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

Last Revised: Tuesday $12^{\rm th}$ May, 2020 – Compiled with LaTeX2e

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 $\frac{1}{x}\sum_{n\leq x} f(n) \sim h(x)$.
$o(f), O_{\alpha}(g), \Omega(h)$	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 $$\limsup_{x\to\infty} \left\lvert \frac{f(x)}{g(x)} \right\rvert > 0.$$ 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 $$C_k(n) := \begin{cases} \varepsilon(n), & \text{if } k = 0; \\ \sum\limits_{d\mid n} \omega(d) C_{k-1}(n/d), & \text{if } k \geq 1. \end{cases}$$
$[q^n]F(q)$	The coefficient of q^n in the power series expansion of $F(q)$ about zero.
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 $\lvert s\rvert < \sigma_a$, 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, $d(n) := \sum_{d\mid n} 1$.

Symbol	**Definition**
$\varepsilon(n)$	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, $f * g(n) := \sum\limits_{d\mid n} f(d)g(n/d)$, 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 $f^{-1}(n) = -\frac{1}{f(1)} \sum\limits_{\substack{d\mid n \\ d>1}} f(d)f^{-1}(n/d)$ 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 $G^{-1}(x) := \sum_{n \leq x} g^{-1}(n)$.
$\mathrm{Id}_k(n)$	The power-scaled identity function, $\mathrm{Id}_k(n) := n^k$ for $n \geq 1$.
$\mathbb{1}_{\mathbb{S}}$, $\chi_{\mathtt{cond}(x)}$	We use the notation $\mathbb{1}, \chi : \mathbb{N} \to \{0, 1\}$ to denote indicator, or characteristic functions. In paticular, $\mathbb{1}_{\mathbb{S}}(n) = 1$ if and only if $n \in \mathbb{S}$, and $\chi_{\mathtt{cond}}(n) = 1$ if and only if n satisfies the condition `cond`.
$\sum_{n \leq x}, \sum_{n \geq m}$	We use the notation $\sum_{n \leq x} f(n) \equiv \sum_{n=1}^{x} f(n) = f(1) + f(2) + \cdots + f(x)$. 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 $n \in [m, \infty)$: $\sum_{n \geq m} f(n) \equiv \sum_{n=m}^{\infty} f(n)$.

Symbol	**Definition**
$\log_*^m(x)$	The iterated logarithm function defined recursively for integers $m \geq 0$ and any $x > 0$, typically taken so that the function is non-negative, 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.
$[\mathtt{cond}]_\delta$	For a boolean-valued `cond`, $[\mathtt{cond}]_\delta$ 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 $\mu(n)$, $M(x) := \sum\limits_{n \leq x} \mu(n)$.
$\nu_p(n)$	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 $p^\alpha \vert\vert n$ (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 $\Omega(n) := \sum_{p^\alpha \vert\vert n} \alpha$. 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 $\omega(1) = \Omega(1) = 0$.
$\pi_k(x), \widehat{\pi}_k(x)$	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 avoid 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.

Symbol	**Definition**
$P(s)$	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, $\sigma_\alpha(n) := \sum\limits_{d\mid n} d^\alpha$, for any $n \geq 1$ and $\alpha \in \mathbb{C}$.
$\left[{n \atop k}\right]$	The unsigned Stirling numbers of the first kind, $\left[{n \atop k}\right] = (-1)^{n-k} \cdot s(n,k)$.
$\sim, \approx, \gtrsim, \lesssim, \gg, \ll$	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.
$\zeta(s)$	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 preconceptions 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 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 an arithmetic function f and some 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 that for a convergent infinite sum we have that

$$\sum_{n\geq 0} \int_n^{n+1} f(x)dx = f(0) + f(1) + f(2) + \cdots \leq \sum_{n\geq 1} f(n) \leq \sum_{n\geq 1} \int_n^{n+1} f(x)dx.$$

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*

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

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 $\overset{\blacktriangle}{\lesssim}$ similarly as $h(x) \overset{\blacktriangle}{\lesssim} r(x)$ if $h \ll r$ as $x \to \infty$. This usage of the notation of $\overset{\blacktriangle}{\gtrsim}, \overset{\blacktriangle}{\lesssim}$ intentionally breaks with the usual conventions 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 $x \to \infty$.

Hence, we emphasize that our new uses of these traditional symbols, $\overset{\blacktriangle}{\gtrsim}, \overset{\blacktriangle}{\lesssim}$, 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

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

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 $\mathrm{ph}(x) \in [\alpha, \beta]$ 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]

$$\begin{aligned} M(x) &= \sum_{n \leq x} \mu(n), \ x \geq 1, \\ &\longmapsto \{1, 0, -1, -1, -2, -1, -2, -2, -2, -1, -2, -2, -3, -2, -1, -1, -2, -2, -3, -3, -2, -1, -2, -2\} \end{aligned}$$

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 $x \rightarrow \infty$:

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

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.

2.0.1 Properties

The well-known approach to evaluating the behavior of $M(x)$ for large $x \rightarrow \infty$ 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

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

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]:

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

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

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

2.0.2 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

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

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 Odlyzyko 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 $M(x)/\sqrt{x} = \Omega_{\pm}(1)$.