Recursive Prime Generating Functions:

a 2-decade mathematical mystery

First Name(s): Binyameen

Surname: Mohamedy

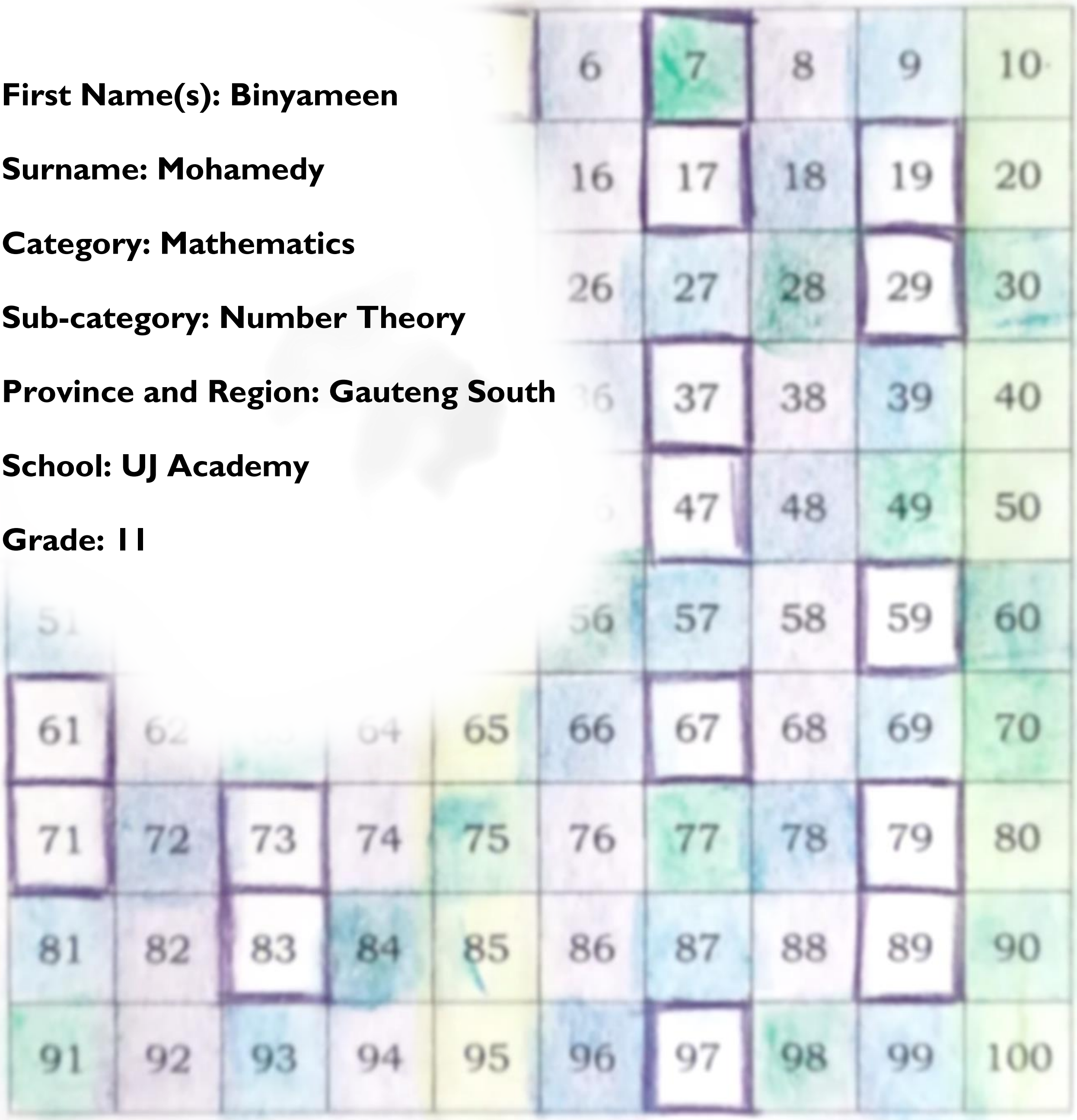
Category: Mathematics

Sub-category: Number Theory

Province and Region: Gauteng South

School: UJ Academy

Grade: 11



PLAGIARISM DECLARATION

PLAGIARISM FORM

DEFINITION OF PLAGIARISM

Plagiarism is the act of presenting someone else's work or ideas as your own; without their consent; by incorporating it into your work without full acknowledgement. This includes all published and unpublished material whether in printed or electronic; in the form of text, computer codes, illustrations, photographs, graphs; from books, the world wide web; journals; data; lectures etc.

I declare that,

- I know what plagiarism is and what constitutes plagiarism,
- this research study is my own work,
- in this research study, other people's work is acknowledged, and all sources of information is referenced.

Signed,

PRINT NAME

1. Binyameen Mohamedy____

SIGNATURE:

1. M.Binyameen____

Date: ___25 July 2024____________________Region: Gauteng South______________________

This Plagiarism declaration document must be signed and presented at the Science Fair, as a compulsory requirement of the Official Document File

ABSTRACT

Primes are central study in mathematics, finding modern applications in encryption algorithms to secure all forms of digital data. In 2003 a function was discovered, now dubbed Rowland's function, that seem to behave in a certain manner. It was conjectured to always produced 1 or prime numbers as output. This proved difficult to solve, as various attempts were made to solve it, however it has remained unsolved since. The author solves this problem due to critical insights gained through analysing numerical and graphical data, the main one being reformulating the problem in terms of proving a certain sum is greater than zero. The general theory of this problem has gained significance due to a discovery linking this theory to several famous unsolved problems in mathematics. This opens up a new road to attack these famous problems.

LIST OF TABLES

LIST OF FIGURES

Note:

All figures and tables were generated by the author.

TERMINOLOGY

gcd(a,b): greatest common divisor of a and b, it can be abbreviated as gcd

Initial Condition[1]: a(h)=k, for some h and k positive integers

Let a(n) = a(n-1) + gcd(a(n-1), n), for every n>h

b(n) = gcd(a(n-1), n)

B(n)=sum of b(k) for all k less than or equal to n.

P(n)=number of primes generated by b(k) for all k less than or equal to n.

max(f(x)) is the largest value of f(x) across some specified range of x.

the floor function $\lfloor N \rfloor$ rounds down the number to the nearest integer thus, $\lfloor 10.9 \rfloor = 10$

the ceiling function, ceil(N) rounds up the integer to the nearest integer, thus ceil(10.3)=10

prime number is an integer only divisible by 1 and itself

RSA encryption, named after its creators R. Rivest, A. Shamir and L. Adleman. It is the standard encryption system.

[1] See initial conditions under Initial Investigation

INTRODUCTION

> Mathematicians have tried in vain to this day to discover some order in the sequence of prime numbers, and we have reason to believe that it is a mystery into which the human mind will never penetrate.
>
> -Leonhard Euler

BACKGROUND

Primes are a central area of study in mathematics. Functions that generate primes usually end up being useless due to a variety of reasons, so mathematicians usually opt to study prime numbers in more indirect ways such as Sieve theory and complex analysis.

In 2003, a new function, now called Rowland's function, was discovered that displayed an interesting behaviour. It was conjectured to only produced 1 or a prime number as its output. This conjecture remained unsolved till now and is the focus of this paper.

The key distinction between this new function and previous functions is the connection formed[2] between the general theory of this function and many famous unsolved problems in Mathematics. This is unusual as these functions normally remain isolated with no applications whatsoever.

In conjunction with that, Rowland's function proved to be a difficult area, in fact thus far, very few unconditional results have been proved for it. That is, for every result proved, the authors were forced to rely on other more difficult, powerful and unproven problems for their result to work. This shows that this is not only a difficult area of research, but it poses a lot of interest for mathematicians due to its connections to other problems in mathematics.

STRUCTURE OF REPORT

Due to mathematics being an abstract field to most, an overview of the structure of the report is given.

The significance of prime numbers, Rowland's function and this project is given under 'Introduction'. An analysis of previous works is given under 'Literature review'. Within, various aspects of Rowland's function are explained through the attempts of other on resolving Rowland's conjecture. It is recommended to read through it.

Hypothesis lists the conjectures this paper focuses on, as well as Rowland's condition. The main proof is given in 'Procedure/Proof', detailing the work of this paper.

Mathematics is a much harder to explain field than others, such as engineering. For this reason, in the Initial Investigation section, a brief account of the authors investigation is given. It also works as an exposition to this theory, explaining various concepts but also gives a few heuristic results under "Analysis on the criteria of r", 'Sequence term', 'identity', and 'Lower bound on Rowland's conjecture', for that reason it is listed under 'Method' and not as an Appendix. It is not necessary to read but might aid the reader.

The work concludes from there with a discussion, future research, conclusion etc.

---

[2] See Cloitre (2018)

DIFFERENCE BETWEEN REPORT AND PLAN

There exist several differences between the original research plan and the final research report. The research plan outlined 4 conjectures whilst the report deals with only 2. This is because of a few discoveries made during the authors research which make Conjecture 3 and 4 far less significant than Conjecture 1 and 2. For this reason conjecture 3 and 4 were removed for not being important enough to study.

INTEREST IN PRIMES

The study of primes is perhaps the most ancient topic in mathematics, second only to geometry. The interest in primes stem from the fact that the problems relating to them are never easy. Mathematicians do not know why their problems are difficult. How ever it is understood that the difficulty arises due to a clash of nature. Primes are inherently multiplicative objects while all our questions for primes are additive, i.e. what type of numbers they form when added up.

The study of primes, aka prime number theory, is akin to Quantum Mechanics. Primes are the atoms of the integers and Quantum Mechanics studies the contents and behaviours of the atom much the same way as prime number theory studies the properties of primes.

Just like Quantum Mechanics with physics, prime number theory holds the deepest and most difficult problems in mathematics, such as:

i. Twin prime conjecture (unsolved for 178 years)
ii. Goldbach's conjecture (unsolved for 282 years)
iii. Legendre's conjectures (unsolved for 112 years)
iv. Finitude of Fermat's primes (unsolved for 165 years)
v. Riemann hypotheses (unsolved for 166 years, prize money of 1 million dollars)

The clay mathematics institute believed 7 problems will define all of mathematics of the 21st century. To convey their importance to the public, they put a 1-million-dollar prize money on each problem. These include the Riemann hypotheses, P vs NP, Navier-stokes equations etc.

In the modern age, prime numbers have found a critical role in internet security. Encryption systems secure all forms of digital data, and current encryption systems, such as RSA encryption, rely largely on prime numbers for its security strength. It is usually taking advantage of a property of prime numbers, namely that decomposing numbers into its prime factors is extremely difficult, whilst multiplying primes into numbers is extremely easy.

From this it is clear that primes are an important area of study, not only for mathematics, but every field.

ROWLAND'S BREAKTHROUGH

In 2003 at the NKS Summer school[3], a group led by Matt Frank explored computing different types of recursive functions and analysed their behaviour. Eric Rowland was one of the participants and got interested in one of the functions they discovered.

[3] Rowland (2008)

The function, b(n), was a recursive function, meaning its value at any term n, depended on the previous term. A example of a simple recursive function would be, f(n) = f(n-1) + 1. This idea is introduced more deeply under 'Initial Investigations'.

The function b(n), now known as Rowland's function, seemed to always output either 1 or prime numbers. They conjectured this was always the case and 5 years later Rowland (2008) provided a proof of this assuming a certain condition held for the initial condition of the function. It is not known whether this condition always holds. Nothing rules out the possible existence of a counterexample, which is why a mathematical proof is needed to show that such a counter example cannot exist.

Other functions were discovered through this event such as the one studied by Cloitre (2011) and Ruiz-Cabello (2017). Through these papers, this field was brought to the attention of many others.

LITERATURE REVIEW

2 main branches of literature exist regarding this topic. The first is papers written on Rowland's function, the study of this paper. It is historically the first function in this field that was studied. Rowland did not start this field, however his work did pull more people into studying this topic.

The second is papers written on alternate functions, these include replacing the gcd(a(n-1), n) with lcm(a(n-1), n) or gcd(a(n-1), n+$(-1)^n$) etc. In this instance, the function b(n), is not always equal to these expressions as well, i.e for the case a(n) = a(n-1) + lcm(a(n-1, n), b(n) is not equal to lcm(a(n-1), n), like how for Rowland's function b(n) = gcd(a(n-1), n). instead $b(n) = \frac{a(n)}{a(n-1)} - 1$.

b(n) always represents the function generating the primes however, in some cases it can generate primes or 1s, the larger value of a twin prime pair[4], etc.

ROWLANDS FUNCTION

Rowland's function is defined as follows.

$$a(n) = a(n-1) + \gcd(a(n-1), n), \text{ for every } n>h \text{ with } a(h)=k,$$

$$b(n) = \gcd(a(n-1), n)$$

Rowland (2008) proved that b(n) is a prime number or 1 for all n >m, under the assumption that an integer m exists that satisfies his assumption:

$$a(m)=rm, \text{ with } r=2 \text{ or } r=3.$$

It is not known whether such an integer m always exists for all initial conditions.

For the initial conditions of a(7), b(n) produces 1 or prime for all n, as a(5)=15=3·5. He was not able to prove that the assumption holds for every initial condition.

[4] Pair of primes such that the difference between them is 2, such as 3 and 5 or 11 and 13.

Rowland's proof was as follows:

By numerical data he knew a(n) generates a prime whenever $\frac{a(n)}{n}$ = 2 or $\frac{a(n)}{n}$ = 3. Assuming $\frac{a(n)}{n}$ is 2 or 3 for some n, he used induction to prove a prime is generated and that they'll be integer m>n such that $\frac{a(m)}{m}$ = 2 or 3. He's proof critically relied on identities derived from the sequence of 1's generated by b(n).

Figure 1 illustrates the structure of his proof.

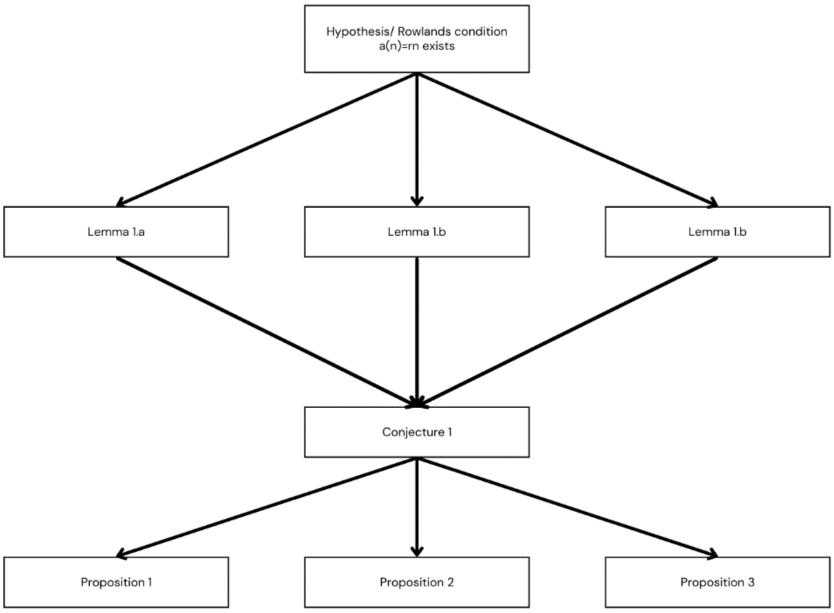


Figure 1: proof structure for Rowland (2008)

Note: By the arrows it is meant that, the block it starts from implies, sometimes partly, the truth of the block it ends at, or rather the block it ends at is deduced by the block it starts at. Thus the truth of Rowlands condition implies the truth of 3 lemmas in Rowlands work. The truth of these 3 lemmas than implies Conjecture 1 to be true.

This condition itself and the value of r may seem arbitrary and random, and Rowland also could not figure out what determined the value of r, and why it was there beyond its existence being critical.

> "The only distinguishing feature of the values r = 2 and r = 3 in the lemma is the guarantee that $\frac{p-1}{r-1}$ is an integer, where p is again the smallest prime divisor of $(r-1)n_1 - 1$."
>
> Rowland (2008).

This project answers these questions. The value rn is actually the absolute maximum value of a(n), and the value of r is determined by a criteria on the initial condition. This will be explained later under Analysis on criteria of r. Thus Rowland's condition really states that if the function achieves the expected maximum value, it will produce a prime number or 1.

Ruiz-Cabello/Chamizo/Raboso (2011) proved Conjecture 1 and 2, again under a certain hypothesis being true, for the special case of h=1, their methods couldn't be generalised for h>1. Their methods are inspired by Cloitre (2011) and differ from the authors. It involves creating an auxiliary function to represent a(n) and prove results on that function that imply it on a(n). It is a somewhat more complicated method than the authors.

RELATED FUNCTIONS:

Ruiz-Cabello (2017) and Shepke (2014) are based on the function b(n) replaced with the lowest common divisor instead, specifically

$$a(n) = a(n-1) + lcm(a(n-1), n);\;\; n > h\; for\; a(h) = k$$
$$with\; b(n) = \frac{a(n)}{a(n-1)} - 1$$

Ruiz-Cabello (2017) conditionally proved conjecture 1,2 and 3 for this alternative function by showing b(p)=p for every prime p except 3. The author initially proved these results for the specific case of h=1, but was able to generalise it to all initial conditions.

Ruiz-Cabello (2017) assumed 2 very strong conjectures, one much stronger than assuming the Riemann hypothesis and the other equally as strong, related to the smallest prime in an arithmetic progression. This is not productive, as the conjectures assumed could be left unsolved for at least another 2 to 3 decades, possible much longer. Thus, his method is generally not the right direction to approach these types of problems.

The method used was similar in structure to Rowland (2008) and from the authors work, it is discovered that both functions have very similar properties but are distinct in few ways. One example would be b(p)=p for every prime p for the function b(n) as defined by Ruiz-Cabello (2017), the authors work discovers b(n) as defined by Rowland (2008) also has the property b(p) =p but only for a specific infinite set of primes p.

Cloitre (2011) created a new formulation of a more general idea of Rowland's work and stated 10 conjectures relating a more general recursive prime generating function to famous Conjectures in Number theory such as the Twin prime Conjecture, Goldbach's Conjecture, Legendre's Conjecture and more. He's work brought this theory to the attention of many other researchers due to its connection to more significant problems.

Shevelev V. (2010) found new recursive functions that generated primes in relation to the famous Twin prime Conjecture. The work is flawed, and a few other papers of the author on the subject are unclear and the included mathematics has glaring problems in the proofs. The conjectures on the function have good numerical data and heuristic evidence from the work however and seem promising.

In summary Conjecture 1- 2 remain unsolved in general. Attempts have been made to prove them, but none succeeded fully. The main limitation exists in that very little information is known about the function. No estimates, or identities are known to be used to develop a full proof. However, in other situations the problem is not that the information is not known, but that the overall proof developed on those data is limited. Ruiz-Cabello (2017) first develops these forms of data, such as identities and estimates, but his overall proofs idea was limited as he had to rely on 2 very different Conjectures.

PROBLEM STATEMENT

Formulated in 2003, Conjecture 1 has remained unsolved since, with partial progress done in Rowland (2008) and Ruiz-Cabello/Chamizo/Raboso (2011), 6 and 9 years later respectively.

The problem thus is to solve Conjectures 1 and 2.

RESEARCH QUESTION

Is Rowland's conjecture, posed in 2003, true? If so, does it produce infinitely many distinct primes?

AIM

To explore the topic of recursive prime generating functions, specifically Rowland's function, by solving or providing significant partial progress on Conjectures 1 and 2.

HYPOTHESIS

CONJECTURES

1) b(n) is prime or 1 for every value of n greater than some fixed value N.[5]
2) b(n) generates infinitely many distinct prime numbers.

Rowlands assumption:

- There exists an integer n such that a(n)=rn for r=2 or r=3
  - Assuming this Rowland provided a conditional proof of Conjecture 1 through induction

[5] See 'Lower bound on Rowlands conjecture' for a proof of what N is atleast and what the author conjectures its value to be

METHOD

MATERIALS AND EQUIPMENT

The Desmos graphing calculator was initially used to graph data however it was weak, slow and sometimes inaccurate. Python with the matpotlib library was used to graph data instead for the rest of the period. A journal was used to do mathematics.

All the code can be accessed at https://github.com/monkeybini/eskom-expo-data-code/blob/main/data.py

No other materials or equipment were used.

DEFINTITIONS AND CONCEPTS

gcd(a,b): greatest common divisor of a and b, it can be abbreviated as gcd

Initial Condition: a(h)=k, for some h and k positive integers

Let $a(n) = a(n-1) + \gcd(a(n-1), n)$, for every $n>h$

$b(n) = \gcd(a(n-1), n)$

B(n)=sum of b(k) for all k less than or equal to n.

PROCEDURE/ PROOF

PROOF OUTLINE: CONJECTURE 1

To prove Conjecture 1, it would suffice to prove Rowland's condition, which states that some integer n always exists such that $a(n) = rn$ for $r = 2$ or $r=3$.

The following expression evaluates to 1, if $a(n)=rn$ and 0 otherwise if $a(n) < rn$.

$$\left\lfloor \frac{a(n)}{rn} \right\rfloor$$

Proving it is greater than 0 for some n would prove Rowland's condition. We do so in the following way.

$$g = \sum_{x \le n \le 2x} \left\lfloor \frac{a(n)}{rn} \right\rfloor$$

If $g>0$, this would imply that Conjecture 1 is true.

Assume $kn \le a(n) \le rn$, for some r and k, and let $O(n)$ be defined as $|O(n)| \le |n|$

Thus, we can rewrite g in the following way.

$$g = \sum_{x \le n \le 2x} \left\lfloor \frac{a(n)}{rn} \right\rfloor$$

$$g = \sum_{x \le n \le 2x} \frac{a(n)}{rn} + O\left(\frac{rn - a(n)}{rn}\right)$$

$$g \ge \sum_{x \le n \le 2x} \frac{a(n)}{rn} - \left(\frac{rn - a(n)}{rn}\right)$$

$$g \ge \sum_{x \le n \le 2x} (2\frac{a(n)}{rn}) - x$$

$$g \ge \frac{2k}{r}x - x$$

Proving that the lower bound is greater than 0 would imply,

$$2k \ge r$$

Thus we can get a restriction for the inequalities needed to prove Conjecture 1.

If $a(n) \ge n$, than $2 \ge r$, and proving that $a(n) \le 2n$, would prove Conjecture 1.

It is a similar situation for the case of r=3. Proving $3n \ge a(n) \ge 2n$ would then also Conjecture 1.

Thus the proof can be split into 2 cases. $a(n) > 2n$ and $2n \ge a(n)$,

In the second case, Conjecture 1 is true if has r=2 and k=n+a(h)-h

$a(n) \ge n + a(h) - h$ is a simple bound to get, by noticing a(n) = a(n-1) + b(n), and that by substituting this formula in place for a(n-1) repeatedly you get

$$\text{a(n) = a(h)+} \sum_{k=h+1}^{n} b(k) \ge \sum_{k=h+1}^{n} 1 = n - h$$

In the first case, if it can be proved that if $a(n) > 2n$, then $3n \ge a(n)$, this would then prove Conjecture 1. Thus the proof of Conjecture 1 can now be rephrased as a problem of deriving an upper bound on a(n) if a lower bound is given.

The proof for this lemma is still being peer reviewed, however initial comments from the reviewer indicate it is flawless. it will be presented at the ISF instead. Certain time limitations on the reviewer and author forced this situation.

PROOF OUTLINE: CONJECTURE 2

Looking at graphical data, it is clear that a(n) has a minimum and maximum value. More interestingly, when a(n) is equal to the minimum value, the next term will produce a prime number and hit the maximum value. Visually this creates a staircase like pattern.

Let r= $\lim_{n\to\infty} sup \frac{a(n)}{n}$, then the author conjectures $\lim_{n\to\infty} inf \frac{a(n)}{n} = \frac{r+1}{2}$

Thus, $2 \lim_{n\to\infty} inf \frac{a(n)}{n} - 1 = \lim_{n\to\infty} sup \frac{a(n)}{n}$

We first prove r exists by showing $\frac{a(n)}{n}$ is less than a constant. A key and elementary identity is used on the gcd and lcm function, namely :

gcd(a, b)·lcm(a, b)=ab

now the proof goes as follows, let $l(n) = lcm(a(n-1), n)$ for clear notation.

$$a(n) = a(n-1) + \text{gcd}\,(a(n-1), n) = a(n-1) + \frac{a(n-1)n}{lcm(a(n-1,n)}$$

$$\frac{a(n)}{n} = a(n-1)(\frac{1}{n} + \frac{1}{l(n)})$$

Now repeatedly substitute that for a(n-1), you get:

$$\frac{a(n)}{n} = a(h) \prod_{k=h+1}^{n} (\frac{1}{k} + \frac{1}{l(k)})$$

Since $l(n) \geq n$, the equation turns into

$$\frac{a(n)}{n} \leq a(h) \prod_{k=h+1}^{n} \frac{2}{k}$$
$$\frac{a(n)}{n} \leq a(h) \frac{2^{n-h}}{\frac{n!}{h!}} = a(h) \frac{2^n}{n!} \frac{h!}{2^h} \leq a(h) \frac{h!}{2^h}$$

Which is a constant depending on h for every function.

Now if $a(n-1) = \frac{r+1}{2} n$

Than it is easy to show that a(n)=rn, by splitting r into the case of r=2 and r=3.

If r=3, than b(n)=n while if r=2, $b(n) = \frac{n}{2}$

This assumes that r=2 or r = 3.

The author's work discovers a conjecture, a criteria on r. it states that the value of r is determined as follows:

1. a(h) ≤ h, r=1. it is trivial case to prove that if a(n)=n than b(n) will always be 1.
2. h+1≤a(h)≤2h+1 but a(h)≠h+2, then r=2
    a. When a(h)=h+2, then $\lim_{n\to\infty} \frac{a(n)}{n} = 1$. This is also a trivial case.
3. 2h+2≤a(h), then r=3

This implies that r=3 almost always, with r=2 being a special family of cases.

Since a(n)=rn, it fulfils Rowland's condition which implies b(n) is prime, meaning n is prime too due. Thus b(p)=p or b(2p)=p for some prime p. Due to Rowland's use of induction, this instance repeats infinitely many times.

This would prove Conjecture 2.

INITIAL INVESTIGATION

The work started out initially by studying Rowland (2008) and then proceeding to look at numerical and graphical data to understand Rowland's function better.

Reading the work 2 main questions come to mind that seemingly Rowland has no answer to either, these are:

I. Why is the value of r only 2 or 3, why was it chosen and why can it not be any other value?
II. When does the function generate a prime?

Before showing the proof, the author will show how he studied and developed his work by answering the above 2 questions.

EXPLANATION ON THE INITIAL CONDITION

Before continuing an explanation is order for what is meant by the initial condition. Rowland's function a(n) is a recursive function meaning that it's value at any point n is some operation done on the previous value.

If a function is defined this way, always depending on the previous value, there has to be some starting value that's not defined recursively otherwise, i.e. if a(1) depends on a(0) and that isn't defined as a specific number one can keep going to negative infinity and have a undefined function.

So every recursive function has a initial condition where a(1) is defined as a specific constant number. For some sequence defining that point at a(1) compared to a(2) or any other term changes the entire sequence, while for others it remains the same.

Examples of recursive functions are:

- Fibonacci sequence: $f(n) = f(n-1) + f(n-2)$ with $f(1)=f(2)=1$
- Powers of 2, $f(n)=f(n-1)\cdot 2$ with $f(1)=2$
- Factorial, $f(n)=f(n-1)\cdot n$ with $f(1)=1$

Whether f(1) or f(2) is set as the starting value the sequence of numbers generated will remain the same as there is no dependence on the value of n in the definition. Table 1 provides an numerical example.

	f(1)	f(2)	f(3)	f(4)	f(5)	f(6)	f(7)	f(8)	f(9)
f(1)=2	2	4	8	16	32	64	128	256	512
f(4)=2	-	-	-	2	4	8	16	32	64

Table 1: f(n)=f(n-1)2

It would give different outputs if f(1)=3 or f(1) was in general any other value. As you can see regardless of whether the sequence starts at f(1) or f(4) the numbers outputted are the same.

But this is not always the case, for example the factorial's sequence will differ because for it f(n) depends on the previous term and the value of n. Table 2 gives a numerical example of this.

	f(1)	f(2)	f(3)	f(4)	f(5)	f(6)	f(7)
f(1)=1	1	2	6	24	120	720	5040
f(2)=1	-	1	3	12	60	360	2520

Table 2: f(n)=f(n-1)n

As you can see, f(n) for the factorial function depends both on h and f(h). changing h or f(h) gives you a different sequence.

This is why for Rowland's function the initial condition of a(h) and h exists.

Table 3 shows an example of b(n)'s property of prime for initial condition of h=1 and a(1)=7.

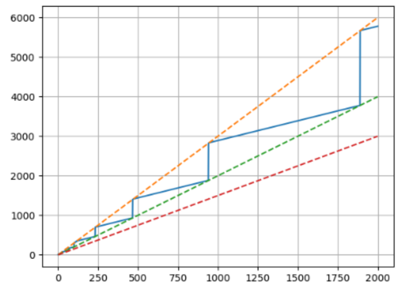
n	*a(n)*	*b(n)*	*B(n)*	*B(n)*
1	7	-	-	-
2	8	1	1	1
3	9	1	1+1	2
4	10	1	1+1+1	3
5	15	5	1+1+1+5	8
6	18	3	1+1+1+5+3	11
7	19	1	1+1+1+5+3+1	12
8	20	1	1+1+1+5+3+1+1	13
9	21	1	1+1+1+5+3+1+1+1	14
10	22	1	1+1+1+5+3+1+1+1+1	15

Table 2:h=1, a(h)=7 -- > a(1)=7

GRAPHING THE FUNCTION

Before plotting the function, try to imagine what the graph would look like. a(n)= a(n-1)+gcd(a(n-1), n), and the gcd of any 2 numbers is at least 1, so if you continuously added 1 to the previous term you should get a straight line graph.

Conjecture 1 states that b(n) [6] is only 1 or a prime number. That can be rephrased as b(n) is 1 and if its greater than 1 it's a prime number. We expect a straight-line graph however if at any point b(n) is prime it should bump up higher than normal. Through this one can expect to see a "bumpy" graph.



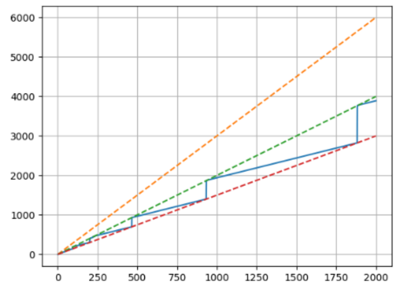


Figure 2: h=1, a(1)=7

Figure 3: h=2, a(2)=5

[6] Remember b(n)= gcd(a(n-1), n)

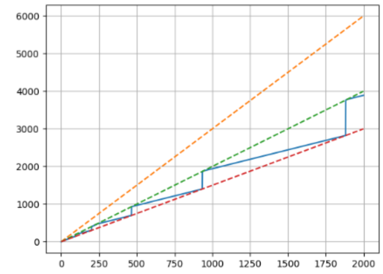


Figure 4: h=4, a(4)=7

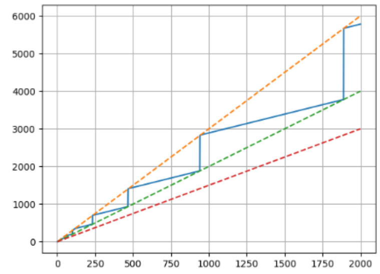


Figure 5: h=3, a(3)=12

The sub graphs are as follows:

- blue zigzag graph is y=a(x)
- orange highest straight-line graph is y=3x
- second highest green straight-line graph is y=2x
- third highest red straight-line graph is y=$\frac{3}{2}$x

ANALYSIS ON THE CRITERIA OF R

we can also see by graphing more data that a(n) always ends up between 2 sets of functions and actually the functions that have r=2 from Rowland's assumptions i.e. have a term such that a(m)=2m for some m, end up between y=2n and y=$\frac{3}{2}$n like Figure 2 and 3. while the functions that have r=3m are in-between y=3n and y=2n like Figure 1 and 4.

Mathematically we can state this as, $\frac{r+1}{2}(n+1)$≤a(n)≤rn. Putting in the valid value of r depending on a(n) you get the respective upper and lower bounds.

In this manner you can see that once a(n) hits it's the bottom graph it generates a prime number large that always makes it touch the upper graph from which b(n) continues to only output 1 in a straight line till a(n) touches the bottom graph again.

In fact we can prove mathematically that if a(n-1) = $\frac{r+1}{2}$ than a(n)= $\frac{r+1}{2}$ n + gcd($\frac{r+1}{2}$n, n).

For the case r=3 we can see the greatest common divisor between 2n and n is obviously n.

For the case r=2, it is not as clear but the gcd of $\frac{3}{2}$n and n will be $\frac{n}{2}$.

Doing a little bit of algebra, you end up getting a(n)=rn.

Seeing this Conjecture 1 can be rephrased slightly into proving a) a(n) has a best possible lower bound in the form above.

This answers question i) but what about question ii)?

Well looking at numerical data you can end up spotting a pattern in the initial condition corresponding to the value of r.

When r=1, a(h) ≤ h, it is trivial case to prove that if a(n)=n than b(n) will always be 1.

When r=2, that is, h+1≤a(h)≤2h+1 but a(h)≠h+2

When a(h)=h+2, then $\lim_{n\to\infty}\frac{a(n)}{n} = 1$. This is also a trivial case.

When r=3, 2h+2≤a(h)

This answers question ii). However, the answers are not satisfactory, it is still not understood why r=2 or 3 and why the primes are generated when they are. For this a mathematical proof would answer these questions much more satisfactory.

However, for question i) we can intuitively understand why a(n) would choose those lower bounds and not just generate a prime randomly. If you imagine the graph of a straight-line but it jumps up every few values, it will grow very quickly. This is not expected for a(n) as it includes n raised to the power of 1 and all other components of its definition are arithmetic. It should grow slower than that.

Imagining it generating primes very rarely, the graph becomes very low. The primes are not very frequent up to any point, but they are not sparse either. This is a weak argument, but it provides some intuition.

STAIRCASE PATTERN

Looking at the graph deeper you can see that any value at which the graph hits the lower function is around twice as much as the point that hits the lower graph. In fact, through a similar heuristic argument an exact expression can be reached. If a(n-1) = $\frac{r+1}{2}n$ than m=ceil($\frac{n}{2}$ - $\frac{1}{r-1}$) and a(m-1) = $\frac{r+1}{2}m$

The argument for that follows by graphing a straight-line graph that approximates 1 sequence of 1's, i.e. 1 straight line segment the a(n) outputs and performing some algebra from there

IDENTITY ON THE LIMITS OF THE SEQUENCE

Another key result discovered is the relationship between $min(\frac{a(n)}{n})$ and $max(\frac{a(n)}{n})$. min denotes the best smallest value that function can achieve, i.e. the largest number the function will always be greater than across all valid n. max denotes the largest possible value the function can achieve, i.e. the smallest number it will always be less than. For both functions, n is taken to be larger than some constant N, detailed in the next section "Lower Bound"

Seeing that $min(\frac{a(n)}{n}) = \frac{r+1}{2}$ while $max(\frac{a(n)}{n}) = r$ we can combine them by solving for r in the min expression and substituting it in the max expression to get $2min(\frac{a(n)}{n}) = max(\frac{a(n)}{n}) + 1$

Analysing the graph more, an expression can be derived that give you all the values of n such that a(n) touches the lower graph with some small error term.

LOWER BOUND ON ROWLANDS CONJECTURE

Rowland (2008) conjectured that Conjecture 1 holds for every n>N, for some number N. It is not known what this N is exactly. In other words, it is unknown when the conjecture begins to take place. A lower bound is given on how large this N must be. The author conjectures that lower bound is in fact the value of N, that is N = $\frac{a(h)-h}{r-1}$

Lemma: N ≥ $\frac{a(h)-h}{r-1}$

Proof:

$$a(n)=a(n-1)+b(n) \qquad \text{eq 1}$$

substituting eq 1 into itself

$$a(n)=a(n-2)+b(n)+b(n-1)$$

repeatedly performing this, eq 1 becomes

$$a(n)=a(h)+\sum_{k=h+1}^{n} b(k)$$

b(n)≥1 thus

$$a(n)\geq a(h)+n-h$$

assuming a(n)=rn,

$$n \geq \frac{a(h)-h}{r-1}$$

■

RESULTS

THEOREM 1:

Let a(n) = a(n-1) + gcd(a(n-1), n), for every n>h, and b(n) = gcd(a(n-1), n)

Then, b(n) is prime or 1 for every n greater than some constant number, if when $a(n) > 2n, then\ a(n) \leq 3n$

THEOREM 2 (HEURISTIC PROOF):

b(n) produces infinitely many distinct prime numbers.

CONJECTURES/HYPOTHESIS:

1) CRITERIA ON R:

Let r= $\lim_{n\to\infty} sup\ \frac{a(n)}{n}$

Then the value of r is determined as follows:

if a(h) ≤ h or a(h) = h+2, then r=1

if h+1≤ a(h) ≤ 2h+1, a(h)≠h+2 then r= 2

If 2h+2 ≤ a(h), then r=3

2) LIMIT IDENTITY ON $\frac{a(n)}{n}$:

$$2\lim_{n\to\infty} inf\ \frac{a(n)}{n} - 1 = \lim_{n\to\infty} sup\ \frac{a(n)}{n}$$

DISCUSSION

The proof of Conjecture 1 ended up being significantly different from previous attempts. A natural progression from here would to attempt to apply the novel methods developed in this paper onto functions explored by other authors such as Ruiz-Cabello (2017) or Cloitre (2011). However it is unlikely, as the foundational ideas differ.

Ruiz-Cabello (2017) is based on alternate version of Rowland's function in which the least common multiple is used instead of the greatest common divisor. It relies on 2 famous and extremely difficult conjectures in analytic number theory, so proving Conjecture 1 for it through the hypothesis it relied on is outside the reach of modern mathematics for now. A possible way would however be to replace its dependence on these hypotheses with a different result that's easier to prove, this seems more possible.

The outlined proof of Conjecture 2 highlights several critical aspects of Rowland's function that was missed by Rowland (2008). The method used in Rowland (2008) skips these aspects entirely, indicating it is the wrong approach for this problem. The right approach would have proven the mentioned relationship between the functions minimum and maximum values, as well as prove the criteria on r alongside the fact b(p) or b(2p) is = p for infinitely many primes p. This approach is best, as it solves most of the mysteries surrounding Rowland's function as well as answers both Conjecture 1 and 2.

In hindsight it is clear Rowland's idea was limited and in the wrong direction.

LIMITATIONS AND ERRORS

TIME CONSTRAINT:

Mathematics is a slow science, the authors work so far has filled the entirety of the journal, more than 100 pages. Rowland (2008) proved his result over the cause of a few years for his PhD thesis. Learning new material took time as they were on research level mathematical topics such as elementary number theory, basic analytic number theory, work in recursive prime generating function etc.

MATERIALS AND EQUIPMENT

Lack of strong enough software for calculating data slowed progress as the author had to program code to calculate the data needed. Python was used for this.

This constraint could potentially be lessened through getting a powerful graphing calculator or digital calculator as powerful as python code but as easy and quick to use as Desmos.

A more powerful laptop would remove the constrained entirely as computing large amounts of data as well as visualising it is taxing on usual office laptops.

No other limitations or constraints were faced.

LIMITATIONS OF THE IDEAS

It would be natural to try and apply the methods developed by the author here onto related functions such as Ruiz-Cabello (2017), but the ideas introduced here rely on ideas specific to Rowlands function, specifically it was by proving Rowland's condition which only exists for Rowland's function. It seems very unlikely to be applied into solving related problems.

However, the Hypothesis formulated from this project, such as the limit identity on $\frac{a(n)}{n}$ hints at similar identities for related functions. This sort of identity would be important forms of data, as proving it for Rowland's function would provide critical information.

A common way to make progress on both Rowland's function and related functions would be finding better approximations to sums and reciprocal sums on the LCM and GCD function, such as :

$$\sum_{k=1}^{n} \frac{1}{gcd(k,n)},$$

$$\sum_{k=1}^{n} \frac{k}{gcd(k,n)},$$

$$\sum_{k=1}^{n} \frac{1}{lcm(k,n)},$$

$$\sum_{k=1}^{n} \frac{k}{lcm(k,n)} \text{ etc}$$

RECOMMENDATIONS FOR FUTURE RESEARCH

The main direction forward would be to prove a few critical features of this function, outlined under 'Proof outline: Conjecture 2'. They were undiscovered till the authors investigations and remain difficult to prove. Proving these features would also provide an alternate proof of Theorem 1, although the author managed to construct a proof that avoids the difficult problem of proving those features and found a simpler method, these features form crucial properties of this function which are undoubtedly needed for any further investigations.

Thus the next step for Rowland's function would be to prove the hypothesis laid out by the author under 'Proof outline: Conjecture 2' and 'Results'. These include an identity on the limits of this function, as well as hypothesis on the types of primes generated and under which conditions.

Beyond Rowland's function, looking at the general theory, the way forward would be to work on alternate functions in the field, such as the ones studied by Ruiz-Cabello (2017) or Cloitre (2011). The former would require some modification of the fundamental argument used by Ruiz-Cabello (2017) or even an entirely new proof.

Another direction of research, a few degrees more difficult and in a different field in related to primes, would be on Carmichael numbers, which are exceptions to Fermat's primality test, which is the fundamental mathematics of the RSA encryption system. It is a topic of great interest, in fact, work on it led to the fourth-place winner of Regeneron 2020.

CONCLUSION

Conjecture 1 is proven to be true, depending on a small lemma, while a heuristic proof is given of Conjecture 2 proving it true also. The method developed was vastly different from previous attempts made. This proves a 2 decade old mathematical conjecture on prime numbers, deepening the theory on Primes, specifically recursive functions generating primes. The theory of recursive prime generating functions is still quite new and is filled with many mathematical mysteries and conjectures. It is not yet understood how to generally approach these problems, and many seem difficult to prove.

Other more minor results were achieved whilst proving Theorem 1/Conjecture 1. These include answering a question in Rowland (2008) on what determines the value of r, as well as discovering the types of primes produced, giving various minor identities on the limit of the function a(n) etc.

None of these results were known before.

SIGNIFICANCE:

This work resolves a 2-decade old Mathematical conjecture. It helps future mathematicians' study related problems and furthers the general theory of recursive prime generating functions. Pure mathematics is an incredibly important field and has a huge but unnoticeable impact on the STEM field, this is due to 2 reasons.

First mathematics develops the underlying theory that other fields use. This can easily be seen in physics, or engineering, it is needed to create any further developments in those fields. Second, mathematics is a slow field, often the mathematical ideas used by other fields, such as Physics and complex analysis, are old ideas in mathematics. They have been carefully developed by mathematicians for decades before being moved on from, at which point their theory is sound and secure enough to be used in practical applications, such as weapons, encryption systems, physics etc. Thus progress in the mathematical field is in the interest of all scientific fields.

The function could potentially improve encryption systems. RSA encryption requires 2 large primes to begin it algorithm. It is normally secure however quantum computers have the potential to decrypt RSA encrypted information. A potential application would be to strengthen RSA encryption by producing larger primes through Rowland's function. It does not modify RSA encryption itself.

South Africa has an image of being undeveloped in the STEM field and contributing nothing to science unlike many western countries. Many events are held to encourage youths to become engineers, chemists, entrepreneurs etc. However, beyond study events for school or financial mathematics, not much has been done to encourage youths to pursue mathematics, and yet mathematics is a foundational tool is every field.

Einstein's work on the general theory of relativity relied on non-Euclidean geometry, a pure mathematics field born from investigations into Euclid's axioms. Prime numbers were famously considered the purest object in mathematics, with no practical application in the real world, however they are now critical components of encryption systems securing all forms of digital data. The Navier-Stokes equations, part of the 7 millennium prize problems with 1-million-dollar prize money, is used in Computational Fluid Dynamics/ CFD simulations.

In this manner, it is clear that, whilst pure mathematics normally has no immediate impact or application, it generally becomes the theoretical foundation over which various physics and engineering is build on. This work encourages youths to pursue subjects such as Mathematics whilst deepening South African mathematics.

ACKNOWLEDGMENTS

Heartfelt gratitude must first go to Nicholos Zhang and Zaahid Sader (friends and former ISEF special award winners), for the encouragement and push to participate in Eskom Expo.

Nicholos Zhang should especially be acknowledged for providing various resources to improve the project, blunt criticism and unwavering support.

Aadila Iyamah should also be thanked for her support, presence and inspiring hard-working attitude throughout her and the authors respective projects.

Appreciation goes to my family and friends for nothing in particular.

Appreciation to UJ Academy for providing a venue, transportation and generally being very supportive of the participation in this event. Sincerest thank you to Ma'am Moshome for sacrificing her time on countless weekends, encouraging students and in general organising this event.

Gratitude goes to Professor Andrew Craig for his advice on the writing aspect.

REFERENCES

Chamizo, F., Raboso, D. and Ruiz-Cabello, S. (2011) '*On Rowland's sequence', The Electronic Journal of Combinatorics*. Available at: https://www.combinatorics.org/ojs/index.php/eljc/article/view/v18i2p10 (Accessed: 25 July 2024).

Cloitre, B. (2011), '*10 conjectures in Additive Number Theory*' *arXiv.org*. Available at: https://arxiv.org/abs/1101.4274 (Accessed: 25 July 2024).

Rowland, E.S. (2008) '*A natural prime-generating recurrence', Journal of Integer Sequences [electronic only]*. Available at: https://eudml.org/doc/45342 (Accessed: 25 July 2024).

Schepke M. (2014), 'On prime generating sequences', Available at: http://www.riemannhypothesis.info/wp-content/uploads/2014/10/schepke_prime_generating_sequences.pdf

SELBERG, A. (1984) 'On an Elementary Method in the Theory of Primes', *The Goldbach Conjecture*, pp. 151–154. Available at: https://doi.org/10.1142/9789814542487_0010.

Serafín Ruiz-Cabello (2017) 'On the use of the least common multiple to build a prime-generating recurrence', *International Journal of Number Theory*, 13(04), pp. 819–833. Available at: https://doi.org/10.1142/s1793042117500439.

Shevelev V. (2010), '*Three theorems on twin primes', arXiv.org*. Available at: https://doi.org/10.48550/arXiv.0911.5478.