New characterizations of the summatory function of the Möbius function

Maxie Dion Schmidt
Georgia Institute of Technology
School of Mathematics

Last Revised: Wednesday 6^{th} January, 2021 @ 01:39:02 – Compiled with LaTeX2e

Abstract

The Mertens function, $M(x) := \sum_{n \leq x} \mu(n)$, is defined as the summatory function of the Möbius function. We prove new characterizations of $M(x)$ using several modern techniques highlighted in the work of Montgomery and Vaughan. The new methods we draw upon connect formulas and recent Dirichlet generating function (or DGF) series expansions related to the canonically additive functions $\Omega(n)$ and $\omega(n)$. The connection between $M(x)$ and the distribution of these core additive functions we prove at the start of the article is an indispensible component to our proofs. The strong additivity of the component sequence primitives leads to regular properties of these subsequences in the new formula for $M(x)$ that include generalizations of Erdös-Kac like theorems satisfied by their distributions.

We characterize the distribution of $M(x)$ at large x by the summatory function $G^{-1}(x) := \sum_{n \leq x} \lambda(n)|g^{-1}(n)|$ where we define a shorthand for the Dirichlet inverse function terms as $g^{-1}(n) := (\omega + 1)^{-1}(n)$ for all $n \geq 1$. The unsigned summands that comprise $G^{-1}(x)$ are proved to have the noted limiting scaled normal tending distribution of values over $n \leq x$ as $x \to \infty$. The interplay of the summatory function $L(x)$ of the Liouville lambda function $\lambda(n)$ is key to interpreting the new exact and asymptotic formulas we prove for $M(x)$. This concrete link relating the limiting behavior of $M(x)$ directly to $L(x)$ makes explicit a typical heuristic that the functions $\mu(n)$ and $\lambda(n)$ are of the same order of magnitude a problem to sum and predict on the average.

Keywords and Phrases: *Möbius function; Mertens function; Dirichlet inverse; Liouville lambda function; prime omega function; prime counting function; Dirichlet generating function (DGF); Erdös-Kac theorem; strongly additive function.*
Math Subject Classifications (MSC 2010): *11N37; 11A25; 11N60; 11N64; and 11-04.*

1 Introduction

1.1 Definitions

We define the *Möbius function* to be the signed indicator function of the squarefree integers in the form of [18, A008683]

$$\mu(n) = \begin{cases} 1, & \text{if } n = 1; \\ (-1)^{\omega(n)}, & \text{if } \omega(n) = \Omega(n) \text{ and } n \geq 2; \\ 0, & \text{otherwise.} \end{cases}$$

The *Mertens function*, or summatory function of $\mu(n)$, is defined on the positive integers as

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

The sequence of slow growing oscillatory values of this summatory function begins as follows [18, A002321]:

$$\{M(x)\}_{x \geq 1} = \{1, 0, -1, -1, -2, -1, -2, -2, -2, -1, -2, -2, -3, -2, -1, -1, -2, -2, -3, -3, -2, -1, -2, \ldots\}.$$

The Mertens function satisfies that $\sum_{n \leq x} M\left(\left\lfloor \frac{x}{n} \right\rfloor\right) = 1$, and is related to the summatory function $L(x) := \sum_{n \leq x} \lambda(n)$ via the relation [5, 9]

$$L(x) = \sum_{d \leq \sqrt{x}} M\left(\left\lfloor \frac{x}{d^2} \right\rfloor\right), x \geq 1.$$

Clearly, a positive integer $n \geq 1$ is *squarefree*, or contains no divisors (other than one) which are squares, if and only if $\mu^2(n) = 1$. A related summatory function which counts the number of *squarefree* integers $n \leq x$ satisfies [4, §18.6] [18, A013928]

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

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

$$\begin{aligned} \mu_+(x) &:= \frac{\#\{1 \leq n \leq x : \mu(n) = +1\}}{x} \xrightarrow{x \to \infty} \frac{3}{\pi^2} \\ \mu_-(x) &:= \frac{\#\{1 \leq n \leq x : \mu(n) = -1\}}{x} \xrightarrow{x \to \infty} \frac{3}{\pi^2}. \end{aligned}$$

1.2 Properties

A conventional approach to evaluating the limiting asymptotic behavior of $M(x)$ for large $x \to \infty$ considers an inverse Mellin transformation of the reciprocal of the Riemann zeta function. In particular, since

$$\frac{1}{\zeta(s)} = \prod_p \left(1 - \frac{1}{p^s}\right) = s \cdot \int_1^\infty \frac{M(x)}{x^{s+1}} dx, \operatorname{Re}(s) > 1,$$

we obtain that

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

The previous two representations lead us to the exact expression of $M(x)$ for any real $x > 0$ given by the next theorem.

Theorem 1.1 (Analytic Formula for $M(x)$, Titchmarsh)**.** *Assuming the Riemann Hypothesis (RH), 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 real* $x > 0$

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

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

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

Under the assumption of the RH, Soundararajan proved new updated estimates bounding $M(x)$ from above for large x in the following forms [19]:

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

1.3 Conjectures on boundedness and limiting behavior

The RH is equivalent to showing that $M(x) = O\left(x^{\frac{1}{2}+\epsilon}\right)$ for any $0 < \epsilon < \frac{1}{2}$. There is a rich history to the original statement of the *Mertens conjecture* which asserts that

$$|M(x)| < C\cdot\sqrt{x},\quad \text{for some absolute constant } C > 0.$$

The conjecture was first verified by Mertens himself for $C = 1$ and all $x < 10000$ without the benefit of modern computation. Since its beginnings in 1897, the Mertens conjecture has been disproven by computational methods with non-trivial simple zeta function zeros with comparitively small imaginary parts in a famous paper by Odlyzko and té Riele [12]. More recent attempts at bounding $M(x)$ naturally consider determining the rates at which the function $M(x)/\sqrt{x}$ grows with or without bound along infinite subsequences, e.g., considering the asymptotics of the function in the limit supremum and limit infimum senses.

We cite that it is only known by computation that [15, *cf.* §4.1] [18, *cf.* A051400; A051401]

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

and

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

Based on work by Odlyzyko and té Riele, it seems probable that each of these limits should evaluate to $\pm\infty$, respectively [12, 7, 8, 6]. Extensive computational evidence has produced a conjecture due to Gonek that in fact the limiting behavior of $M(x)$ satisfies [11]

$$\limsup_{x\to\infty} \frac{|M(x)|}{\sqrt{x}\cdot(\log\log\log x)^{\frac{5}{4}}} = O(1).$$

2 A concrete new approach to characterizing $M(x)$

The main interpretation to take away from the article is that we have rigorously motivated an equivalent *alternate characterization* of $M(x)$ by constructing combinatorially relevant sequences related to the distribution of the primes and to standard strongly additive functions that have not yet been studied in the literature surrounding the Mertens function. This new perspective offers equivalent characterizations of $M(x)$ by formulas involving the summatory functions $G^{-1}(x) := \sum_{n\leq x} g^{-1}(n)$ and the prime counting function $\pi(x)$ given in Section 6.

The proofs of key properties of these new sequences bundles with it a scaled normal tending probability distribution for the unsigned magnitude of $|g^{-1}(n)|$ that is similar in many ways to the Erdös-Kac theorems for $\omega(n)$ and $\Omega(n)$. Moreover, since $\operatorname{sgn}(g^{-1}(n)) = \lambda(n)$, it follows that we have a new probabilistic perspective from which to express distributional features of the summatory functions $G^{-1}(x)$ as $x \to \infty$ in terms of the properties of $|g^{-1}(n)|$ and $L(x) := \sum_{n\leq x} \lambda(n)$. Note that formalizing the properties of the distribution of $L(x)$ is typically viewed as a problem on par with, or equally as difficult in order to understanding the distribution of $M(x)$ well as $x \to \infty$. The results in this article concretely connect the distributions of $L(x)$, a well defined scaled normally tending probability distribution, and $M(x)$ as $x \to \infty$.

The new sequence $g^{-1}(n)$ defined precisely below and $G^{-1}(x)$ are crucially tied to standard, canonical examples of strongly and completely additive functions, e.g., $\omega(n)$ and $\Omega(n)$, respectively. As such, it is not surprising that we are able to relate the distributions of these functions by limiting probabilistic normal distributions which are similar to the celebrated results given by the Erdös-Kac theorems for the prime omega function variants. Using the definition of $g^{-1}(n)$, we are able to reinterpret and reconcile exact formulas for $M(x)$ naturally by an easy-to-spot relationship to the distinct primes in the factorizations of $n \leq x$. The prime-related combinatorics at hand here are discussed in more detail by the remarks given in Section 4.3.

2.1 Summatory functions of Dirichlet convolutions of arithmetic functions

Theorem 2.1 (Summatory functions of Dirichlet convolutions)**.** *Let $f, h : \mathbb{Z}^+ \to \mathbb{C}$ 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 and h, respectively, and that $F^{-1}(x) := \sum_{n\leq x} f^{-1}(n)$ denotes the summatory function of the Dirichlet inverse of f for any $x \geq 1$. We have the following exact expressions for the summatory function of the convolution $f * h$ for all integers $x \geq 1$:*

$$\begin{aligned}
\pi_{f*h}(x) &:= \sum_{n\leq x}\sum_{d|n} f(d)h(n/d) \\
&= \sum_{d\leq x} f(d) H\left(\left\lfloor \frac{x}{d}\right\rfloor\right) \\
&= \sum_{k=1}^{x} H(k)\left[F\left(\left\lfloor \frac{x}{k}\right\rfloor\right) - F\left(\left\lfloor \frac{x}{k+1}\right\rfloor\right)\right].
\end{aligned}$$

Moreover, for all $x \geq 1$

$$\begin{aligned}
H(x) &= \sum_{j=1}^{x} \pi_{f*h}(j)\left[F^{-1}\left(\left\lfloor \frac{x}{j}\right\rfloor\right) - F^{-1}\left(\left\lfloor \frac{x}{j+1}\right\rfloor\right)\right] \\
&= \sum_{k=1}^{x} f^{-1}(k)\cdot \pi_{f*h}\left(\left\lfloor \frac{x}{k}\right\rfloor\right).
\end{aligned}$$

Corollary 2.2 (Convolutions arising from Möbius inversion)**.** *Suppose that h is an arithmetic function such that $h(1) \neq 0$. Define the summatory function of the convolution of h with μ by $\widetilde{H}(x) := \sum_{n\leq x}(h*\mu)(n)$. Then the Mertens function is expressed by the sum*

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

Corollary 2.3 (A motivating special case)**.** *We have that for all $x \geq 1$*

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

2.2 An exact expression for $M(x)$ in terms of strongly additive functions

Fix the notation for the Dirichlet invertible function $g(n) := \omega(n) + 1$ and define its inverse with respect to Dirichlet convolution by $g^{-1}(n) = (\omega+1)^{-1}(n)$. We can compute exactly that (see Table T.1 starting on page 35)

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

There is not a simple meaningful direct recursion between the distinct values of $g^{-1}(n)$, except through auxiliary function sequences whose properties we will discuss in detail. The distribution of distinct sets of prime exponents is still clearly quite regular since $\omega(n)$ and $\Omega(n)$ play a crucial role in the repitition of common values of $g^{-1}(n)$. The following observation is suggestive of the quasi-periodicity of the distribution of distinct values of this inverse function we notice below over $n \geq 2$:

Heuristic 2.4 (Symmetry in $g^{-1}(n)$ in the prime factorizations of $n \leq x$)**.** Suppose that $n_1, n_2 \geq 2$ are such that their factorizations into distinct primes are given by $n_1 = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ and $n_2 = q_1^{\beta_1} \cdots q_r^{\beta_r}$ for $= \omega(n_i) \geq 1$. If $\{\alpha_1, \ldots, \alpha_r\} \equiv \{\beta_1, \ldots, \beta_r\}$ as multisets of prime exponents, then $g^{-1}(n_1) = g^{-1}(n_2)$. For example, g^{-1} has the same values on the squarefree integers with exactly one, two, three, and so on prime factors.

Conjecture 2.5 (Characteristic properties of the inverse sequence)**.** *We have the following properties characterizing the Dirichlet inverse function $g^{-1}(n)$:*

(A) *For all $n \geq 1$, $\operatorname{sgn}(g^{-1}(n)) = \lambda(n)$;*

(B) *For all squarefree integers $n \geq 1$, we have that*

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

(C) *If $n \geq 2$ and $\Omega(n) = k$, then*

$$2 \leq |g^{-1}(n)| \leq \sum_{j=0}^{k} \binom{k}{j} \cdot j!.$$

We illustrate the conjecture clearly using the computation of initial values of this inverse sequence in Table T.1. The signedness property in (A) is proved precisely in Proposition 3.1. A proof of (B) in fact follows from Lemma 4.1 stated on page 13. The realization that the beautiful and remarkably simple combinatorial form of property (B) in Conjecture 2.5 holds for all squarefree $n \geq 1$ motivates our pursuit of simpler formulas for the inverse functions $g^{-1}(n)$ through sums of auxiliary subsequences of arithmetic functions denoted by $C_k(n)$ (see Section 4). That is, we observe a familiar formula for $g^{-1}(n)$ at many integers and then seek to extrapolate and prove there are regular tendencies of this sequence viewed more generally at any $n \geq 2$.

An exact expression for $g^{-1}(n)$ through a key semi-diagonal of these subsequences is given by

$$g^{-1}(n) = \lambda(n) \times \sum_{d|n} \mu^2\left(\frac{n}{d}\right) C_{\Omega(d)}(d), n \geq 1,$$

where the sequence $\lambda(n) C_{\Omega(n)}(n)$ has DGF $(P(s)+1)^{-1}$ for $\operatorname{Re}(s) > 1$ (see Proposition 3.1). In Corollary 5.5, we prove that the approximate average order mean of the unsigned sequence satisfies

$$\mathbb{E}|g^{-1}(n)| \asymp (\log n)^2 \sqrt{\log\log n}, \text{ as } n \to \infty.$$

In Section 5, we also prove the next variant of an Erdös-Kac theorem like analog for a component sequence $C_{\Omega(n)}(n)$. This leads us to conclude the following statement for $\mu_x(C) := \log\log x + \hat{a} - \frac{1}{2}\log\log\log x$, $\sigma_x(C) := \sqrt{\mu_x(C)}$, $\hat{a}$ an absolute constant, and any $y \in \mathbb{R}$ (see Corollary 5.7):

$$\frac{1}{x} \cdot \#\{2 \leq n \leq x : |g^{-1}(n)| - \mathbb{E}|g^{-1}(n)| \leq y\} = \Phi\left(\frac{\frac{\pi^2}{6}y - \mu_x(C)}{\sigma_x(C)}\right) + O\left(\frac{1}{\sqrt{\log\log x}}\right), \text{ as } x \to \infty.$$

Thus, the regularity and quasi-periodicity we have alluded to in the remarks above are actually quantifiable in so much as $|g^{-1}(n)|$ for $n \leq x$ tends to its average order with a non-central normal tendency depending on x as $x \to \infty$. That is, if x is sufficiently large and we pick any integer $n \in [2, x]$ uniformly at random, then each of the following statements holds:

$$\mathbb{P}\left(|g^{-1}(n)| - \frac{6}{\pi^2}\mathbb{E}|g^{-1}(n)| \leq 0\right) = o(1) \tag{A}$$

$$\mathbb{P}\left(|g^{-1}(n)| - \frac{6}{\pi^2}\mathbb{E}|g^{-1}(n)| \leq \frac{6}{\pi^2}\mu_x(C)\right) = \frac{1}{2} + o(1). \tag{B}$$

Moreover, for any positive real $\delta > 0$ we have that

$$\mathbb{P}\left(|g^{-1}(n)| - \frac{6}{\pi^2}\mathbb{E}|g^{-1}(n)| \leq \frac{6}{\pi^2}\mu_x(C)^{1+\delta}\right) = 1 + o_\delta(1), \text{ as } x \to \infty. \tag{C}$$

A consequence of (A) and (C) in the probability estimates above is that for any fixed $\delta > 0$ and $n \in \mathcal{S}_1(\delta)$ taken within a set of asymptotic density one we have that

$$\frac{6}{\pi^2}\mathbb{E}|g^{-1}(n)| \leq |g^{-1}(n)| \leq \frac{6}{\pi^2}\mathbb{E}|g^{-1}(n)| + \frac{6}{\pi^2}\mu_x(C)^{\frac{1}{2}+\delta}.$$

Hence when we integrate over a sufficiently spaced set of (e.g., set of wide enough) disjoint consecutive intervals containing large enough integer values, we can assume that an asymptotic lower bound on the contribution of $|g^{-1}(n)|$ is given by the function's average order, and an upper bound is given by the related upper limit above for any fixed $\delta > 0$. In particular, observe that by Corollary 5.7 and Corollary 5.5 we can see that

$$\frac{\pi^2}{6 \cdot \sigma_x(C)} \times \int_{-\infty}^{\infty} z \cdot \Phi'\left(\frac{\frac{\pi^2}{6}z - \mu_x(C)}{\sigma_x(C)}\right) dz = \frac{6}{\pi^2} \cdot \sigma_x(C) = o\left(\mathbb{E}|g^{-1}(x)|\right).$$

Remark 2.6 (Uniform asymptotics from certain bivariate counting DGFs)**.** We emphasize the recency of the method demonstrated by Montgomery and Vaughan in constructing their original proof of Theorem 3.5 (stated below). To the best of our knowledge, this textbook reference is one of the first clear-cut applications documenting something of a hybrid DGF-and-OGF type approach to enumerating sequences of arithmetic functions and their summatory functions. This interpretation of certain bivariate DGFs offers a window into the best of both generating function series worlds. It combines the additivity implicit to the coefficients indexed by a formal power series variable formed by multiplication of these structures, while coordinating the distinct DGF-best property of the multiplicativity of distinct prime powers invoked by taking powers of a reciprocal Euler like product. Another set of proofs constructed based on this type of hybrid power series DGF is utilized in Section 5 when we prove the Erdös-Kac theorem like analog that holds for the component sequence $C_{\Omega(n)}(n)$, which is crucially related to $|g^{-1}(n)|$ by the results in Section 4.

2.2.1 Formulas illustrating the new characterizations of $M(x)$

Let $G^{-1}(x) := \sum_{n \leq x} g^{-1}(n)$ for integers $x \geq 1$. We prove that (see Proposition 6.3)

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

$$= G^{-1}(x) + \sum_{p \leq x} G\left(\left\lfloor \frac{x}{p} \right\rfloor\right), x \geq 1.$$

This formula implies that we can establish new *lower bounds* on $M(x)$ along large infinite subsequences by bounding appropriate estimates of the summatory function $G^{-1}(x)$. The take on the regularity of $|g^{-1}(n)|$ is imperative to our argument formally bounding the growth of $M(x)$ through its new characterizations by $G^{-1}(x)$. A more combinatorial approach to summing $G^{-1}(x)$ for large x based on the distribution of the primes is outlined in our remarks in Section 4.3.

In the proofs given in Section 6, we begin to use these new equivalent characterizations to relate the distributions of $|g^{-1}(n)|$, $G^{-1}(x)$, $\lambda(n)$ and its often classically studied summatory function $L(x)$, to $M(x)$ as $x \to \infty$. In particular, Proposition 6.1 proves that like the known bound for $M(x)$, we have that $G^{-1}(x) = o(x)$ as $x \to \infty$. The results in Corollary 6.2 prove that for almost every sufficiently large x

$$G^{-1}(x) = O\left(\max_{1 \leq t \leq x} |L(t)| \cdot \mathbb{E}|g^{-1}(x)|\right).$$

Moreover, if the RH is true, then we have the following result for any $\varepsilon > 0$ and almost every integer $x \geq 1$:

$$G^{-1}(x) = O\left(\frac{\sqrt{x} \cdot (\log x)^{\frac{5}{2}}}{(\log\log x)^{2+\varepsilon}} \times \exp\left(\sqrt{\log x} \cdot (\log\log x)^{\frac{5}{2}+\varepsilon}\right)\right).$$

By applying Corollary 6.5, we have that as $x \to \infty$

$$M(x) = O\left(G^{-1}(x) + G^{-1}\left(\frac{x}{2}\right) + \frac{\sqrt{x} \cdot G^{-1}(\sqrt{x})}{\log x} \pm (\log\log x) \times \max_{\sqrt{x} < k < \frac{x}{2}} \frac{|G^{-1}(k)|}{k}\right).$$

Moving forward, a discussion of the properties of the summatory functions $G^{-1}(x)$ motivates more study in the future to exhaust the full range of possibilities for this new method.

7 Conclusions

We have identified and precisely defined a key sequence, $\{g^{-1}(n)\}_{n\geq 1}$, which corresponds to the Dirichlet inverse of the additive function derivative, $g := \omega + 1$. In general, we find that the Dirichlet inverse of any arithmetic function f such that $f(1) \neq 0$ is expressed at each $n \geq 2$ as a signed sum of m-fold convolutions of f with itself for $1 \leq m \leq \Omega(n)$. The strong additivity of $\omega(n)$ and its known limiting distribution stated via the Erdös-Kac theorem then provides a predictable foundation on which $|g^{-1}(n)|$ is evaluated on average for $n \leq x$ as $x \to \infty$.

As we discussed in the remarks in Section 4.3, it happens that there is a natural combinatorial interpretation to the distribution of distinct values of $|g^{-1}(n)|$ for $n \leq x$ involving the primes $p \leq x$ at large x. In particular, the magnitude of $|g^{-1}(n)|$ depends only on the pattern of the exponents of the prime factorization of n in so much as $|g^{-1}(n_1)| = |g^{-1}(n_2)|$ whenever $\omega(n_1) = \omega(n_2)$, $\Omega(n_1) = \Omega(n_2)$, and where the is a one-to-one correspondence $\nu_{p_1}(n_1) = \nu_{p_2}(n_2)$ between the distinct primes $p_1|n_1$ and $p_2|n_2$. The signedness of $g^{-1}(n)$ is given by $\lambda(n)$ for all $n \geq 1$. This leads to a familiar unpredictability and dependence of the summatory functions $G^{-1}(x)$ on the distribution of the function $L(x)$, the summatory function of the Liouville lambda function.

What we prove in the last results in Section 6 provides an equivalent characterization of the limiting properties of $M(x)$ by exact formulas and asymptotic relations involving $G^{-1}(x)$ and $L(x)$. We emphasize that our new work on the Mertens function proved within this article is significant in providing a new lense through which we can approach bounding $M(x)$, rather than in proving explicit new best known bounds on the classical function at this point. The computational data generated in Table T.1 suggests numerically, especially when compared to the initial values of $M(x)$, that the distribution of $|G^{-1}(x)|$ may be easier to work with that those of $|M(x)|$ or $|L(x)|$. The remarks given in Section 4.3 about the direct relation of the distinct (and repitition of) values of $|g^{-1}(n)|$ for $n \leq x$ to the distribution of the primes and their distinct powers are also suggestive that bounding a main term for $G^{-1}(x)$ should be fairly regular along some infinitely tending subsequences of the integers.

One topic that we do not touch on in the article is to consider the limiting correlation between $\lambda(n)$ and the unsigned sequence of $|g^{-1}(n)|$ whose limiting distribution we have identified and rigorously proved. An analysis of the new rephrasing of exact formulas for $M(x)$ we give through $G^{-1}(x)$ and $L(x)$ that is probabilistic in grounding may be fruitful towards proving better than currently known bounds on $M(x)$ moving forward. That is, much in the same way that variants of the Erdös-Kac theorem are proved by defining the random variables related to $\omega(n)$, we suggest an analysis of the summatory function $G^{-1}(x)$ by scaling the provably nicely distributed $|g^{-1}(n)|$ for $n \leq x$ as $x \to \infty$ by its signed weight of $\lambda(n)$ using an initial heuristic along these lines.

An experiment illustrated in the online supplementary computational reference [17] suggests that for many, if not most sufficiently large x, we may consider replacing the summatory function with terms weighted by $\lambda(n)$

$$G^{-1}(x) := \sum_{n\leq x} \lambda(n)|g^{-1}(n)|, x \geq 1,$$

by alternate sums that average these sequences differently while still preserving the original order of $|G^{-1}(x)|$ up to a bounded constant. For example, each of the following three summatory functions offers a unique interpretation of an average of sorts that "mixes" the values of $\lambda(n)$ with the unsigned sequence $|g^{-1}(n)|$ over $1 \leq n \leq x$:

$$\begin{aligned}
G_{*}^{-1}(x) &:= \sum_{n\leq x} \frac{1}{2\gamma - 1 + \log n} \times \sum_{d|n} \lambda\left(\frac{n}{d}\right) |g^{-1}(d)| \\
G_{**}^{-1}(x) &:= \sum_{n\leq x} \frac{1}{2\gamma - 1 + \log n} \times \sum_{d|n} \lambda\left(\frac{n}{d}\right) g^{-1}(d) \\
G_{***}^{-1}(x) &:= \sum_{n\leq x} \frac{1}{2\gamma - 1 + \log n} \times \sum_{d|n} g^{-1}(d).
\end{aligned}$$

Then based on preliminary numerical results, a large proportion of the $y \leq x$ for large x satisfy

$$\left|\frac{G_{*}^{-1}(y)}{G^{-1}(y)}\right|, \left|\frac{G_{**}^{-1}(y)}{G^{-1}(y)}\right|, \left|\frac{G_{***}^{-1}(y)}{G^{-1}(y)}\right| \in (0, 3].$$

Variants of this type of summatory function identity exchange are suggested for future work on these topics.

Acknowledgements

We thank and give much appreciation to the following professors that offered discussion, feedback and correspondence while the article was written: Paul Pollack, Steven Miller, Ernie Croot, Bruce Reznick and Michael Lacey. The work on the article was supported in part by relief funding made available within the School of Mathematics at the Georgia Institute of Technology in 2020. Without this combined support, the article would not have been possible.