2019

Ruth-Aaron Numbers: An Exploration in Analytic Number Theory

Madeleine Farris
mfarris@wellesley.edu

Ruth-Aaron Numbers: An Exploration in Analytic Number Theory

Madeleine Farris
advisor: Professor Alex Diesl

submitted in partial fulfillment of the prerequisite for honors

Department of Mathematics
Wellesley College
April 2019

Abstract

If we let $S(n)$ be the sum of the prime factors (with multiplicity) of an integer n, then we define a Ruth-Aaron Number to be any n such that $S(n) = S(n+1)$. The main results regarding Ruth-Aaron numbers that Erdős and Pomerance proved show that the Ruth-Aaron numbers have density zero and the sum of their reciprocals converges. We extend their results by replacing the function S (which sums the prime powers of a number n) with other functions f on the prime factors. In the process we take a look at the history of analytic number theory, and review some classical results from the field.

Acknowledgements

First and foremost I thank my advisor, Professor Diesl, not only for his assistance with the mathematics but also for providing me room and guidance to grow and learn. Thanks to my academic advisor Andy for his mentorship and friendship throughout these past 4 years. To Professor Volić and James for lending a listening ear to both my complaints and conundrums. Thanks to Maggie Rivers for going on this rollercoaster ride of thesis and grad school together with me. And last (but certainly not least) I thank Amy, not only for being my constant companion in mathematical pursuits, but also for her unending grace proffered in cups of tea and late night Ramen.

Contents

Chapter 1

Introduction to Analytic Number Theory

When I began working on the project that would eventually evolve into my senior thesis, I knew very little about the field of analytic number theory. Throughout the course of my work on this project I have learned quite a bit about not only the research process and Ruth-Aaron numbers, but also about analytic number theory as a whole: its purpose, its tools, and many classical results. In this document I present all of these things to you, not just the progress towards new proofs but some history and context of the project as well. My goal is to give you the (condensed) experience that I had working this year: a window into analytic number theory through the lens of one particular problem.

To begin the journey we will take a brief look at the history of analytic number theory, as well as proofs of some of the classical results in the field. I then delve more deeply into the background behind my particular project, something referred to as "Ruth-Aaron Numbers". To help demonstrate the tools and proof style of the field I provide a reading guide for [4], the main paper that my project grew out of. Finally, I state and prove two new results and demonstrate progress towards an even larger result.

1.1 History

One attempt at defining analytic number theory would be "a field in mathematics that applies tools from analysis to questions and objects in number theory." However, this definition is about as enlightening as the phrase "analytic number theory" itself. But what are the questions and problems of analytic number theory, and what tools are used to tackle them?

To answer the first question in as most generality as possible, analytic number theory involves problems about the integers (often more specifically the primes) and tends to be focused on what occurs in the "long run" or for extremely large values. One of the first and biggest problems in analytic number theory involves the prime-counting function.

Definition 1. Define the **prime counting function**, notated $\pi(x)$, to be the number of primes less than or equal to x for all real numbers x.

In general, understanding the behavior of $\pi(x)$ is hugely important to understanding the behavior of primes. Some of the earliest roots of analytic number theory come from trying to approximate $\pi(x)$ for large values of x.

Herein begins the answer to the second question: frequently analytic number theory seeks to provide estimations of various quantities (often expressed as arithmetic functions). These estimations come in many forms, for example upper and lower bounds and asymptotic analyses (things which will be discussed in more detail later). Despite the connotation of "estimation" as inexact, many of the estimations obtained in analytic number theory are quite tight and prove to be extremely useful results. As an example, the first big result in analytic number theory is the Prime Number Theorem which provides an asymptotic estimate of the prime-counting function.

Theorem 1.1. *(Prime Number Theorem) If $\pi(x)$ is the prime counting function, then*

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1$$

The first conjecture towards this result was given by Adrien-Marie Legendre in either 1797 or 1798, suggesting that $\pi(x)$ is well approximated by $x/(A\ln(x)+B)$ for some constants A, B. He later made a stronger statement that he believed A to be 1 and B to be -1.08366. Peter Gustav Lejeune Dirichlet made a similar conjecture, this time using the logarithmic integral. Later in 1848, Pafnuty Chebyshev made some progress on the problem. He showed that there exist constants such that $Ax/\log(x) \leq \pi(x) \leq Bx/\log(x)$ and that if the limit of $\pi(x)/(x/\ln(x))$ exists then it must be equal to 1. Using these same methods he proved Bertrand's postulate which states that there exists a prime number between n and $2n$ for any $n \geq 2$.

Despite these results found by Chebyshev, a proof of the Prime Number Theorem in its entirety was still elusive. However, around the same time other results in analytic number theory were first being developed. In 1837 Dirichlet published a result known as "Dirichlet's theorem on arithmetic progressions" which said that for any two, positive coprime integers a, d there are infinitely many primes of the form $a+nd$ where n is some non-negative integer. Another way of looking at this is to say that there are infinitely many primes congruent to a mod d. The reason this is a theorem about arithmetic progressions (which is not immediately obvious) is that another rephrasing of what Dirichlet proved is that the numbers $a+nd$ form an arithmetic progression that contains an infinite number of prime numbers.

Beyond the consequences of this theorem on its own, Dirichlet's work is extremely vital to the history of analytic number theory because in its proof it introduced new ideas and methods which are now central to the field. Chief among them was the introduction of complex functions into analytic number theory in the form of Dirichlet characters and L-functions. We would be remiss if we did not talk about the importance of complex analysis in analytic number theory, however we will not require complex analysis for the work done later in the thesis, and so we will only briefly touch on it. Dirichlet introduced two new definitions:

Definition 2. A Dirichlet character is any function χ from the integers $\mathbb{Z}$ to the complex numbers $\mathbb{C}$ such that χ has the following properties:

1. There exists a positive integer k such that $\chi(n) = \chi(n+k)$ for all n

2. If $\gcd(n, k) > 1$ then $\chi(n) = 0$; if $\gcd(n, k) = 1$ then $\chi(n) \neq 0$.

3. $\chi(mn) = \chi(m)\chi(n)$ for all integers m and n.

Definition 3. If χ is a Dirichlet character and s is a complex variable with real part greater than 1 then the meromorphic continuation of the function

$$L(\chi, s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

is a **Dirichlet L-function**.

Up until now we haven't seen any complex numbers anywhere, and so the introduction of a complex function to analytic number theory may seem to come out of the blue. To explain this, first let us talk briefly about the Riemann zeta function, which is the $L-$function with $\chi(n) = 1$. The Riemann zeta function is an extremely famous function, known most because of its starring role in the Riemann Hypothesis. We can more easily see the connection between the Riemann zeta function and number theory through the Euler product formula, which was proved by Leonhard Euler in 1737:

Theorem 1.2 (Euler Product Formula)**.**

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

Here we can see a bit better the relationship between these $L-$functions and prime numbers. On the left hand side we have a sum over natural numbers and on the right a product of primes. This relationship allows us to swap out information about the natural numbers, which are well understood, to the prime numbers. In fact, Bernhard Riemann further investigated this in 1859 in a paper titled “On the Number of Primes Less Than a Given Magnitude,” in which he made some important connections between the Riemann zeta function and the distribution of primes (as well as stating the famous Riemann Hypothesis). It was also in this paper that the idea of applying complex analysis to solving number theoretic problems was introduced. Indeed, this idea was so groundbreaking that in 1896 Jacques Hadamard and Charles Jean de la Vallée Poussin simultaneously proved the long sought after Prime Number Theorem by extending some methods used by Riemann. This closed the chapter on the question about the prime counting function, but opened the door to a huge field of mathematics.

1.2 Groundwork

To help us further understand the tools and objects used in the field we first need to define some terms. We stated earlier that analytic number theory involves problems about the integers, but the way it tends to ask and answer those questions is through arithmetic functions:

Definition 4. Any function $f(n)$ which takes in the natural numbers and whose outputs are a subset of the complex numbers is an **arithmetic function**.

We have already seen an example of an arithmetic function: the prime counting function. We can also further characterize arithmetic functions by how they interact with addition and multiplication:

Definition 5. If $m, n \in \mathbb{N}$ and f is an arithmetic function such that $f(mn) = f(m) + f(n)$ for coprime m, n then we say that f is **additive**. However, if $f(mn) = f(m) + f(n)$ for all m, n then we say that f is **completely additive**.

Definition 6. Similarly, if $m, n \in \mathbb{N}$ and f is an arithmetic function such that $f(mn) = f(m)f(n)$ for coprime m, n then we say that f is **multiplicative**. However, if $f(mn) = f(m)f(n)$ for all natural numbers m, n then f is **completely multiplicative**

It turns out the $\pi(x)$ is neither multiplicative nor additive. An example of a completely additive function is the function that counts the number of prime divisors of n with multiplicity. On the other hand an example of a multiplicative function is the sum of the divisors of n. All of these functions are ones that are frequently investigated or utilized in analytic number theory.

Since we would like to investigate such functions' behavior "in the long run" it will be very useful to have a tool to compare one function to another. The common way of doing this is called Landau Notation. The following definitions are taken from [6] pages 44-45.

Definition 7. We say that two functions f and g satisfy the relationship $f(x) \ll g(x)$ as $x \to \infty$ if there exists constants C and x_0 such that $f(x) \leq Cg(x)$ for all $x > x_0$

It is important to note that while this is the formal definition of this notation, we often do not explicitly state the constants C and x_0. This notation is on the one hand helpful because it allows us to suppress unnecessary constants, on the other hand it also creates a useful ordering on functions. This is because the relation "$\ll$" is both reflexive and transitive. However it is not symmetric, i.e. $f \ll g$ does not imply $g \ll f$, and so we still need a tool to talk about when two functions are about equal.

Definition 8. For functions f, g, and h we say that $f(x) = g(x) + O(h(x))$ if $|f(x) - g(x)| \ll h(x)$

This relation, on the other hand, is truly an equivalence relation. A word of warning though, we will often write that $f(x) = O(g(x))$ when $f(x) \ll g(x)$, however this is not an equivalence relation (since it does not satisfy symmetry, as above), while it may be tempting to treat it as such. There is also one other notation that we will use occasionally, given as follows.

Definition 9. For functions f, g we say that f is $o(g(x)$ if for all constants C there exists x_0 such that $f(x) < Cg(x)$ for all $x > x_0$.

The difference between the definitions of these terms is slight, however they are indeed quite different. In particular, little-o notation is "stronger" in that little-o notation implies big-o notation. We often think of little-o notation as being a measure of when two functions are asymptotic to each other. Now, with these tools, we can begin looking a bit deeper at analytic number theory.

1.3 Chebyshev's Methods

To help us better understand exactly what these "estimating" methods are and how we use them, we'll work through a couple of canonical results in analytic number theory using Chebyshev's methods. Many of these problems are based off of the exercises from Section 3.1 of [5]. To start let us define some useful functions that were first introduced by Chebyshev.

Definition 10. Chebyshev started by defining the functions $\theta(n)$ and $\psi(n)$ which have gotten quite a bit of use beyond just this result.

$$\theta(n) = \sum_{p \leq n} \log(p)$$

$$\psi(n) = \sum_{p^\alpha \leq n} \log(p)$$

We will also need the following lemma which gives us a useful bound on one of Chebyshev's functions. The proof of this result can be found in 3.1 of [5].

Lemma 1.3. $\theta(n) \leq 2n \log(2)$

Now we are ready to prove Chebyshev's result.

Theorem 1.4. *There are positive constants A and B such that*

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

Proof. We first claim that

$$(\pi(x) - \pi(\sqrt{x})) \log(\sqrt{x}) \leq \theta(x) - \theta(\sqrt{x}) \tag{1.5}$$

To see this we observe that

$$\sum_{\sqrt{x} < p < x} 1 = \pi(x) - \pi(\sqrt{x})$$

Then we can estimate this sum as follows

$$\log(\sqrt{x}) \sum_{\sqrt{x} < p < x} 1 \leq \sum_{\sqrt{x} < p < x} \log(p)$$

Using Lemma 1.3 we can rewrite (1.5) as

$$\pi(x) \leq \frac{4x \log(2)}{\log(x)} + \pi(\sqrt{x})$$

However $\pi(\sqrt{x}) = O(\sqrt{x})$ and also

$$\frac{4x\log(2)}{\log(x)} + O(\sqrt{x}) = O\left(\frac{x}{\log(x)}\right)$$

which gives us that

$$\pi(x) \leq O\left(\frac{x}{\log(x)}\right).$$

Therefore we have shown the upper bound. For the lower bound, we start by considering $\psi(x) - \theta(x)$, which will allow us to translate information about one of these functions into the other. We would like to replace this difference of sums with a product of sums (which we can then estimate). To do so we represent it as a sum of $\log(p)$ for each $p^2 < x$ and multiply by a sum which will help us account for the other possible powers (we do not account for primes raised to the first power, because they are precisely what we are subtracting out). To account for these higher power primes we multiply by a second sum, which accounts for the multiplicity of the $\log(p)$, i.e. the power of the prime. We find that

$$\begin{aligned}\psi(x) - \theta(x) &= \sum_{p^\alpha \leq x} \log(p) - \sum_{p \leq x} \log(p)\\ &= \sum_{p^2 \leq x} \log(p) \sum_{2 \leq m \leq \frac{\log(x)}{\log(p)}} 1\end{aligned}$$

We can then make this one sum and estimate using Lemma 1.3

$$\begin{aligned}\sum_{p^2 \leq x} \log(p) \sum_{2 \leq m \leq \frac{\log(x)}{\log(p)}} 1 &\leq \sum_{p^2 \leq x} \log(p) \left\lceil \frac{\log(x)}{\log(p)} \right\rceil\\ &\leq \sqrt{x}\log(x)\end{aligned}$$

Thus we have shown that $\psi(x) - \theta(x) = O(\sqrt{x}\log(x))$. We now claim that $x \ll \psi(x)$. To see this take $x = p_1^{a_1} \cdots p_n^{a_n}$. Then we have

$$\begin{aligned}\log(x) &= \log(p_1^{a_1} \cdots p_n^{a_n})\\ &\leq \sum_{p^\alpha \leq x} \log(p)\\ &= \psi(x)\end{aligned}$$

which demonstrates our claim. Since $\psi(x) - \theta(x) = O(\sqrt{x}\log(x))$ we have that $x \ll \theta(x)$ as well. We observe that

$$\sum_{\sqrt{x} < p < x} \log(p) \leq \pi(x)\log(x)$$

and conclude that $x \ll \pi(x)\log(x)$. This implies the following

$$\pi(x) \gg \frac{x}{\log(x)}$$

□

This was Chebyshev's main result, but we can use his methods to prove a few other useful results. We would like to work towards an estimation of a couple of common sums of primes: the sum of primes as well as $\log(p)/p$ and $1/p$, which are sums that shows up frequently in various problems (and in fact as we'll see later on they are used multiple times throughout the proofs of the new results). We start with the sum of $\log(p)/p$. Before we are able to prove this, however, we need a formula for $\log(n)$:

Lemma 1.6.

$$\sum_{n \leq x} \log(n) = x\log(x) - x + O(\log(x))$$

The proof of this result is not difficult, but is fairly technical. For the interested reader, you can use the Euler-Maclaurin summation formula with $k = 1$ and $f(x) = \log(n)$. We are now ready to take a look at our first summation.

Theorem 1.7.

$$\sum_{p \leq x} \frac{\log(p)}{p} = \log(x) + O(1)$$

Proof. We will actually prove a slightly different result, which will imply our claim. Instead of looking at $\log(n)$ we will consider the closely related Von Mangoldt function, defined as follows:

$$\Lambda(n) = \begin{cases} \log(p) & \text{if } n = p^\alpha \text{ for some prime } p \text{ and integer } \alpha \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

Consider the following sum

$$\sum_{n \leq x} \frac{\Lambda(n)}{n}$$

Then observe that

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \sum_{p^\alpha} \frac{\log(p)}{p}.$$

But since $\log(p^\alpha) = \alpha \log(p)$ for any given p it follows that

$$\sum_{p \leq x} \frac{\log(p)}{p} = O\left(\sum_{n \leq x} \frac{\Lambda(n)}{n}\right)$$

and therefore we can estimate the sum over the Von Mangoldt function and that bound will apply to our original sum as well. To do this, we will evaluate something better known in two separate ways, one of which gives us a useful bound, and one of which relates to what we actually want to know about. In our case we observe that

$$\begin{aligned}\sum_{n\le x}\log(n) &= \sum_{d|n}\Lambda(d)\\ &= \sum_{d\le x}\Lambda(d)\left\lfloor\frac{x}{d}\right\rfloor\\ &= x\sum_{d\le x}\frac{\Lambda(d)}{d}+O(\psi(x))\end{aligned}$$

But from the proof of Theorem 1.4 we know we can replace $\psi(x)$ with $O(x)$. On the other hand from Lemma 1.6 we know that

$$\sum_{n\le x}\log(n) = x\log(x)-x+O(\log(x))$$

Putting these results together and dividing the entire expression by x we find that

$$\sum_{d\le x}\frac{\Lambda(d)}{d}+O(1)=\log(x)+O(1)$$

which then implies our claim. □

To evaluate the sum of $1/p$ we will need a technique referred to as partial summation, which is extremely useful in evaluating sums.

Theorem 1.8 (Partial Summation)**.** *Suppose* $\{a_n\}_{n=1}^{\infty}$ *is a sequence of complex numbers and* $f(t)$ *is a continuously differentiable function on* $[1,x]$*. Set*

$$A(t)=\sum_{n\le t}a_n$$

Then

$$\sum_{n\le x}a_nf(n)=A(x)f(x)-\int_1^x A(t)f'(f)dt$$

Now we are ready to evaluate the sum of reciprocals of primes:

Theorem 1.9.

$$\sum_{p\le x}\frac{1}{p}=\log\log(x)+O(1)$$

Proof. We will first put our summand into a slightly different form in order to utilize partial summation, considering:

$$\sum_{p\le x}\frac{\log(p)}{p}\frac{1}{\log(p)}$$

then in the formula for partial summation we take

$$A(t)=\sum_{p\le x}\frac{\log(p)}{p}$$

and $f(p)=1/\log(p)$. Using partial summation and Theorem 1.7 we find that

$$\begin{aligned}\sum_{p\le x}\frac{1}{p}&=(\log(x)+O(1))\frac{1}{x}+\int_2^x\frac{\log(t)+O(1)}{x\log^2(x)}dt\\&=O(1)+\int_2^x\frac{\log(t)}{t\log^2(t)}dt+\int_2^x\frac{O(1)}{t\log^2(t)}dt\\&=O(1)+\log\log(x)+O(1)-\frac{1}{\log(x)}+O(1)\\&=\log\log(x)+O(1)\end{aligned}$$

as desired. □

The last sum we will look at is the sum of all primes up to some x. The estimation given below can be found using partial summation similar to above.

Theorem 1.10.

$$\sum_{p\le x}p=\frac{x^2}{2\log(x)}+O\left(\frac{x^2}{\log^2(x)}\right)$$

These estimations will serve us well in the coming chapters as we evaluate various sums of primes.

Chapter 2

Previous Work in Ruth-Aaron Numbers

2.1 Background

Now we will move on to discussing the main problem of this thesis: a generalization of some results regarding "Ruth-Aaron Numbers." The Ruth-Aaron Numbers have a slightly strange bit of history to them: they are named for the baseball players Babe Ruth and Hank Aaron. In 1974 Hank Aaron hit his 715th major league homerun, beating Babe Ruth's previous record of 714. Due to this event the numbers 714 and 715 received quite a bit of publicity, and Carl Pomerance observed that they have some interesting properties. The first one noticed is that their product is the product of the first 7 primes (conjectured to be the largest consecutive pair of numbers whose product is the product of the first k primes for some k). However, these numbers have another fascinating property, but before we talk about that, we define a useful function.

Definition 11. For $n = p_1^{a_1} \cdots p_k^{a_k}$ for distinct primes p_i, define a function $S(n)$ by

$$S(n) = \sum_{i=1}^{k} a_i p_i$$

Suppose we take 714 whose prime factorization is $2 \cdot 3 \cdot 7 \cdot 17$. Then we have $S(714) = 2+3+7+17 = 29$. On the other hand for $715 = 5 \cdot 11 \cdot 13$ we have $S(715) = 5+11+13 = 29$. So $S(714) = S(715)$, a somewhat surprising result that was observed first by one of Pomerance's students named Jeremy Jordan. Then Pomerance along with David E. Penney and Carol Nelson published these observations about 714 and 715 in a recreational math journal (see [3]). In this paper they established what a "Ruth-Aaron Number" is, i.e any positive integer n such that $S(n) = S(n+1)$. The first few such n are 5, 8, 15, 77, 125, 714, 948 and 1330. Through an explicit construction they showed that if we assume Schintzel's Hypothesis H (see [3]) then there are infinitely many Ruth-Aaron Numbers.

They also stated in the paper that "The numerical data suggest that Aaron numbers are rare. We suspect they have density 0, but we cannot prove this." That is, however, until Paul Erdős joined the picture and together he and Pomerance published a paper in 1974 in which they proved that the density of the Ruth-Aaron Numbers is indeed 0 (see [2], pgs. 1-2). Then much later in 2002 Pomerance published a paper in memory of Paul Erdős in which he proved an even tighter result: that the sum of reciprocals of Ruth-Aaron Numbers is bounded (see [4]).

To help us understand exactly what these results mean, let's first establish a few definitions.

Definition 12 (Ruth-Aaron Number)**.** Any positive integer n such that $S(n) = S(n+1)$ is a **Ruth-Aaron Number**.

Definition 13 (Density)**.** If A is a subset of the natural numbers $\mathbb{N}$ and we define $a(n)$ as the number of elements of A less than or equal to n then A has **density** α if

$$\lim_{n\to\infty} \frac{a(n)}{n} = \alpha$$

Density is essentially a measure of how frequently a set of numbers appears within the natural numbers. A good way of thinking about density of the Ruth-Aaron numbers is to draw an analogy to the prime numbers. We think of the prime numbers becoming more and more rare the further along the number line we get, and a more precise way to say that is to say that they have density zero.

Lemma 2.1. *The prime numbers have density 0.*

Proof. We consider the following limit

$$\lim_{x\to\infty} \frac{\pi(x)}{x}$$

and using the Prime Number Theorem we replace $\pi(x)$ with $x/\ln(x)$

$$\begin{aligned}\lim_{x\to\infty} \frac{\pi(x)}{x} &= \lim_{x\to\infty} \frac{1}{\ln(x)} \\ &= 0\end{aligned}$$

□

As a brief side note, let us admire the power of the Prime Number Theorem, which turns this into quite a short proof. Without it, proving this result is much longer (although not impossible).

In this respect the Ruth-Aaron numbers are similar to the prime numbers: they become less frequent as you get further along in the natural numbers. However the second big result, that the sum of reciprocals of the Ruth-Aaron Numbers converges, is different from the prime numbers. Indeed, Euler proved in 1737 that the sum of the reciprocals of prime numbers diverges, which is a strengthening of Euclid's proof that there are an infinite number of prime numbers. However, the sum of reciprocals of the Ruth-Aaron numbers does converge, and so in some ways we think of them as being "less frequent" than the prime numbers.

This analogy to the prime numbers not only helps us understand the significance of the results we have found, but also encapsulates a common strategy used in analytic number theory: take something you don't understand well and swap it out for something that you do. Recall that this is precisely what Euler's Product Formula did for us, swapping out information about the natural numbers to the primes. We will see this quite a bit in the following sections as we swap out Ruth-Aaron numbers with prime numbers (in the form of largest prime factors).

We seek to extend these previous results to more general forms of Ruth-Aaron Numbers. To do so, we consider the sums of the prime factors after they have been put through some nice arithmetic function. We considered two such alterations. The first is as follows

Definition 14. For $n = p_1^{a_1} \cdots p_d^{a_d}$ for distinct primes p_i, define a function $S_k(n)$ by

$$S_k(n) = \sum_{i=1}^{k} a_i p_i^k$$

Then we define a **K-th Power Ruth-Aaron Number** to be any number n such that $S_k(n) = S_k(n+1)$.

For the second variation we first define a new arithmetic function as follows:

Definition 15. The **Euler Totient Function** is the function $\varphi(n)$ that counts the number of positive integers up to n that are relatively prime to n.

Then we consider the following set of numbers

Definition 16. For $n = p_1^{a_1} \cdots p_d^{a_d}$ for distinct primes p_i, define a function $S_\varphi(n)$ by

$$S_\varphi(n) = \sum_{i=1}^{k} a_i \varphi(p_i)$$

Then we define a **Euler-Totient Ruth-Aaron Number** to be any number n such that $S_\varphi(n) = S_\varphi(n+1)$.

We believe that these extensions of the Ruth-Aaron number should behave similarly to that of the Ruth-Aaron numbers, namely that they have density zero and that the sum of their reciprocals converges. We are able to prove such a bound on the Euler-Totient Ruth-Aaron Numbers, but have only been able to show that the desnity of the $K-$th Power Ruth-Aaron numbers is 0. To prove these results we utilized the same process as Pomerance does in [4]. Therefore we start by working through the proof in [4] in close detail.

2.2 A Guided Tour Through Pomerance's Ruth-Aaron Numbers

Before we get into the new results, we will spend some time with [4], Pomerance's second paper on Ruth-Aaron numbers, from which our results grew out of. For someone not familiar with the field of analytic number theory the paper may seem somewhat mystifying. To give us a chance to see some tools of analytic number theory in action, we will work through this paper, pulling some of the less obvious math that is happening in the background to the forefront. This will also then allow us to streamline some of our proofs later on, that draw directly from this one. The theorem that Pomerance proves is as follows:

Theorem 2.2. *The number of integers $n \leq x$ with $S(n) = S(n+1)$ is*

$$O\left(\frac{x(\log\log x)^4}{(\log x)^2}\right)$$

Before we get into the proof we will first walk through an outline of it as well as prove some preliminary results. First, some notation. Let $P(n)$ denote the largest prime factor of n and suppose we have a Ruth-Aaron pair, i.e. that $S(n) = S(n+1)$. Then we invoke our largest prime factors that will help us throughout the proof: write $n = pk$ and $n+1 = qm$ where $p = P(n)$ and $q = P(n+1)$. We also write $k = rl$ where $r = P(k)$. Now we can take a look at the proof.

The proof is split into cases. In each case we will count the number of Ruth-Aaron numbers that exist within those constraints, and we show that that number is less than the bound in Theorem 2.2. There are quite a few cases, and they become more and more nested as we go along, and so we provide a table below which outlines them. In the second column we list what the constraints of that case are, and in the third what the bound on the number of Ruth-Aaron numbers is in that case. Observe that in each case the bound given in the final column is less than or equal to the bound stated in our theorem.

Case	Constraints	Bound
1	$p, q > x^{1/2}\log(x)$	$x/\log^2 x$
2	$S(k), S(m) < p/(\log x)^2$	$x\log\log x/(\log x)^2$
3	$S(k), S(m) > p/(\log x)^2$	
3.1	$p \leq x^{1/3}$	$x(\log\log x)^4/(\log x)^2$
3.2	$p > x^{1/3}$	
3.2.1	$p \geq x^{2/5}$	$x^{4/5+o(1)}$
3.2.2	$x^{1/3} < p < x^{2/5}$	
3.2.2a	$P(l), P(m) < x^{1/6}$	$x^{29/30}/(\log x)^2$
3.2.2b	$P(l) \geq x^{1/6}$	
3.2.2b(i)	$p < x^{1/3}(\log x)^{c+5}$	$x\log\log x/(\log x)^2$
3.2.2b(ii)	$p > x^{1/3}(\log x)^{c+5}$	$x/\log(x)^2$
3.2.2c	$P(m) \geq x^{1/6}$	$x/(\log x)^{2c+2}$

Before we get into explaining the proofs of these results, we need to prove a couple lemmas. In each case, the lemma is something that Pomerance uses in the proof, but that he provides a very short explanation or reference for. We expand on these things and prove them in their entirety. The first result is one that puts a bound on p and q. In the paper, Pomerance references a useful result from [1] as proof, but to help us understand this better we will work through the details.

Lemma 2.3. *The number of $n \leq x$ such that $p \leq x^{1/\log\log x}$ and $q \leq x^{1/\log\log x}$ is $O(x/(\log x)^2)$.*

Proof. Let $\psi(x, y)$ denote the number of positive integers not exceeding x which contain no prime factors greater than y. We wish to estimate $\psi(x, y)$ for $y = x^{1/\log\log(x)}$. Define $u := \log(x)/\log(y)$, and observe that since $\log(y) = \log(x)/\log\log(x)$ it follows that $u = \log\log(x)$. We will then use Theorem 2 from [1] which tells us that:

$$\log(\psi(x,y)) \;\leq\; \log(x\rho(u)) + 1/2\log(1+u) + O(\log\log(y)) + O\left(\frac{\log(x)^2}{y}\right) + O(R)$$

Where we define $\rho(u)$ as follows

$$\rho(u) \;=\; \exp\left[-u\left(\log(u) + \log\log(u) - 1 + \frac{\log\log(u)}{\log(u)} - \frac{1}{\log(u)} + O\left(\frac{(\log\log(u))^2}{(\log u)^2}\right)\right)\right]$$

and R is given in Section 1 of [1]. In this section de Bruijin shows that $O(R)$ is dominated by the first term, and therefore it can subsumed in the bound on $\log(x\rho(u))$ which we will find later. We want to ultimately estimate $\psi(x, y)$ as being $O(x/(\log(x))^2$ and to do so we will estimate each summand of this expression, starting with the terms in $\rho(u)$. First we will examine the last term, and observe that $\log\log(u)^2 \leq \log(u)$ since

$$\begin{aligned}(\log\log(u))^2 &= (\log\log(\log\log(x)))^2 \\ &\leq \log\log\log(x) \\ &= \log(u)\end{aligned}$$

Therefore we have that

$$\begin{aligned}O\left(\frac{\log\log(u)^2}{\log(u)^2}\right) \;&\leq\; O\left(\frac{1}{\log(u)}\right) \\ &=\; O\left(\frac{1}{\log\log\log(x)}\right)\end{aligned}$$

However $1/\log\log\log(x)$ is $O(1)$ and thus this entire term can be approximated using $O(1)$. The term $\log\log(u)/\log(u)$ becomes $\log\log\log\log(x)/\log\log(x)$ and we can estimate that this is approximately $O(1)$ as well. Similarly -1 and $-1/\log(u)$ are also $O(1)$, and we can put this altogether to get that:

$$\frac{\log\log(u)}{\log(u)} - 1 + \frac{\log\log(u)}{\log(u)} - \frac{1}{\log(u)} + O\left(\frac{\log\log(u)^2}{\log(u)^2}\right) = O(1)$$

Therefore we have that

$$\begin{aligned}\rho(u) \;&=\; \exp[-u(\log(u) + \log\log(u) + O(1)] \\ &=\; \exp[\log\log(x)[\log\log\log(x) + O(\log\log\log\log(x))]\end{aligned}$$

But for sufficiently large x we know $O(\log\log\log\log(x)$ is positive and we have

$$\begin{aligned}x\rho(u) &\leq x\,\exp(-\log\log(x)\log\log\log(x)) \\ &= \frac{x}{\exp(\log\log\log(x)\log\log(x)} \\ &= \frac{x}{\log(x)^{\log\log\log(x)}} \\ &\leq \frac{x}{\log(x)^2}\end{aligned}$$

To estimate $(1/2)\log(1+1/u)$ we use the Taylor Series approximation of $\log(u)$. We get that

$$\begin{aligned}\frac{1}{2}\log\left(1+\frac{1}{u}\right) &= \frac{1}{2}\left(\frac{1}{u}-\frac{1}{2u^2}+\dots\right)\\ &= \frac{1}{2}\left(\frac{1}{\log\log(x)}-\dots\right)\\ &\leq O\left(\frac{1}{\log\log(x)}\right)\end{aligned}$$

We then observe that $O(\log\log(y)) = O(\log\log(y)) = O(\log\log(x) - \log\log\log(x))$ We also find that

$$O\left(\frac{(\log(x)^2)}{y}\right) = O\left(\frac{\log(x)^2}{x^{1/\log\log(x)}}\right)$$

is dominated by the previous term. Therefore we have that

$$\begin{aligned}\psi(x,y) &\leq \frac{x}{\log(x)^2}+\frac{\log\log\log(x)}{2}+O\left(\frac{1}{\log\log(x)}\right)+O(\log\log(x)-\log\log\log(x))\\ &\leq O\left(\frac{x}{\log(x)^2}\right)\end{aligned}$$

□

The next lemma was established in [2], and we will recreate the proof in entirety here for clarity. We start by proving a small lemma which will be utilized later.

Lemma 2.4. *The function $t/\log(t)$ is increasing for all $t \geq e$.*

Proof. We will show that the derivative of $t/\log(t)$ is positive for all $t \geq e$. Taking the derivative yields

$$\frac{d}{dt}\left(\frac{t}{\log(t)}\right) = \frac{\log(t)-1}{(\log(t))^2}$$

Observe that $\log(t)^2$ is positive for all t. Moreover $log(t) - 1$ is positive for all t such that $log(t) > 1$, and therefore for all $t > e$. □

We now prove the result originally established in [2].

Lemma 2.5. *For all $P(n) \geq 5$ we have that*

$$P(n) \leq S(n) \leq \frac{P(n)\log n}{\log P(n)}$$

Proof. Let us write $n = \prod_{i=1}^{k} p_i^{a_i}$ where each p_i is prime and $p_i > p_j$ for all $i < j$. Since we know that $t/\log(t)$ is increasing for $t \geq e$ then for any $P(n) \geq e$ we have that

$$\begin{aligned} P(n) &\leq S(n) \\ &= \sum_{i=1}^{k} a_i p_i \\ &= \sum_{i=1}^{k} \frac{a_i p_i \log(p_i)}{\log(p_i)} \\ &\leq \sum_{i=1}^{k} \frac{a_i p_1 \log(p_i)}{\log(p_1)} \end{aligned}$$

But then we can pull out the factors of p_1 and rewrite our sum to get that

$$\begin{aligned} \sum_{i=1}^{k} \frac{a_i p_1 \log(p_i)}{\log(p_1)} &= \frac{p_1}{\log(p_1)} \sum_{i=1}^{k} a_i \log(p_i) \\ &= \frac{p_1}{\log(p_1)} \sum_{i=1}^{k} \log(p_i^{a_i}) \\ &= \frac{P(n)\log(n)}{\log(P(n))} \end{aligned}$$

□

With these results we are now ready to look at the proof of Theorem 2.2

Proof of Theorem 2.2. Pomerance begins by putting upper bounds on the largest prime factors p and q by using Lemma 2.3. The lemma tells us that the number of n such that $p \leq x^{1/\log\log x}$ is less than $O(x/\log^2 x)$. Since this is less than the bound we are trying to achieve then it is negligible and so we can assume otherwise, i.e. that

$$p > x^{1/\log\log x} \quad , \quad q > x^{1/\log\log x} \tag{2.6}$$

We also observe that but for a negligible amount of n we have that $P(n), P(n+1) \geq 5$ and so we assume that Lemma 2.5 holds for all $P(n)$.

We would like to be able to use information regarding prime factors in order to count the number of Ruth-Aaron numbers, and so Pomerance has us solve for p, q in terms of their smaller factors k, m. We know that $pk = n$ and $qm = n+1$ and therefore $pk + 1 = qm$. We also have that $S(n) = p + S(k)$ and $S(n+1) = q + S(m)$ and thus since $S(n) = S(n+1)$ then $p + S(k) = q + S(m)$. Then from these two relationships we solve for p and q:

$$p = \frac{(S(k) - S(m))m - 1}{k - m} \quad q = \frac{(S(k) - S(m))k - 1}{k - m} \tag{2.7}$$

This allows Pomerance to conclude that k, m completely determine p, q. Using this information, Pomerance finds some bounds on p, q given certain constraints on k, m.

Here we enter into our first case, referring back to the table this is Case 1. If $k, m < x^{1/2}/\log(x)$ then it follows that the number of choices for p, and hence also for n, is the number of choices for k times the number of choices of m. This is then at most $x/(\log(x))^2$. Therefore if both $k, m < x^{1/2}/\log(x)$, then the number of n is negligible, and we may assume otherwise, so we can conclude that $k > x^{1/2}/\log(x)$ or $m > x^{1/2}/\log(x)$. If $m > x^{1/2}/\log(x)$ then that means that $k \leq x^{1/2}/\log(x)$. Then we want to find a condition on p under this condition on k such that we still have that $n \leq x$. Therefore we can conclude that $p \leq x^{1/2}\log(x)$. We follow a parallel argument for the case when $k > x^{1/2}/\log(x)$ and we conclude that

$$p \ \leq \ x^{1/2}\log(x) \quad \textit{or} \quad q \ \leq \ x^{1/2}\log(x) \tag{2.8}$$

Pomerance seeks to resolve this "or" statement into an "and" statement. Suppose that $p > x^{1/2}\log(x)$. Then from Lemma 2.5 it follows that

$$\begin{aligned} p &\leq S(n) \\ &= S(n+1) \\ &\leq \frac{P(n+1)\log(n+1)}{\log(P(n+1))} \\ &= \frac{q\log(n+1)}{\log(q)} \end{aligned}$$

Then since $p > x^{1/2}\log(x)$ it follows that $q \leq x^{1/2}\log(x)$ so that

$$\begin{aligned} \frac{q\log(n+1)}{\log(q)} &\leq \frac{x^{1/2}\log(x)\log(n+1)}{\log(x^{1/2}\log(x))} \\ &= \frac{x^{1/2}\log(x)\log(n+1)}{1/2\log(x)+\log\log(x)} \\ &< \frac{x^{1/2}\log(x)\log(n+1)}{1/2\log(x)} \\ &= 2x^{1/2}\log(n+1) \\ &\leq 2x^{1/2}\log(x) \end{aligned}$$

We can do a similar procedure for the case when $q > x^{1/2}\log(x)$ and thus we have that

$$p \ < \ 2x^{1/2}\log(x) \quad \text{and} \quad q \ < \ 2x^{1/2}\log(x) \tag{2.9}$$

This concludes Case 1, and we move on to Case 2. In this case we consider when $S(k)$ and $S(m)$ are sufficiently small. Suppose that

$$S(k) < \frac{p}{(\log(x))^2} \quad , \quad S(m) < \frac{q}{(\log(x))^2} \tag{2.10}$$

Then since $p + S(k) = q + S(m)$ it follows that $p - q = S(m) - S(k)$ and thus we find that

$$\begin{aligned} |p-q| &= |S(m) - S(k)| \\ &\leq |S(m)| + |S(k)| \\ &< \left|\frac{q}{\log(x)^2}\right| + \left|\frac{p}{\log(x)^2}\right| \\ &= \frac{p+q}{\log(x)^2} \end{aligned}$$

Therefore Pomerance concludes that

$$|p-q| < \frac{p+q}{\log(x)^2} \tag{2.11}$$

Here is where we will translate all of this information gathered about the largest prime factors into information about Ruth-Aaron Numbers. Pomerances tells us that for p satisfying (2.6), the number of primes q such that (2.12) holds is at most $O(p \log\log x / \log^3 x)$ and the sum of $1/q$ for such primes q is $O(\log\log x / \log^3 x)$. He also observes that for a given choice of p, q the number of $n \leq x$ with $p|n$ and $q|(n+1)$ is at most $1 + x/pq$. Thus if (2.11) holds, then we can count the number of n that satisfy these constraints as follows:

$$\sum_{\substack{p,q \text{ subject to} \\ (2.6),(2.9),(2.11)}} \left(1 + \frac{x}{pq}\right)$$

Pomerance estimates this sum quickly, and so we will work out the details here. We first split this sum into two parts and consider

$$\sum_{\substack{p,q \text{ subject to} \\ (2.6),(2.9),(2.11)}} 1$$

Since the number of primes q that satisfy (2.11) and with p satisfying Equation (2.6) is less than $p \log\log(x)/\log(x)^3$ then we have that

$$\sum_{\substack{p,q \text{ subject to} \\ (2.6),(2.9),(2.11)}} 1 \leq \sum_{\substack{p,q \text{ subject to} \\ (2.6),(2.9),(2.11)}} \frac{p \log\log(x)}{\log(x)^3}$$

Using Theorem 1.10 we find that

$$\begin{aligned} \sum_{p < 2x^{1/2}\log(x)} p &= \frac{(2x^{1/2}\log(x))^2}{2\log(2x^{1/2}\log(x))} + O\left(\frac{(2x^{1/2}\log(x))^2}{\log(2x^{1/2}\log(x))}\right) \\ &\leq \frac{4x\log(x)^2}{2\log(x)} + O\left(\frac{4x\log(x)^2}{\log(x)}\right) \\ &= O(2x\log(x)) \end{aligned}$$

And therefore we conclude that

$$\begin{aligned}\sum_{p<2x^{1/2}\log(x)} \frac{p\log\log(x)}{\log(x)^2} &= O\left(\frac{2x\log(x)\log\log(x)}{\log(x)^3}\right)\\ &= O\left(\frac{2x\log\log(x)}{\log(x)^2}\right)\\ &= O\left(\frac{x\log\log(x)}{\log(x)^2}\right)\end{aligned}$$

Now we consider the second half of the sum and since we know from Pomerance that the sum of $1/q$ for q satisfying (2.11) is $\log\log(x)/p\log(x)^3$ then we have that

$$\sum_{\substack{p,q \text{ subject to}\\ (2.6),(2.9),(2.11)}} \frac{x}{pq} \leq \sum_{\substack{p,q \text{ subject to}\\ (2.6),(2.9),(2.11)}} \frac{x\log\log(x)}{p\log(x)^3}$$

Using Theorem 1.9 we find that

$$\begin{aligned}\sum_{p<2x^{1/2}\log(x)} \frac{1}{p} &= \log\log(2x^{1/2}\log(x))\\ &= \log\log((4x)^{1/2}) + \log\log\log(x)\\ &= \frac{\log\log(x)}{2} + \log\log\log(x)\end{aligned}$$

Therefore we have

$$\sum_{p<2x^{1/2}\log(x)} \frac{x\log\log(x)}{p\log(x)^3} \leq \frac{x\log\log(x)\log\log(x)}{2\log(x)^3} + \frac{x\log\log(x)\log\log\log(x)}{\log(x)^3}$$

but we observe

$$\begin{aligned}\frac{x\log\log(x)\log\log(x)}{2\log(x)^3} &< \frac{x\log\log(x)\log(x)}{2\log(x)^3}\\ &= \frac{x\log\log(x)}{2\log(x)^2}\\ &\leq O\left(\frac{x\log\log(x)}{\log(x)^2}\right)\end{aligned}$$

and for the second half we have

$$\frac{x\log\log(x)\log\log\log(x)}{\log(x)^3} \leq O\left(\frac{x\log\log(x)}{\log(x)^2}\right)$$

Therefore we have shown that

$$\sum_{p<2x^{1/2}\log(x)} \frac{x\log\log(x)}{p\log(x)^3} \leq O\left(\frac{x\log\log(x)}{\log(x)^2}\right)$$

Thus we conclude that

$$\sum_{\substack{p,q \text{ subject to} \\ (2.6),(2.9),(2.11)}} \left(1 + \frac{x}{pq}\right) \ll \frac{x \log\log(x)^4}{(\log(x)^2}$$

We have found that the number of n that satisfy the conditions in (2.10) is sufficiently small and so we can assume that (2.10) does not hold. This concludes the second case in the earlier table, and so we move on to Case 3, which will further split into subcases. Since the arguments for $S(k)$ and $S(m)$ are parallel, Pomerance only gives the details for the first case, so we assume that

$$S(k) \geq \frac{p}{\log(x)^2} \tag{2.12}$$

We write $k = rl$ where $r = P(k)$. Here, Pomerance provides bounds on q and p. We obtain these bounds using Lemma 2.5:

$$\begin{aligned} q &= P(n+1) \\ &\leq S(n+1) \\ &= S(n) \\ &\leq \frac{P(n)\log(n)}{\log(P(n))} \\ &= \frac{p\log(n)}{\log(p)} \\ &\leq \frac{p\log(x)}{\log(p)} \end{aligned}$$

and similarly we find that

$$\begin{aligned} p &= P(n) \\ &\leq S(n) \\ &= S(n+1) \\ &\leq \frac{P(n+1)\log(n+1)}{\log(P(n+1))} \\ &= \frac{q\log(n+1)}{\log(q)} \\ &\leq \frac{q\log(x+1)}{\log(q)} \end{aligned}$$

Therefore we conclude that

$$q \leq \frac{p\log(x)}{\log(p)} \quad , \quad p \leq \frac{q\log(x+1)}{\log(q)} \tag{2.13}$$

Now, Pomerance tells us that Equation (2.13) implies that $\log(q) \geq 2/3 \log p$ which gives $q \geq (1/2)p\log(p)/\log(x)$. Since these claims take a bit of proof, we will state and prove them as a lemma.

Lemma 2.14. $\log(q) \geq \frac{2}{3}\log(p)$ *and* $q \geq \frac{p\log(p)}{2\log(x)}$

Proof. We use the second inequality from (2.13) to find that

$$\log(p) \;\leq\; \log(q) + \log\log(x+1) - \log\log(q)$$

On the other hand we also know that $q > x^{1/\log\log(x)}$ which implies that $\log\log\log(x) > \log\log(x) - \log\log(q)$. Putting these together we find that $\log(p) \leq \log(q) + \log\log(x+1) - \log\log(q)$, and this implies that

$$\log(p) - \log\log(x) \;\leq\; \log(q)$$

We now want to show that $2/3\log(p) \leq \log(p) - \log\log(x)$. We observe that for sufficiently large x we have that

$$\begin{aligned} 3\log\log\log(x)\log\log(x) \;&<\; 3\log\log(x)\log\log(x) \\ &<\; \log(x) \end{aligned}$$

and therefore $\log(x) > 3\log\log\log(x)\log\log(x)$. This implies that $1/3\log(p) - \log\log\log(x) > 0$, and therefore $2/3\log(p) + (1/3\log(p) - \log\log\log(x)) \geq 2/3\log(p)$ and

$$\frac{2}{3}\log(p) \;\leq\; \log(p) - \log\log(x)$$

Therefore we have the first inequality. For the second, we use the second inequality from (2.13) to get that

$$\frac{p\log(p)}{\log(x+1)}\frac{\log(q)}{\log(p)} \;\leq\; q$$

and using our previous result we have

$$\frac{\log(q)}{\log(p)} \;\geq\; \frac{2}{3}$$

Then we also find that

$$\begin{aligned} \frac{p\log(p)}{\log(x+1)}\frac{2}{3} \;&\geq\; \frac{1}{2}\frac{p\log(p)}{\log(x+1)} \\ &\geq\; \frac{p\log(p)}{\log(x)} \end{aligned}$$

Putting all of these together we get

$$\begin{aligned} q \;&\geq\; \frac{p\log(p)}{\log(x+1)}\frac{\log(q)}{\log(p)} \\ &\geq\; \frac{2}{3}\frac{p\log(p)}{\log(x+1)} \\ &\geq\; \frac{1}{2}\frac{p\log(p)}{\log(x)} \end{aligned}$$

□

This then allows us to conclude with the next bound given by Pomerance:

$$\frac{p\log(p)}{2\log(x)} \;\le\; q \;\le\; \frac{p\log(x)}{\log(p)} \tag{2.15}$$

Pomerance gives a similar bound for r. We know from Lemma 2.5 that

$$\frac{r\log(k)}{\log(r)} \;\ge\; \frac{p}{\log(x)^2}$$

Which implies that

$$\frac{p\log(p)}{2(\log(x))^3} \;\le\; r \;\le\; p \tag{2.16}$$

Then Pomerance tells us, similar to before, that for a given choice of p, r, q the number of $n \le x$ with $pr|n$ and $q|(n+1)$ is at most $1 + x/(prq)$. Pomerance splits this up into another case, labelled Case 3.1 in the table, supposing $p \le x^{1/3}$. This will allow us to evaluate our sum more easily, and we will deal with the other cases later. The number of n in this case is then at most

$$\sum_{\substack{p,q,r \text{ subject to } (2.6),(2.15),(2.16)\\ p\le x^{1/3}}} 1 + \frac{x}{prq}$$

Pomerance provides some bounds on this sum, saying that

$$\begin{aligned}\sum_{\substack{p,q,r \text{ subject to } (2.6),(2.15),(2.16)\\ p\le x^{1/3}}} 1 + \frac{x}{prq} \;&\ll\; \frac{x}{\log^3 x} + \sum_{p > x^{1/\log\log x}} \frac{x}{p}\frac{\log\log x}{\log p}\frac{\log\log x}{\log p} \\ &\ll\; \frac{x(\log\log x)^4}{(\log x)^2}.\end{aligned}$$

To see how he got this, we first observe that we can start by splitting this up into two sums, of 1 and x/prq. To bound the first, we will use the Prime Number Theorem. Since $p, q, r \le x^{1/3}$ then we have that

$$\begin{aligned}\sum_{p,q,r\le x^{1/2}} 1 &= \left(\frac{x^{1/3}}{\log(x^{1/3})}\right)^3 \\ &\ll \frac{x^3}{\log^2(x)}\end{aligned}$$

For the second summation, we want to turn the sum over p, q and r into just a sum over p. To do so we consider this as a triple sum, and consider each individually. Firstly the sum of $1/q$. From (2.15) we have upper and lower bounds on q, and so we can use the fact that the sum over the reciprocal of primes is $\log\log(x)$. Using this we get that

$$\sum_{q \text{ subject to } (2.15)} \frac{1}{q} \le \log\log\left(\frac{p\log(x)}{\log(p)}\right) - \log\log\left(\frac{p\log(p)}{2\log(x)}\right)$$

which Pomerance simplifies to being simply less than $\log\log(x)/\log(p)$. Similarly, for r we find using (2.16) that

$$\sum_{r \text{ subject to (2.16)}} \frac{1}{r} \leq \log\log(p) - \log\log\left(\frac{p\log(p)}{2\log^3(x)}\right)$$

which Pomerance bounds similarly. Now we can turn our sum into a sum over p only, considering

$$\sum_{p \geq x^{1/\log\log(x)}} \frac{x}{p}\frac{\log\log(x)}{\log(p)}\frac{\log\log(x)}{\log(p)}$$

But this is essentially negligible. As x gets bigger the sum is over fewer p, and the sum of $1/p\log(p)^2$ is certainly convergent. Therefore for sufficiently large x we can make this sum as small as we want. We conclude that the number of Ruth-Aaron numbers for $p \leq x^{1/3}$ is less than $O(x(\log\log x)^4/(\log x)^2)$, and thus Case 3.1 is complete.

Now, for Case 3.2 we assume $p > x^{1/3}$ (this will later be split into further subcases). Here, Pomerance finds a useful relationship between our prime factors. From Equation (2.15) we have that

$$\frac{p}{6} \leq q \leq 3p$$

and from (2.16) that

$$\frac{p}{6\log(x)^2} \leq r \leq p$$

Using (2.7) and substituting $k = rl$ we find that

$$p(rl - m) = (r + S(l) - S(m))m - 1$$

Then we pull out the r from the right hand side, move it to the left hand side, and multiply the entire equation by l to get that

$$pl(rl - m) - rml = (S(l) - s(m))ml - l$$

and then we can factor on the left hand side to find that

$$(pl - m)(rl - m) = (S(l) - S(m))ml - l + m^2 \tag{2.17}$$

Pomerance uses the divisor function to estimate the number of choices we have for r. The divisor function is denoted $\tau(x)$ and it counts the number of divisors of some integer x. Since $rl - m$ is a factor of the left hand side of (2.17) then anytime there is a divisor of the right hand side, that corresponds with one possible option for r. Not all divisors will result in a value for r, but we can still use this as an upper bound. So we have that the number of choices of r (and therefore for n), is at most

$$\tau((S(l) - S(m))ml - l + m^2)$$

It is known that $\tau(n) \leq n^{o(1)}$ and so for sufficiently large x we have that $(S(l) - S(m))ml - l + m^2 \leq x$ and therefore the number of choices for r is

$$\tau((S(l) - S(m))ml - l + m^2) \leq x^{o(1)}$$

Pomerance now splits this into a subcase, i.e Case 3.2.1 in the table, when $p \geq x^{2/5}$. We hope to show that the number of n in this case is sufficiently small so we can discount it and bound p from both above and below. We have that $l \leq x/(pr)$ and from above that $r \geq p/(6\log(x)^2)$ and therefore we conclude that $l \leq x\log(x)^2/p^2$. Similarly we have that $m = (n+1)/q$ and that $p/6 < q$ so that $m \leq x/p$. We then conclude that the number of choices for n is at most $x^{4/5+o(1)}$. Hence we can assume that

$$x^{1/3} < p < x^{2/5}. \tag{2.18}$$

This assumption begins Case 3.2.2 which we will further split up into subcases. We now consider Case 3.2.2a where Pomerance assumes that

$$P(l) < x^{1/6} \quad , \quad P(m) < x^{1/6} \tag{2.19}$$

Here, Pomerance does some quick estimations of these values so that we can assume the above equation in fact doesn't hold. He tells us that Equation (2.19) implies $p + r = q + O(x^{1/6})$. This is because $p + r = q + S(m) - S(l)$ and since $S(m)$ and $S(l)$ are the sums of the smaller prime factors they are mostly dominated by $P(l)$ and $P(m)$. Given p, r it follows that the number of choices for q is $O(x^{1/6})$. But the number of choices for p, r with $r \leq p$ and (2.18) holding is $O(x^{4/5})$. Thus the number of triples p, q, r is $O(x^{29/30})$. But $prq \gg x/(\log(x))^2$, so the number of choices for n given p, r, q is $O(\log(x)^2)$. It follows that but for $O(x^{29/20}(\log(x)^2)$ choices for $n \leq x$ we have that (2.19) does not hold. This concludes Case 3.2.2a.

Pomerance now considers Case 3.3.3b, i.e that $P(l) \geq x^{1/6}$. Write $l = sj$ where $s = P(l)$. He rewrites (2.17) as

$$(psj - m)(rsj - m) = mjs^2 + ((S(j) - S(m))mj - j)s + m^2 \tag{2.20}$$

We want to fix a choice for j, m and sum over choices for s which will help us count the number of n in this case. Here, Pomerance proves a lemma that he uses throughout the rest of the proof.

Lemma 2.21. *Suppose A, B, C are integers with $gcd(A, B, C) = 1, D := B^2 - 4AC \neq 0, A \neq 0$. Suppose the maximum value of $|At^2 + Bt + C|$ on the interval $[1, x]$ is M_0. Let $M = max\{M_0, |D|, x\}$, let $\mu = \lceil \log(M)/\log(x) \rceil$ and assume that $\mu \leq (1/7)\log\log(x)$. Then*

$$\sum_{n \leq x} \tau(|An^2 + Bn + C|) \leq x(\log(x))^{2^{3\mu+1}+4}$$

holds uniformly for $x \geq x_0$. (We interpret $\tau(0)$ as 0. The number x_0 is an absolute constant independent of the choice of A, B, C.)

Pomerance applies the lemma with $A = mj$, $B = (f(j) - f(m) - 2)mj - j$ and $C = m^2$. We need to check that these satisfy the constraints of the lemma, Pomerance supplies some explanation and we will simply flesh out a few different aspects of it.

Since j comprises the small prime factors of n and m comprises the small prime factors of $n+1$, we have that $\gcd(j,m) = 1$. Then clearly $\gcd(A,C) = m$, and $\gcd(A,B,C) = \gcd(B,(A,C)) = \gcd(B,m)$. Suppose for the sake of contradiction that α is a divisor of m such that $a|B$. Then $\alpha k = (S(j) - S(m))mj - j$ for some $k \in \mathbb{Z}$. Then $k = (S(j) - S(m))j(m/\alpha) - j/\alpha$, but $j/\alpha \notin \mathbb{Z}$, a contradiction. Thus $\gcd(A,B,C) = 1$. Note that $4AC = 4m^3j$, that $j^2|B^2$ and that $B^2 \equiv j^2 \mod m$. Thus if $D = 0$, then $B^2 = 4m^3j$. Therefore since $j^2|B^2$ then $j|4m^3$ and thus $j|4$. Also $4m^3j \equiv j^2 \mod m$, and thus $j^2 \equiv 0 \mod m$. Therefore $m = 1$. Then $j = 2$ or $j = 4$. This gives the triples $2,2,1$ or $4,6,1$ for A,B,C and neither choice has $D = 0$. Thus $D \neq 0$. Pomerance also assumes further that $j < 6x^{1/6}(\log x)^2, m \ll x^{2/3}$, and $s \leq 6x^{1/3}\log x)^2/j$ which gives that the maximum of $|As^2 + Bs + C|$ for the range of s is $\ll x^{4/3}(\log x)^2$. Then using the lemma we find that

$$\sum_{s \leq 6x^{1/3}(\log x)^2/j} \tau(|As^2 + Bs + C|) \leq \left(\frac{1}{j}\right) x^{1/3}(\log x)^c \tag{2.22}$$

for some positive constant c.

We now split into Case 3.2.2b(i), i.e we suppose that $x^{1/3} < p \leq x^{1/3}(\log x)^{c+5}$, in which case Pomerance tells us that the number of n is at most

$$\sum_{p \asymp q} \left(1 + \frac{x}{pq}\right) \ll x^{2/3}(\log x)^{2c+10} + \frac{x}{\log x}\sum \frac{1}{p} \ll \frac{x \log\log x}{(\log x)^2}.$$

Thus, we move into Case 3.2.2b(ii) in the table, i.e supposing that $p > x^{1/3}(\log x)^{c+5}$. Then $m \ll x^{2/3}/(\log x)^{c+5}$, so that summing (2.22) over all choices for m, j we get a quantity that is $\ll x/(\log x)^2$, which is negligible.

Finally, Pomerance considers the remaining case in (2.19), i.e. Case 3.2.2c in the table, when $P(m) \geq x^{1/6}$. Let $m = tu$ where $t = P(m)$. Then, from (2.17), we get

$$\begin{aligned}(pl - m)(rl - m) &= (S(l) - S(m))ml - l + m^2 \\ &= (S(l) - t - S(u))(tu)l - l + (tu)^2 \\ &= t^2(u^2 - ul) + t(S(l)ul - S(u)ul) - l,\end{aligned}$$

and thus

$$(pl - tu)(rl - tu) = t^2(u^2 - ul) + t(ulS(l) - ulS(u)) - l. \tag{2.23}$$

Pomerance then applies the lemma once again, this time to the quadratic polynomial with $A = u^2 - ul, B = ul(f(l) - f(u)), C = -l$. As before, we may assume that $p \geq x^{1/3}(\log x)^{c+5}$ so that $l \leq 6x^{1/3}/(\log x)^{2c+3}$. We have $u \ll x^{1/2}$, and $t \ll (1/u)x^{2/3}$. Summing the number of divisors of the right side of (2.23) for t, u, l ranging as stated, we get an estimate that is $\ll x/(\log x)^{2c+2}$, which is negligible. This completes the proof.

□

Chapter 3

New Results

In this section we will extend the results about Ruth-Aaron numbers to the Euler-Totient Ruth-Aaron Numbers and the k-th Power Ruth-Aaron Numbers. For the Euler-Totient Ruth-Aaron numbers we were able to achieve the bound in [4] using similar methods. For the k-th power Ruth-Aaron numbers we are able to show that the density is 0 but we cannot yet prove that the sum of reciprocal converges. Instead, we discuss some attempts at proving this and provide support for the conjecture that such a bound is possible.

3.1 Euler-Totient Ruth-Aaron Numbers

We show in this section that the sum of the reciprocals of Euler-Totient Ruth-Aaron Numbers is bounded. This is not a hugely surprising result, as the Euler-Totient function has a very small effect on prime factors. Therefore we use the same methods as in [4], which we provided a reading guide for above. As such, we move through this proof at a slightly faster pace, particularly in places where it is analogous to the proof in [4]. The strategy is the same: consider the largest prime factors of n and $n+1$, and working in cases gather information about them that can be translated into information about n and $n+1$ themselves.

Theorem 3.1. *The number of integers $n \le x$ with $S_\varphi(n) = S_\varphi(n+1)$ is*

$$O\left(\frac{x(\log\log x)^4}{(\log x)^2}\right)$$

In particular, the sum of the reciprocals of the Euler-Totient Ruth-Aaron numbers is bounded.

Proof. Let $P(n)$ denote the largest prime factor of n. Say $n \le x$ and $S_\varphi(n) = S_\varphi(n+1)$. Write $n = pk$ and $n+1 = qm$ with $p = P(n)$ and $q = P(n+1)$.

We first note that from 2.3 we may assume that

$$p > x^{1/\log\log x} \quad \text{and} \quad q > x^{1/\log\log x} \tag{3.2}$$

We establish a useful result for $P(n) \ge 5$. We use Lemma 2.4 which tells us that $t/\log t$ is increasing for $t \ge e$ and the fact that $2/\log 2 < 5/\log 5$. Observe that

$$\varphi(p_i) \;=\; p_i - 1 \;\le\; P(n) - 1 \;=\; \varphi(P(n))$$

for all distinct primes p_i. Additionally $\varphi(P(n)) = P(n) - 1 \le P(n)$. Thus

$$\frac{\varphi(p_i)}{\log \varphi(p_i)} \le \frac{P(n)}{\log P(n)}$$

Then we have

$$\begin{aligned} S_\varphi(n) &= \sum_{i=1}^{d} a_i \varphi(p_i) \\ &= \sum_{i=1}^{d} \frac{a_i \varphi(p_i) \log \varphi(p_i)}{\log \varphi(p_i)} \\ &\le \sum_{i=1}^{d} \frac{a_i p_1 \log \varphi(p_i)}{\log p_1} \\ &= \frac{p_1}{\log p_1} \sum_{i=1}^{d} a_i \log \varphi(p_i) \end{aligned}$$

Observe that

$$\begin{aligned} \varphi(p_1)^{a_1} \cdots \varphi(p_d)^{a_d} &= (p_1 - 1)^{a_1} \cdots (p_d - 1)^{a_d} \\ &\le p_1^{a_1} \cdots p_d^{a_d} \\ &= n \end{aligned}$$

Thus finally

$$\begin{aligned} \frac{p_1}{\log p_1} \sum_{i=1}^{d} a_i \log \varphi(p_i) &\le \frac{p_1 \log n}{\log p_1} \\ &= \frac{P(n) \log n}{\log P(n)} \end{aligned}$$

Therefore we have the following result

$$P(n) \le S_\varphi(n) \le \frac{P(n) \log n}{\log P(n)} \tag{3.3}$$

In light of (3.2), we may assume $P(n), P(n+1) \ge 5$, so that (3.3) holds for n and $n+1$.

We next note that the numbers k, m determine the primes p, q. Since $n = pk$ and $n + 1 = qm$ we have that $pk + 1 = qm$. Additionally from $S_\varphi(n) = S_\varphi(n+1)$ we get $\varphi(p) + S_\varphi(k) = \varphi(q) + S_\varphi(m)$ and since $\varphi(p) = p - 1$ and $\varphi(q) = q - 1$. Therefore we have the following two equations

$$pk + 1 = qm \quad , \quad p + S_\varphi(k) = q + S_\varphi(m)$$

Therefore by substituting and solving for p and q we get that

$$p = \frac{(S_\varphi(k) - S_\varphi(m))m - 1}{k - m} \quad , \quad q = \frac{(S_\varphi(k) - S_\varphi(m))k - 1}{k - m} \tag{3.4}$$

Thus, the number of choices for n corresponding to choices of k, m with $k, m < x^{1/2}/\log x$ is at most $x/(\log x)^2$. We hence may assume that

$$p \leq x^{1/2} \log x \quad \text{or} \quad q \leq x^{1/2} \log x \tag{3.5}$$

Suppose $p > x^{1/2} \log x$. Then (3.3) and (3.5) imply that

$$\begin{aligned} p \;\leq\; S_\varphi(n) \;=\; S_\varphi(n+1) \;&\leq\; \frac{q \log(n+1)}{\log q} \\ &\leq\; \frac{x^{1/2} \log x \log(n+1)}{\log(x^{1/2} \log x)} \\ &=\; \frac{2x^{1/2} \log x \log(n+1)}{\log x + \log\log x} \\ &\leq\; \frac{2x^{1/2} \log x \log(n+1)}{\log x} \\ &\leq\; 2x^{1/2} \log x \end{aligned}$$

A similar inequality holds if $q > x^{1/2} \log x$. We conclude that

$$p < 2x^{1/2} \log x \quad \text{and} \quad q < 2x^{1/2} \log x \tag{3.6}$$

Suppose that

$$S_\varphi(k) < \frac{p}{(\log x)^2} \quad , \quad S_\varphi(m) < \frac{q}{(\log x)^2} \tag{3.7}$$

Then since $p + S_\varphi(k) = q + S_\varphi(m)$, we have

$$\begin{aligned} |p - q| \;&=\; |f(m) - f(k)| \\ &\leq\; |f(m)| + |f(k)| \\ &<\; \frac{p}{(\log x)^2} + \frac{q}{(\log x)^2} \end{aligned}$$

Therefore we have

$$|p - q| < \frac{p + q}{(\log x)^2} \tag{3.8}$$

For p satisfying (3.2), the number of primes q such that (3.8) holds is $O(p \log\log x/(\log x)^3)$ and the sum of $1/q$ for such primes q is $O(\log\log x/(\log x)^3)$. Now, for a given choice of p, q the number of $n \leq x$ with $p|n$ and $q|(n+1)$ is at most $1 + x/(pq)$. Thus if (3.7) holds, the number of n that we are counting is at most

$$\begin{aligned} \sum_{p,q \text{ subject to } (3.2),(3.6),(3.8)} 1 + \frac{x}{pq} \;&\ll\; \sum_{p < 2x^{1/2} \log x} \frac{p \log\log x}{\log^3 x} + \frac{x \log\log x}{p(\log^3 x)} \\ &\ll\; \frac{x \log\log x}{\log^2 x} \end{aligned}$$

We thus may assume that (3.7) does not hold. The arguments for the cases $S_\varphi(k) \geq p/(\log x)^2$ and $S_\varphi(m) \geq q/(\log x)^2$ are parallel, so we shall only give the details for the first case. That is, we shall assume that

$$S_\varphi(k) \;\geq\; \frac{p}{(\log x)^2} \tag{3.9}$$

Write $k = rl$ where $r = P(k)$. As in the proof of (3.6), the inequality (3.3) gives us

$$q \leq p\frac{\log x}{\log p} \quad , \quad p \leq q\frac{\log x + 1}{\log q}$$

The second inequality and (3.2) imply that $\log q \geq 2/3 \log p$, so that $q \geq p \log p/2 \log x$. That is, we have

$$p\frac{\log p}{2\log x} \;\leq\; q \;\leq\; p\frac{\log x}{\log p} \tag{3.10}$$

Similarly, (3.9) gives us

$$\frac{p\log p}{2(\log x)^3} \;\leq\; r \;\leq\; p. \tag{3.11}$$

For a given choice of p, r, q Pomerance tells us that the number of $n \leq x$ with $pr|n$ and $q|n+1$ is at most $1 + x/prq$. Now, we suppose $p \leq x^{1/3}$. The number of possible n in this case is following sum:

$$\sum_{\substack{p,q,r \text{ subject to } (3.2),(3.9),(3.11) \\ p \leq x^{1/3}}} 1 + \frac{x}{prq} \;\ll\; \frac{x}{\log^3 x} + \sum_{p > x^{1/\log\log x}} \frac{x}{p}\frac{\log\log x}{\log p}\frac{\log\log x}{\log p}$$
$$\ll\; \frac{x(\log\log x)^4}{(\log x)^2}$$

Thus, we may assume that $p > x^{1/3}$. It follows from (3.10) that $p/6 < q < 3p$, and it follows from (3.11) that $p/6(\log x)^2 \leq r \leq p$.

Using (3.4), we have that

$$p(k - m) \;=\; p(rl - m) \;=\; (S_\varphi(k) - S_\varphi(m))m - 1$$

and since $k = rl$ we also have

$$S_\varphi(k) \;=\; S_\varphi(rl) \;=\; S_\varphi(r) + S_\varphi(l) \;=\; \varphi(r) + f(l) = r - 1 + f(l)$$

Hence

$$p(rl - m)l - rml \;=\; (S_\varphi(l) - S_\varphi(m) - 1)ml - l$$

and we conclude that

$$(pl - m)(rl - m) \;=\; (S_\varphi(l) - S_\varphi(m) - 1)ml - l + m^2 \tag{3.12}$$

Thus, given l, m the number of choices of r, and hence for n, is at most

$$\tau((f(l) - f(m) - 1)ml - l + m^2) \leq x^{o(1)}$$

where τ denotes the divisor function. Suppose that $p \geq x^{2/5}$. Since $l \leq x/pr \ll x(\log x)^2/p^2$ and $m \ll x/p$, we conclude that the number of choices for n is at most $x^{4/5+o(1)}$. Hence we may assume that

$$x^{1/3} < p < x^{2/5} \tag{3.13}$$

Suppose that

$$P(l) < x^{1/6} \quad , \quad P(m) < x^{1/6} \tag{3.14}$$

Then $p + r = q + O(x^{1/6})$. Given p, r it follows that the number of choices for q is $O(x^{1/6})$. But the number of choices for p, r with $r \leq p$ and (3.13) holding is $O(x^{4/5})$. Thus the number of triples p, q, r is $O(x^{29/30})$. But $prq \gg x/(\log x)^2$, so the number of choices for n given p, r, q is $O((\log x)^2)$. It follows that but for $O(x^{29/30}(\log x)^2)$ choices for $n \leq x$ we have that (3.14) does not hold.

We first consider the case that $P(l) \geq x^{1/6}$. Write $l = sj$ where $s = P(l)$. We rewrite (3.12) as

$$(psj - m)(rsj - m) \;=\; ((S_\varphi(j) - S_\varphi(m) - 2)mj - j)s + m^2 + mjs^2 \tag{3.15}$$

We shall fix a choice for j, m and sum over choices for s.

We now apply Lemma 2.21 with $A = mj$, $B = (S_\varphi(j) - S_\varphi(m) - 2)mj - j$ and $C = m^2$. Since j comprises the small prime factors of n and m comprises the small prime factors of $n+1$, we have that $\gcd(j, m) = 1$ and so $\gcd(A, B, C) = 1$. Similar to in the previous section we have $D \neq 0$. We then assume that $j < 6x^{1/6}(\log(x))^2, m \ll x^{2/3}$, and $s \leq 6x^{1/3}(\log(x))^2/j$, we have that the maximum of $|As^2 + Bs + C|$ for the range of s is $\ll x^{4/3}(\log(x))^2$. It follows from the lemma that

$$\sum_{s \leq 6x^{1/3}(\log x)^2/j} \tau(|As^2 + Bs + C|) \;\leq\; \left(\frac{1}{j}\right) x^{1/3}(\log x)^c \tag{3.16}$$

for some positive constant c.

If $x^{1/3} < p \leq x^{1/3}(\log x)^{c+5}$, the number of n in this case is at most

$$\sum_{p \asymp q} \left(1 + \frac{x}{pq}\right) \;\ll\; x^{2/3}(\log x)^{2c+10} + \frac{x}{\log x} \sum \frac{1}{p} \;\ll\; \frac{x \log\log x}{(\log x)^2}$$

Thus, we may assume that $p > x^{1/3}(\log x)^{c+5}$. Then $m \ll x^{2/3}/(\log x)^{c+5}$, so that summing (3.16) over all choices for m, j we get a quantity that is $\ll x/(\log x)^2$, which is negligible.

Finally, we consider the remaining case in (3.14) when $P(m) \geq x^{1/6}$. Let $m = tu$ where $t = P(m)$. Then, from (3.12), we obtain

$$\begin{aligned}(pl - m)(rl - m) \;&=\; (S_\varphi(l) - S_\varphi(m) - 1)ml - l + m^2 \\ &=\; (S_\varphi(l) - t + 1 - S_\varphi(u) - 1)(tu)l - l + (tu)^2 \\ &=\; t^2(u^2 - ul) + t(S_\varphi(l)ul - S_\varphi(u)ul) - l\end{aligned}$$

and thus

$$(pl - tu)(rl - tu) = t^2(u^2 - ul) + t(ulS_\varphi(l) - ulS_\varphi(u)) - l \tag{3.17}$$

We apply the Lemma to the quadratic polynomial with $A = u^2 - ul, B = ul(S_\varphi(l) - S_\varphi(u)), C = -l$. As before, we may assume that $p \geq x^{1/3}(\log x)^{c+5}$ so that $l \leq 6x^{1/3}/(\log x)^{2c+3}$. We have $u \ll x^{1/2}$, and $t \ll (1/u)x^{2/3}$. Summing the number of divisors of the right side of (3.17) for t, u, l ranging as stated, we get an estimate that is $\ll x/(\log x)^{2c+2}$, which is negligible. This completes the proof. □

This theorem then finally gives us the promised result regarding the sum of reciprocals.

Corollary 3.18. *The sum of reciprocals of the Euler-Totient Ruth-Aaron Numbers is bounded.*

3.2 K-th Power Ruth-Aaron Numbers

We now move to considering the kth Power Ruth-Aaron Numbers. We first establish that they have density 0, using methods from [2]. Later we will move to a discussion on finding a tighter bound as in [4].

3.2.1 Density

The proof presented here is a generalization of the result in [2], and follows the same strategy. We start by considering how close $P(n)$ and $P(n+1)$ tend to be. To do this we use a result proved in [2] that tells us that in general, they are fairly far apart. We then consider how the largest prime factors relate to $S_k(n)$ and $S_k(n+1)$ themselves. Essentially, we argue that the largest prime factors are dominating enough for us to use them as a good approximation for $S_k(n)$ and $S_k(n+1)$. Then from our first result this allows us to conclude that k-th Power Ruth-Aaron Numbers are indeed quite rare.

Theorem 3.19. *The k-th Power Ruth-Aaron Numbers have density zero.*

Proof. From [2] we have the following theorem which demonstrates the relationship between the largest prime factors of n and $n+1$ which we will use to analyze the density of the k-th Power Ruth-Aaron Numbers.

Theorem 3.20. *For all $\varepsilon > 0$ there is a $\delta > 0$ such that for sufficiently large x, the number of $n \leq x$ with*

$$\frac{1}{x^\delta} < \frac{P(n)}{P(n+1)} < x^\delta \tag{3.21}$$

is less than εx.

Theorem 2 in [2] is the $k = 1$ case, and so we will establish an analagous theorem for the general k case.

Theorem 3.22. *For all $\varepsilon > 0$ there exists a $\delta > 0$ such that for sufficiently large x there are at least $(1-\varepsilon)x$ choices for $n \geq x$ such that*

$$P(n)^k \;<\; S_k(n) \;<\; (1+x^{-\delta})P(n)^k \tag{3.23}$$

Proof. Since any integer $n \leq x$ is divisible by at most $\log x/\log 2$ primes, we have for large x and composite $n \leq x$

$$\begin{aligned} S_k(n) &= P(n)^k + f_k\left(\frac{n}{P(n)}\right)^k \\ &\leq P(n)^k + P\left(\frac{n}{P(n)}\right)^k \frac{\log x}{\log 2} \\ &< P(n)^k + P\left(\frac{n}{P(n)}\right)^k x^{\delta} \end{aligned} \tag{3.24}$$

If Theorem 3.22 fails, then, other than $o(x)$ choices of $n \leq x$ we have

$$S_k(n) \;\geq\; (1+x^{-\delta})P(n)^k \tag{3.25}$$

and thus it follows that

$$P\left(\frac{n}{P(n)}\right)^k \;>\; \frac{P(n)^k}{x^{k\delta}} \tag{3.26}$$

Now let $\varepsilon > 0$. From Theorem A of Erdős and Pomerance there is a $\delta_0 = \delta_0(\varepsilon) > 0$ such that for large x, the number of $n \leq x$ with $P(n) < x^{\delta_0}$ is at most $\varepsilon x/3$. For each pair of primes p, q the number of $n \leq x$ with $P(n)^k = p^k$ and $P(n/P(n))^k = q^k$ is at most $[x/pq]$. Hence for large x the number of $n \leq x$ for which Theorem 3.22 fails is at most

$$\begin{aligned} o(x) + \frac{\varepsilon x}{3} + \sum_{\substack{x^{\delta_0} \leq p \\ x^{-2\delta} p < q \leq p}} \left[\frac{x}{pq}\right] &< \frac{\varepsilon x}{2} + x \sum_{\substack{x^{\delta_0} \leq p \\ x^{-2\delta} p < q \leq p}} \frac{1}{p}\frac{1}{q} \\ &< \frac{\varepsilon x}{2} + \frac{4\delta x}{\delta_0} \end{aligned} \tag{3.27}$$

taking $\delta = \delta_0 \varepsilon/8$ completes the proof.

□

As Theorem 3.20 implies that $P(n)$ and $P(n+1)$ are usually not close, but Theorem 3.22 suggests that for most n we have $S_k(n) \approx P(n)^k$ and $S_k(n+1) \approx P(n+1)^k$. Thus $S_k(n)$ and $S_k(n+1)$ are usually not very close, and in particular we usually have $S_k(n) \neq S_k(n+1)$. Therefore we have established that the density of the k-th Power Ruth-Aaron Numbers is 0.

□

3.2.2 Sum of Reciprocals

The sum of reciprocals of the k-th Power Ruth-Aaron Numbers turned out to be more challenging than the previous results. The methods in [4] do not easily extend to the k-th power case. The k-th Power Ruth-Aaron Numbers do not satisfy a nice linear relationship between their largest prime factors and smaller prime factors the way the Ruth-Aaron Numbers do. In particular, Pomerance uses (2.7) repeatedly. This result, however, does not hold for the k-th power case, and in fact gets increasingly more complicated the higher k is. While this does preclude a good generalization of the methods previously used, we do not think this suggests that such a bound cannot be found.

The intuition for a tighter bound on the k-th Power Ruth-Aaron Numbers comes from considering the largest prime factor again. With Ruth-Aaron Numbers, in general we consider the largest prime factor because it is often the "dominating term," i.e. it is often large enough (in comparison to the smaller prime factors) to be a good estimate for $S(n)$. In the case of the k-th Power Ruth-Aaron Numbers, this is in some ways even more true. The bigger k is, the more the difference between the largest prime factors and the lower prime factors, and the more $S_k(n)$ is dominated by $P(n)$. Since it is uncommon for the largest prime factors of n and $n+1$ to be very close, this would then suggest that the same is true for $S_k(n)$ and $S_k(n+1)$.

Our first attempt to support this with proof was using the same methods we previously did. This works up to a point. In particular we can show that if n is a k-th Power Ruth-Aaron Number and $P(n) = p$ then the number of such n where $p \leq x^{1/3}$ is sufficiently small. The only real concern in this part is the above observation about non-linearity. However for these results we actually only need that the smaller prime factors determine the larger prime factors, which holds regardless of linearity. From there we can do a straightforward extension of the proof from Pomerance in [4] for $p \leq x^{1/3}$. For larger p however, Pomerance makes an argument regarding divisors which requires the elusive linearity from above.

For these reasons, we considered new ways that we could analyze the k-th Power Ruth-Aaron Numbers. Because we are looking at the largest prime factor, we decided to think a little bit about situations where the largest prime factors of n and $n+1$ are more likely to be very different. Our first thought was that if n and $n+1$ have the same number of prime factors, we would except $P(n)$ and $P(n+1)$ to be quite different. We started with a very small case to see if we could gather anything from it and generalize to a larger case, so we investigated what happens when n and $n+1$ have only two prime factors. It turns out that there are no k-th Power Ruth-Aaron Numbers of this form.

Theorem 3.28. *Suppose $n = p_1p_2$ and $n+1 = q_1q_2$ with $p_1 \geq p_2$ and $q_1 \geq q_2$. Then $S_2(n) \neq S_2(n+1)$ for all n.*

Proof. Suppose first that n is even. Then we have that

$$p_1 = \frac{q_1q_2 - 1}{2} \tag{3.29}$$

Then by definition of $S_k(n)$ and (3.29) we have that

$$4(S_k(n) - S_k(n+1)) = 4^k + (q_1q_2 - 1)^k - 2^kq_1^k - 2^kq_2^k \tag{3.30}$$

But then since $q_1 \geq q_2$ we have that

$$4(S_k(n) - S_k(n+1)) \geq 4^k + (q_2^2 - 1)^k - 2^{k+1} q_2^k \tag{3.31}$$

For simplicity, let $q_2 = x$. Since $4^k > 0$ we consider the last two terms, and factor, finding that

$$(x^2-1)^k - 2^{k+1}x^k = (x^2 - 1 - 2\sqrt[k]{2}x)((x^2-1)^{k-1} + (x^2-1)^{k-2}(2\sqrt[k]{2}x) + \cdots + (x^2-1)(2\sqrt[k]{2}x)^{k-2} + (2\sqrt[k]{2}x)^{k-1}) \tag{3.32}$$

We first observe that every summand in the second factor is positive. We now analyze the first factor. Define $f(x) := x^2 - 1 - 2\sqrt[k]{2}x$. Since $\sqrt[k]{2} \leq \sqrt{2}$ we observe that

$$\begin{aligned} f(5) &= 25 - 1 - 10\sqrt[k]{2} \\ &> 24 - 10\sqrt{2} \\ &> 0 \end{aligned} \tag{3.33}$$

Taking the derivative we find that

$$f'(x) = 2x - 2\sqrt[k]{2}$$

By inspection we see that the only root of $f'(x)$ is at $\sqrt[k]{2}$ and that $f'(x)$ is positive for all $x > \sqrt[k]{2}$. Therefore it follows that $f(x)$ is increasing for all $x > \sqrt[k]{2}$. Together with Equation (3.33) we conclude that $f(x) > 0$ for all $x \geq 5$. Thus $S_k(n) > S_k(n+1)$ for even n and $q_2 \geq 5$.

Suppose that $q_2 = 5$ and n is even. Then

$$p_1 = \frac{3q_1 - 2}{2}$$

and therefore if we let $q_1 = x$ it follows that

$$S_k(n) - S_k(n+1) = 4^k + (3x-1)^k - 9 - x^k$$

Then we factor and find that

$$(3x-1)^k - x^k = (3x - 1 - x)((3x-1)^{k-1} + (3x-1)^{k-2}(x) + \cdots + (3x-1)x^{k-2} + x^k$$

We first observe that $(3x-1)^{k-1} > 1$ for all positive x, and every other summand in the second factor is positive. Therefore the second factor is greater than 1. Additionally $3x - 1 - x \geq 5$ for all $x > 2$. Thus $(3x-1)^k - x^k > 5$ and also $4^k \geq 4$. Therefore $(3x-1)^k - x^k + 4^k > 9$ and we conclude that $S_k(n) - S_k(n+1) > 0$ in this case. Thus this holds for all even n.

We now suppose that n is odd. Then we find that

$$4^k(S_k(n+1) - S_k(n)) \geq 4^k + (p_2^2 + 1)^k - 2^{k+1}p_2^k \tag{3.34}$$

Let $p_2 = x$. Since 4^k is always positive we consider the last two terms and factor:

$$(x^2+1)^k - 2^{k+1}x^k = (x^2 + 1 - 2\sqrt[k]{2}x)((x^2+1)^{k-1} + (x^2+1)^{k-2}(2\sqrt[k]{2}x) + \cdots + (x^2+1)(2\sqrt[k]{2}x)^{k-2} + (2\sqrt[k]{2}x)^{k-1}) \tag{3.35}$$

All summands in the second factor are positive and so we consider the first factor. Define $f(x) := (x^2 + 1 - 2\sqrt[k]{2}x)$. Then observe that

$$\begin{aligned} f(3) &= 10 - 2\sqrt[k]{2}(3) \\ &\geq 10 - 6\sqrt{2} \\ &> 0 \end{aligned}$$

Taking the derivative we have that

$$f'(x) = 2x - \sqrt[k]{2}$$

Therefore the only root of $f'(x)$ is $\sqrt[k]{2}$ and that $f'(x)$ is positive for all $x > 2$. Therefore $f(x)$ is increasing for all $x > 2$. Since n is odd we know that $p_2 > 2$ and so we conclude that $S_k(n+1) > S_k(n)$ for all odd n. □

This approach however does not generalize well to the situation where n has more prime factors. In the above proof we can narrow down the expression we are looking at to being in one variable which allows us to take the derivative. However, when we increase the number of prime factors we do not have enough information to do so, and the analogous tools we have for multivariable functions do not provide the information we would need to do a similar analysis. However, the conjecture still seems fairly likely, and perhaps with more time and thought new approaches can be developed to handle these issues.

3.3 Further Research

The k-th Power Ruth-Aaron Numbers are certainly an area that could warrant more study since we are fairly confident that a tighter bound can be obtained for them. It might be possible to prove that there exists some large enough k such that the k-th Power Ruth-Aaron Numbers will be well bounded. Further research might work towards tightening the already existing bounds, in many cases we estimated using fairly loose bounds. Another area of research is to examine other arithmetic functions' behavior on the Ruth-Aaron Numbers. Perhaps there is a way to generalize the characteristics that such a function would need to find tight bounds on the number of Ruth-Aaron Numbers. We could also consider behavior of Ruth-Aaron-like numbers, such as comparing $S(n)$ and $S(n+k)$ for various k. We expect that the numbers being immediately adjacent is not an altogether essential aspect of their behavior, and so it would not be difficult to generalize. Alternatively we could consider when $S(n) = S(n+1) = S(n+2)$ and so on, something that is not well understood. Pomerance and Erdős conjecture in [2] that for all k there exists an n such that $S(n) = S(n+1) = \cdots = S(n+k)$, but this has not been proven. In all cases we would hope that new approaches beyond those used here would shed more light on Ruth-Aaron numbers and their generalizations.

Bibliography

[1] de Bruijn N. G., *On the number of positive integers $\leq x$ and free of prime factors $> y$.*, Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen: Series A: Mathematical Sciences **69(3)** (1966), 239-247.

[2] P. Erdős and C. Pomerance, *On the largest prime factors of n and $n+1$*, Aequationes Mathematicae **17** (1978), 311-321.

[3] D. E. Penney C. Nelson C. Pomerance, *714 and 715*, Journal of Recreational Mathematics **7.2** (1974), 87-89.

[4] C. Pomerance, *Ruth-Aaron Numbers Revisited*, Paul Erdős and his Mathematics **I** (2002), 569-579.

[5] M. Ram Murty, *Problems in Analytic Number Theory*, Springer, 2008.

[6] Jeffrey Stopple, *A Primer of Analytic Number Theory: From Pythagoras to Riemann*, Cambridge University Press, 2003.