\log\log x)^2\right).$$

We also simplify the term involved in the resulting bound by writing

$$\left(\frac{\log\log x}{\log\log\log\log x}\right)^{\log\log\log\log x}=(\log\log\log x)^{\frac{\log\log x}{\log\log\log\log x}}=\exp\left(\log\log x\right)=\log x.$$

$$= o(\sqrt{x}),$$

where $B_{\ell,2}$ can be taken as an indefinite, but still absolute constant with respect to u_0. The maximum in the previous equation is clearly attained by taking $t := \sqrt{x}$. The bound follows, and will be good enough to dispense with this term as of the negligible limiting form of $o(1)$ when we scale the function as $|M(x)|/\sqrt{x}$.

Next, we have to prove a related bound on the second portion of the interval from $\sqrt{x} \leq t \leq x/2$:

$$-\int_{\sqrt{x}}^{x/2} \frac{2x}{t^2 \log(x)} \cdot G_\ell^{-1}(t)dt \overset{\blacktriangle}{\gtrsim} \frac{2\sqrt{x}}{\log x} \cdot \left(\max_{\sqrt{x} \leq t \leq x/2} G_\ell^{-1}(t) \right)$$

$$= 2C_{\ell,1}\sqrt{x} \cdot \frac{(\log\log\log\sqrt{x})^{5+2\log 2 + \frac{1}{3\log 2}}}{\log\log\log\log\sqrt{x}} + o(1).$$

Finally, since $G_\ell^{-1}(x) = o(\sqrt{x})$, we obtain in total that as $x \to \infty$ along this infinite subsequence:

$$\frac{|M(x)|}{\sqrt{x}} \overset{\blacktriangle}{\gtrsim} 2C_{\ell,1} \cdot \frac{(\log\log\log\sqrt{x})^{5+2\log 2 + \frac{1}{3\log 2}}}{\log\log\log\log\sqrt{x}} + o(1),$$

The above expression tends to $+\infty$ as $x \to \infty$, however, only extremly (logarithmically) slowly and along the defined infinite subsequence of asymptotically very large x defined by nested powers of the exponential function. $\qquad\square$

8   Conclusions

8.1   Summary

- Using average order bounds, summatory functions, and the $\overset{\blacktriangle}{\sim}$-type relations for lower bounds.

- Somewhat oddly, we did not need substantially improved bounds on $L_0(x) := \sum_{n \leq x} \lambda(n)$ than what is already known in upper bound form to obtain our new bounds on the Mertens function, aka, summatory function of the "testier" Möbius function.

8.2   Future research and work that still needs to be done

- Refinements of these bounds to find the tightest possible lower (limit supremum) bounds, e.g., proofs of an optimal version of Gonek's original conjecture.

- Generalizations to weighted Mertens functions of the form $M_\alpha(x) := \sum_{n \leq x} \mu(n) n^{-\alpha}$.

- Indications of sign changes and exceptionally small, or zero values of $M(x)$.

- What our more combinatorial approach to bounding $M(x)$ effectively suggests about necessary, but unproved, zeta zero bounds that have historically formed the basis for arguments bounding $M(x)$ using Mellin inversion.

- Evaluate alternate strategies and approaches using different Dirichlet convolution functions besides g and $g^{-1}(n)$ (corresponding to $\pi(x)$) with Theorem 3.1.

8.3   Motivating a general technique towards bounding the summatory functions of arbitrary arithmetic f

8.4   The general construction using Theorem 3.1

8.4.1   A proposed generalization

For each $n \geq 1$, let $A(n) \subseteq \{d : 1 \leq d \leq n, d|n\}$ be a subset of the divisors of n. We say that a natural number $n \geq 1$ is A-primitive if $A(n) = \{1, n\}$. Under a list of assumptions so that the resulting A-convolutions are *regular convolutions*, we get a generalized multiplicative Möbius function [15, §2.2]:

$$\mu_A(p^\alpha) = \begin{cases} 1, & \alpha = 0; \\ -1, & p^\alpha > 1 \text{ is } A\text{-primitive}; \\ 0, & \text{otherwise.} \end{cases}$$

We also define the functions $\omega_A(n) := \#\{d|n : \text{d is an A} - \text{primitive factor of n}\}$ and $\Omega_A(n) := \#\{p^\alpha||n : \text{p is an A} - \text{primitive factor of n}\}$. Then the characteristic function of the set $A := \cup_{n \geq 1} A(n)$ is given by $\chi_A(n) = [n \in A]_\delta$. By Möbius inversion, we have that $\chi_A = \omega_A *_A \mu_A$. Moreover, for the A-counting function, $\pi_A(x)$, defined by

$$\pi_A(x) := \#\{n \leq x : n \in A\},$$

we can define a corresponding notion of a generalized A-Mertens function, $M_A(x) := \sum_{n \leq x} \mu_A(n)$. This function then satisfies (by Theorem 3.1) the relation that

$$M_A(x) = \sum_{k=1}^{x} (\omega_A + 1)^{-1}(k) \cdot \pi_A(x/k),$$

where the inverse function, $(\omega_A + 1)^{-1}(n)$, is defined with respect to A-convolution. We conjecture, but do not prove here, that $\operatorname{sgn}((\omega_A + 1)^{-1}(n)) = \lambda_A(n) =: (-1)^{\Omega_A(n)}$.

Using formulas similar in construction to (12), we can differentiate to find expressions for $\pi_A(x)$. The significance of this is that provided we can prove sufficiently large bounds for $M_A(x)$ along the same lines as we have done for $M(x)$, the resulting formula may be able to speak towards the density, or even infinitude in special cases, of the set A.

8.5 Working / TODO

(i) The average order, $\mathbb{E}[\omega(n)] = \log \log n$, imparts an iterated logarithmic structure to our expansions, which many have conjectured we should see in limiting bounds on $M(x)$, but which are practically elusive in most non-conjectural known formulas I have seen proved rigorously in print.

(ii) The additivity of $\omega(n)$ dictates that the sign of $g^{-1}(n) = (\omega + 1)^{-1}(n)$ is $\operatorname{sgn}(g^{-1}(n)) = \lambda(n)$ (see Proposition 4.2). The corresponding weighted summatory functions of $\lambda(n)$ have more established predictable properties, such as known sign biases and upper bounds. These summatory functions are generally speaking more regular and easier to work with than traditional approaches to summing $M(x)$ and its complicating summand terms of the Möbius function. Note that our proof is essentially much different than what is known about sums of consecutive values of $\mu(n)$ over short intervals, both in interpretation and methodology.

More generally, we have that for f a non-negative additive arithmetic function that vanishes at one, $\operatorname{sgn}((f+1)^{-1}) = \lambda(n) = (-1)^{\Omega(n)}$. We can state similar properties for the common case of multiplicative f in the form of the following result: If $f(n) > 0$ for all $n \geq 1$ and f is multiplicative, then $\operatorname{sgn}(f^{-1}(n)) = (-1)^{\omega(n)}$.

References

[1] T. M. Apostol. *Introduction to Analytic Number Theory.* Springer–Verlag, 1976.

[2] P. Erdös. On the integers having exactly k prime factors. *Annals of Mathematics*, 40(1):53–66, 1946.

[3] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009 (Third printing 2010).

[4] J. Friendlander and H. Iwaniec. *Opera de Cribero.* American Mathematical Society, 2010.

[5] C. E. Fröberg. On the prime zeta function. *BIT Numerical Mathematics*, 8:87–202, 1968.

[6] H. W. Gould and T. Shonhiwa. A catalog of interesting Dirichlet series. *Missouri J. Math. Sci.*, 20(1):2–18, 2008.

[7] G. H. Hardy and E. M. Wright, editors. *An Introduction to the Theory of Numbers.* Oxford University Press, 2008 (Sixth Edition).

[8] G. Hurst. Computations of the Mertens function and improved bounds on the Mertens conjecture. *https: // arxiv. org/ pdf/ 1610. 08551/* , 2017.

[9] T. Kotnik and H. te Riele. The Mertens conjecture revisited. *Algorithmic Number Theory*, 7th International Symposium, 2006.

[10] T. Kotnik and J. van de Lune. On the order of the Mertens function. *Exp. Math.*, 2004.

[11] H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory: I. Classical Theory.* Cambridge, 2006.

[12] A. M. Odlyzko and H. J. J. te Riele. Disproof of the Mertens conjecture. *J. REINE ANGEW. MATH*, 1985.

[13] Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, editors. *NIST Handbook of Mathematical Functions.* Cambridge University Press, 2010.

[14] P. Ribenboim. *The new book of prime number records.* Springer, 1996.

[15] J. Sándor and B. Crstici. *Handbook of Number Theory II.* Kluwer Academic Publishers, 2004.

[16] M. D. Schmidt. A catalog of interesting and useful Lambert series. *arXiv/math.NT(2004.02976)*, 2020.

[17] N. J. A. Sloane. The Online Encyclopedia of Integer Sequences, 2020.

[18] K. Soundararajan. Partial sums of the Möbius function. *Annals of Mathematics*, 2009.

A   Appendix: Supplementary tables and data

T.1 Table: Computations with a highly signed Dirichlet inverse function

n	Primes		Sqfree	PPower	$\bar{\mathbb{S}}$		$g^{-1}(n)$	$\lambda(n)\,\text{sgn}(g^{-1}(n))$	$\lambda(n)g^{-1}(n) - \hat{f}_1(n)$	$\lambda(n)g^{-1}(n) - \hat{f}_2(n)$		$G^{-1}(n)$	$G_+^{-1}(n)$	$G_-^{-1}(n)$
1	1^1	–	Y	N	N	–	1	1	0	0	–	1	1	0
2	2^1	–	Y	Y	N	–	-2	1	0	0	–	-1	1	-2
3	3^1	–	Y	Y	N	–	-2	1	0	0	–	-3	1	-4
4	2^2	–	N	Y	N	–	2	1	0	-1	–	-1	3	-4
5	5^1	–	Y	Y	N	–	-2	1	0	0	–	-3	3	-6
6	$2^1 3^1$	–	Y	N	N	–	5	1	0	-1	–	2	8	-6
7	7^1	–	Y	Y	N	–	-2	1	0	0	–	0	8	-8
8	2^3	–	N	Y	N	–	-2	1	0	-2	–	-2	8	-10
9	3^2	–	N	Y	N	–	2	1	0	-1	–	0	10	-10
10	$2^1 5^1$	–	Y	N	N	–	5	1	0	-1	–	5	15	-10
11	11^1	–	Y	Y	N	–	-2	1	0	0	–	3	15	-12
12	$2^2 3^1$	–	N	N	Y	–	-7	1	2	-2	–	-4	15	-19
13	13^1	–	Y	Y	N	–	-2	1	0	0	–	-6	15	-21
14	$2^1 7^1$	–	Y	N	N	–	5	1	0	-1	–	-1	20	-21
15	$3^1 5^1$	–	Y	N	N	–	5	1	0	-1	–	4	25	-21
16	2^4	–	N	Y	N	–	2	1	0	-3	–	6	27	-21
17	17^1	–	Y	Y	N	–	-2	1	0	0	–	4	27	-23
18	$2^1 3^2$	–	N	N	Y	–	-7	1	2	-2	–	-3	27	-30
19	19^1	–	Y	Y	N	–	-2	1	0	0	–	-5	27	-32
20	$2^2 5^1$	–	N	N	Y	–	-7	1	2	-2	–	-12	27	-39
21	$3^1 7^1$	–	Y	N	N	–	5	1	0	-1	–	-7	32	-39
22	$2^1 11^1$	–	Y	N	N	–	5	1	0	-1	–	-2	37	-39
23	23^1	–	Y	Y	N	–	-2	1	0	0	–	-4	37	-41
24	$2^3 3^1$	–	N	N	Y	–	9	1	4	-3	–	5	46	-41
25	5^2	–	N	Y	N	–	2	1	0	-1	–	7	48	-41
26	$2^1 13^1$	–	Y	N	N	–	5	1	0	-1	–	12	53	-41
27	3^3	–	N	Y	N	–	-2	1	0	-2	–	10	53	-43
28	$2^2 7^1$	–	N	N	Y	–	-7	1	2	-2	–	3	53	-50
29	29^1	–	Y	Y	N	–	-2	1	0	0	–	1	53	-52
30	$2^1 3^1 5^1$	–	Y	N	N	–	-16	1	0	-4	–	-15	53	-68
31	31^1	–	Y	Y	N	–	-2	1	0	0	–	-17	53	-70
32	2^5	–	N	Y	N	–	-2	1	0	-4	–	-19	53	-72
33	$3^1 11^1$	–	Y	N	N	–	5	1	0	-1	–	-14	58	-72
34	$2^1 17^1$	–	Y	N	N	–	5	1	0	-1	–	-9	63	-72
35	$5^1 7^1$	–	Y	N	N	–	5	1	0	-1	–	-4	68	-72
36	$2^2 3^2$	–	N	N	Y	–	14	1	9	1	–	10	82	-72
37	37^1	–	Y	Y	N	–	-2	1	0	0	–	8	82	-74
38	$2^1 19^1$	–	Y	N	N	–	5	1	0	-1	–	13	87	-74
39	$3^1 13^1$	–	Y	N	N	–	5	1	0	-1	–	18	92	-74
40	$2^3 5^1$	–	N	N	Y	–	9	1	4	-3	–	27	101	-74
41	41^1	–	Y	Y	N	–	-2	1	0	0	–	25	101	-76
42	$2^1 3^1 7^1$	–	Y	N	N	–	-16	1	0	-4	–	9	101	-92
43	43^1	–	Y	Y	N	–	-2	1	0	0	–	7	101	-94
44	$2^2 11^1$	–	N	N	Y	–	-7	1	2	-2	–	0	101	-101
45	$3^2 5^1$	–	N	N	Y	–	-7	1	2	-2	–	-7	101	-108
46	$2^1 23^1$	–	Y	N	N	–	5	1	0	-1	–	-2	106	-108
47	47^1	–	Y	Y	N	–	-2	1	0	0	–	-4	106	-110
48	$2^4 3^1$	–	N	N	Y	–	-11	1	6	-4	–	-15	106	-121

Table T.1: *Computations of* $\mathbf{g^{-1}(n)} \equiv (\omega + 1)^{-1}(\mathbf{n})$ *for small* $1 \leq \mathbf{n} \leq 48$.

The column labeled `Primes` *provides the prime factorization of each n so that the values of $\omega(n)$ and $\Omega(n)$ are easily extracted. The columns labeled, respectively,* `Sqfree`, `PPower` *and* $\bar{\mathbb{S}}$ *list inclusion of n in the sets of squarefree integers, prime powers, and the set $\bar{\mathbb{S}}$ that denotes the positive integers n which are neither squarefree nor prime powers. The next two columns provide the explicit values of the inverse function $g^{-1}(n)$ and indicate that the sign of this function at n is given by $\lambda(n) = (-1)^{\Omega(n)}$.*

The next two columns show the small-ish magnitude differences between the unsigned magnitude of $g^{-1}(n)$ and the summations $\hat{f}_1(n) := \sum_{k \geq 0} \binom{\omega(n)}{k} \cdot k!$ and $\hat{f}_2(n) := \sum_{k \geq 0} \binom{\omega(n)}{k} \cdot \#\{d|n : \omega(d) = k\}$. Finally, the last three columns show the summatory function of $g^{-1}(n)$, $G^{-1}(x) := \sum_{n \leq x} g^{-1}(n)$, deconvolved into its respective positive and negative components: $G_+^{-1}(x) := \sum_{n \leq x} g^{-1}(n) \left[g^{-1}(n) > 0\right]_\delta$ and $G_-^{-1}(x) := \sum_{n \leq x} g^{-1}(n) \left[g^{-1}(n) < 0\right]_\delta$.