Combinatorial methods for approximating the Mertens function and sums of the Möbius function

Maxie D. Schmidt

School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332

maxieds@gmail.com
mschmidt34@gatech.edu

Last Revised: September 19, 2019

Abstract

The Mertens function, $M(x) = \sum_{n \leq x} \mu(n)$, is classically defined to be the summatory function of the Möbius function $\mu(n)$. In some sense, the Móbius function can be viewed as a signed indicator function of the squarefree integers which have asymptotic density of $6/\pi^2 \approx 0.607927$ and a corresponding well-known asymptotic average order formula. The signed terms in the sums in the definition of the Mertens function introduce complications in the form of semi-randomness and cancellation inherent to the distribution of the Möbius function over the natural numbers. The Mertens conjecture which states that $|M(x)| < C \cdot \sqrt{x}$ for all $x \geq 1$ has a well-known disproof due to Odlyzko et. al. It is widely believed that $M(x)/\sqrt{x}$ is an unbounded function which changes sign infinitely often and exhibits a negative bias over all natural numbers $x \geq 1$. We focus on obtaining new bounds for $M(x)$ by methods that generalize to handle other related cases of special number theoretic summatory functions.

Keywords. *Mertens function; Möbius function; summatory function; Ramanujan's sum.*
MSC (2010). *11A25; 11N37; 11N56.* .

1 Introduction

1.1 Summatory functions of the Móbius function

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 given by

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

There are many known variants and special properties of the Möebius function and its generalizations [21, *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

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

A related function which counts the number of *squarefree* integers than x sums the average order of the Möbius function as

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

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 − even over asymptotically large and variable intervals. As we consider next, these local oscillations and irregularity in growth lead to many natural unsolved questions about the eventual boundedness (or lack thereof) along subsequences of the natural numbers.

We define the notion of a *generalized, or weighted, Mertens summatory function* for fixed $\alpha \in \mathbb{C}$ as

$$M_\alpha^*(x) = \sum_{n \leq x} n^\alpha \mu(n),\ x \geq 1, \tag{1}$$

where the special case of $M_0^*(x)$ corresponds to the definition of the classical Mertens function $M(x)$ given above. The plots shown in Figure 1.1 illustrate the chaotic behavior of the growth of these functions for x in small intervals when $\alpha \in \{-1, 0, 1, 2\}$. Related questions are often posed in relation to the strikingly similar properties of the summatory functions over the Liouville lambda function, $L_\alpha(x) := \sum_{n \leq x} \lambda(n) n^{-\alpha}$ [?, ?].

1.2 Properties and bounds on $M(x)$

1.2.1 Exact formulae

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

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

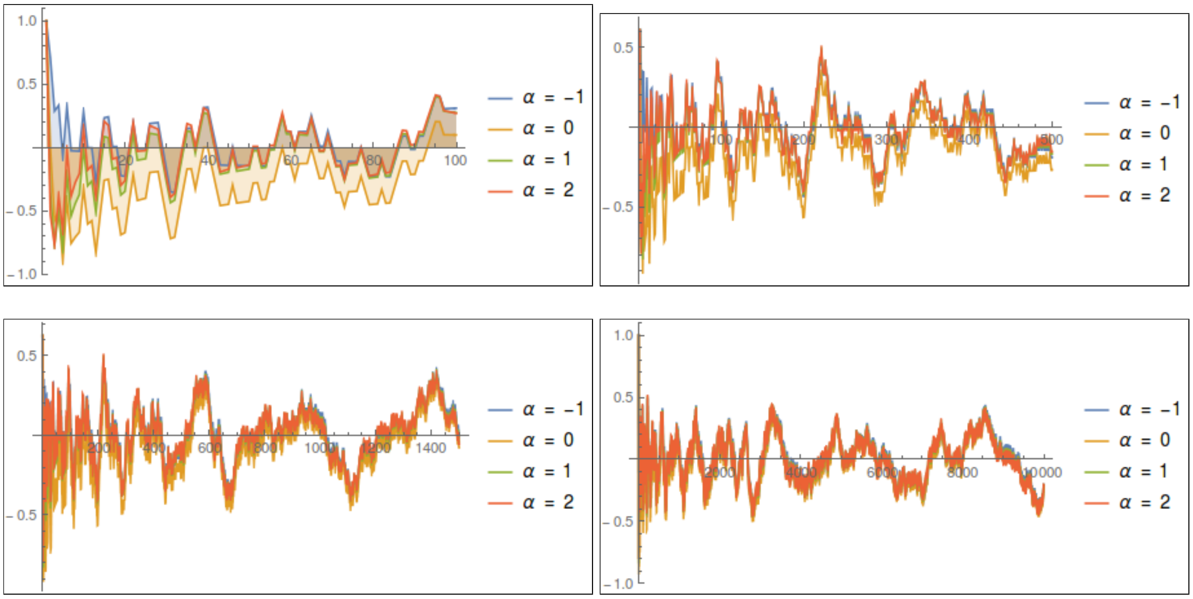

Figure 1.1: Comparison of the Mertens Summatory Functions $M_\alpha(x)/x^{\frac{1}{2}+\alpha}$ for Small x and α

we then obtain that

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

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 1 (Analytic Formula for $M(x)$). If the RH is true, the 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 $x \in \mathbb{R}_{>0}$

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

1.2.2 Explicit bounds for large x

An unconditional bound on the Mertens function due to Walfisz [?] states that there is an absolute constant $C > 0$ such that

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

Under the assumption of the RH, Soundarajan proved new updated estimates bounding $M(x)$ for large x in 2007 of the following forms:

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

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

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

1.3 Open problems

The Riemann Hypothesis is equivalent to showing that $M(x) = O\left(x^{1/2+\varepsilon}\right)$ for any $0 < \varepsilon < \frac{1}{2}$. For $\operatorname{Re}(\alpha) < 1$, we know the limiting absolute behavior of these functions as $x \to \infty$ as the Dirichlet generating function

$$\frac{1}{\zeta(\alpha)} = \lim_{x \longrightarrow \infty} M_{-\alpha}^{*}(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^{\alpha}},$$

which is definitively bounded for all large x. It is still unresolved whether

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

although computational evidence suggests that this is a likely conjecture [9, 7]. 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.

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 in so much as $M(x) < 0$ more frequently than $M(x) > 0$ over all $x \in \mathbb{N}$[1] 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. A probabilistic proof along these lines has been given by Ng in 2008, though to date an exact rigorous proof (rather than somewhat heuristic argument) that $M(x)/\sqrt{x}$ is unbounded still remains elusive. We cite that prior to this point it is known that [16, *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 [14, 10, 9, 7]. It is also known that $M(x) = \Omega_{\pm}(\sqrt{x})$ and $M(x)/\sqrt{x} = \Omega_{\pm}(1)$.

1.4 A summary of the constructions of our new exact formulas for $M(x)$

The construction and disproof due to Odlyzko et. al. relies on relatively low-laying zeros of the Riemann zeta function to establish lower bounds on the limit supremum of $M(x)/\sqrt{x}$. Since this ground-breaking work which showed a new approach to breaking classical bounds known for the function, however, much of the recent work in this direction has been experimental or computational in nature. The work started in this article began as a happy accident working with divisor sums involving the Möbius function that were eventually realized to be related to this classical function by a database search for the resulting sequences. The background research for where these "special" or at least noteworthy started sums have been found is available in the references. It is a central point of mine that while this seems to be conjured from outer space, the more "combinatorial", or rather

[1] See for example the discussion in the following thread: https://mathoverflow.net/questions/98174/is-mertens-function-negatively-biased.

less analytic in flavor than traditional approaches, offers a wide range of new techniques to explore in the context of the Mertens function and some of its most famous unsolved problems.

A key component to writing down the new identities and formulas for $M(x)$ based on this new "novel" starting point is in identifying inversion relations between triangular sequences which naturally fit into a matrix construction[2]. The problem of exactly identifying the inverse sequence – even given a solid and well founded orthogonal sequence for it – is still challenging at best. In some senses due to these limitations elementary techniques and transforms such as Abel and partial summation are the name of the game to extracting new properties about $M(x)$. I have not by any means familiarized myself enough with the rigor of random matrix theory and its connections to zeta and L-functions to advance these results towards that end. Instead, the more self, and externally imposed description, of the basis for the new results in this thesis article are roughly speaking combinatorial in their footing – at least from the typical perspective that Mellin transforms and zero sums are *the* (only) vehicle to explore new interesting mathematics in number theory.

With that preface in mind, there are clearly, very clearly, understated relations between the objects defined in the next sections and deep number theoretic underpinnings. My goal, and philosophy here (so far) is to remain exact and precise with the formulas until it is absolutely – and unequivocally necessary – to truncate in the name of issuing theorems on asymptotic properties which are relevant. I hope that the new spirit of these formulations and the interesction of combinatorial number theory in classically analytic problems is of absorbable utility to others who are still not too old to do what they used to like to think about. Thank you to those with an open mind and notable under-pessimism on moving forward – whatever that means.

2 Introducing a more common language: Problems we are interested in approximating

The Mertens function case is the motivation for exploring the setup of a more general construction for relating summatory functions of special sequences via natural recurrence relations which arise for these functions. As it turns out, all of the machinery we will need to formulate new bounds for the classical Mertens function case can really just be re-worked and re-stated up front to provide theorems for analogous new bounds on other summatory functions. We might as well sketch out this more general scenario before diving head first into the initially very, very interesting classical special case. Then, later, once we have rigorously divined proofs of new bounds for this exceptionally interesting case for $M(x)$, we will have a loose path backwards which will allow us to cover the existence of related bounds for other summatory functions which we can show show up in a variety of contexts.

This is all just wordiness to say that upon some reflection, the mechanics that seem to work well in driving our solution (so to speak) to the classical Mertens function problems at hand, are really just specializations of a more general theorem framework which we will strive to prove here. Our loose requirements on the form of a general summatory function is that it satisfies a special (but natural, as we shall see) recurrence relation for $x \geq x_0$, and that the functions forming the coefficients and non-homogeneous offsets of this summatory function have "reasonable-to-aproximate" bounds dominating their asymptotic growth for large $x \gg 1$.

2.1 Introducing several key problems of interest to our study

Before we dive into definitions that formulate the necessary complication that this more general framework brings with it, let's be precise in stating the nature of the results about a general summatory function, denoted $M^*(x)$, which we are interested in investigating. Here, the assumption is that $M^*(x)$ inherits some signed oscillatory nature from its constituent summands – just as $M(x)$ classically varies according to localized behaviors of the Möbius function. Thus our problem is substantially more interesting (and difficult) than only estimating and best-error-bounding a monotone non-decreasing sum of arithmetic functions.

Problem of Interest 2.1 (Proving Scaled Unboundedness). We let the smooth positive function $\delta(x)$ take the place of $\sqrt{x}$ in the classical Mertens function problem statements. The question at hand is whether for a fixed, appropriately (and application dependent) chosen $\delta(x)$, we have unboundedness towards both $\pm\infty$. That

[2] Depending on background and context, triangles of inversion transformation coefficients can be (in a broad sense) viewed as generalized Möbius functions. In [17, §2], this observation is explicitly motivated, as are the corresponding objects of generalized Mertens functions which form summatory functions over these triangles.

is to say, is it true that each of

$$\limsup_{x\to\infty} \frac{M^*(x)}{\delta(x)} = +\infty,$$
$$\liminf_{x\to\infty} \frac{M^*(x)}{\delta(x)} = -\infty,$$

hold in the limiting case? If these two questions assert a "YES" given some optimally chosen function $\delta(x)$, we next should ask along which subsequences of the positive integers is this unboundedness achieved. That is to say, given any real $M > 0$, can we identify (natural) infinite subsequences $(x_n(M))_{n\geq 1}, (y_m(M))_{m\geq 1}$ such that

$$\frac{M^*(x_n(M))}{\delta(x_n(M))} > M, \forall n \geq 1,$$
$$\frac{M^*(y_m(M))}{\delta(y_m(M))} < -M, \forall m \geq 1?$$

Is there a limiting measure for the form of these subsequences, i.e., is there a limiting distribution ν on $\mathbb{R}$ for either of $M^*(x_n(M))/\delta(x_n(M))$ or $M^*(y_m(M))/\delta(y_m(M))$ (*cf.* Ng's probabilistic approach to $M(x)$)?

Problem of Interest 2.2 (Proving Best Possible Limit Supremum and Infimum Growth Rates). This is essentially the generalized form of verifying Gonek's conjecture for $M(x)$. That is to say, we seek positive smooth functions $\delta(x)$ and $\xi(x)$, and some absolute constants $C_1, C_2; D_\ell, D_u > 0$, such that

$$\limsup_{x\to\infty} \frac{M^*(x)}{\delta(x) \times \xi(x)^{C_1}} = D_u,$$
$$\liminf_{x\to\infty} \frac{M^*(x)}{\delta(x) \times \xi(x)^{C_2}} = D_\ell.$$

Is it true that we should expect that $C_1 = C_2$, or otherwise any symmetry between the limiting constants D_ℓ and D_u?

Problem of Interest 2.3 (Expected Time to Exceed a Preset Bound).

Problem of Interest 2.4 (Frequency of Visitation to a Preset Bound).

Problem of Interest 2.5 (Distribution of Zeros of the Summatory Function).

2.2 Definitions for the general problem statement

2.2.1 Requirements One.

The procedure here begins by selecting a number theoretically relevant function, $\omega^* : \mathbb{Z}^+ \to \mathbb{C}$, such that its summatory function satisfies "nice" (separable) bounds of the form

$$S_{\omega^*}(x) := \sum_{n\leq x} \omega^*(n) = \alpha_{\omega^*}(x) + O\left(\beta_{\omega^*}(x)\right).$$

We define the separability, or distinguishability, of the main term from the error bound in the previous equation according to the condition that

$$\mathrm{TODO}_{1,\ell}(x) \ll \left|\frac{\alpha_{\omega^*}(x)}{\beta_{\omega^*}(x)}\right| \ll \mathrm{TODO}_{1,u}(x).$$

For example, depending on how our results turn out, it may be desirable to require the upper and lower bounds on the main-term-to-error ratio above to take the form $(\log x)^{1+\varepsilon}, (\log\log x)^{1/2+\varepsilon}, Cx^{\varepsilon}$, etc., though we are not yet sure what these optimal constraints should be set to just yet. It may be later prudent to be able to assure matrix invertibility by enforcing the condition that $\omega^*(n) \neq 0$ for all $n \geq x_0$.

2.2.2 Requirements Two.

Suppose that we have a recurrence relation (or functional-ish equation) for $M^*(x)$ of the form

$$M^*(x) = \sum_{d=1}^{x} \omega^*(d) M^*\left(\left\lfloor \frac{x}{j} \right\rfloor\right) + \pi^*(x), \forall x \geq x_0. \tag{2}$$

Here, we prescribe some non-rigorously constraining bounds on the growth of the non-homogeneous function term, $\pi^*(x)$, of one of the following forms:

(A) We require that $\pi^* : \mathbb{Z}^+ \to \mathbb{C}$ is *monotone* (of application specific type). Futhermore, we have "nice" enough formulas describing the change of this arithmetic function as:

$$\Delta[\pi^*](k) = \mathrm{TODO}_2(k).$$

(B) We require explicit known (rough, or semi-precise) upper and lower bounds governing the growth of this function for sufficiently large $x \gg 1$:

$$\gamma(x) - \rho_\ell(x) \leq \pi^*(x) \leq \gamma(x) + \rho_u(x), \forall x \geq \mathcal{X}_0.$$

Then we have an exact formula for the summatory function $M^*(x)$ which holds for all $x \geq x_0$ prescribed according to the next theorem. The bounds imposed on all of the functions used to define this problem constriction so far obviously impact the precise growth rates of the summatory functions, $M^*(x)$.

Theorem 2 (Exact Summatory Function Formulas). Suppose that the summatory function $M^*(x)$ satisfies (2) for all $x \geq x_0$. Then we have an exact formula of the form

$$M^*(x) = \sum_{k=1}^{x} t_{x,j}(M) \left[\pi^*(k) + \Delta_\pi(k)\right],$$

where the lower triangular coefficients $t_{x,j}(M)$ are defined as the entries of the inverse of the matrix

$$(t_{i,j})_{1 \leq i,j \leq x} = \left(S_{\omega^*}\left(\left\lfloor \frac{2i+1}{j} \right\rfloor\right) - S_{\omega^*}\left(\left\lfloor \frac{2i+1}{j+1} \right\rfloor\right) + \Delta_\omega(i,j)\right)_{1 \leq i,j \leq x}^{-1}, \tag{3}$$

and where the functions $\Delta_\pi(k)$ and $\Delta_\omega(i,j)$ are adjustment factors that are non-zero only when $x_0 > 1$. In particular, we have precisely that

$$\Delta_\pi(k) = TODO\left[1 \leq k \leq x_0\right]_\delta,$$
$$\Delta_\omega(i,j) = TODO\left[1 \leq i,j \leq x_0\right]_\delta.$$

Remarks on the Proof. The proof we give in the special case of the Mertens function, $M(x)$, is constructive. We will wait until after we build on that result in Section 3 to complete the proof for the cases under this abstraction of notation. Notice that while this formula is certainly exact, the matrix entries, $t_{x,j}(M)$, that dictate the growth of the formula for $M^*(x)$ are generally signed, oscillatory, and more or less unpredictable (as we ought expect them to be). Thus, while this result is a good starting point for formulating new bounds on $M^*(x)$, in practice we should expect there to be truncations and approximation that enters the mix as x grows large – or approaches certain clearly defined subintervals of the positive integers. In other words, the limiting approach starts with this theorem, proceeds by applying a non-trivial lemma about the growth properties satisfied by $t_{x,j}(M)$ where we need key shrinkage or growth, followed by plug-and-chug with the resulting asymptotic estimates. $\square$

Remark 2.6 (Simplifying Assumptions Moving Forward). We would like to apply this construction for "nice-enough" number theoretic functions $\omega^*(n)$ such that its corresponding summatory function has a reasonable enough asymptotic formula (with bounded error term) for us to work with. In these cases, the ordinary (non-inverse) matrices defined on the right-hand-side of (3) ought have manageable formulas. In particular, we notice that since the ordinary matrices are lower triangular with ones on the diagonal, for each fixed column index $j \geq 1$, the sequence of shifted coefficients, $\{t_{x+j-1,j}^{-1}\}_{x \geq 1}$, has a Dirichlet inverse.

Now if we can obtain a plausible formula for these sequences of Dirichlet inverse functions, $h_j(x)$ satisfying

$$\sum_{d|x} t_{x+j-1,j}^{-1} \cdot h_j(x/d) = \delta_{x,1},$$

then 1) since we can relate these sequences to the matrix inverse sequences, $t_{x,j}$; and 2) if we can assemble workable Dirichlet series over the component functions – we are able to apply an integral formula to extract the coefficients of the reciprocal Dirichlet series. That is to say that, if $D_A(s) := \sum_{n\geq 1} a(n)/n^s$ is absolutely convergent whenever $\mathrm{Re}(s) > \sigma_A$, then we can recover the Dirichlet inverse, $a^{-1}(n)$, function values by integrating:

$$a^{-1}(x) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \frac{x^{\sigma+\imath t}}{D_A(\sigma+\imath t)} dt, \forall x \in \mathbb{Z}^+; \sigma > \sigma_A.$$

Then we relate these sequences to the matrix inverse terms, $t_{x+j-1,j}$, and proceed to apply Theorem 2 to bound $M^*(x)$ according to the methodology in one of our key problems of interest stated in Section 2.1. There is of course more subtlety to the execution of this procedure, but this plan of attack should suffice to illustrate the methods forward to new readers at this point.

2.3 Other motivating, non-classical applications

Note that Section 5 goes into significantly more background and preparation to generalize the procedure for the classical Mertens function $M(x)$ to Mertens (and Möbius) functions which are constructed in a slightly different manner. In addition to this solid application, there are a couple of more instances which are worth enumerating where our general approach and the construction offered by Theorem 2 arise in useful applications. Let's enumerate some attention one example case by case below.

2.3.1 Example I: Sums over Dirichlet convolution divisor sums

Let the summatory function $G(x) := \sum_{n\leq x} g(n)$. We then consider the summatory functions of the Dirichlet convolution between two arithmetic functions, f, g: $(f * g)(n) = \sum_{d|n} f(d)g(n/d)$. Suppose that this summatory function is given by $\hat{\pi}(x)$. Then we see that

$$\begin{aligned} \hat{\pi}(x) &= \sum_{n\leq x} \sum_{d|n} f(d)g(n/d) \\ &= \sum_{d=1}^{x} f(d) G\left(\left\lfloor \frac{x}{d} \right\rfloor\right). \end{aligned}$$

Note that this is actually a special case of the Mertens function recurrences which lead to the form of Theorem 2. When the left-hand-side of the previous equations is not easily expressed in terms of the summatory function $G(x)$, we have the following related theorem which allows us to exactly express the formula for $G(x)$ in terms of the forms of f and $\hat{\pi}$.

Theorem 3. Let the arithmetic functions f, g be given with respective summatory functions $F(x)$ and $G(x)$ where $f(1) \neq 0$. Suppose that $\hat{\pi}(x)$ is a non-identically zero function defined on the positive integers. Furthermore, suppose that we have a recurrence relation for G of the form

$$\hat{\pi}(x) = \sum_{d=1}^{x} f(d) G\left(\left\lfloor \frac{x}{d} \right\rfloor\right), \forall x \geq 1.$$

Then we have the exact expressions

$$G(x) = \sum_{k=1}^{x} t_{x,k}(f) \hat{\pi}(k), \forall x \geq 1,$$

where the lower triangular sequence $t_{x,j}(f)$ depends only on f (and F) as the inverse matrix entries of

$$(t_{i,j}(f))_{1\leq i,j\leq x} = \left(F\left(\left\lfloor \frac{i}{j} \right\rfloor\right) - F\left(\left\lfloor \frac{i}{j+1} \right\rfloor\right)\right)_{i\leq i,j\leq x}^{-1}.$$

Since $f(1) \neq 0$, it's Dirichlet inverse, $f^{-1}(n)$, exists and provides us with an expansion of the inverses of $(t_{i,j}(f))$ of the form

$$t_{x,j}^{-1}(f) = \sum_{k=\left\lfloor \frac{x}{j+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{j} \right\rfloor} f^{-1}(k) = F^{-1}\left(\left\lfloor \frac{x}{j} \right\rfloor\right) - F^{-1}\left(\left\lfloor \frac{x}{j+1} \right\rfloor\right),$$

if $F^{-1}(x)$ denotes the summatory function of f^{-1}.

Given the interpretation of the summatory functions over an arbitrary Dirichlet convolution (and the vast number of such identities for special number theoretic functions – *cf.* [?]), it is not surprising that this formulation of the first theorem may well provide many fruitful applications, indeed! In addition to those cited in the compendia of the catalog reference, we have notable identities of the form: $(f*1)(n) = [q^n]\sum_{m\geq 1} f(m)q^m/(1-q^m)$, $\sigma_k = \mathrm{Id}_k * 1$, $\mathrm{Id}_1 = \phi * \sigma_0$, $\chi_{\mathrm{sq}} = \lambda * 1$ (see sections below), $\mathrm{Id}_k = J_k * 1$, $\log = \Lambda * 1$, and of course $2^\omega = \mu^2 * 1$.

The result in Theorem 3 is natural and displays a quite beautiful form of symmetry between the initial matrix terms,

$$t_{x,j}(f) = \sum_{k=\left\lfloor \frac{x}{j+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{j} \right\rfloor} f(k),$$

and the corresponding inverse matrix,

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

as expressed by the duality of f and its Dirichlet inverse function f^{-1}. Since the recurrence relations for the summatory functions $G(x)$ arise naturally in applications where we have established bounds on sums of Dirichlet convolutions of arithmetic functions, we will go ahead and prove it here before moving along to some motivating examples of the use of this theorem.

Proof of Theorem 3. Let h, g be arithmetic functions where $g(1) \neq 1$ 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 $S_{g,h}(x)$ 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:

$$\begin{aligned} S_{g,h}(x) &:= \sum_{n=1}^{x}\sum_{d|n} g(n)h(n/d) = \sum_{d=1}^{x} g(d)H\left(\left\lfloor \frac{x}{d} \right\rfloor\right) \\ &= \sum_{i=1}^{x}\left[G\left(\left\lfloor \frac{x}{i} \right\rfloor\right) - G\left(\left\lfloor \frac{x}{i+1} \right\rfloor\right)\right] H(i). \end{aligned}$$

Thus we have an implicit statement of a recurrence relation for the summatory function H, weighted by g and G, whose non-homogeneous term is the summatory function, $S_{g,h}(x)$, of the Dirichlet convolutions $g * h$. 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

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

Then 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

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

It is a nice round 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 invertibly 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:

$$\begin{aligned}(g_{x,j}) &= (I-U^T)^{-1}\left(g\left(\frac{x}{j}\right)[j|x]_{\delta}\right)(I-U^T) \\ (g_{x,j})^{-1} &= (I-U^T)^{-1}\left(g^{-1}\left(\frac{x}{j}\right)[j|x]_{\delta}\right)(I-U^T) \\ &= \left(\sum_{k=1}^{\left\lfloor \frac{x}{j}\right\rfloor} g^{-1}(k)\right)(I-U^T) \\ &= \left(\sum_{k=1}^{\left\lfloor \frac{x}{j}\right\rfloor} g^{-1}(k) - \sum_{k=1}^{\left\lfloor \frac{x}{j+1}\right\rfloor} g^{-1}(k)\right).\end{aligned}$$

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

$$\begin{aligned}H(x) &= \sum_{k=1}^{x} g_{x,k}^{-1} \cdot S_{g,h}(k) \\ &= \sum_{k=1}^{x}\left(\sum_{j=\left\lfloor \frac{x}{k+1}\right\rfloor+1}^{\left\lfloor \frac{x}{k}\right\rfloor} g^{-1}(j)\right)\cdot S_{g,h}(k).\end{aligned}$$

It stands to reason that explicit (accurate) bounds on the initial summatory function $S_{g,h}(x)$ ought dictate quite a bit about limiting behaviors and tendencies of the function $H(x)$, such as unboundedness in the limit supremum or infimum sense, when $x \gg 1$. We will tackle these topics in later sections. □

There are many natural ways given the above theorem for expressing new formulas for the Mertens function, $M(x)$. We will simply motivate our theorem's methodology and approach by giving one example, of which many others, and some perhaps more optimal, exist.

Corollary 4 (Convolutions Arising From Möbius Inversion). *Suppose that g is an arithmetic functions 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*

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

Remark 2.7. Ideally, we can use the corollary above to express a definite, usable formula by which we may attempt to derive new bounds for the Mertens function. If this is the goal in applying Corollary 4, the the Dirichlet inverse of g should not depend on μ in any obvious ways, so as to avoid recursion in computing $M(x)$ exactly. This excludes some better known divisor sum convolution identities whose left-hand-side summatory functions have well established bounds with error terms, for example, $1 = d * \mu$, $\phi = \mathrm{Id}_1 * \mu$, and $\sigma = \phi * d$. However, the following two special cases illustrate how our new results allow us to express $M(x)$ in terms of classically well known, and well studied interesting function cases:

- Using $\lambda = \chi_{\mathrm{sq}} * \mu$, where χ_{sq} denotes the characteristic function of the squares, we obtain that $\widetilde{G}(x) \equiv L_0(x)$, the non-weighted summatory function of $\lambda(n)$. The corresponding inverse sequence, g^{-1} for $g(n) := \chi_{\mathrm{sq}}$, can be calculated numerically, leading to the following sequence values:

  $\left\{\chi_{\mathrm{sq}}^{-1}\right\}_{n\geq 1} = \{1,0,0,-1,0,0,0,0,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,-1,0,0,0,0,0,0,0,0,0,0,1,\ldots\}.$

  It appears that the inverse function depends on the values of the Möebius function, so we move along.

- Using $\chi_{\mathbb{P}} + \varepsilon = (\omega + 1) * \mu$ where $\chi_{\mathbb{P}}$ is the characteristic function of the primes, we have that $\widetilde{G}(x) = \pi(x) + 1$, and can compute the inverse sequence of $g(n) := \omega(n) + 1$ 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 apparently dictated by $\lambda(n) = (-1)^{\Omega(n)}$, though no formula for the unsigned magnitudes is immediately obvious.

Example 2.8 (A Recurrence Relating Squares of the Mertens Function). Let
$$\hat{B}_k(x) := \sum_{n \leq x} n^k, k \in \mathbb{R}.$$
When $k \geq 0$ is integer-valued, Faulhaber's formula provides an explicit summation in powers of x to express the function $\hat{B}_k(x)$ involving the Bernoulli numbers. For $k < 1$, the sequence of $\hat{B}_k(x)$ corresponds to the partial sums of a convergent Riemann zeta function value.

We have a known divisor sum convolution for the k^{th} powers of the identity function given by $\mathrm{Id}_k = \sigma_k * \mu$, where $\sigma_k(n) = \sum_{d|n} d^k$ denotes the generalized sum-of-divisors function. It has a known Dirichlet inverse function expressed by the convolution
$$\sigma_k^{-1}(n) = \sum_{d|n} d^k \mu(d) \mu\left(\frac{n}{d}\right).$$
By Theorem 3, we obtain that
$$\hat{B}_k(x) = \sum_{d=1}^{x} \sigma_k(d) M\left(\left\lfloor \frac{x}{d} \right\rfloor\right) \qquad \Longrightarrow$$
$$M(x) = \sum_{m=1}^{x} \left(\sum_{j=\left\lfloor \frac{x}{m+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{m} \right\rfloor} \sigma_k^{-1}(j) \right) \hat{B}_k(m).$$
The inner summation terms in the last equation correspond to a difference of summatory function inputs for $S_k^{-1}(x) := \sum_{n \leq x} \sigma_k^{-1}(n)$. Since the Dirichlet inverse of $\sigma_k(n)$ is given as a divisor sum, we can further transform it with the theorems at hand as
$$\begin{aligned} S_k^{-1}(x) &= \sum_{d=1}^{x} d^k \mu(d) M\left(\left\lfloor \frac{x}{d} \right\rfloor\right) \\ &= \sum_{i=1}^{x} \left[M_k\left(\left\lfloor \frac{x}{i} \right\rfloor\right) - M_k\left(\left\lfloor \frac{x}{i+1} \right\rfloor\right) \right] M(i). \end{aligned}$$
Now by shifting the index of summation, and re-interpreting with our sum of divisor sums construction, we re-expand the last formula for $M(x)$ by writing
$$\begin{aligned} M(x) &= \sum_{i=1}^{x} \left[\hat{B}_k\left(\left\lfloor \frac{x}{i} \right\rfloor\right) - \hat{B}_k\left(\left\lfloor \frac{x}{i+1} \right\rfloor\right) \right] S_k^{-1}(i) \\ &= \sum_{j=1}^{x} \sum_{i=j}^{x} \left[\hat{B}_k\left(\left\lfloor \frac{x}{i} \right\rfloor\right) - \hat{B}_k\left(\left\lfloor \frac{x}{i+1} \right\rfloor\right) \right] \left[M_k\left(\left\lfloor \frac{i}{j} \right\rfloor\right) - M_k\left(\left\lfloor \frac{i}{j+1} \right\rfloor\right) \right] M(j). \end{aligned}$$
We can of course take $k := 0$ in this setup to see a two-index, square-like dependence of the classical Mertens function terms. In this special case, we can simplify according to the observation that $\hat{B}_0(x) = x$ for all $x \geq 1$.

Proposition 5 (Expressing the Mertens Function by Any Arithmetic Function). *Suppose that f is any Dirichlet invertible multiplicative arithmetic function. Let $\widetilde{F}(x)$ be the summatory function of $f * 1$, where by convention we denote $F(x) = \sum_{n \leq x} f(n)$ and $F^{-1}(x) = \sum_{n \leq x} f^{-1}(n)$. Then we have that*

(A) For all $x \geq 1$,

$$M(x) = \sum_{k=1}^{x} \left(\sum_{j=\left\lfloor \frac{x}{k+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{k} \right\rfloor} (\mu * f^{-1})(j) \right) F(k);$$

(B) For all $x \geq 1$, we have a recurrence relation for $M(x)$ given by

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

(C) Moreover, we can write

$$M(x) = \sum_{k=1}^{x} h_{x,k}^{-1} \cdot M(k), \forall x \geq 1,$$

where the inverse matrices $(h_{x,j})^{-1}$ have the formula

$$\begin{aligned} h_{x,j}^{-1} &= \sum_{r=1}^{\left\lfloor \frac{x}{j+1} \right\rfloor} f^{-1}(r) \left[f(j) - F\left(\left\lfloor \frac{x}{r(j+1)} \right\rfloor \right) \right] \\ &\quad + \sum_{r=\left\lfloor \frac{x}{j+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{j} \right\rfloor} f^{-1}(r) \left[F\left(\left\lfloor \frac{x}{rj} \right\rfloor \right) - F\left(\left\lfloor \frac{x}{r(j+1)} \right\rfloor \right) + F(j-1) \right], \forall 1 \leq j \leq x. \end{aligned}$$

Proof. □

2.3.2 Example II: Generalized α-power scaled Mertens functions

This is really just a way of re-writing the results of either Theorem 2 or Theorem 3 to express a formula for the weighted Mertens functions, $M_\alpha^*(x)$, defined by (1). In particular, observe that for any arithmetic function f_0, we can expand

$$\begin{aligned} S_{0,\alpha}(x) &:= \sum_{n \leq x} \frac{1}{n^\alpha} \sum_{d|n} f_0(d) \mu(n/d) \\ &= \sum_{d=1}^{x} \frac{f_0(d)}{d^\alpha} M_\alpha^*\left(\left\lfloor \frac{x}{d} \right\rfloor \right). \end{aligned}$$

Now to put these sums in the class of recurrence relations we handle by Theorem 2, we must identify functions such that

$$S_{0,\alpha}(x) := M_\alpha^*(x) - \pi_\alpha^*(x), \forall x \geq 1.$$

Notice that given a fixed choice of f_0, the function π_α^* can be recovered from the formula

$$f_0(x) = \sum_{d|x} x^\alpha \cdot \nabla[S_{0,\alpha}](d), x \geq 1.$$

Otherwise, if $S_{0,\alpha}(x)$ does not depend on the weighted summatory function for general x, we are firmly in the territory covered by the case of Theorem 3. Notice that the free parameter in the form of the function $f_0(n)$ not only determines the non-homogeneous function in the recurrence relation above, but it allows us some flexibility to play with the bounds on the summatory functions in question.

2.3.3 Example III: Weighted summatory functions for the Liouville lambda function

We define the weighted summatory function of the *Liouville lambda function*, $\lambda(n) = (-1)^{\Omega(n)}$ where $\Omega(n)$ counts the total number of prime factors of n (counting multiplicity), for a parameter $\alpha \geq 0$ as follows:

$$L_\alpha(x) := \sum_{n \leq x} \frac{\lambda(n)}{n^\alpha}, x \geq 1.$$

The special cases of the functions $L(n) \equiv L_0(n)$ and $T(n) \equiv L_1(n)$ are well studied in the literature [**?**, **?**, **?**]. For example, it is known that

$$L_0(x) = O\left(\sqrt{x} \exp\left(C \cdot \log^{1/2}(x) \left(\log\log x\right)^{5/2+\varepsilon}\right)\right),$$

for any $\varepsilon > 0$; and that

$$L_\alpha(x) = O\left(x^{1-\alpha} \exp\left(-C_\alpha \frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}\right)\right),$$

for absolute constants $C_\alpha > 0$ which depend only on the parameter $\alpha \in \mathbb{R}_{\geq 0}$. There are analogous expressions for these summatory functions as we gave for $M(x)$ in Section 1 which involve sums over the non-trivial zeros of the Riemann zeta function. Such formulas result from an inverse Mellin transform and the known Dirichlet series representations over $\lambda(n)$:

$$\int_1^\infty \frac{L_\alpha(x)}{x^{s+1}} dx = \frac{\zeta(2\alpha+2s)}{s \cdot \zeta(\alpha+s)}, \operatorname{Re}(s) > 1-\alpha.$$

For our purposes, we have the fundamental identity that $\chi_{\mathrm{sq}}(n) = (\lambda * 1)(n)$ where $\chi_{\mathrm{sq}}(n)$ is the characteristic function of the squares, i.e., $\chi_{\mathrm{sq}}(n) = [\sqrt{n} \in \mathbb{N}]_\delta$. Now by summing over $n^{-\alpha} \times \chi_{\mathrm{sq}}(n)$ from $1 \leq n \leq x$, we can similarly obtain the following recursion, which corresponds to a case of Theorem 3:

$$H_x^{(2\alpha)} = \sum_{d \leq x^2} \frac{1}{d^\alpha} \cdot L_\alpha\left(\left\lfloor \frac{x^2}{d} \right\rfloor\right), x \geq 1.$$

The left-hand-side sequence corresponds to the generalized 2α-order harmonic number sequence for $\alpha \geq 0$. Some special case formulas for these sequences are given by $H_x^{(0)} = x$, $H_x^{(1)} = \log x + \gamma + O(x^{-1})$. For $\alpha \neq 1/2$, we can approximate these sequences by the integral

$$H_x^{(2\alpha)} \approx \int_1^x \frac{dt}{t^{2\alpha}} = \frac{1}{2\alpha-1}\left(1 - x^{1-2\alpha}\right),$$

or perhaps more precisely when $\operatorname{Re}(\alpha) > 1/2$ by the Riemann zeta function minus its tail as

$$H_x^{(2\alpha)} = \zeta(2\alpha) + O(x^{-2\alpha}).$$

We can use our new result in Theorem 3 to enumerate some almost immediate properties and corollaries that tie $L_\alpha(x)$ to the corresponding weighted Mertens functions, $M_{-\alpha}(x)$. In particular, since the Dirichlet inverse of $\mathrm{Id}_\beta^{-1}(n) = \mu(n)\,\mathrm{Id}_\beta(n)$ for any fixed $\beta \in \mathbb{C}$, we quickly work our theorem magic to obtain that

$$\begin{aligned} L_\alpha(x) &= \sum_{k=1}^{x} \left(\sum_{j=\left\lfloor \frac{x}{k+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{k} \right\rfloor} \frac{\mu(j)}{j^\alpha} \right) H_{\sqrt{k}}^{(2\alpha)} \\ &= \sum_{k=1}^{x} \left(\sum_{j=\left\lfloor \frac{x}{k+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{k} \right\rfloor} \frac{1}{j^\alpha} \right) M_{-\alpha}(k). \end{aligned} \tag{4}$$

We will hit on a few notable special cases of α to make a representative point about what the theorem provides us information on off the bat. Namely, we will consider $\alpha := 1$ (later we should try $\alpha = 0, 1/2$).

Corollary 6 (Relating $L_1(x)$ to the Mertens Function $M_{-1}(x)$). *For all sufficiently large $x \geq 1$, we have that*

$$L_1(x) \approx \frac{\pi^2}{6} M_{-1}(x) + o(1).$$

Thus if one of these two functions grows without bound, or approaches zero, then so must the other.

Proof. We use the first form of (4) together with the bound

$$H_{\sqrt{k}}^{(2)} = \frac{\pi^2}{6} + O\left(\frac{1}{k}\right),$$

to write that

$$L_1(x) = \frac{\pi^2}{6} M_{-1}(x) + O\left(\sum_{k=1}^{x} \left[M_{-1}\left(\left\lfloor \frac{x}{k} \right\rfloor\right) - M_{-1}\left(\left\lfloor \frac{x}{k+1} \right\rfloor\right)\right] \frac{1}{k}\right).$$

In the previous equation we make use of the fact that the intervals

$$\bigcup_{1 \leq k \leq x} \left[\left\lfloor \frac{x}{k+1} \right\rfloor + 1, \left\lfloor \frac{x}{k} \right\rfloor\right] \cap \mathbb{Z} = [1, x] \cap \mathbb{Z}.$$

Now to change the order of summatory functions we are considering, notices that if $S(k) = \frac{1}{k}$, $S(k) - S(k-1) \approx -\frac{1}{k^2}$, so that is our new approximate recurrence weight function when we exchange the sums. We now bound the error term by summing

$$\begin{aligned} E_1(x) &= \left| \sum_{k=1}^{x} \left[\sum_{j=\left\lfloor \frac{x}{k+1} \right\rfloor + 1}^{\left\lfloor \frac{x}{k} \right\rfloor} \frac{1}{j^2} \right] M_{-1}(k) \right| \\ &\approx \left| \sum_{k=1}^{x} \frac{2k+1}{x^2} M_{-1}(k) \right|. \end{aligned}$$

If we perform summation by parts on the numerator of the error sum, we will see that it has lesser order than x^2. So we obtain that the error is $o(1)$. □

2.3.4 Example IV: An initially motivating observation relating special functions to $M(x)$

2.3.5 Example V: A discussion of the key differences and limitations of Theorem 2

The form of Theorem 2 occurs in applications if we have a recurrence for a summatory function whose non-homogeneous term also depends on that summatory function. As in the previous example section four, we arrived at cancellation in the matrix weighting the terms of $M(x)$, which prevented us from applying our nicer basic result in Theorem 3. This can arise in interesting formulations of problems. However, the structure inherited from the matrices implicit to the first theorem form are substantially more complicated in nature. In fact, as is stands, there does not appear to be a clear formula for the corresponding inverse sequence in the general case.

To better explain the complications that arise in inverting matrices of this structural class – and in particular, why the similar construction that led to the symmetric formulas in Theorem 3 does not work here – we observe the following irregular factoid about the differences of the floor functions $\left\lfloor \frac{2x+1}{j} \right\rfloor$ over successive x:

$$\left\lfloor \frac{2x+1}{j} \right\rfloor - \left\lfloor \frac{2x-1}{j} \right\rfloor = \begin{cases} 2, & j = 1; \\ 1, & j \geq 3 \text{ odd and } (2x+1, j) = j \vee (x, j) = j; \\ 1, & j \equiv 0 \bmod 4, \left(2x+1, \frac{j}{2}\right) = \frac{j}{2} \vee \left(x, \frac{j}{2}\right) = \frac{j}{2}; \\ 1, & j \equiv 2 \bmod 4 \wedge \frac{j}{2} \mid x; \\ 0, & \text{otherwise.} \end{cases}$$

We let $\hat{\chi}_j(x)$ denote the indicator of those indices (x, j) such that this difference is non-zero. Thus the difference of summatory functions at these inputs corresponds to

$$G\left(\left\lfloor \frac{2x+1}{j} \right\rfloor\right) - G\left(\left\lfloor \frac{2x-1}{j} \right\rfloor\right) = g\left(\left\lfloor \frac{2x+1}{j} \right\rfloor\right) \hat{\chi}_j(x) + g(2x) \left[j = 1\right]_\delta.$$

If the complicated divisibility structure by which we were forced to discover the analog to the proof technique that worked so elegantly in Theorem 3, does not disuade the reader immediately, we will get to work on that point shortly.

In constructing the recurrence relations common to Theorem 2, it is often also not possible to directly shift the matrix defining the difference of summatory functions directly without introducing a shift matrix parameter.

Let's consider the matrices $\hat{G} := (G\left(\left\lfloor \frac{x}{j} \right\rfloor\right))$, then the matrix problem inherited by those functions satisfying this type of functional recurrence relation looks like the following for some non-identically-zero matrix Δ:

$$(g_{x,j})^{-1} = (I - U^T)^{-1} \left(\left(g\left(\left\lfloor \frac{2x+1}{j} \right\rfloor \right) \hat{\chi}_j(x) + g(2x) \left[j = 1 \right]_{\delta} \right) + (I - U^T)\Delta \right)^{-1} (I - U^T).$$

The middle matrix in the right-hand-side is what kills the beautful result we have from the second variant of the theorem. For example, for $1 \leq x, j \leq 7$, where $*$ has been used as a placeholder for terms with expansions too long to fit on a page, the inverse of this middle matrix term resembles the following:

$$\begin{pmatrix} \frac{1}{g(2)+g(3)+2} & 0 & 0 & 0 \\ -\frac{g(4)+g(5)-2}{(g(2)+1)(g(2)+g(3)+2)} & \frac{1}{g(2)+1} & 0 & 0 \\ -\frac{g(4)+g(5)-g(3)(g(4)+g(5)-2)+g(2)g(6)+g(6)+g(2)g(7)+g(7)-2}{(g(2)+1)^2(g(2)+g(3)+2)} & \frac{1-g(3)}{(g(2)+1)^2} & \frac{1}{g(2)+1} & 0 \\ * & \frac{g(3)^2-2g(3)-(g(2)+1)g(4)+1}{(g(2)+1)^3} & \frac{1-g(3)}{(g(2)+1)^2} & \frac{1}{g(2)+1} \\ * & * & \frac{1-g(3)}{(g(2)+1)^3} & \frac{1}{(g(2)+1)^2} \\ * & * & \frac{-g(4)(g(2)+1)^2+g(3)^2(g(2)+1)-(g(2)+2)g(3)+1}{(g(2)+1)^4} & \frac{1-(g(2)+1)g(3)}{(g(2)+1)^3} \\ * & * & * & \frac{1-2(g(2)+1)g(3)}{(g(2)+1)^4} \end{pmatrix}$$

Performing the pre and post multiplication by the shift matrices in the expression for $(g_{x,j})^{-1}$ only complicates matters and obscures any formula from reach. I have concluded after months of looking the wrong way at this problem – that is to say, by requiring a dead-end matrix substructure instead of picking a prettier formulation of an equivalent problem – that we ought instead defer our attention to cases of divisor sum identities that fit the mold and beautful symmetry in the statement of Theorem 3.