Eigenvalues of the Redheffer Matrix and Their Relation to the Mertens Function

Will Dana

June 7, 2015

Contents

1 Introduction

The Riemann hypothesis, the statement that the nontrivial zeros of the Riemann zeta function have real part $\frac{1}{2}$, is usually viewed as one of the most important unsolved problems in mathematics, in no small part due to the remarkably wide variety of number-theoretic consequences that can be deduced from it. One of the simplest and best-known of these is an equivalent statement about the asymptotics of the Mertens function, the cumulative sum of the Möbius function.

While the Mertens function appears erratic at first glance, it can be represented fairly simply as the determinant of a matrix (the Redheffer matrix) defined in terms of divisibility, and this has spurred some research into the eigenvalues of the matrix. Some early results were presented by Barrett, Forcade, and Follington in their paper "On the Spectral Radius of a (0, 1) Matrix Related to Mertens' Function"[1], which gave an easy proof that most of the eigenvalues are 1, and showed asymptotics on the largest one.

This paper will outline the properties and importance of the Möbius and Mertens functions, then introduce the Redheffer matrix and describe its properties proved in the paper of Barrett et al.

1.1 Notations

This paper uses the following notations:

- $a \mid b$ means b is divisible by a. $a \nmid b$ is the negation.
- $\lfloor x \rfloor$, the **floor** of x, is the greatest integer less than x.
- $\Re(z)$ is the real part of z.

2 Number-Theoretic Background

2.1 Motivation: The Riemann Hypothesis

Recall that the **Riemann zeta-function** is defined for $\Re(s) \geq 1$, $s \neq 1$ by the **Dirichlet series**

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

The function is analytic on this domain, and it can be analytically continued to the entire complex plane with the exception of a simple pole at $s = 1$.

The zeta function's connection to number theory comes from its factorization into the **Euler Product Form**:

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}$$

Bernhard Riemann famously used this representation to obtain an explicit formula for the quantity of primes less than a given number, using the zeroes of the zeta function. As a result, information on the zeroes of ζ translates into information on the distribution of the prime numbers.

It is known that, besides the **trivial zeroes** of ζ, the ones occurring at negative even integers, the zeroes must lie in the **critical strip** $\{s : 0 \leq \Re(s) \leq 1\}$, and that they are symmetrical about the line $\Re(s) = 1/2$. In the same paper in which he derived his formula, Riemann also made the following, much stronger conjecture:

The Riemann Hypothesis. *All of the nontrivial zeroes of the Riemann zeta function have real part* $1/2$*.*

In combination with Riemann's and other estimates for number theoretic functions based on the zeta function, the Riemann hypothesis would effectively allow the tightest possible bounds to be obtained on the error of these estimates.

(These facts, and several other properties of the ζ function cited in this paper, come from E. C. Titchmarsh's *The Theory of the Riemann Zeta-function*. [4])

3 The Möbius and Mertens functions

3.1 The Möbius Function

The Möbius function $\mu : \mathbb{N} \to \{-1, 0, 1\}$ is given by

$$\mu(n) = \begin{cases} 0 & n \text{ has a repeated prime factor (that is, a square factor)} \\ 1 & n \text{ has an even number of nonrepeated prime factors} \\ -1 & n \text{ has an odd number of nonrepeated prime factors} \end{cases}$$

So in particular, $\mu(1) = 1$; $\mu(p) = -1$ for any prime p; $\mu(6) = 1$, since it has an even number of prime factors; $\mu(12) = 0$, since it is divisible by a square.

While this definition may seem somewhat arbitrary, μ has a variety of useful properties, which relate to the following basic fact of the Riemann ζ function. [4]

Theorem 1.

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}$$

Proof (sketch). Given the Euler product form of ζ:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}$$

then

$$\frac{1}{\zeta(s)} = \prod_{p \text{ prime}} (1 - p^{-s})$$

Consider the series formed by multiplying out this infinite product; each of the terms produced will be of the form $(-p_1^{-s})(-p_2^{-s}) \ldots (-p_k^{-s})$ for some collection of *distinct* primes $p_1, \ldots, p_k$. Then the coefficient of the term n^{-s} in the resulting sum will be 0 if it has any repeated prime factors (it will not appear at all in the product), 1 if it has an even number of distinct prime factors, and -1 if it has an odd number. This is the definition of the Möbius function. $\square$

From this, and an important lemma on Dirichlet series, a more directly number-theoretic property of μ can be derived:

Lemma 1 ([2]). *Let*

$$c_n = \sum_{d|n} a_d b_{n/d}$$

Then

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s} \sum_{n=1}^{\infty} \frac{b_n}{n^s} = \sum_{n=1}^{\infty} \frac{c_n}{n^s}$$

Proof. The terms in the product of the two series will be given by $\frac{a_j b_k}{(jk)^s}$. As such, the terms with denominator n^s will have coefficient $a_d b_{n/d}$, where d is an arbitrary divisor of n. Summing these terms gives the result. $\square$

Corollary 1.

$$\sum_{i|n} \mu(i) = \begin{cases} 1 & n = 1 \\ 0 & \textit{otherwise} \end{cases}$$

Proof (sketch). On one hand, the product $\frac{1}{\zeta(s)}\zeta(s)$ is clearly 1. On the other hand, it has a Dirichlet series which can be calculated using Theorem 1 and Lemma 1, with coefficients given by

$$\sum_{i|n} \mu(i)$$

Comparing this with the coefficients found above (1 if $n = 1$, 0 otherwise) gives the result. $\square$

While there exist alternative proofs of this result, this method of proof is remarkably quick and lends itself to a wide variety of identities. It will return in section 6.

3.2 The Mertens Function

The Mertens function $M : \mathbb{N} \to \mathbb{Z}$ is defined to be the cumulative sum of the Möbius function:

$$M(n) = \sum_{k=1}^{n} \mu(k)$$

The most important feature of the Mertens function is its connection with the Riemann hypothesis:

Theorem 2. *The Riemann hypothesis is true if and only if*

$$M(n) = O(n^{1/2+\epsilon})$$

for any $\epsilon > 0$.

The "if" direction of this equivalence is shown in Appendix A. The "only if" direction uses more complicated zeta-function machinery, which is detailed in Titchmarsh [4, p. 370].

Roughly speaking, this result states that an equivalent condition for the Riemann hypothesis is that the Mertens function does not grow significantly faster than the square root. As such, results on the growth of the Mertens function are useful to the general study of prime numbers and the zeta function.

4 The Redheffer Matrix

4.1 Definition

The Redheffer matrix $A_n = \{a_{ij}\}$ is defined by $a_{ij} = 1$ if $j = 1$ or $i \mid j$, and $a_{ij} = 0$ otherwise. For example, the 6×6 matrix is

$$A_6 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

To account for the initial column of 1s, A_n is sometimes represented as a sum $A_n = C_n + D_n$: the matrix $D_n = \{d_{ij}\}$, with $d_{ij} = 1$ if and only if $i \mid j$, and the matrix $C_n = \{c_{ij}\}$, with $c_{ij} = 1$ if and only if $j = 1$ and $i \neq 1$:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

4.2 The Determinant of $\mathbf{A_n}$

Barrett et al.'s calculation of the determinant of the Redheffer matrix depends on a minor lemma regarding the inverse of the matrix D_n.

Lemma 2. *The inverse matrix* $D_n^{-1} = \{\delta_{ij}\}$ *is given by*

$$\delta_{ij} = \begin{cases} \mu(j/i) & \text{if } i \mid j \\ 0 & \text{otherwise} \end{cases}$$

For example,

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & -1 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Proof. The ijth entry of the product of the matrices, p_{ij}, is

$$p_{ij} = \sum_{k=0}^{n} d_{ik}\delta_{kj}.$$

The term $d_{ik}\delta_{kj}$ is 0 unless $i \mid k$ and $k \mid j$, so in particular, if $i \nmid j$, $p_{ij} = 0$. If $i \mid j$,

$$p_{ij} = \sum_{k \text{ with } i|k, k|j} \mu\left(\frac{j}{k}\right) = \sum_{k'|(j/i)} \mu\left(\frac{j/i}{k'}\right) = \begin{cases} 1 & j/i = 1 \\ 0 & \text{otherwise} \end{cases}$$

where this last equality follows from Corollary 1. But then $p_{ij} = 1$ exactly when $i = j$, so $D_n D_n^{-1} = I_n$, as required. □

Theorem 3.

$$\det(A_n) = M(n)$$

Proof. By Lemma 2, $\det(D_n^{-1}) = 1$. Thus

$$\det(A_n) = \det(D_n^{-1}A_n) = \det(D_n^{-1}(C_n + D_n)) = \det(D_n^{-1}C_n + I_n)$$

All of the columns of C_n but the first are identically 0, so all of the columns of the product $D_n^{-1}C_n$ but the first are identically 0. Thus $D_n^{-1}C_n$ is lower triangular, and so is $D_n^{-1}C_n + I_n$; then its determinant is given by the product of its diagonal entries. All of these entries but the first are 1, since only I_n contributes to them. It is easily seen that the first diagonal entry of $D_n^{-1}C_n$ is $\sum_{k=2}^{n} \mu(k)$, which means that the first diagonal entry, and thus the determinant, of the matrix is $M(n)$. □

In particular, this means that the product of the eigenvalues of A_n is $M(n)$. So information about the eigenvalues could be used to obtain information on $M(n)$.

5 The Eigenvalues of $\mathrm{A_n}$

5.1 A Graph-Theoretic Interpretation

Instead of directly using the characteristic polynomial $p_n(x) = \det(xI_n - A_n)$ to obtain information on the eigenvalues of A_n, Barrett et al. investigate the closely related polynomial $q_n(x) = \det(xI_n + (A_n - I_n))$. For example, $q_6(x)$ is the determinant of

$$\begin{pmatrix} x & 1 & 1 & 1 & 1 & 1 \\ 1 & x & 0 & 1 & 0 & 1 \\ 1 & 0 & x & 0 & 0 & 1 \\ 1 & 0 & 0 & x & 0 & 0 \\ 1 & 0 & 0 & 0 & x & 0 \\ 1 & 0 & 0 & 0 & 0 & x \end{pmatrix}$$

Note that

$$\begin{aligned} p_n(x) &= \det(xI_n - A_n) = (-1)^n \det(A_n - xI_n) \\ &= (-1)^n \det((1-x)I_n + A_n - I_n) = (-1)^n q_n(1-x) \end{aligned}$$

As such, the eigenvalues $\{r_j\}$ of A_n are simply given by $\{1 - r'_j\}$, where the r'_js are the roots of $q_n(x)$. It will thus suffice to consider $q_n(x)$ for the remainder of the paper, converting the roots to eigenvalues when necessary.

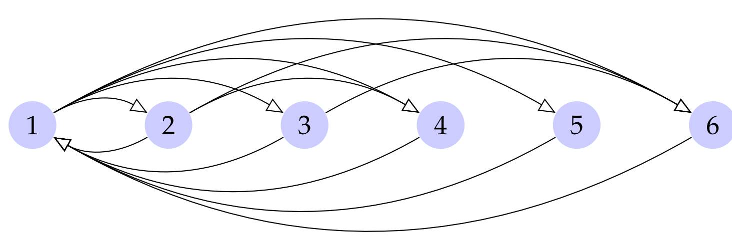
A key insight into determining the properties of this polynomial comes from viewing the matrix $A_n - I_n$ as corresponding to a certain graph.

Definitions. *A* ***directed graph*** *consists of a set of* ***vertices*** *and a set of ordered pairs of vertices, denoting* ***edges*** *between them.*

Given any square matrix, all of whose entries are 0 or 1, a directed graph can be associated with it by assigning a vertex to each row (or column, equivalently) and including an edge between the ith and jth vertices if and only if $a_{ij} = 1$. For example, $A_6 - I_6$

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

corresponds to the graph



Note that in the case of the A_ns, subtracting I_n has the effect of removing edges that loop from a node to itself in the graph.

Returning to the polynomial $q_n(x) = \det(xI_n + (A_n - I_n))$, the determinant calculation allows facilitation by a formulation in terms of permutations:

Lemma 3 ([5]). *Given an $n \times n$ matrix $A = \{a_{ij}\}$,*

$$\det(A) = \sum_{\sigma} a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)} \text{sign}(\sigma)$$

where the sum is taken over all permutations of n elements, and $\text{sign}(\sigma)$ *is 1 if the permutation is formed by an even number of transpositions, -1 if it is formed by an odd number.*

Now, the result of applying this formula to $xI_n + A_n - I_n$ can be broken down by looking at the cycles of the permutation. Given a permutation of a finite set, a **cycle** is a sequence of elements of the permuted set $e_1, e_2, \dots, e_n$ such that $\sigma(e_1) = e_2, \sigma(e_2) = e_3, \dots,$ and $\sigma(e_n) = e_1$. A permutation can then be written down in terms of its cycles. For example, given the permutation that sends $1, 2, 3, 4, 5, 6$ to $2, 4, 6, 1, 5, 3$, its cycles are given by $1 \to 2 \to 4 \to 1$, $3 \to 6$, and $5 \to 5$. This cyclic form of the permutation is commonly notated as

$$(1\ \ 2\ \ 4)(3\ \ 6)(5)$$

From the perspective of cycles, we can make some general statements about the terms in the above form of the determinant. The key point is that the term in the above sum corresponding to a specific permutation σ will vanish unless all of the terms $a_{i,\sigma(i)}$ are nonzero. And in the specific case of $xI_n + A_n - I_n$, $a_{i,\sigma(i)}$ will be nonzero exactly when $\sigma(i) = 1$ or $i \mid \sigma(i)$. So, given a permutation σ,

- If σ has a nontrivial cycle that does not contain 1, then the term is 0. Any nontrivial cycle must contain a number i such that $\sigma(i) < i$ (since if $\sigma(i) = i$ for any number, the cycle is trivial, and if $\sigma(i) > i$ for every i in the cycle, it would never return to the beginning). However, if $\sigma(i) < i$, it cannot be that $i \mid \sigma(i)$, and since the cycle does not contain 1, $\sigma(i) \neq 1$, so $a_{i,\sigma(i)} = 0$ and the whole term is 0.

- If σ has no nontrivial cycles at all, then it is just the identity permutation. The sign is 1, and for any i, $a_{i,\sigma(i)} = a_{i,i} = x$, so the identity permutation contributes the term x^n to the determinant.

- The remaining possibility is that σ has exactly one nontrivial cycle, and it contains 1, since any other nontrivial cycle would not contain 1. For this permutation to contribute a nonzero term, it must be that for each i in the cycle, either $i \mid \sigma(i)$ or $\sigma(i) = 1$. But this is equivalent to saying that for each i in the cycle, there is an edge going from i to $\sigma(i)$ in the directed graph corresponding to $A_n - I_n$ constructed above. This means that the permutations producing nonzero terms in the determinant formula are the ones whose nontrivial cycle containing 1 corresponds to a cycle in the graph, a sequence of vertices with edges pointing from each one to the next and looping around, containing 1.

Because of the way the graph is constructed, its cycles are chains of the form $1 \to a \to ab \to abc \to \ldots \to 1$, where each number is a multiple of the previous one, and the cycle can return to 1 at any time.

Then consider the permutation σ having such a cycle as its only nontrivial cycle (acting as the identity elsewhere). For i in the cycle, $a_{i,\sigma(i)} = 1$, and for all other i, $a_{i,\sigma(i)} = a_{i,i} = x$. In general, a permutation with a single nontrivial cycle of length k can be obtained as a sequence of $k-1$ transpositions (by exchanging the first element of the cycle with each of the following ones in turn) so the sign of the permutation is $(-1)^{k-1}$. So the term that the cycle contributes is

$$a_{1,\sigma(1)}a_{2,\sigma(2)}\ldots a_{n,\sigma(n)}\text{sign}(\sigma) = (-1)^{k-1}x^{n-k}$$

Putting this all together with the term produced by the identity permutation, we get a useful form for the polynomial q_n [1]:

Theorem 4.

$$q_n(x) = x^n + \sum_{k=1}^{n}(-1)^{k-1}c(n,k)x^{n-k}$$

where $c(n,k)$ is the number of cycles of length k in the graph corresponding to $A_n - I_n$.

Once again, we consider the example of $n = 6$:

- $c(6,1) = 0$, since as pointed out above, the graph contains no self-loops, or equivalently, there are no 1s on the diagonal of $A_n - I_n$. In general, $c(n,1) = 0$.

- $c(6,2) = 5$. Any 2-cycle can be portrayed as starting from 1, going to any other number (since 1 will divide it) and returning to 1, and there are 5 other numbers. In general, $c(n,2) = n-1$.

- $c(6,3) = 3$. A 3-cycle will consist of 1, some other number, and then some multiple of that number, which then returns to 1. These cycles are given by $1,2,4$; $1,2,6$; and $1,3,6$.

- $c(6,k) = 0$ for $k > 3$. A cycle of length 4 would have to contain 4 numbers, each a nontrivial multiple of the last, but the smallest possible 4th number in the cycle would be 8 $(1,2,4,8)$ which is greater than n.

Thus $q_6(x) = x^6 - 5x^4 + 3x^3$. The roots of this polynomial are -2.49, 0, 0, 0, 0.66, and 1.83 (approximately); thus the eigenvalues of A_6 are -0.83, 0.34, 1, 1, 1, and 3.49.

The last bullet point above extends to give one of the important results of the paper of Barrett et al.

Theorem 5. *The matrix A_n has eigenvalue 1 with multiplicity $n - \lfloor \log_2 n \rfloor - 1$.*

Proof. In general, in the graph corresponding to some n, the longest cycle is $1 \to 2 \to 4 \to \ldots \to 2^{\lfloor \log_2 n \rfloor} \to 1$, since every entry in the cycle must be a nontrivial multiple of the previous one, and multiplying by 2 each time grows the cycle's elements as slowly as possible. This cycle contains $\lfloor \log_2 n \rfloor + 1$ numbers, which means that $c(n, k) = 0$ for $k > \lfloor \log_2 n \rfloor + 1$. Then the lowest power of x occurring in q_n is $x^{n - \lfloor \log_2 n \rfloor + 1}$, meaning q_n has the root 0 with multiplicity $n - \lfloor \log_2 n \rfloor + 1$. The correspondence between the roots of q_n and p_n gives the result. $\square$

5.2 The Spectral Radius

After establishing that most of the eigenvectors of A_n are 1, Barrett et al. go on to consider the **spectral radius** of the matrix: the maximum absolute value of any of its eigenvalues. The Perron-Frobenius theorem [3] states that a matrix with nonnegative entries and the additional property of irreducibility has a positive eigenvalue whose absolute value is the spectral radius. The Redheffer matrix satisfies this irreducibility condition, and thus it has the largest positive eigenvalue given by the theorem. Now consider the polynomial

$$f_n(x) = p_n(x+1) = (-1)^n q_n(-x) = x^n - \sum_{k=1}^{n} c(n,k) x_n^{n-k}$$

Clearly the largest root of p_n will correspond to the largest root of f_n, so say x_n is this root of f_n; then $\rho_n = x_n + 1$ is the largest eigenvalue. It suffices to consider x_n. First,

$$x_n^2 = \sum_{k=1}^{n} c(n,k) x_n^{2-k}$$

(obtained from simply dividing both sides by x_n^{n-2}).

One can then leverage this equality to prove the following asymptotic result about the largest eigenvalue of A_n:

Theorem 6.

$$\lim_{n \to \infty} \frac{\rho_n}{\sqrt{n}} = 1$$

This result is interesting in comparison with Theorem 2's restatement of the Riemann Hypothesis, because it suggests that the Hypothesis is equivalent to the rough statement that the product of all the other eigenvalues of A_n grows very slowly with n—that is, it is $O(n^\epsilon)$ for any $\epsilon > 0$.

Barrett et al. prove this result using the squeeze theorem. The lower bound $x_n^2 > n - 1$ simply follows from the fact that $c(n, 2) = n - 1$, so

$$x_n^2 = n - 1 + \text{additional positive terms}$$

An upper bound on x_n^2 then comes from the following bound on the $c(n, k)$:

Lemma 4.

$$c(n,k) < n\frac{(\log n)^{k-2}}{(k-2)!}$$

for $k \geq 3$*.*

Proof (sketch). By induction. The base case can be determined from the definition; recall as discussed above that 3-cycles in the divisibility graph are triplets $1 \to a \to ab$ where $ab \leq n$. The number of cycles starting with a fixed a is $\lfloor n/a \rfloor - 1$, the number of nontrivial multiples less than n. So

$$c(n,3) = \sum_{a=2}^{n-1} \left(\left\lfloor \frac{n}{a} \right\rfloor - 1 \right) \leq \left(n \sum_{a=2}^{n-1} \frac{1}{a} \right) - n + 2 \leq n \log n$$

which follows from bounding the sum using the corresponding integral. Then, assume the induction hypothesis for a fixed k. We can obtain a recursive formula for $c(n, k+1)$ by expressing the number of cycles of length $k+1$ in terms of those of length k. Consider the $k+1$-cycles having a fixed penultimate element w. Then such a cycle can be built up by choosing a k-cycle ending on w, and adding a multiple vw, where v can be chosen from among $2, 3 \ldots, \lfloor n/w \rfloor$. The number of k-cycles ending on w is given by $c(w,k) - c(w-1,k)$, so putting this together,

$$c(n, k+1) = \sum_{w=2^{k-1}}^{n} (c(w,k) - c(w-1,k)) \left(\left\lfloor \frac{n}{w} \right\rfloor - 1 \right)$$

(The sum is taken from $w = 2^{k-1}$ since no k-cycle can end up on a number any lower, as discussed in the previous section.) Rearranging the sum and applying the induction hypothesis, Barrett et al. get

$$\sum_{w=2^{k-1}}^{n} (c(w,k) - c(w-1,k)) \left(\left\lfloor \frac{n}{w} \right\rfloor - 1 \right) \leq \frac{n}{(k-2)!} \sum_{w=2^{k-1}}^{n-1} \frac{(\log w)^{k-2}}{w}$$

They then take advantage of where the summand is increasing and decreasing to bound it above by the integral

$$\frac{n}{(k-2)!} \int_1^n \frac{(\log w)^{k-2}}{w} dw = n \frac{(\log n)^{k-1}}{(k-1)!}$$

which completes the proof.

□

Having proven this, they can conveniently say that

$$\begin{aligned} x_n^2 = \sum_{k=1}^{n} c(n,k) x_n^{2-k} &< \sum_{k=3}^{\infty} n \frac{(\log n)^{k-2}}{(k-2)!} \sqrt{n-1}^{2-k} \\ &= n(e^{(\log n)/\sqrt{n-1}} - 1) = n(n^{1/\sqrt{n-1}} - 1) \end{aligned}$$

which also tends asymptotically to n.

6 Observations on the Eigenvectors of $\mathbf{A_n}$

The natural alternative path towards learning about the eigenvalues of A_n is to consider its eigenvectors. This section collects some miscellaneous observations in that area.

First, a quick statement can be made regarding eigenvectors with the multiple eigenvalue 1.

Theorem 7. *The eigenspace of A_n with eigenvalue 1 has dimension $\lceil n/2 \rceil - 1$.*

In particular, A_n is non-diagonalizable for $n \geq 5$.

Proof. This can be seen simply by row-reducing $A_n - I_n$ to determine its nullity. Because of the way $A_n - I_n$ is constructed, it is already almost in echelon form. For reference, consider again the matrix $A_6 - I_6$:

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Note that in general, the $(\lfloor n/2 \rfloor + 1)$th row of $A_n - I_n$ and all rows below it consist of a 1 followed by 0s, since all nontrivial multiples of $\lfloor n/2 \rfloor + 1$ are greater than n. Then this row can be subtracted from all but the first to remove the other 1s in the first column, and then permuted to the top of the matrix, so that it looks like this:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

For any n, the resulting matrix will be in echelon form because of the way A_n is constructed: after the second row, the first nonzero entry in each row will occur two columns after the previous row. Since all the rows beneath the $(\lfloor n/2 \rfloor + 1)$th are zeroed out, the required nullity is $n - (\lfloor n/2 \rfloor + 1) = \lceil n/2 \rceil - 1$. □

However, it is easier to make further statements about the eigenvectors of A_n^T (or equivalently, the left/row eigenvectors of A_n). These are still relevant because the eigenspaces of A_n and A_n^T are isomorphic, and they have an interpretation in terms of important number-theoretic operations, stemming from the following description:

Lemma 5. *$(a_1, a_2, \ldots, a_n)$ is an eigenvector of A_n^T with eigenvalue λ if and only if:*

- *for $n \geq 2$, $\sum_{d|n} a_d = \lambda a_n$, and*

- $\sum_{k=1}^{n} a_k = \lambda a_1$

Proof. This follows directly from the definitions of A_n and eigenvectors. $\square$

Note that if an eigenvalue $\lambda \neq 1$, then given a value for a_1, the remaining a_js are uniquely determined by the lemma, since they can be computed recursively as

$$a_j = \frac{1}{\lambda - 1} \sum_{d|j, d\neq j} a_d$$

This result immediately follows:

Theorem 8. *Every eigenspace of A_n with eigenvalue not equal to 1 has dimension 1.*

So now define the function $v_\lambda : \mathbb{N} \to \mathbb{R}$ for an arbitrary $\lambda \neq 1$ using this recursive formula:

$$v_\lambda(n) = \begin{cases} 1 & n = 1 \\ \frac{1}{\lambda-1} \sum_{d|n, d\neq n} v_\lambda(d) & \text{otherwise} \end{cases}$$

Then Lemma 5 can be reformulated as a statement about eigenvalues.

Theorem 9. *λ is an eigenvalue of A_n if and only if*

$$\sum_{k=1}^{n} v_\lambda(k) = \lambda$$

Proof. If λ is an eigenvalue, then in particular there is a corresponding eigenvector with first component 1. By the discussion above, its coefficients a_k must be given by $v_\lambda(k)$, so by Lemma 5, they satisfy the given relation. And, if the $v_\lambda(k)$ satisfy the given relation, then by Lemma 5, they are the components of an eigenvector with eigenvalue λ. $\square$

6.1 The Dirichlet Series of $\mathbf{v}_\lambda$

A major advantage of v_λ's definition in terms of sums over divisors is an interesting interaction between its Dirichlet series

$$V_\lambda(s) = \sum_{n=1}^{\infty} \frac{v_\lambda(n)}{n^s}$$

and the ζ function. (Note that in the following proof sketches using this series, it will be manipulated strictly algebraically. To make the proofs fully rigorous would require a consideration of convergence. However, the convergence and analyticity of Dirichlet series are well understood. Some general results on convergence can be found in Gamelin [2, p. 376-378].)

Theorem 10.

$$V_\lambda(s) = \frac{\lambda - 1}{\lambda - \zeta(s)}$$

Proof. By Lemma 1,

$$V_\lambda(s)\zeta(s) = \sum_{n=1}^{\infty}\left(\sum_{d|n} v_\lambda(d)\right)\frac{1}{n^s} = \left(\sum_{n=1}^{\infty}\frac{\lambda v_\lambda(n)}{n^s}\right) - \lambda + 1 = \lambda V_\lambda(s) - \lambda + 1$$

since the identity $\sum_{d|n} v_\lambda(d) = \lambda v_\lambda(n)$ holds for all n except 1, thus accounting for the additional constant term. Solving for $V_\lambda(s)$ gives the result. □

Theorem 11. *Let $a_k(n)$ be the number of ways of expressing n as a product of k factors, order mattering. Then if $\lambda > 1$,*

$$v_\lambda(n) = \left(1 - \frac{1}{\lambda}\right)\sum_{k=0}^{\infty}\frac{a_k(n)}{\lambda^k}$$

Proof. Rearranging Theorem 10,

$$V_\lambda(s) = \left(1 - \frac{1}{\lambda}\right)\frac{1}{1 - \frac{\zeta(s)}{\lambda}}$$

For s sufficiently large, $1 < \zeta(s) < \lambda$, and

$$V_\lambda(s) = \left(1 - \frac{1}{\lambda}\right)\sum_{k=0}^{\infty}\frac{\zeta(s)^k}{\lambda^k}$$

Now note $a_k(n)$ satisfies the recursive definition

$$a_k(n) = \begin{cases} 1 & k = 0, n = 1 \\ 0 & k = 0, n \neq 1 \\ \sum_{d|n} a_{k-1}(n/d) & k > 0 \end{cases}$$

since the number of expressions of n as a product of k factors with a specific first factor d is the number of ways of expressing n/d as a product of $k-1$ factors. Then $\sum_{n=1}^{\infty}\frac{a_0(n)}{n^s} = 1$, and applying Lemma 1 repeatedly gives

$$\zeta(s)^k = \sum_{n=1}^{\infty}\frac{a_k(n)}{n^s}$$

This implies

$$V_\lambda(s) = \left(1 - \frac{1}{\lambda}\right)\sum_{k=0}^{\infty}\frac{1}{\lambda^k}\sum_{n=1}^{\infty}\frac{a_k(n)}{n^s} = \left(1 - \frac{1}{\lambda}\right)\sum_{n=1}^{\infty}\left(\sum_{k=0}^{\infty}\frac{a_k(n)}{\lambda^k}\right)\frac{1}{n^s}$$

and comparing coefficients gives the result. □

Combining this result with Theorem 9 and bounds on the $a_k(n)$ could lead to an alternate proof of the result on the spectral radius in the paper of Barrett et al.

7 Conclusion

The Redheffer matrix provides an interesting alternate perspective on questions of divisibility. Like the Riemann zeta-function (to which it is intimately connected), it effectively bundles together the divisibility properties of many numbers into a single object, which can be studied using the techniques of linear algebra. The readily attainable results on the eigenvalues presented here suggest that the matrix provides a fruitful line of inquiry.

A The Relationship Between the Mertens Function and the Riemann Hypothesis

Theorem 12. *If*

$$M(n) = O(n^{1/2+\epsilon})$$

for all $\epsilon > 0$, then all zeroes of $\zeta(s)$ have real part $1/2$.

Proof. The idea is to show that, if the given bound on the Mertens function holds, the series $\sum_{k=1}^{\infty} \frac{\mu(k)}{k^s}$ converges to an analytic function for $\Re(s) > 1/2$. By Theorem 1, this function is equal to $1/\zeta(s)$, so if the series converges for a given s, $\zeta(s) \neq 0$. The zeroes of ζ in the critical strip are mirrored about the line $\Re(s) = 1/2$, so having shown there are no nontrivial zeroes with $\Re(s) > 1/2$, there can be no nontrivial zeroes with $\Re(s) < 1/2$ either.

Let $s = 1/2+\epsilon$ for $0 < \epsilon < 1/2$. Then recall the summation by parts formula

$$\sum_{k=1}^{m} a_k b_k = \left(\sum_{k=1}^{m-1} (a_k - a_{k+1}) B_k \right) + a_m B_m$$

where $B_k = \sum_{j=1}^{k} b_j$. Then in particular,

$$\sum_{k=1}^{m} \frac{\mu(k)}{k^s} = \left(\sum_{k=1}^{m-1} \left(\frac{1}{k^s} - \frac{1}{(k+1)^s} \right) M(k) \right) + \frac{M(m)}{m^s}$$

Assuming the given condition on the Mertens function, there is a constant C such that $|M(k)| < Ck^{1/2+\epsilon/2}$ for all k. Then

$$\left| \left(\frac{1}{k^s} - \frac{1}{(k+1)^s} \right) M(k) \right| = \left| \frac{(k+1)^s - k^s}{(k^2+k)^s} M(k) \right| < |(k+1)^s - k^s| \left| \frac{Ck^{1/2+\epsilon/2}}{k^{2s}} \right|$$
$$< sk^{s-1} \frac{Ck^{1/2+\epsilon/2}}{k^{1+2\epsilon}} = Csk^{-1-\epsilon/2} < Ck^{-1-\epsilon/2}$$

Since $\sum C/k^{1+\epsilon/2}$ converges for $\epsilon > 0$, the series converges. The uniform convergence over the required domain can now be established using a general theorem on convergence of Dirichlet series [2, p. 377]:

Theorem 13. *If a series $\sum a_n/n^s$ converges at $s = s_0$, then it converges uniformly in any sector $\{|\arg(s - s_0)| \leq \pi/2 - \epsilon\}$ for $\epsilon > 0$.*

Here, under the assumption, $\sum \mu(n)/n^s$ converges for any $s > 1/2$, and by the theorem, for any point s with $\Re(s) > 1/2$, there is a set containing s on which it converges uniformly. Thus $1/\zeta(s)$ is analytic on the required domain, and $\zeta(s)$ can have no zeroes there. □

This theorem is interesting in that a weaker big-O statement about $M(n)$ leads to a weaker statement about the zeroes of ζ:

Theorem 14. *If*

$$M(n) = O(n^{\alpha+\epsilon})$$

for some fixed real α and any $\epsilon > 0$, and s is a nontrivial zero of ζ, then $1 - \alpha \leq \Re(s) \leq \alpha$.

This follows from simply replacing "1/2" in the above proof with α.

References

[1] W. W. Barrett, R. W. Forcade, and A. D. Follington, On the Spectral Radius of a (0, 1) Matrix Related to Mertens' Function, *Linear Alg. and its Appl.*, **107**(1987), 151–159

[2] T. Gamelin, *Complex Analysis*, p. 377. Springer, New York, 2001.

[3] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*, second edition. Academic Press, 1985. 570 pages.

[4] E. C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, second edition. Oxford University Press, Oxford, 1986

[5] S. Treil, *Linear Algebra Done Wrong*, self-published, Brown University, 2014