Generalizing Ruth-Aaron Numbers

Y. Jiang and S.J. Miller

January 16, 2021

Abstract - Let $f(n)$ be the sum of the prime divisors of n, counted with multiplicity; thus $f(2020) = f(2^2 \cdot 5 \cdot 101) = 110$. *Ruth-Aaron* numbers, or integers n with $f(n) = f(n+1)$, have been an interest of many number theorists since the famous 1974 baseball game gave them the elegant name after two baseball stars. Many of their properties were first discussed by Erdős and Pomerance in 1978. In this paper, we generalize their results in two directions: by raising prime factors to a power and allowing a small difference between $f(n)$ and $f(n+1)$. We prove that the number of integers up to x with $f_r(n) = f_r(n+1)$ is $O\left(\frac{x(\log\log x)^3 \log\log\log x}{(\log x)^2}\right)$, where $f_r(n)$ is the *Ruth-Aaron* function replacing each prime factor with its $r-$th power. We also prove that the density of n remains 0 if $|f_r(n) - f_r(n+1)| \leq k(x)$, where $k(x)$ is a function of x with relatively low rate of growth. Moreover, we further the discussion of the infinitude of *Ruth-Aaron* numbers and provide a few possible directions for future study.

Keywords : Ruth-Aaron numbers; largest prime factors; multiplicative functions; rate of growth

Mathematics Subject Classification (2020) **:** 11N05; 11N56; 11P32

1 Introduction

On April 8, 1974, Henry Aaron hit his 715th major league home run, sending him past Babe Ruth, who had a 714, on the baseball's all-time list. As the event received much advance publicity, the numbers 714 and 715 were mentioned by millions of people including mathematicians at the time, whose attention likely deviated from the phenomenal baseball game and was attracted by the beautiful properties of the two consecutive numbers.

They first noticed that

$$714 \cdot 715 \;=\; 510510 \;=\; 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \;=\; P_7,$$

where P_k denotes the product of the first k primes. Without too much effort, we can find expressions for P_1, P_2, P_3, P_4 as the product of two consecutive integers. However, after 714 and 715, no more products turned up for integer pairs below 10^{6021}. They thus conjectured that 714 and 715 are the largest consecutive integer pair to be written as the product of the first k primes.

This conjecture is just the beginning of the beauty of the two integers. In fact, let the unique prime factorization of an integer n be $p_1^{a_1}p_2^{a_2}\cdots p_k^{a_k}$, and define

$$f(n) \;:=\; a_1p_1 + a_2p_2 + \cdots + a_kp_k. \tag{1}$$

Nelson, Penney, and Pomerance found that $f(714) = f(715)$ [24].

We call the function $f(n)$ the *Ruth-Aaron* function, an integer n with the property $f(n) = f(n+1)$ a *Ruth-Aaron* number, and the pair n, $n+1$ a *Ruth-Aaron* pair. A function f is completely additive if $f(ab) = f(a) + f(b)$ holds for all integers a, b. It easily follows that the *Ruth-Aaron* function has the nice property of being completely additive. A computer search for all the *Ruth-Aaron* numbers not exceeding 50000 found just 42 values (with the difference between adjacent numbers growing), suggesting that their density is 0.

n	$f(n) = f(n+1)$	n	$f(n) = f(n+1)$	n	$f(n) = f(n+1)$
5	5	5405	75	26642	193
8	6	5560	150	26649	66
15	8	5959	160	28448	144
77	18	6867	122	28809	117
125	15	8280	40	33019	149
714	29	8463	54	37828	211
948	86	10647	39	37881	93
1330	33	12351	205	41261	64
1520	32	14587	532	42624	57
1862	35	16932	107	43215	118
2491	100	17080	79	44831	480
3248	44	18490	93	44891	82
4185	45	20450	421	47544	299
4191	141	24895	401	49240	1242

Table 1: *Ruth-Aaron* numbers not exceeding $50,000$.

In fact, in 1978, only a few years after the famous baseball game, Erdős and Pomerance proved this result of density 0 [11]. They also established that when x is sufficiently large, the number of *Ruth-Aaron* numbers is at most $C \cdot \frac{x}{(\log x)^{1-\epsilon}}$ for every $0 < \epsilon < 1$, where $C = C(\epsilon)$ is a constant dependent upon ϵ.

In this paper, we extend the results obtained by Erdős and Pomerance. As an arithmetic function bearing certain similarities to the Sigma function and the Prime Omega function (See Appendix A), $f(n)$ renders several natural directions for generalization, one of which is to raise the prime factors to a power. Hence, we first introduce the $r-$th power *Ruth-Aaron* numbers.

Definition 1.1 *An integer $n = \prod_{i=1}^{k} p_i^{a_i}$ is a $r-$th power Ruth-Aaron number if*

$$f_r(n) \;=\; f_r(n+1), \tag{2}$$

where $f_r(n) = \sum_{i=1}^{k} a_i p_i^r$.

We prove an upper bound on the number of $r-$th power *Ruth-Aaron* numbers up to x in Section 4 (we will state the result later in the introduction). Our result improved that of Erdős and Pomerance by a factor of $(\log\log x)^3 \log\log\log x/\log x$. Moreover, inspired by Cohen, Cordwell, Epstein, Kwan, Lott, and Miller's study of near perfect numbers [6], we introduce the concept of *$k(x)$-near-Ruth-Aaron* numbers.

Definition 1.2 *An integer n is $k(x)$-near-Ruth-Aaron when*

$$|f_r(n) - f_r(n+1)| \;\leq\; k(x). \tag{3}$$

Obviously, when $k(x) = 0$, n is a $r-$th power Ruth-Aaron number.

As *Ruth-Aaron* numbers *seem* extremely rare, we weaken the condition and investigate how an absolute difference of a small amount between $f_r(n)$ and $f_r(n+1)$ affect the density in Section 3; in particular, rather than requiring $f_r(n)$ to equal $f_r(n+1)$, we merely require them to be "close." Moreover, Nelson, Penney, and Pomerance [24] proved the infinitude of *Ruth-Aaron* numbers under Schinzel's Conjecture, which provides us another direction for generalization. In Section 5, we prove that there are infinitely many real r such that there is a $r-$th power *Ruth-Aaron* number.

As our results are concerning only upper bounds for these numbers, we can and do absorb any set of numbers in the arguments below that is small relative to this bound into our error terms.

For future study, we can place *Ruth-Aaron* numbers in linear equations such as the Fibonacci sequence. We can initiate the study of *Rabonacci* numbers, or integer n with $f(n) = f(n-1) + f(n-2)$. Another possibility is to expand the equation $f(n) = f(n+1)$ to $k-$tuples. Inspired by Erdős [10], we conjecture that for every integer $k \geq 1$, there exists $n, n+1, \ldots, n+k$ such that

$$f(n) \;=\; f(n+1) \;=\; \cdots \;=\; f(n+k). \tag{4}$$

In fact, a computer search [25] tells us that even when $k = 2$, solutions are extremely rare.

n	$f(n) = f(n+1) = f(n+2)$
417 162	533
6 913 943 284	5428

Table 2: List of *Ruth-Aaron* triples below 10^{10}.

Due to the rarity of *Ruth-Aaron* triples, following the previous conjecture, we propose that the number of solutions to (4) is finite for any $k \geq 2$.

Although *Ruth-Aaron* numbers are named after two baseball stars (instead of a mathematician like most functions and theorems are) and thus have a more or less recreational origin, their study leads us to a variety of great mathematics, which all play key roles in this paper:

- the sieve method [16],
- the Prime Number Theorem [32],
- the Chinese Remainder Theorem [8],
- De Bruijn's estimation on integers with regards to the size of their largest prime factors [5],
- the Catalan Conjecture [29],
- the linear independence of radicals, and
- inequalities such as Cauchy-Schwarz and Jensen's.

The *Ruth-Aaron* numbers' profound connection to such mathematics not only once again proves the beauty of their properties, but should also justify why they merit a closer inspection.

1.1 Notations and Definitions

We set some notation and define the key objects of this paper.

Definition 1.3 *Numbers $f_1, f_2, \ldots, f_n$ are linearly independent over $\mathbb{Z}$ if, when coefficients $a_i \in \mathbb{Z}$, the following equation*

$$a_1 f_1 + a_2 f_2 + \cdots + a_n f_n \;=\; 0 \tag{5}$$

holds only when $a_i = 0$.

We use the following notations.

- We adopt a notation similar to the $\mathcal{A}_{\mathbf{a}}$ used by Luca and Stănică [18], to denote linear equations with the *Ruth-Aaron* function.
  Let $k \geq 1$ be a positive integer, and let $\mathbf{a} = (a_0, a_1, \ldots, a_k)$ be a vector with integer components not all zero. Put $\mathcal{R}^r_{\mathbf{a}}(x)$ for the set of all integers n not exceeding x such that

  $$\sum_{i=0}^{k} a_i f_r(n+i) \;=\; 0. \tag{6}$$

  Then it's not hard to notice that all integers n up to x with $f_r(n) = f_r(n+1)$ coincide with $\mathcal{R}^r_{(\mathbf{1},-\mathbf{1})}(x)$.

- We use $P(n)$ to denote the largest prime factor of integer n.
- We use $A(x,t)$ to denote the number of $n \le x$ with $P(n) \ge x^t$, and $a(x,t)$ to denote the fraction of $n \le x$ with $P(n) \ge x^t$.
- We use $\Psi(x,y)$ to denote the number of n up to x with prime factors no greater than y.
- Big O notation: We write $k(x) = O(g(x))$ or $k(x) \ll g(x)$ if there exists a positive constant C such that $k(x) < C \cdot g(x)$ for all sufficiently large x.
- Little o notation: We write $k(x) = o(g(x))$ if $\lim_{x\to\infty} k(x)/g(x) = 0$.
- We write $k(x) \sim g(x)$ if $\lim_{x\to\infty} k(x)/g(x) = 1$.
- We denote by E the constant

$$\begin{aligned} E &= \sum_{k=1}^{\infty} \frac{(\mu(k))^2}{k\varphi(k)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)} \\ &= 1.9435964368\ldots, \end{aligned} \tag{7}$$

  where $\zeta(x)$ is the Riemann Zeta function, $\mu(k)$ the Möbius function, and $\varphi(x)$ the Euler Totient function.

Remark 1.4 Our estimations of the number of n at most x that satisfy certain conditions are mostly expressed as $O(g(x))$, where $g(x)$ is a function of x. At the expense of tedious constant chasing we could make all the multiplicative constants explicit, but as we are concerned with the decay rate with respect to x there is no need to do so, and we choose big O notation for the sake of readability. Many approximations and scalings presented are not appear optimal, as there is no need since they smaller than our main term.

1.2 Main Results

We obtained the following main results as well as a few others which are not listed below but are introduced later with the proofs of the theorems. These results will be proved using lemmas from Section 2 and important theorems and conjectures in Appendix B.

Theorem 1.5 *For real $r \ge 1$ and every ϵ with $0 < \epsilon < 1$, let $\delta_0 = \delta_0(\epsilon)$ be a constant dependent upon ϵ. Let δ be subject to the following conditions:*

$$\begin{aligned} \delta &\le \delta_0 r\epsilon/14 \\ 0 < \delta &< \delta_0^2\epsilon/4E^2A \\ \delta &< \delta_0/4, \end{aligned} \tag{8}$$

where A (see [11]) is a fixed constant around 8. Then $k(x)$-near-Ruth-Aaron numbers have density 0 for any $k(x)$ such that

$$k(x) \le (x^{r\delta} - x^{-\delta} - 1)x^{r/\log\log x}. \tag{9}$$

This result indicates that when x is sufficiently large, if the difference between $f_r(n)$ and $f_r(n+1)$ is essentially less than $x^{\delta'}$, where δ' is arbitrarily small, then the density of n under this condition is still 0. Hence, not only are *Ruth-Aaron* numbers rare, $k(x)$-*near-Ruth-Aaron* numbers are also rare for $k(x)$ at most a small power of x. In particular, if $k(x)$ is a constant or a power of logarithm, $k(x)$-*near-Ruth-Aaron* numbers are likewise very rare.

Moreover, recall that $\#\mathcal{R}^r_{(1,-1)}(x)$ denotes the number of integers up to x with $f_r(n) = f_r(n+1)$. We are able to obtain the following new results of $r-$th power *Ruth-Aaron* number:

Theorem 1.6

- *When $r = -1$,*

$$\#\mathcal{R}^{-1}_{(1,-1)}(x) \ \ll \ x^{2(\log\log x/\log x)^{1/2}} \ = \ \exp\left(2(\log\log x \log x)^{1/2}\right), \tag{10}$$

  which means for every $\epsilon > 0$, we can find x sufficiently large such that $\#\mathcal{R}^{\epsilon r}_{(1,-1)}(x) \ll x^{\delta}$. In fact, when r is negative

$$\#\mathcal{R}^r_{(1,-1)}(x) \ \ll \ x^{O((\log\log x/\log x)^{r/r-1})}. \tag{11}$$

- *When r is rational but not an integer,*

$$\#\mathcal{R}^r_{(1,-1)}(x) \ = \ 0. \tag{12}$$

- *When $r \geq 1$ is real,*

$$\#\mathcal{R}^r_{(1,-1)}(x) \ = \ O\left(\frac{x\log\log\log x(\log\log x)^3}{(\log x)^2}\right). \tag{13}$$

Erdős and Pomerance [11] conjectured that there are infinitely many *Ruth-Aaron* numbers. While their conjecture is still open, we prove a related problem, namely the infinitude of r for which $\#\mathcal{R}^r_{(1,-1)}(x) > 0$.

Theorem 1.7 *There are infinitely many real numbers r such that $\#\mathcal{R}^r_{(1,-1)}(x) > 0$.*

1.3 Outline

In Section 2 we present a few preliminary results that provide a general overview for the problems we study, and which will be used extensively throughout the paper. Then, we discuss the proofs of Theorems 1.5, 1.6, and 1.7 (Sections 3, 4, and 5). We conclude with possible future research directions in Section 6.

2 Preliminary Results

We begin by generalizing some results from Erdős and Pomerance [11] which will be useful in proving our main theorems. Lemma 2.1 and Corollary 2.2 are introduced for the sake of proving Lemma 3.1, as two cases in terms of the size of $P(n)$ are taken care of by the corollary. Lemma 2.3 and Lemma 2.5 are frequently used in the sieve method to bound various sums over reciprocals of primes. Lemma 2.7 and Lemma 2.8 introduce an upper bound for $f_r(n)$ with regards to $P(n)^r$ and a lower bound for the size of $P(n)$. These results are used extensively throughout the paper.

Erdős and Pomerance [11] first introduced a well-known result due to Dickman [7] and others which bounds how often the larges prime factor of $n \leq x$ is at least n^t.

Lemma 2.1 *For every $x > 0$ and every t, $0 \leq t \leq 1$, let $A(x,t)$ denote the number of $n \leq x$ with $P(n) \geq x^t$. The function*

$$a(t) \ := \ \lim_{x\to\infty} x^{-1}A(x,t) \tag{14}$$

is defined and continuous on $[0,1]$.

Corollary 2.2 *From Lemma 2.1 we obtain that there exists $\delta_0 = \delta_0(\epsilon)$ sufficiently small $(0 < \delta_0 \leq \frac{1}{4})$ such that for large x, the number of $n \leq x$ with*

$$P(n) \ < \ x^{\delta_0} \text{ or } x^{1/2-\delta_0} \ \leq P(n) \ < \ x^{1/2+\delta_0} \tag{15}$$

is less than $\epsilon x/3$.

Proof. Let $\epsilon > 0$. From Lemma 2.1, $a(t) = \lim_{x\to\infty} x^{-1}A(x,t)$, which means the fraction of $n \leq x$ such that $P(n) < x^{\delta_0}$ converges to $a(0) - a(\delta_0)$ when $x \to \infty$. Similarly, the fraction of n which satisfy $x^{1/2-\delta_0} \leq P(n) < x^{1/2+\delta_0}$ converges to $a(1/2-\delta_0) - a(1/2+\delta_0)$.

As defined, $a(x,t)$ is the fraction of $n \leq x$ with $P(n) \geq x^t$. Consider $P(n) \leq x^{\delta_0}$ first. We can find δ_1 and X_1 such that $\forall\ \delta_0 \leq \delta_1$ and $x \geq X_1$, we have $a(x,0) - a(x,\delta_0) \leq \epsilon/8$ and within $\epsilon/2020$ of $a(0) - a(\delta_0)$. For the second condition, because $a(t)$ is continuous, given any ϵ we can always find δ_2 and X_2 such that if δ_0 is at most δ_2 and x is at least X_2 then $a(x,1/2-\delta_0) - a(x,1/2+\delta_0)$ is at most $\epsilon/8$ and within $\epsilon/2020$ of $a(1/2-\delta_0) - a(1/2+\delta_0)$. We take $\delta = \min(\delta_1,\delta_2)$ and $X = \max(X_1,X_2)$, then the fraction of $n \leq x$ satisfying one of the two conditions is no greater than $\epsilon/3$, which means the number of n is no greater than $\epsilon x/3$. □

Lemmas 2.3 and 2.5 are used frequently in later sections to bound various sums over reciprocals of primes.

Lemma 2.3 *We have*

$$\begin{aligned} \sum_{e^u \le p \le e^v} \frac{1}{p} &< \log(v/u) + \frac{C}{u} \\ \sum_{u_0 \le p \le v_0} \frac{1}{p} &< \frac{C + \log(v_0/u_0)}{\log u_0}. \end{aligned} \tag{16}$$

Proof. We use the Abel Partial Summation Formula:

$$\sum_{1 \le x \le n} c_x f(x) = c_n f(n) - \int_1^n c_t f'(t) dt, \tag{17}$$

and the Prime Number Theorem (weaker version): if $\pi(n)$ is the number of at most n, then

$$\pi(n) = \frac{n}{\log n} + o\left(\frac{n}{(\log n)^2}\right). \tag{18}$$

First of all, we prove that

$$\sum_{p \le n} \frac{\log p}{p} = \log n + O(1). \tag{19}$$

We know that the power of p in $n!$ equals $\lfloor \frac{n}{p} \rfloor + \lfloor \frac{n}{p^2} \rfloor + \lfloor \frac{n}{p^3} \rfloor + \cdots$. Thus, we have

$$\begin{aligned} \log(n!) &= \sum_{p \le n} \log p \left(\left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \cdots \right) \\ &\le n \cdot \sum_{p \le n} \frac{\log p}{p} + \sum_{p \le n} \log p \left(\frac{n}{p^2} + \frac{n}{p^3} + \cdots \right) \\ &= n \cdot \sum_{p \le n} \frac{\log p}{p} + n \sum_{p \le n} \frac{\log p}{p(p-1)}. \end{aligned} \tag{20}$$

Since $\sum_{p \le n} \frac{\log p}{p(p-1)} < \sum_{k=2}^{n} \frac{\log k}{k(k-1)}$ converges, we have $n \sum_{p \le n} \frac{\log p}{p(p-1)} = O(n)$. Therefore,

$$\log(n!) = n \cdot \sum_{p \le n} \frac{\log p}{p} + O(n). \tag{21}$$

On the other hand,

$$\begin{aligned} \log(n!) &= \sum_{k=1}^{n} \log k = \int_1^n \log x \; dx + O(\log n) \\ &= n \log n - n + O(\log n) \\ &= n \log n + O(n). \end{aligned} \tag{22}$$

Comparing the two results, we obtain (19):

$$\sum_{p\leq n}\frac{\log p}{p} \;=\; \log n + O(1). \tag{23}$$

We now use the Abel Partial Summation Formula. Let $f(x) = \frac{1}{\log x}$ and $c_n = \sum_{p\leq n}\frac{\log p}{p}$, then we have

$$\begin{aligned}\sum_{p\leq n}\frac{1}{p} &= \frac{c_n}{\log n} + \int_1^n \frac{c_t}{t(\log t)^2}dt \\ &= 1 + \frac{O(1)}{\log n} + \int_2^n \frac{1}{t\log t}dt + \int_2^n \frac{O(1)}{t(\log t)^2}dt \\ &= \log\log n + O\left(\frac{1}{\log n}\right) + 1. \end{aligned}\tag{24}$$

Then for primes e^u, e^v,

$$\begin{aligned}\sum_{e^u\leq p\leq e^v}\frac{1}{p} &= \sum_{p\leq e^v}\frac{1}{p} - \sum_{p\leq e^u}\frac{1}{p} + \frac{1}{e^u} \\ &= \log(v/u) + O\left(\frac{1}{u}\right) \\ &\leq \log(v/u) + \frac{O(1)}{u}. \end{aligned}\tag{25}$$

Therefore, we have

$$\sum_{e^u\leq p\leq e^v}\frac{1}{p} \;\leq\; \frac{O(1) + u\log(v/u)}{u}. \tag{26}$$

Let $u_0 = e^u, v_0 = e^v$, then

$$\sum_{u_0\leq p\leq v_0}\frac{1}{p} \;\leq\; \log\left(\frac{\log v_0}{\log u_0}\right) + \frac{C}{\log u_0}. \tag{27}$$

We now prove that $\log\left(\frac{\log v_0}{\log u_0}\right) + \frac{C}{\log u_0} < \frac{C+\log(v_0/u_0)}{\log u_0}$. Since we have $t < e^{t-1}$ for any $t > 1$,

$$\begin{aligned}\frac{v}{u} &< e^{v/u-1} \\ \log(v/u) &< \frac{v}{u} - 1 \\ u\log(v/u) &< (v-u) \\ \log u_0 \log\left(\frac{\log v_0}{\log u_0}\right) &< \log(v_0/u_0). \end{aligned}\tag{28}$$

Therefore, we have

$$\sum_{u_0 \le p \le v_0} \frac{1}{p} \;<\; \frac{C + \log(v_0/u_0)}{\log u_0}. \tag{29}$$

□

Remark 2.4 We have two inequalities in Lemma 2.3; (29) is weaker than (26) as we replace u/v by $e^{u/v} - 1$, and the following theorems apply mostly the result from (29). In fact, it turns out that in this paper the inequality in (29) is tight enough and applicable in most cases, and we will adopt the inequality in (26) otherwise.

Lemma 2.5 *We have*

$$\sum_{p \ge t} \frac{1}{p \log p} \;=\; \frac{1}{\log t} + O\left(\frac{1}{(\log t)^2}\right). \tag{30}$$

Proof. We use Abel's Partial Summation Formula. Let

$$\begin{aligned} S(t) \;&=\; \sum_{p \ge t} \frac{1}{p \log p} = \sum_{p \ge t} \frac{\log p}{p} \cdot (\log p)^{-2}. \\ A(x) \;&= \sum_{p \le x} \frac{\log p}{p} - \sum_{p \le t} \frac{\log p}{p} = \log x - \log t + O(1). \\ f(t) \;&=\; \frac{1}{(\log t)^2},\; f'(t) \;=\; -\frac{2}{t(\log t)^3}. \end{aligned} \tag{31}$$

Because $A(x)$ concerns primes in the interval $(t, x]$, the following sum differs from the one in (31) by at most the term $\frac{1}{t \log t}$, which can be absorbed in the error; thus, it is sufficient to study $S(t)$:

$$\begin{aligned} S(t) \;&=\; \frac{1}{t \log t} + \lim_{x \to +\infty} A(x) f(x) - \int_t^\infty A(x) f'(x)\, dx \\ &=\; \frac{1}{t \log t} + 2 \int_t^\infty (\log x - \log t + O(1)) \frac{1}{x (\log x)^3}\, dx \\ &=\; \frac{1}{t \log t} + 2 \int_t^\infty \frac{1}{x (\log x)^2}\, dx - 2 \int_t^\infty (\log t - O(1)) \frac{1}{x (\log x)^3}\, dx \\ &=\; \frac{1}{t \log t} - 2 \frac{1}{\log x}\bigg|_t^\infty + \frac{\log t}{(\log x)^2}\bigg|_t^\infty - \frac{O(1)}{(\log x)^2}\bigg|_t^\infty \\ &=\; \frac{2}{\log t} - \frac{1}{\log t} + O\left(\frac{1}{(\log t)^2}\right) \\ &=\; \frac{1}{\log t} + O\left(\frac{1}{(\log t)^2}\right), \end{aligned} \tag{32}$$

which completes our proof. □

Remark 2.6 We also have

$$\sum_{p\geq t}\frac{1}{p(\log p)^2} \;<\; \frac{1}{(\log t)^2}+O\left(\frac{1}{(\log t)^3}\right). \tag{33}$$

Proof. Let

$$\begin{aligned} S(t) &= \sum_{p\geq t}\frac{1}{p(\log p)^2} \\ A(x) &= \sum_{t\leq p\leq x}\frac{1}{p} \;<\; \frac{C+\log(x/t)}{\log t} \\ f(t) &= \frac{1}{(\log t)^2},\; f'(t) \;=\; -\frac{2}{t(\log t)^3}. \end{aligned} \tag{34}$$

Then we have

$$\begin{aligned} S(t) &= \lim_{x\to+\infty}A(x)f(x)-\int_t^\infty A(x)f'(x)\;dx \\ &< 2\int_t^\infty\left(\frac{C+\log x-\log t}{\log t}\right)\frac{1}{x(\log x)^3}\;d(x) \\ &= 2\left(\int_t^\infty\left(\frac{C}{\log t}-1\right)\frac{1}{x(\log x)^3}\;dx+\int_t^\infty\frac{1}{x(\log x)^2\log t}\;dx\right) \\ &= \frac{1}{(\log x)^2}\bigg|_t^\infty-\frac{O(1)}{\log t(\log x)^2}\bigg|_t^\infty-\frac{2}{\log x\log t}\bigg|_t^\infty \\ &= -\frac{1}{(\log t)^2}+\frac{O(1)}{(\log t)^3}+\frac{2}{(\log t)^2} \\ &= \frac{1}{(\log t)^2}+O\left(\frac{1}{(\log t)^3}\right). \end{aligned} \tag{35}$$

□

Lemma 2.7 *If $P(n)\geq 5$ and $r\geq 1$, we have*

$$f_r(n) \;\leq\; P(n)^r\cdot\frac{\log n}{\log P(n)}. \tag{36}$$

Proof. Consider the function $g(x)=\frac{x^r}{\log x}$, where r is a real number and $x\geq e^{1/r}$, then

$$\begin{aligned} g'(x) &= \frac{rx^{r-1}\log x-x^r\cdot 1/x}{(\log x)^2} \\ &= \frac{x^{r-1}(r\log x-1)}{(\log x)^2} \\ &> 0, \end{aligned} \tag{37}$$

which means $g(x)$ increases when $x \geq e^{1/r}$. Without the loss of generality, let $p_1 = P(n)$, then we have

$$\begin{aligned} f_r(n) &= \sum_{i=1}^{k} a_i p_i^r \\ &\leq \sum_{i=1}^{k} a_i p_1^r \cdot \frac{\log p_i}{\log p_1} \\ &= \frac{P(n)^r}{\log P(n)} \cdot \sum_{i=1}^{k} \log p_i^{a_i} \\ &= P(n)^r \cdot \frac{\log n}{\log P(n)}, \end{aligned} \tag{38}$$

which completes our proof. □

Lemma 2.8 *The number of n up to x not satisfying*

$$P(n) > x^{1/\log\log x} \text{ and } P(n+1) > x^{1/\log\log x} \tag{39}$$

is at most $O\left(\frac{x}{(\log x)^2}\right)$.

Proof. Let $\Psi(x, y)$ denote the number of n up to x with prime factors no greater than y, and set $u = \log x / \log y$. A result from De Bruijn [5] states that if $(\log x)^2 < y \leq x^{1/3}$ then

$$\log \Psi(x, y) < x(\log y)^2 \exp\left(-u \log u - u \log\log u + O(u)\right). \tag{40}$$

We replace y with $x^{1/\log\log x}$ and find

$$\begin{aligned} \Psi(x, y) &< x \left(\frac{\log x}{\log\log x}\right)^2 \cdot \exp\left(O(\log\log x) - \log\log x(\log\log\log x + \log\log\log\log x)\right) \\ &< x \left(\frac{\log x}{\log\log x}\right)^2 \cdot \exp\left(O(-\log\log x \log\log\log x)\right) \\ &< O\left(x \left(\frac{\log x}{\log\log x}\right)^2 \cdot \frac{1}{(\log x)^{\log\log\log x}}\right) \\ &< O\left(\frac{x}{(\log x)^2}\right). \end{aligned} \tag{41}$$

Therefore, the number of n for which

$$P(n) > x^{1/\log\log x} \text{ and } P(n+1) > x^{1/\log\log x}, \tag{42}$$

doesn't hold is $O(x/(\log x)^2)$. □

Remark 2.9 We hence know that for the majority of integers no greater than x, $P(n) > x^{1/\log\log x}$, which means $P(n)$ is typically larger than $\log x$ to any power. Moreover, De Koninck and Ivić [20] showed that the sum of largest prime factors of integers up to x is of size $x^2/\log x$, which means the average largest prime factor of integers up to x is of size $x/\log x$. When x is sufficiently large, $x/\log x > x^\delta$ for any $0 < \delta < 1$. Therefore, the average largest prime factor is greater than x to any power less than 1, indicating that a considerable number of $P(n)$ are very large.

3 Proof of Theorem 1.5

In this section we prove Theorem 1.5. We first introduce the following two lemmas generalized from Erdős and Pomerance [11], the first of which indicates that the largest prime factors of two consecutive integers are usually far apart, and the second proves that $P(n)^r$ is often the dominating element that determines the size of $f_r(n)$.

Lemma 3.1 *For each $0 < \epsilon < 1$, there is a $\delta > 0$ such that for sufficiently large x, the number of $n \le x$ with*

$$x^{-\delta} \ < \ P(n)/P(n+1) \ < \ x^{\delta} \tag{43}$$

is less than ϵx.

Proof. We know from Corollary 2.2 that $\exists\ \delta_0 = \delta_0(\epsilon)$ sufficiently small $(0 < \delta_0 \le 1/4)$ such that for large x, the number of $n \le x$ with

$$P(n) \ < \ x^{\delta_0} \text{ or } x^{1/2-\delta_0} \ \le P(n) \ < \ x^{1/2+\delta_0} \tag{44}$$

is less than $\epsilon x/3$. Now we consider the remaining cases:

$$\text{(i) } x^{\delta_0} \ \le \ P(n) \ < \ x^{1/2-\delta_0}$$
$$\text{(ii) } P(n) \ \ge \ x^{1/2+\delta_0}.$$

We will show that for every $0 < \epsilon < 1$, there exists δ such that such that the number of $n \le x$ satisfying one of (i) and (ii) while (43) holds is less than $\epsilon x/3$.

We consider (i) first. We know that for each pair of primes p, q, there are at most $1 + \lfloor \frac{x}{pq} \rfloor$ choices[1] of $n \le x$ for which $P(n) = p,\ P(n+1) = q$. For inequality (43) to hold, $px^{-\delta} < q < px^{\delta}$. Then for large x, the number of $n \le x$ in case (43) holds is (we may

[1]This is because the number of $n \le x$ such that $P(n) = p, P(n+1) = q$ is bounded by the number of $n \le x$ such that $n \equiv 0 \pmod p$ and $n \equiv -1 \pmod q$. By Chinese Remainder Theorem, $n \equiv bp \pmod{pq}$ which means there are at most $1 + \left\lfloor \frac{x}{pq} \right\rfloor$ such $n \le x$.

assume $0 < \delta < \delta_0/4$)

$$\begin{aligned}
\sum_{\substack{x^{\delta_0} \le p < x^{1/2-\delta_0} \\ px^{-\delta} < q < px^{\delta}}} \left(1 + \left\lfloor \frac{x}{pq} \right\rfloor\right) &< x^{1-2\delta_0+\delta} + x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p} \sum_{px^{-\delta} < q < px^{\delta}} \frac{1}{q} \\
&< x^{1-2\delta_0+\delta} + x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p} \cdot \frac{C + \log(px^{\delta}/px^{-\delta})}{\log(px^{-\delta})} \text{ (Lemma 2.3)} \\
&< x^{1-2\delta_0+\delta} + x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p} \cdot \frac{C + \log x^{2\delta}}{\log(px^{-\delta})}. \qquad (45)
\end{aligned}$$

Moreover, since $p \ge x^{\delta_0} > x^{4\delta} > x^{3\delta}$, we have $x^{-\delta} > p^{-1/3}$,

$$\begin{aligned}
x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p} \cdot \frac{C + \log(x^{2\delta})}{\log(px^{-\delta})} &< x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p} \cdot \frac{C + \log(x^{2\delta})}{\log(p^{2/3})} \\
&= x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p} \cdot \frac{\frac{3}{2}C + \log(x^{3\delta})}{\log p}. \qquad (46)
\end{aligned}$$

When x is sufficiently large, we have

$$\begin{aligned}
x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p} \cdot \frac{C + \log(x^{2\delta})}{\log(px^{-\delta})} &< x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p} \cdot \frac{1.1 \log(x^{3\delta})}{\log p} \\
&< \frac{3.3}{3} \cdot 3\delta x \log x \sum_{x^{\delta_0} \le p < x^{1/2-\delta_0}} \frac{1}{p \log p} \\
&< \frac{4}{3} \cdot 3\delta x \log x \frac{1}{\delta_0 \log x} \text{ (Lemma 2.5)} \\
&< 4x\delta/\delta_0. \qquad (47)
\end{aligned}$$

Therefore, we have

$$\sum_{\substack{x^{\delta_0} \le p < x^{1/2-\delta_0} \\ px^{-\delta} < q < px^{\delta}}} \left(1 + \left\lfloor \frac{x}{pq} \right\rfloor\right) < x^{1-2\delta_0+\delta} + 4\delta x/\delta_0 \qquad (48)$$

If we choose δ such that

$$\delta < \delta_0 \epsilon/13, \qquad (49)$$

then (48) implies there are less than $\epsilon x/3$ choices of n.

Now we consider case (ii). Let $a = n/P(n)$ and $b = (n+1)/P(n+1)$. Then $a < x/x^{1/2+\delta_0} = x^{1/2-\delta_0}$, and because (3.1) holds, $b \le \lfloor x^{1/2-\delta_0+\delta} \rfloor + 1$ ($b = x^{1/2-\delta_0+\delta}$ is possible

only when $x^{1/2-\delta_0+\delta} \in \mathbb{Z}$)[2], and $x^{-\delta}/2 < a/b = nP(n+1)/(n+1)P(n) < 2x^{\delta}$. Meanwhile, when a, b are fixed, the number of $n \le x$ for which $n = aP(n),\ n+1 = bP(n+1)$ is at most the number of primes $p \le x/a$ such that $(ap+1)/b$ is a prime. Bézout's Identity (or the Euclidean algorithm) gives that for integers a, b, there always exists $m, n \in \mathbb{Z}$ such that $ma + nb = \gcd(a, b)$. Now, because $\gcd(a, b) = 1$ and $2 \mid ab$, the number of p with $bq - ap = 1$ is greater than 1. Moreover, given $(ap+1)/b$ is an integer, all p are in the same residue class modulo b. Let $p = kb + c$, where $k, c \in \mathbb{Z}_+$ and $c \in [0, b-1]$. Let $d = (ac+1)/b$. Then we are counting positive integer k not exceeding x/ab with primes $kb + c = P(n)$, $ka + d = P(n+1)$. Let the number of such k be Pairs(x). By Brun's Method of sifting by a thin set of primes [16], we have

$$\begin{aligned} \text{Pairs}(x) &\le \frac{Ax}{ab\log^2(x/ab)} \prod_{p \mid ab} \left(1 - \frac{1}{p}\right)^{-1} \\ &= \frac{Ax}{\varphi(a)\varphi(b)\log^2(x/ab)}, \end{aligned} \tag{50}$$

where A is a constant of size around 8 and φ is Euler's totient function, or the number of integers up to n that are relatively prime to n. Because we are investigating only the $x-$dependent components and not the multiplicative constants, our only concern here is the size of $1/\log^2(x/ab)$ in relation to the change of a, b. In particular, as all summations are positive, no cancellation is involved, and thus it suffices to show this sum is of the same size for all a, b in our ranges.

We now bound above and below $1/\log^2(x/ab)$. Because $a, b \in \mathbb{N}$,

$$\frac{1}{\log^2(x/ab)} > \frac{1}{(\log x)^2} \tag{51}$$

Meanwhile,

$$\begin{aligned} \frac{1}{\log^2(x/ab)} &< \frac{1}{\log^2(x/(x^{2\cdot(1/2-\delta_0)} \cdot 2x^{\delta}))} \\ &= \frac{1}{\log^2(2x^{2\delta_0-\delta})} \\ &< \frac{2}{(2\delta_0-\delta)^2(\log x)^2}. \end{aligned} \tag{52}$$

This shows us that we can remove ab at a cost of a multiplicative change in the result. We now use the result of Landau [17]: if $E = \zeta(2)\zeta(3)/\zeta(6)$, then $\sum_{n \le x} 1/\varphi(n) = E \log x +$

[2]This is because $b = \frac{n+1}{P(n+1)} < \frac{x+1}{P(n)} \cdot x^{\delta} < \frac{x+1}{x^{1/2+\delta_0}} \cdot x^{\delta} = \frac{x+1}{x} \cdot x^{1/2-\delta_0+\delta}$, meaning that $b \le \lfloor \frac{x+1}{x} x^{1/2-\delta_0+\delta} \rfloor = \lfloor x^{1/2-\delta_0+\delta} + x^{-1/2-\delta_0+\delta} \rfloor$. Because $-1/2 - \delta_0 + \delta < -1/2$, when x is sufficiently large, we have $\lfloor x^{1/2-\delta_0+\delta} + x^{-1/2-\delta_0+\delta} \rfloor \le \lfloor x^{1/2-\delta_0+\delta} \rfloor + 1$; therefore, $b \le \lfloor x^{1/2-\delta_0+\delta} \rfloor + 1$.

$o(1)$. Therefore, using (52),

$$\begin{aligned}
\text{Pairs}(x) &< \frac{2Ax}{(2\delta_0-\delta)^2(\log x)^2}\sum_{a\le x^{1/2-\delta_0}}\frac{1}{\varphi(a)}\sum_{ax^{-\delta}/2<b<2ax^{\delta}}\frac{1}{\varphi(b)}\\
&< \frac{2Ax}{(2\delta_0-\delta)^2(\log x)^2}\sum_{a\le x^{1/2-\delta_0}}\frac{E\log(2ax^{\delta}/\frac{ax^{-\delta}}{2})+o(1)}{\varphi(a)}\\
&< \frac{Ax}{(2\delta_0-\delta)^2(\log x)^2}\sum_{a\le x^{1/2-\delta_0}}\frac{3E\log x^{2\delta}+o(1)}{\varphi(a)}\\
&< \frac{7EA\delta x}{(2\delta_0-\delta)^2\log x}\sum_{a\le x^{1/2-\delta_0}}\frac{1}{\varphi(a)}\\
&< \frac{8E^2A\delta x(1/2-\delta_0)\log x}{(2\delta_0-\delta)^2\log x}\\
&< \frac{4E^2A\delta x}{(2\delta_0-\delta)^2}. \qquad (53)
\end{aligned}$$

Let

$$0 \;<\; \delta \;<\; \delta_0^2\epsilon/4E^2A \text{ and } \delta \;<\; \delta_0/4, \qquad (54)$$

then (53) implies there are fewer than $\epsilon x/3$ choices for such n. Thus, if we choose δ such that both (49) and (54) hold, then there are less than ϵx choices of n for every sufficiently large x, completing our proof. □

Remark 3.2 For our purposes, the estimation in inequalities (45) and (46) is sufficient, but if we substitute $\sum_{px^{-\delta}<q<px^{\delta}} 1/q$ with $\log\frac{\log px^{\delta}}{\log px^{-\delta}}+C/\log px^{-\delta}$, in other words, if we use inequality (27) rather than (29), then with a bit more work we could get $\delta < \frac{\epsilon\delta_0}{6.12\log(1/2\delta_0)}$.

Remark 3.3 Moreover, we can easily extend the result to $r \ge 1$. From (49) and (54) we know δ depends on ϵ, and because ϵ is arbitrary between 0 and 1, δ can be very small. For every $0 < \epsilon < 1$, we find $\delta' = r\cdot\delta$ (where δ satisfies (49) and (54); hence δ' can be very small) such that for sufficiently large x, the number of $n \le x$ with

$$x^{-\delta'} \;<\; P(n)^r/P(n+1)^r \;<\; x^{\delta'} \qquad (55)$$

is less than ϵx.

Lemma 3.4 *When $r \ge 1$, for every $\epsilon > 0$, let $\delta = \delta_0 r\epsilon/14$, then for sufficiently large x there are at least $(1-\epsilon)x$ choices for composite integer $n \le x$ such that*

$$P(n)^r \;\le\; f_r(n) \;<\; (1+x^{-\delta})P(n)^r. \qquad (56)$$

Proof. We know that any n at most x is divisible by at most $\log x/\log 2$ primes. By the Prime Number Theorem, the number of primes up to x is $O(x/\log x)$. Then for any $0 < \epsilon_0 < 1$, we can always find sufficiently large x such that $O(x/\log x) = \epsilon_0 x$, which means the number of prime n up to x is $o(x)$ and can be absorbed. Thus, we have for sufficiently large x and composite n:

$$\begin{aligned} f_r(n) &= P(n)^r + f_r(n/P(n) \\ &\leq P(n)^r + P(n/P(n))^r \cdot \log x/\log 2 \\ &< P(n)^r + P(n/P(n))^r \cdot x^{\delta} \end{aligned} \tag{57}$$

for any fixed δ. We prove that there are at most ϵx choices of $n \leq x$ such that

$$f_r(n) \ \geq \ (1 + x^{-\delta})P(n)^r \tag{58}$$

holds for all except $o(x)$ choices of $n \leq x$. Then, for such n, if (58) holds, from (57) we have

$$P(n/P(n))^r \ > \ x^{-2\delta} \cdot P(n)^r. \tag{59}$$

Let $\epsilon > 0$. We know from Corollary 2.2 that there exists $\delta_0 = \delta_0(\epsilon)$ such that for large x the number of $n \leq x$ with $P(n) < x^{\delta_0}$ is at most $\epsilon x/3$. Meanwhile, we know for each pair of primes p, q, there are at most $\lfloor \frac{x}{pq} \rfloor$ choices of $n \leq x$ with $P(n)^r = p^r$ and $P(n/P(n))^r = q^r$. Hence, from (59), for large x the number of $n \leq x$ for which (59) doesn't hold is at most (assume $0 < \delta < \delta_0/3$):

$$\begin{aligned} o(x) + \frac{\epsilon x}{3} + \sum_{\substack{x^{\delta_0} < p \\ x^{-2\delta/r} p < q \leq p}} \left\lfloor \frac{x}{pq} \right\rfloor &< \frac{\epsilon x}{2} + x \sum_{x^{\delta_0} < p} \frac{1}{p} \sum_{x^{-2\delta/r} p < q \leq p} \frac{1}{q} \\ &< \frac{\epsilon x}{2} + x \sum_{x^{\delta_0} < p} \frac{1}{p} \cdot \frac{C + (2\delta \log x)/r}{\log(x^{-2\delta/r} p)} \\ &< \frac{\epsilon x}{2} + \frac{6\delta}{r} \cdot x \log x \sum_{x^{\delta_0} < p} \frac{1}{p \log p} \\ &< \frac{\epsilon x}{2} + \frac{7\delta x}{r\delta_0} \text{ (Lemma 2.5)} \end{aligned} \tag{60}$$

where $o(x)$ accounts for all $n \leq x$ of the form $n = P(n)$. We take $\delta = \delta_0 r\epsilon/14$, then (60) is no greater than ϵx. Therefore, the number of $n \leq x$ such that $P(n)^r < f_r(n) < (1 + x^{-\delta} P(n)^r)$ is at least $(1 - \epsilon)x$, completing the proof. □

We recall Theorem 1.5.

Theorem 1.5 *For real $r \geq 1$ and every ϵ with $0 < \epsilon < 1$, let $\delta_0 = \delta_0(\epsilon)$ be a constant dependent upon ϵ. Let δ be subject to the following conditions:*

$$\begin{aligned} \delta &\leq \delta_0 r\epsilon/14 \\ 0 < \delta &< \delta_0^2\epsilon/4E^2A \\ \delta &< \delta_0/4, \end{aligned} \tag{61}$$

where A (see [11]) is a fixed constant around 8. Then $k(x)$-near-Ruth-Aaron numbers have density 0 for any $k(x)$ such that

$$k(x) \leq (x^{r\delta} - x^{-\delta} - 1)x^{r/\log\log x}. \tag{62}$$

Now, we know from Lemma 3.1 that there are less than ϵx choices of n which satisfy

$$x^{-r\delta} < \frac{P(n)^r}{P(n+1)^r} < x^{r\delta}. \tag{63}$$

Without loss of generality, let $P(n) > P(n+1)$; the other case is handled similarly[3]. Then there are less than ϵx choices of n with

$$(x^{-r\delta} - 1)P(n+1)^r < P(n)^r - P(n+1)^r < (x^{r\delta} - 1)P(n+1)^r. \tag{64}$$

Because $r\delta > 0$, the LHS of (64) is negative, which means there are less than ϵx choices of n with

$$0 < P(n)^r - P(n+1)^r < (x^{r\delta} - 1)P(n+1)^r. \tag{65}$$

Then for at least $(1-\epsilon)x$ choices of n, we have

$$P(n)^r - P(n+1)^r > (x^{r\delta} - 1)P(n+1)^r. \tag{66}$$

Meanwhile, we know from Lemma 3.4 that for all but ϵx choices of n we have

$$\begin{aligned} P(n)^r < f_r(n) &< (1 + x^{-\delta})P(n)^r \\ P(n+1)^r < f_r(n+1) &< (1 + x^{-\delta})P(n+1)^r. \end{aligned} \tag{67}$$

Therefore, there are more than $(1-\epsilon)x$ choices of n with

$$\begin{aligned} f_r(n) - f_r(n+1) &> P(n)^r - (1 + x^{-\delta})P(n+1)^r \\ &> (x^{r\delta} - x^{-\delta} - 1)P(n+1)^r \\ &> (x^{r\delta} - x^{-\delta} - 1)x^{r/\log\log x}, \end{aligned} \tag{68}$$

which means the density of n with

$$|f_r(n) - f_r(n+1)| < (x^{r\delta} - x^{-\delta} - 1)x^{r/\log\log x} \tag{69}$$

is 0. In fact, when x is sufficiently large, the RHS of (69) is greater than x^δ, $(\log x)^k$, for any k, or $O(1)$, which means the density of n up to x with $|f_r(n) - f_r(n+1)| < k(x)$, where $k(x)$ is one of the functions above, is also 0.

[3] One of the Erdős-Turán conjectures asserts that the asymptotic density of $n \leq x$ with $P(n) > P(n+1)$ is $\frac{1}{2}$ [11]. Erdős and Pomerance showed that the number of n up to x with $P(n) > P(n+1)$ is greater than $0.0099x$ [11]. This result was recently improved by Lü and Wang, who proved that the density is larger than 0.2017 [19].

4 Proof of Theorem 1.6

Recall that the notation $\#\mathcal{R}^r_{(1,-1)}(x)$ denotes the number of integers up to x with $f_r(n) = f_r(n+1)$. Recall Theorem 1.6.

Theorem 1.6

- *When $r = -1$,*

$$\#\mathcal{R}^{-1}_{(1,-1)}(x) \ \ll \ x^{2(\log\log x/\log x)^{1/2}} \ = \ \exp\left(2(\log\log x \log x)^{1/2}\right), \tag{70}$$

which means for every $\epsilon > 0$, we can find x sufficiently large such that $\#\mathcal{R}^{\text{`}r}_{(1,-1)}(x) \ll x^{\delta}$. In fact, when r is negative

$$\#\mathcal{R}^r_{(1,-1)}(x) \ \ll \ x^{O((\log\log x/\log x)^{r/r-1})}. \tag{71}$$

- *When r is rational but not an integer,*

$$\#\mathcal{R}^r_{(1,-1)}(x) \ = \ 0. \tag{72}$$

- *When $r \geq 1$ is real,*

$$\#\mathcal{R}^r_{(1,-1)}(x) \ = \ O\left(\frac{x\log\log\log x(\log\log x)^3}{(\log x)^2}\right). \tag{73}$$

We first show that the number of $n \leq x$ when $r = -1$ is less than $x^{2(\log\log x/\log x)^{1/2}}$, which means for a fixed δ arbitrarily small and x sufficiently larger, $\#\mathcal{R}^{-1}_{(1,-1)}(x) \leq x^{\delta}$. Then, we will show, using linear independence, that when r is a non-integer rational, $r-$th power *Ruth-Aaron* numbers do not exist. Last, we will present an initial result by Erdős and Pomerance regarding the number of *Ruth-Aaron* numbers [11] before generalizing it to $r \geq 1$.

4.1 Negative r

Proof. First, we prove that when $r = -1$, $r-$th power *Ruth-Aaron* number n exists only when, in the unique prime factorization of n and $n+1$, the power of each prime is a multiple of the prime. In other words,

$$\begin{aligned} n \ &= \ p_1^{a_1} \cdot p_2^{a_2} \cdots p_k^{a_k} \\ n+1 \ &= \ q_1^{b_1} \cdot q_2^{b_2} \cdots q_l^{b_l}, \end{aligned} \tag{74}$$

where $p_i | a_i$ and $q_i | b_i$.

Let $a_i', b_i' \in \mathbb{Q}$, and $a_i' = \frac{a_i}{p_i}$, $b_i' = \frac{b_i}{q_i}$, where $a_i, b_i \in \mathbb{N}$. Then we have

$$\begin{aligned} a_1' + a_2' + \cdots + a_k' \ &= \ f_{-1}(n) \ = \ f_{-1}(n+1) \ = \ b_1' + b_2' + \cdots + b_l' \\ \frac{a_1}{p_1} + \frac{a_2}{p_2} + \cdots + \frac{a_k}{p_k} \ &= \ \frac{b_1}{q_1} + \frac{b_2}{q_2} + \cdots \frac{b_l}{q_l} \\ ((a_1 p_2 \cdots p_k) + \cdots + (a_k p_1 \cdots p_{k-1}))q_1 \cdots q_l \ &= \ ((b_1 q_2 \cdots q_l) + \cdots + (b_l q_1 \cdots q_{l-1}))p_1 \cdots p_k. \end{aligned} \tag{75}$$

It's obvious that the left hand side has to be divisible by $p_1 p_2 \cdots p_k$. Since $\gcd(n, n+1) = 1$, $\gcd(q_i, p_j) = 1$, which means

$$\begin{aligned} a_1 p_2 \cdots p_k + \cdots + a_k p_1 \cdots p_{k-1} &\equiv 0 \pmod{p_i} \\ a_i p_1 \cdots p_{i-1} p_{i+1} \cdots p_k &\equiv 0 \pmod{p_i} \\ a_i &\equiv 0 \pmod{p_i}, \end{aligned} \tag{76}$$

which means each $p_i | a_i$. Similarly, $q_i | b_i$.

Next, we rewrite n as

$$n = p_1^{a_1} \cdot p_2^{a_2} \cdots p_t^{a_t} \cdot \prod_{i>t} p_i^{a_i}. \tag{77}$$

Without loss of generality, let n be odd. If n is even, then we instead analyze $n+1$. Let $p_1 < p_2 < \cdots < p_k$, $q_1 < q_2 < \cdots < q_l$, then $\log_{p_i} x < \log x$ and $p_i \geq 3$. We have

$$\begin{aligned} p_i^{a_i' p_i} &\leq n \\ p_i \log p_i \leq a_i' p_i \log p_i &\leq \log n \\ p_i < \log n &\leq \log x. \\ a_i' \leq (\log_{p_i} n)/p_i &< (\log_{p_i} x)/3, \end{aligned} \tag{78}$$

which means for each p_i, where $i = 1, 2, \ldots, t$, we can choose a_i from $\{0, 1, \ldots, \log_{p_i} x/3\}$. Because $p_i \geq 3$, there are at most $(\log_{p_i} x)/3 + 1 \leq \log x/3$ choices of each a_i, hence the number of choices of the first t prime powers is at most $((\log x)^2/3)^t$. Moreover, because $p_i > t$ for $i > t$, we have

$$\begin{aligned} \left(\prod_{i>t}^{k} p_i^{a_i'} \right)^t &= \prod_{i>t}^{k} p_i^{a_i' t} \\ &< \prod_{i>t}^{k} p_i^{a_i' p_i} = \prod_{i>t}^{k} p_i^{a_i} \\ &< x. \end{aligned} \tag{79}$$

By the Fundamental Theorem of Arithmetic, no two products $\prod_{i>t}^{k} p_i^{a_i}$ can be equal, which means the number of $\prod_{i>t}^{k} p_i^{a_i'}$, hence $\prod_{i>t}^{k} p_i^{a_i}$, is at most $x^{1/t}$. Thus, the number of n up to x with $f_{-1}(n) = f_{-1}(n+1)$ is at most $(\log x)^{2t} \cdot x^{1/t}$. Let $s(t) = (\log x)^{2t} \cdot x^{1/t}$, and we choose $t = (\log x / 2 \log\log x)^{1/2}$. Then

$$\begin{aligned} s(t) &= (\log x)^{2t} \cdot x^{1/t} \\ &= x^{2t \log\log x / \log x + \frac{1}{t}} \\ &= x^{2(2 \log\log x / \log x)^{1/2}}. \end{aligned} \tag{80}$$

Thus, the number of n up to x with $f_{-1}(n) = f_{-1}(n+1)$ is at most $x^{2(2\log\log x/\log x)^{1/2}}$. Since $(\log\log x/\log x)^{1/2} \ll \epsilon$ for every $\epsilon > 0$, we can find x sufficiently large such that

$$x^{2(2\log\log x/\log x)^{1/2}} \ \ll \ x^{\epsilon}. \tag{81}$$

□

Remark 4.1 We can give a similar proof when r is a negative integer less than -1. Let $r = -m$, where m is a positive integer. With an approach similar to that of (75), we have $p_i^m | a_i$. Then, similar to the argument in (78),

$$\begin{aligned} p_i^{a_i' p_i^m} &\leq n \leq x \\ a_i' p_i^m \log p_i &\leq \log x \\ p_i &< (\log x)^{1/m} \\ a_i' &< (\log x)/p^m \\ &< (\log x)/3^m, \end{aligned} \tag{82}$$

which means the number of choices of the first t prime powers that can divide n is at most $(\log x)^{2t}$. Meanwhile, similar to the argument in (79),

$$\begin{aligned} \left(\prod_{i>t}^{k} p_i^{a_i'}\right)^{t^m} &= \prod_{i>t}^{k} p_i^{a_i' t^m} \\ &\leq \prod_{i>t}^{k} p_i^{a_i' p_i^m} \\ &< x, \end{aligned} \tag{83}$$

which means the number of choices of $\prod_{i>t}^{k} p_i^{a_i}$ is at most x^{1/t^m}. Therefore, the number of n is bounded by

$$\begin{aligned} s(t) &= (\log x)^{2t} \cdot x^{1/t^m} \\ &= x^{2t\log\log x/\log x + 1/t^m}. \end{aligned} \tag{84}$$

We choose $t = (m\log x/(2\log\log x))^{1/m+1}$, then

$$\begin{aligned} S(t) = \ & x^{O\left((2\log\log x/(m\log x))^{m/(m+1)}\right)} \\ = \ & x^{O\left((\log\log x/\log x)^{r/(r-1)}\right)}. \end{aligned} \tag{85}$$

Therefore, the number of n up to x is at most $x^{O\left((\log\log x/\log x)^{r/(r-1)}\right)}$.

Next, we look at the case where r is a non-integer rational, which means $f_r(n)$ and $f_r(n+1)$ are summations of distinct radicals of prime powers. We prove that there are no *Ruth-Aaron* numbers in this case.

4.2 Non-integer Rational r

Proof. Let $r = x/y$ be a non-integer rational, then $f_r(n) = f_r(n+1)$ is equivalent to

$$a_1 p_1^{x/y} + a_2 p_2^{x/y} + \cdots + a_k p_k^{x/y} \;=\; b_1 q_1^{x/y} + b_2 q_2^{x/y} + \cdots + b_l q_l^{x/y} \tag{86}$$

where $a_i, b_i \in \mathbb{Z}$. We must show that (86) holds only if $a_i = b_i = 0$. In other words, $p_i^{x/y}$ and $q_j^{x/y}$, where $i = 1, \ldots, k$, $j = 1, \ldots, l$, are linearly independent over $\mathbb{Z}$.

Boreico shows that when n_i are distinct $k-$th power-free integers, the sum $S = \sum a_i \sqrt[k]{n_i}$, where $a_i \in \mathbb{Z}$ are not all zero, is non-zero [4]. In this case, let $n_i = p_i^x$. Because all of p_i, q_i are distinct, n_i are distinct integers; meanwhile, because $\gcd(x, y) = 1$, n_i are $y-$th power free. Using Boreico's result we can easily show that when $\sum a_i p_i^{x/y} - \sum b_i q_i^{x/y} = 0$, we must have $a_i = b_i = 0$, thus completing the proof. □

Remark 4.2 Above discusses only the circumstance of $r > 0$. Likewise, when r is a negative non-integer rational, we can multiply both sides of equation (86) by a factor of $\left(\prod_{i=1}^k p_i \prod_{j=1}^l q_j\right)^{x/y}$, so n_i remain distinct and $y-$th power-free, and we can still apply Boreico's result to get $a_i = b_i = 0$.

4.3 Positive $r \geq 1$

Next, we look at the case where $r \geq 1$. As mentioned, we first introduce a result by Erdős and Pomerance [11]. As we generalize many of these arguments, we give an expanded version of their argument.

Theorem 4.3 *For every $0 < \epsilon < 1$,*

$$\#\mathcal{R}_{(1,-1)}(x) \;=\; O\left(\frac{x}{(\log x)^{1-\epsilon}}\right). \tag{87}$$

Now we look at our generalization to $r \geq 1$. Due to the length of the proof, we divide it into three sections. In Section 4.3.1, we will introduce a general bound on the size of the largest prime factor of an integer; Sections 4.3.2 and 4.3.3 discuss the two cases in terms of the size of $f_r(n/P(n))$ and $f_r((n+1)/P(n+1))$, and conclude with an estimation of $\#R^r_{(1,-1)}(x)$ in each case. Eventually, we absorb Case (i) into Case (ii) and arrive at our final result. We adapt and advance an approach of Pomerance [27], and we are able to improve Erdős and Pomerance's $O(x/(\log x)^{1-\epsilon})$ [11] by $(\log\log x)^3 \log\log\log x/\log x$, hence a refinement to Pomerance's result [27].

4.3.1 We show that $x^{1/\log\log x} \leq P(n), P(n+1) \leq 2^{1/r} x^{1/2} \log x$.

Proof. Since $n+1$ exceeds x only when $n = x$, in general we assume that $n+1 \leq x$. Let $p = P(n)$ and $q = P(n+1)$, and write $n = p \cdot k,\ n+1 = q \cdot m$. By Lemma 2.8, we may assume

$$P(n) \;>\; x^{1/\log\log x} \text{ and } P(n+1) \;>\; x^{1/\log\log x}. \tag{88}$$

We know from Lemma 2.7 that for all $P(n) \geq 5$, we have

$$P(n)^r \ \leq f_r(n) \ \leq P(n)^r \cdot \frac{\log n}{\log P(n)}. \tag{89}$$

In order to apply (42), we assume $P(n), P(n+1) \geq 5$, so that (89) holds for both n and $n+1$. Next, we give an upper bound on the size of p, q using the fixed values of k, m. We show that given k, m, primes p, q are determined uniquely. In fact, from the two equations

$$\begin{aligned} pk + 1 &= qm \\ p^r + f_r(k) &= q^r + f_r(m) \end{aligned} \tag{90}$$

we have

$$\begin{aligned} p &= \frac{qm-1}{k} \\ \left(\frac{qm-1}{k}\right)^r - q^r &= f_r(m) - f_r(k). \end{aligned} \tag{91}$$

Let

$$\begin{aligned} g(q) &= (qm-1)^r - (qk)^r - k^r \cdot (f_r(m) - f_r(k)) \\ g'(q) &= rm(qm-1)^{r-1} - kr(qk)^{r-1} \\ &= r(m(qm-1)^{r-1} - k(qk)^{r-1}). \end{aligned} \tag{92}$$

Meanwhile,

$$g\left(\frac{1}{m}\right) \ = \ -\left(\frac{k}{m}\right)^r - k^r \cdot (f_r(m) - f_r(k)). \tag{93}$$

Because $q \geq \frac{1}{m}$ and $r \geq 1$, if $m > k$, then from (90)) we know $p > q$ and $f_r(m) > f_r(k)$; thus, $qm - 1 > qk$ and $g'(q) > 0$ which means $g(q)$ always increases. Moreover, because $g\left(\frac{1}{m}\right) < 0$, there exists only one q with $g(q) = 0$. Similarly, if $m < k$, then $g(q)$ always decreases and $g\left(\frac{1}{m}\right) > k^r - \left(\frac{k}{m}\right)^r > 0$, which also means there exists only one $q > 0$ with $g(q) = 0$. Therefore, q is uniquely expressible by k, m, which means p, q are determined by k, m. Thus, the number of choices for n determined by the choices of k, m when $k, m < x^{1/2}/\log x$ is at most $x/(\log x)^2$. Hence, we may assume

$$k \ \geq \ x^{1/2} \log x \ \text{ or } \ m \ \geq \ x^{1/2} \log x. \tag{94}$$

Because $n = p \cdot k \leq x, \ n + 1 = q \cdot m \leq x$, we thus may assume

$$p \ \leq \ x^{1/2} \log x \ \text{ or } \ q \ \leq \ x^{1/2} \log x. \tag{95}$$

Suppose $p > x^{1/2} \log x$, then $q \leq x^{1/2} \log x$. Let $h(x) = x^r / \log x$, then

$$h'(x) = \frac{x^{r-1}(r \log x - 1)}{(\log x)^2} \tag{96}$$

Because $r \geq 1$, $h(x)$ increases on $x \geq 3$. We have

$$\begin{aligned} p^r \;\leq\; f_r(n) \;=\; f_r(n+1) \;&\leq\; q^r \cdot \frac{\log(n+1)}{\log q} \\ &\leq\; \frac{(x^{1/2}\log x)^r \cdot \log(n+1)}{\log(x^{1/2}\log x)} \\ &<\; 2 \cdot (x^{1/2}\log x)^r \\ p \;&<\; 2^{1/r} \cdot x^{1/2}\log x. \end{aligned} \tag{97}$$

A similar inequality can be obtained for $q > x^{1/2}\log x$. Therefore, we have

$$p \;<\; 2^{1/r} \cdot x^{1/2}\log x \;\text{ and }\; q \;<\; 2^{1/r} \cdot x^{1/2}\log x. \tag{98}$$

Now we look at $f_r(k)$ and $f_r(m)$. In terms of the upper bound of $f_r(k), f_r(m)$, we have the following two cases:

$$\begin{aligned} &\text{Case (i): } f_r(k) \;<\; \frac{p^r}{\log^{2r} x},\; f_r(m) \;<\; \frac{q^r}{\log^{2r} x} \\ &\text{Case (ii): (WLOG) } f_r(k) \;\geq\; \frac{p^r}{\log^{2r} x}. \end{aligned}$$

□

4.3.2 Case (i) Discussion.

Proof. We consider Case (i) first. Consider a function $v(x) = (x+1)^r - x^r - 1$, then

$$\begin{aligned} v'(x) \;&=\; r((x+1)^{r-1} - x^{r-1}) \\ &>\; 0, \end{aligned} \tag{99}$$

and the function has a root at $x = 0$, which means $(x+1)^r > x^r + 1$ for all $x > 0$. Applying this result, we have $p^r + q^r < (p+q)^r$ and $|p-q|^r \;<\; |p^r - q^r|$ since $p \neq q > 1$. Meanwhile, by the assumption $f_r(k) < \frac{p^r}{\log^{2r} x}, f_r(m) < \frac{q^r}{\log^{2r} x}$, we have

$$\begin{aligned} |p^r - q^r| \;&=\; |f_r(m) - f_r(k)| \\ &<\; \frac{p^r + q^r}{\log^{2r} x}. \end{aligned} \tag{100}$$

Then

$$|p-q|^r \;<\; |p^r - q^r| \;<\; \frac{p^r + q^r}{\log^{2r} x} \;<\; \frac{(p+q)^r}{\log^{2r} x}. \tag{101}$$

Therefore,

$$\begin{aligned} |p-q| \;&<\; \frac{p+q}{(\log x)^2} \\ p \cdot \frac{(\log x)^2 - 1}{(\log x)^2 + 1} \;<\; q \;&<\; p \cdot \frac{(\log x)^2 + 1}{(\log x)^2 - 1}. \end{aligned} \tag{102}$$

For all p satisfying (42), the number of primes q such that (102) holds is

$$\sum_{p\frac{(\log x)^2-1}{(\log x)^2+1}<q<p\frac{(\log x)^2+1}{(\log x)^2-1}} 1 \ll \pi\left(p\frac{(\log x)^2+1}{(\log x)^2-1}\right)-\pi\left(p\frac{(\log x)^2-1}{(\log x)^2+1}\right). \tag{103}$$

We apply an explicit form of the Brun-Titchmarsh theorem [23] which states that

$$\pi(x+y)-\pi(x) \le \frac{2y}{\log y} \tag{104}$$

for all integers $x \ge 1, y \ge 2$. Then, using $p > x^{1/\log\log x}$, we have:

$$\begin{aligned}\pi\left(p\frac{(\log x)^2+1}{(\log x)^2-1}\right)-\pi\left(p\frac{(\log x)^2-1}{(\log x)^2+1}\right) &\le \frac{2p\frac{4(\log x)^2}{\log^4 x-1}}{\log\left(p\frac{4(\log x)^2}{\log^4 x-1}\right)} \\ &\ll O\left(\frac{\frac{p}{(\log x)^2}}{\log p}\right) \\ &< O\left(\frac{p}{(\log x)^2\cdot\frac{\log x}{\log\log x}}\right) \\ &< O\left(\frac{p\log\log x}{(\log x)^3}\right). \end{aligned} \tag{105}$$

Thus, when x is sufficiently large, using $p < 2^{1/r}x^{1/2}\log x$, the number of q satisfying $p\frac{(\log x)^2-1}{(\log x)^2+1} < q < p\frac{(\log x)^2+1}{(\log x)^2-1}$ is at most

$$\begin{aligned}\sum_{p\frac{(\log x)^2-1}{(\log x)^2+1}<q<p\frac{(\log x)^2+1}{(\log x)^2-1}} 1 &= O\left(\frac{p\log\log x}{(\log x)^3}\right) \\ &\le O\left(\frac{x^{1/2}\log\log x}{(\log x)^2}\right). \end{aligned} \tag{106}$$

Meanwhile, for sufficiently large x, the sum of $\frac{1}{q}$ for such primes q is

$$\begin{aligned}\sum_{p\frac{(\log x)^2-1}{(\log x)^2+1}<q<p\frac{(\log x)^2+1}{(\log x)^2-1}} \frac{1}{q} &< O\left(\frac{p\log\log x}{(\log x)^3}\right)\cdot\frac{1}{p\frac{(\log x)^2-1}{(\log x)^2+1}} \\ &= O\left(\frac{\log\log x}{(\log x)^3}\right). \end{aligned} \tag{107}$$

Therefore, in Case (i), using the result from (106) and (107), $\#\mathcal{R}^r_{(1,-1)}(x)$ is at most

$$\begin{aligned}\sum_{\substack{p\frac{(\log x)^2-1}{(\log x)^2+1}<q<p\frac{(\log x)^2+1}{(\log x)^2-1}\\ x^{1/\log\log x}<p<2^{1/r}x^{1/2}\log x}} \left(1+\left\lfloor\frac{x}{pq}\right\rfloor\right) &< \left(\frac{x^{1/2}\log\log x}{(\log x)^2}\right)\cdot\pi\left(2^{1/r}x^{1/2}\log x\right)+\sum\left\lfloor\frac{x}{pq}\right\rfloor \\ &\ll \left(\frac{x^{1/2}\log\log x}{(\log x)^2}\right)\cdot\left(x^{1/2}\right)+\frac{x\log\log x}{(\log x)^3}\sum_{\substack{x^{1/\log\log x}<\\ p<2^{1/r}x^{1/2}\log x}}\frac{1}{p} \\ &< O\left(\frac{x\log\log x}{(\log x)^2}\right)+O\left(\frac{x(\log\log x)^2}{(\log x)^3}\right) \text{ (Lemma 2.3)} \\ &= O\left(\frac{x\log\log x}{(\log x)^2}\right). \qquad (108)\end{aligned}$$

□

4.3.3 Case (ii) Discussion

Proof. Now we consider Case (ii). Write $k = t\cdot u$, where $t = P(k)$, and we have

$$\begin{aligned}p^r \le f_r(n) = f_r(n+1) &\le q^r\cdot\frac{\log(n+1)}{\log q} \\ &\le q^r\cdot\frac{\log(x+1)}{\log q} \\ &\le q^r\cdot\frac{\log(x+1)}{\log x^{1/\log\log x}} \\ &\le q^r\cdot 2\log\log x. \qquad (109)\end{aligned}$$

Similarly,

$$\begin{aligned}q^r \le f_r(n+1) = f_r(n) &\le p^r\cdot\frac{\log n}{\log p} \\ &\le p^r\cdot\frac{\log x}{\log p} \\ &\le p^r\cdot\log\log x. \qquad (110)\end{aligned}$$

Therefore, we have

$$\frac{p}{2^{1/r}(\log\log x)^{1/r}} \le q \le p\cdot(\log\log x)^{1/r}. \qquad (111)$$

Using the result from Lemma 2.8 again, the number of k for which we don't have $t > k^{1/\log\log k}$ is $O(k/(\log k)^2)$; thus, we may assume $t > k^{1/\log\log k}$ and $t \ge 5$. Applying results

from Lemma 2.7,

$$\begin{aligned}\frac{p^r}{\log^{2r} x} \;\le\; f_r(k) \;&\le\; t^r \cdot \frac{\log k}{\log t} \\ &\le t^r \cdot \log\log k \\ &< \; t^r \cdot \log\log x. \end{aligned} \tag{112}$$

We get that

$$\frac{p}{(\log x)^2 (\log\log x)^{1/r}} \;<\; t \;\le\; p. \tag{113}$$

This result implies that $P(n/P(n))$ and $P(n)$ are relatively close. Moreover, in terms of the size of $P(n)$, we have the following two cases:

$$\begin{aligned}&\text{Case (ii.1): } p \;\le\; x^{1/3} \\ &\text{Case (ii.2): } p \;>\; x^{1/3}.\end{aligned}$$

Consider Case (ii.1) first. The number of $n \le x$ with $pt|n$ and $q|n+1$ in this case is at most

$$\begin{aligned}&\sum_{\substack{x^{1/\log\log x} < p \le x^{1/3} \\ p/(2^{1/r}(\log\log x)^{1/r}) < \\ q < p(\log\log x)^{1/r} \\ p/((\log x)^2(\log\log x)^{1/r}) < t \le p}} \left(1 + \left\lfloor \frac{x}{ptq} \right\rfloor\right) \\ &\ll\; \pi\left(x^{1/3}\right) \cdot \pi\left(x^{1/3}(\log\log x)^{1/r}\right) \cdot \pi\left(x^{1/3}\right) + \sum \left\lfloor \frac{x}{ptq} \right\rfloor \\ &\ll\; O\left(\frac{x^{1/3}}{\log x} \cdot \frac{x^{1/3}\log\log x}{\log x} \cdot \frac{x^{1/3}}{\log x}\right) + \sum \left\lfloor \frac{x}{ptq} \right\rfloor \\ &<\; O\left(\frac{x\log\log x}{(\log x)^3}\right) + x \sum \frac{1}{p} \sum \frac{1}{t} \sum \frac{1}{q} \\ &\ll\; O\left(\frac{x\log\log x}{(\log x)^3}\right) + O\left(x \sum \frac{1}{p} \sum \frac{1}{t} \cdot \frac{\log\log\log x}{\log p}\right) \text{ (Lemma 2.3)} \\ &\ll\; O\left(\frac{x\log\log x}{(\log x)^3}\right) + O\left(x\log\log\log x \sum \frac{1}{p\log p} \sum \frac{1}{t}\right) \\ &\ll\; O\left(\frac{x\log\log x}{(\log x)^3}\right) + O\left(x\log\log\log x \sum \frac{1}{p\log p} \cdot \frac{\log\log x}{\log p}\right) \\ &\le\; O\left(\frac{x\log\log x}{(\log x)^3}\right) + O\left(\frac{x\log\log\log x(\log\log x)^3}{(\log x)^2}\right) \text{ (Lemma 2.5)} \\ &=\; O\left(\frac{x\log\log\log x(\log\log x)^3}{(\log x)^2}\right). \end{aligned} \tag{114}$$

Recall that in Case (i), $\#\mathcal{R}^r_{(1,-1)}(x) = O\left(\frac{x \log\log x}{(\log x)^2}\right)$, which means we can absorb Case (i) into error estimate. Now we turn to Case (ii.2), where $p > x^{1/3}$. We have

$$\begin{aligned} q^r &\leq p^r \cdot \frac{\log x}{\log p} \\ &< 3 \cdot p^r. \end{aligned} \tag{115}$$

Meanwhile, because $p > x^{1/\log\log x}$, when x is sufficiently large we have

$$\begin{aligned} p^r &\leq q^r \cdot \frac{\log(x+1)}{\log q} \\ r \log p &\leq r \log q + \log \frac{\log(x+1)}{\log p} \\ &\leq r \log q + 2 \log\log\log x \\ &\leq r \log q + \frac{1}{2} \log x^{1/\log\log x} \\ \log p &\leq \frac{3}{2} \cdot \log q. \end{aligned} \tag{116}$$

Therefore, we have

$$\begin{aligned} p^r &\leq q^r \cdot \frac{\log(x+1)}{\log q} \\ &\leq q^r \cdot \frac{\log(x+1)}{2\log p/3} \\ &< q^r \cdot \frac{9\log(x+1)}{2\log x} \\ &< 6q^r. \end{aligned} \tag{117}$$

Therefore, $p/6^{1/r} < q < 3^{1/r}p$. Moreover, recall (112); we have

$$\frac{p^r}{\log^{2r} x} \leq f_r(k) \leq t^r \cdot \frac{\log k}{\log t}. \tag{118}$$

Suppose $t < p^{1/2}$, then the number of $n \leq x$ is at most

$$\begin{aligned} \sum_{\substack{t<p^{1/2} \\ x^{1/3}<p<2^{1/r}x^{1/2}\log x}} 1 &\ll \pi\left(\left(2^{1/r}x^{1/2}\log x\right)^{1/2}\right)\pi\left(2^{1/r}x^{1/2}\log x\right) \\ &\ll \frac{x^{1/4}\log^{1/2} x}{\log x} \cdot \frac{x^{1/2}\log x}{\log x} \\ &= O\left(\frac{x^{3/4}}{(\log x)^{1/2}}\right), \end{aligned} \tag{119}$$

which we can absorb into error estimate. Therefore, $t > p^{1/2}$. Using $p > x^{1/3}$, we have

$$\begin{aligned}
\frac{p^r}{\log^{2r} x} &\leq t^r \cdot \frac{\log k}{\log t} \\
&< t^r \cdot \frac{2\log x}{\log p} \\
\frac{p}{(\log x)^2} &< t \cdot \left(\frac{2\log x}{\log x/3}\right)^{1/r} \\
p &< t \cdot 6^{1/r}(\log x)^2. \qquad (120)
\end{aligned}$$

Therefore, we have

$$\frac{p}{6^{1/r}(\log x)^2} < t < p. \qquad (121)$$

Recall $f_r(k) = t \cdot u$ and $p/6^{1/r} < q < 3^{1/r}p$. If $p \geq x^{2/5}$, then

$$\begin{aligned}
u \leq \frac{x}{pt} \leq \frac{6^{1/r}x(\log x)^2}{p^2} &= O(x^{1/5}(\log x)^2) \\
m \leq \frac{x}{q} \ll \frac{x}{p} &\leq x^{3/5}, \qquad (122)
\end{aligned}$$

which means that the number of uniquely determined n is at most $O(x^{4/5}(\log x)^2)$, a number absorbable by error estimate. Therefore, we need to consider only

$$x^{1/3} < p < x^{2/5}. \qquad (123)$$

Now, recall that $k = t \cdot u$. Let $u = v \cdot w$, where $v = P(u)$. We have the following equations:

$$\begin{aligned}
p \cdot k + 1 &= q \cdot m \\
p^r + f_r(k) &= q^r + f_r(m). \qquad (124)
\end{aligned}$$

Then we have

$$\begin{aligned}
p^r + t^r + f_r(u) &= \left(\frac{pk+1}{m}\right)^r + f_r(m) \\
(upt+1)^r - (mp)^r - (mt)^r &= m^r(f_r(u) - f_r(m)). \qquad (125)
\end{aligned}$$

When $r = 1$, Pomerance [27] demonstrated the following.

$$\begin{aligned}
upt - mp - mt &= m(f(u) - f(m)) - 1 \\
(up - m)(ut - m) &= um(f(u) - f(m)) - u + m^2. \qquad (126)
\end{aligned}$$

Given u, m, the number of choices of t, and thus for n, is determined by the number of divisors of $um(f(u) - f(m)) - u + m^2$, and is at most

$$\tau\left(um(f(u) - f(m)) - u + m^2\right) \leq x^{o(1)}, \qquad (127)$$

where $\tau(x)$ denotes the divisor function. Let $m = s \cdot z$, where $s = P(m)$. Recall $t = P(k)$. We first show that the number of n with $t, s \le x^{1/6}$ can be absorbed into error estimate. In fact, the number of triples p, q, t is at most $O(x^{2/5} \cdot x^{2/5} \cdot x^{1/6}) = O(x^{29/30})$, and since

$$\begin{aligned} pqt &\ge x^{1/3} \cdot 6^{1/r} x^{1/3} \cdot \frac{x^{1/3}}{6^{1/r}(\log x)^2} \\ &= \frac{x}{(\log x)^2}, \end{aligned} \tag{128}$$

the number of n up to x for which $t, s \le x^{1/6}$ is at most $O(x^{29/30}(\log x)^2)$. Thus, we may assume that at least one of s, t is greater than $x^{1/6}$.

Recall $u = v \cdot w$ where $v = P(u)$. First, we consider $v > x^{1/6}$. We can rewrite (126) as

$$(vwp - m)(vwt - m) = wmv^2 + ((f(w) - f(m))mw - w)\, v + m^2. \tag{129}$$

Pomerance [27] proved the following.

Lemma 4.4 *Let A, B, C be integers with $\gcd(A, B, C) = 1$, $D := B^2 - 4AC \neq 0$, and $A \neq 0$. Let M_0 be the maximum value of $|At^2 + Bt + C|$ on the interval $[1, x]$. Let $M = \max\{M_o, |D|, x\}$ and let $\mu = \lceil \log M / \log x \rceil$ and assume $\mu \le \frac{1}{7} \log\log x$, then*

$$\sum_{n \le x} \tau\left(|An^2 + Bn + C|\right) \le x(\log x)^{2^{3\mu+1}+4} \tag{130}$$

holds uniformly for $x \ge x_0$, where x_0 is an absolute constant.

In our case, let $A = wm, B = ((f(w) - f(m))mw - w), C = m^2$. Both $\gcd(A, B, C) = 1$ and $D \neq 0$ are easily verified since $\gcd(w, m) = 1$. Moreover, recall that $t > \frac{p}{6^{1/r}(\log x)^2}$, we have

$$\begin{aligned} u \le \frac{x}{pt} \le \frac{6x(\log x)^2}{p^2} &\le 6x^{1/3}(\log x)^2 \\ w = \frac{u}{v} \le \frac{u}{x^{1/6}} &\le 6x^{1/6}(\log x)^2 \\ v &\le \frac{6x^{1/3}(\log x)^2}{w} \\ m \le \frac{x}{q} &\ll x^{2/3}. \end{aligned} \tag{131}$$

Then $M_0 \ll x^{4/3}(\log x)^2$. It follows from the lemma that

$$\sum_{v \le (6x^{1/3}(\log x)^2)/w} \tau(|Av^2 + Bv + C|) \ll \frac{6x^{1/3}(\log x)^c}{w}, \tag{132}$$

where c is a positive constant. Moreover, if $x^{1/3} < p \le x^{1/3}(\log x)^{c+5}$, $\#\mathcal{R}_{(1,-1)}(x)$ is at most

$$\begin{aligned}\sum_{\substack{x^{1/3}<p\le x^{1/3}(\log x)^{c+5}\\ p/6<q<3p}} \left(1+\frac{x}{pq}\right) &\ll O\left(x^{2/3}(\log x)^{2c+10}\right) + x \sum_{x^{1/3}<p\le x^{1/3}(\log x)^{c+5}} \frac{1}{p} \sum_{p/6<q<3p} \frac{1}{q} \\ &\ll O(x^{2/3}(\log x)^{2c+10}) + O\left(\frac{x\log\log x}{(\log x)^2}\right) \\ &\ll O\left(\frac{x\log\log x}{(\log x)^2}\right), \end{aligned} \tag{133}$$

which is small enough to be absorbed by error estimate. If $p > x^{1/3}(\log x)^{c+5}$, then $m \ll \frac{x}{p} \ll x^{2/3}/(\log x)^{c+5}$ and $w \le \frac{u}{v} \le \frac{u}{x^{1/6}} \le 6x^{1/6}/(\log x)^{2c+8}$. Then summing (132) over all choices of w, m the quantity is less than $O(x/(\log x))^2$, which is indeed negligible.

Finally, recall that $m = s \cdot z$ where $s = P(m)$. We consider the case where $s > x^{1/6}$. From (126) we have

$$(pu - sz)(tu - sz) \;=\; (z^2 - uz)s^2 + (f(u) - f(z))uzs - u. \tag{134}$$

Similarly, let $p \ge x^{1/3}(\log x)^{c+5}$, then $u \le \frac{x}{pt} \le \frac{6x(\log x)^2}{p^2} \le x^{1/3}/(\log x)^{2c+8}$. Moreover,

$$\begin{aligned} z \;\le\; \frac{x}{qs} &\ll \frac{x^{2/3}}{x^{1/6}} \ll x^{1/2} \\ s \;\le\; \frac{m}{z} &\ll x^{2/3}/z. \end{aligned} \tag{135}$$

Therefore, considering the ranges of s, z, u, the right hand side of (134) has less than $O(x/(\log x)^2)$ choices, suggesting that this case is also negligible.

Recall Lemma 3.1 and Lemma 3.4: when $r \ge 1$, $P(n)^r$ is the dominating term of $f_r(n)$, and the larger r is, the closer $f_r(n)$ is to $P(n)^r$; meanwhile, $P(n)$ and $P(n+1)$ are usually at least a factor of x^δ apart, suggesting the increasing rarity of $r-$th power *Ruth-Aaron* numbers as r increases. Therefore, although we are unable to generalize the quadratic formula $|Ax^2 + Bx + C|$ to a higher power[4] at this point, we can substantiate firmly that when $r \ge 1$, the number of $r-$th power *Ruth-Aaron* numbers up to x with $x^{1/3} < p < x^{2/5}$ is much less than the estimated result followed by Lemma 4.4. Therefore, when $r \ge 1$,

$$\#\mathcal{R}^r_{(1,-1)}(x) \;=\; O\left(\frac{x(\log\log x)^3(\log\log\log x)}{(\log x)^2}\right). \tag{136}$$

□

[4]For example, when $r = 2$, we have a factorization similar to (126): $\left(u^2pt + u - m^2 + ump - umt\right) \cdot \left(u^2pt + u - m^2 - ump + umt\right) = (um)^2(f_2(u) - f_2(m)) + m^2(m^2 - 2u)$. However, to tackle $r = 2$ with an approach similar to the one introduced in this section, a result for the summation of divisors of quartic functions will appear necessary.

Remark 4.5 In our proof of Theorem 1.6, one crucial result we applied is that of De Bruijn [5] in Lemma 2.8. This result states that for all but $O(x/(\log x)^2)$ of $n \leq x$ we have $P(n), P(n+1) > x^{1/\log\log x}$, lays the ground work for major arguments in the rest of the proof. It serves as a major lower bound for p, and helps us obtain a tighter bound on q and t in relation to p, which we relied on when counting the number of n up to x.

We present a table of 2−nd power *Ruth-Aaron* numbers below $5 \cdot 10^7$:

n	$f_2(n) = f_2(n+1)$
6371184	40996
16103844	22685
49214555	103325

Table 3: 2−nd power *Ruth-Aaron* numbers not exceeding $5 \cdot 10^7$.

5 Proof of Theorem 1.7

In order to prove Theorem 1.7, we need to first introduce a special case of the Catalan Conjecture which we can prove directly.

Theorem 5.1 (Catalan Conjecture) *The only natural number solution* (a, b, x, y) *of*

$$a^x - b^y = 1 \tag{137}$$

for $a, b > 1$, $x, y > 0$ *is* $(3, 2, 2, 3)$.

The Catalan Conjecture was proved by Mihăilescu in 2002 [21].

Lemma 5.2 (Special case of the Catalan Conjecture) *The number of* $n \leq x$ *with* $P(n) \leq 3$ *is* $O((\log x)^2)$. *Meanwhile, the largest* n *with* $P(n), P(n+1) \leq 3$ *is 8.*

Proof. First, we prove that the number of n up to x with $n = 2^a \cdot 3^b$, where a, b are nonnegative integers, is $O((\log x)^2)$. In fact, since the number of powers of 2 and 3 no greater than x are $\log_2 x$ and $\log_3 x$ respectively, there are at most $\log_2 x \cdot \log_3 x = O((\log x)^2)$ number of $n \leq x$ with $n = 2^a \cdot 3^b$. Meanwhile, because $\gcd(n, n+1) = 1$, when $P(n), P(n+1) \leq 3$, either $n = 2^a, n+1 = 3^b$, or $n = 3^b, n+1 = 2^a$.

Case (i): $3^b = 2^a + 1$. Let the order of 2 modulo 3^b be d. We have,

$$\begin{aligned} 2^a &\equiv -1 \pmod{3^b} \\ 2^{2a} &\equiv 1 \pmod{3^b} \\ 2^d &\equiv 1 \pmod{3^b}. \end{aligned} \tag{138}$$

Then we have

$$\begin{aligned} &d \mid \varphi(3^b) = 2 \cdot 3^{b-1} \\ &d \nmid a,\ d \mid 2a. \end{aligned} \tag{139}$$

It's obvious that $2 \mid d$, then (139) tells us that $d = 2a$ (or $d = 2$, then there is no solution), and that $d = 2 \cdot 3^k$ for some nonnegative integer $k \leq b-1$. Thus, $a = 3^k$. When $k = 0$, $(a, b) = (1, 1)$; when $k \geq 1$, we have

$$\begin{aligned} 2^{3^k} + 1 &= (2^{3^{k-1}} + 1)(2^{2 \cdot 3^{k-1}} - 2^{3^{k-1}} + 1) \\ 2^{2 \cdot 3^{k-1}} - 2^{3^{k-1}} + 1 &\equiv \left((-1)^2\right)^{3^{k-2}} - (-1)^{3^{k-2}} + 1 \\ &\equiv 1 + 1 + 1 \\ &\equiv 3 \pmod 9, \end{aligned} \tag{140}$$

which means $2^{3^k} + 1$ is a power of 3 only when $2^{2 \cdot 3^{k-1}} - 2^{3^{k-1}} + 1$ is 3. Thus, the largest solution to $3^b = 2^a + 1$ is $(a, b) = (3, 2)$.

Case (ii): $2^a = 3^b + 1$. Then $2^a \equiv 1 \pmod 3$, which means $2|a$, since 2 is the order of 2 modulo 3; thus, let $a = 2a_0$, where a_0 is a nonnegative integer. Then $2^a - 1 = (2^{a_0} + 1)(2^{a_0} - 1)$. Because $2^{a_0} - 1$ and $2^{a_0} + 1$ are relatively prime unless $a_0 = 0$ or 1, the largest solution in this case is $(a, b) = (2, 1)$. Thus, the largest n with $P(n), P(n+1) \leq 3$ is 8. □

Now we recall Theorem 1.7.

Theorem 1.7 *There are infinitely many real numbers r such that $\#\mathcal{R}^r_{(1,-1)}(x) > 0$.*

Proof. It suffices to show there exists an infinite decreasing series of positive r. Let n be of the form $n = p^2 - 1$, where p is a prime. Then $n + 1 = p^2$. We claim that for any n_0 that is an r_0-th power *Ruth-Aaron* number, we can always find $n > n_0$ that is an $r-$th power *Ruth-Aaron* number where $0 < r < r_0$.

We define a function of r for any fixed n:

$$g_n(r) = f_r(n+1) - f_r(n). \tag{141}$$

We want to show that $g_n(r)$ increases on $r > 0$ and has one and only one root for a fixed n. We will use the function's continuity to show that the equation has at least one root, and its monotonicity and values at $r = 0$ and $r = 1$ to prove that the function has one and only one root, and it's between 0 and 1. First, we have for prime p,

$$\begin{aligned} g_n(r) &= f_r(n+1) - f_r(n) \\ &= 2 \cdot p^r - \sum_{q \mid p^2 - 1} q^r \\ &= 2 \cdot p^r - 2 \cdot 2^r - \sum_{q \mid (p^2-1)/4} q^r. \end{aligned} \tag{142}$$

It's obvious that $g_n(r)$ is continuous. Meanwhile,

$$\begin{aligned} g_n(0) &= 2 - 2 - \sum_{q \mid (p^2-1)/4} 1 \\ &< 0. \end{aligned} \tag{143}$$

Because the function $\frac{x}{\log x}$ increases on $x \geq e$, for all except $O((\log x)^2)$ number of $n \leq x$, $f_1(n) \leq P(n)\log n/\log P(n) \leq n$ holds. Meanwhile, for any n of the form $2^a \cdot 3^b$, since there are $O(x/\log x)$ different primes up to x, we will drop such an n and consider the next prime p and the respective $n = p^2 - 1$. We thus have

$$\begin{aligned} g_n(1) &= 2p - f_1(p-1) - f_1(p+1) \\ &> 2p - (p-1) - (p+1) \\ &= 0. \end{aligned} \tag{144}$$

Therefore, due to the continuity of function $g_n(r)$, for any fixed n, there must exist at least one r with $f_r(n+1) = f_r(n)$. We proceed to prove that $g_n(r) > 0$ on $r \geq 1$. If this claim holds, then all *Ruth-Aaron* r for $n = p^2 - 1$ must satisfy $r \in (0, 1)$. In fact, let $\frac{p-1}{2} = \prod_{i=1}^{k_1} s_i^{a_i}$ and $\frac{p+1}{2} = \prod_{i=1}^{k_2} t_i^{b_i}$, where s_i, t_i are distinct prime factors of $\frac{p-1}{2}$ and $\frac{p+1}{2}$ respectively. Because $\frac{p-1}{2}, \frac{p+1}{2}$ are consecutive thus relatively prime, we have

$$\begin{aligned} g_n(r) &= f_r(n+1) - f_r(n) \\ &= 2 \cdot p^r - 2 \cdot 2^r - \sum_{i=1}^{k_1} a_i s_i^r - \sum_{i=1}^{k_2} b_i t_i^r \\ g_n'(r) &= 2p^r \log p - 2 \cdot 2^r \log 2 - \sum_{i=1}^{k_1} (a_i s_i^r \log s_i) - \sum_{i=1}^{k_2} (b_i t_i^r \log t_i) \\ &> 2p^r \log p - 2 \cdot 2^r \log 2 - \sum_{i=1}^{k_i} \left(\left(\frac{p-1}{2} \right)^r a_i \log s_i \right) - \sum_{i=1}^{k_2} \left(\left(\frac{p+1}{2} \right)^r b_i \log t_i \right) \\ &= 2p^r \log p - 2 \cdot 2^r \log 2 - \left(\frac{p-1}{2} \right)^r \log \frac{p-1}{2} - \left(\frac{p+1}{2} \right)^r \log \frac{p+1}{2} \\ &> 2p^r \log p - 2 \cdot 2^r \log 2 - \left(\frac{p-1}{2} \right)^r \log p - \left(\frac{p+1}{2} \right)^r \log p. \end{aligned} \tag{145}$$

Consider the function $v(r) = (1+x)^r - x^r - 1$, where $r \geq 1$ is a real number and $x > 0$. Because $v'(r) = (1+x)^r - x^r - 1 > x^r(\log(x+1) - \log x) > 0$, we have $(1+x)^r > 1 + x^r$ when $r \geq 1$. Thus,

$$\begin{aligned} \left(\frac{p-1}{2} \right)^r + \left(\frac{p+1}{2} \right)^r &= \left(\frac{p-1}{2} \right)^r \left(1 + \left(\frac{p+1}{p-1} \right)^r \right) \\ &< \left(\frac{p-1}{2} \right)^r \left(\frac{2p}{p-1} \right)^r \\ &= p^r. \end{aligned} \tag{146}$$

Therefore,

$$\begin{aligned} g_n'(r) &> 2p^r \log p - 2 \cdot 2^r \log 2 - p^r \log p \\ &= p^r \log p - 2 \cdot 2^r \log 2 \\ &> 0 \end{aligned} \tag{147}$$

when $p \geq 3$. Therefore, when $r \geq 1$, $g_n(r)$ reaches its minimum at $r = 1$, and $g_n(1) > 0$, which means all *Ruth-Aaron r* with $n = p^2 - 1$ are between 0 and 1. We continue to show that $g_n(r)$ increases on $(0, 1)$. If this holds, then there exists exactly one *Ruth-Aaron r* for a fixed n. From (145) we have

$$g_n'(r) \;>\; p^r \log p - 2 \cdot 2^r \log 2 - \left(\frac{p-1}{2}\right)^r \log \frac{p-1}{2} - \left(\frac{p+1}{2}\right)^r \log \frac{p+1}{2}. \tag{148}$$

We want to show that $\left(\frac{p-1}{2}\right)^r \log \frac{p-1}{2} + \left(\frac{p+1}{2}\right) \log \frac{p+1}{2} < 2 \cdot p^r \log \frac{p}{2}$. From Cauchy-Schwarz inequality [32],

$$\left(\sum_{i=1}^{m} a_i^2\right) \cdot \left(\sum_{i=1}^{m} b_i^2\right) \;\geq\; \left(\sum_{i=1}^{m} a_i b_i\right)^2, \tag{149}$$

and equality holds only when $a_1/b_1 = a_2/b_2 = \cdots = a_m/b_m$. Applying this result, we have

$$\begin{aligned} &\left(\left(\frac{p-1}{2}\right)^r \log \frac{p-1}{2} + \left(\frac{p+1}{2}\right)^r \log \frac{p+1}{2}\right)^2 \\ &\quad\leq \left(\left(\frac{p-1}{2}\right)^{2r} + \left(\frac{p+1}{2}\right)^{2r}\right)\left(\log^2 \frac{p-1}{2} + \log^2 \frac{p+1}{2}\right). \end{aligned} \tag{150}$$

Meanwhile, from Jensen's inequality [32], if a function f is concave and $k_1, k_2, \ldots, k_m$ are positive reals summing up to 1, then

$$f\left(\sum_{i=1}^{m} k_i x_i\right) \;\geq\; \sum_{i=1}^{m} k_i f(x_i), \tag{151}$$

and equality holds if and only if $x_1 = x_2 = \cdots = x_m$. Because the function x^r is concave on $r < 1$, we have

$$\begin{aligned} \left(\frac{p-1}{2}\right)^{2r} + \left(\frac{p+1}{2}\right)^{2r} &< 2\left(\frac{((p-1)/2)^2 + ((p+1)/2)^2}{2}\right)^r \\ &< 2\left(\frac{p-1}{2} + \frac{p+1}{2}\right)^{2r} \\ &= 2p^{2r}. \end{aligned} \tag{152}$$

Moreover, because the function $(\log x)^2$ is concave on $x > e$, we have for most x,

$$\begin{aligned} \log^2 \frac{p-1}{2} + \log^2 \frac{p+1}{2} &< 2\log^2 \frac{\left(\frac{p-1}{2} + \frac{p+1}{2}\right)}{2} \\ &= 2\log^2 \frac{p}{2}. \end{aligned} \tag{153}$$

Therefore, from (150),

$$\begin{aligned}\left(\left(\frac{p-1}{2}\right)^r \log\frac{p-1}{2} + \left(\frac{p+1}{2}\right)\log\frac{p+1}{2}\right)^2 &< 2p^{2r}\cdot 2\log^2\frac{p}{2}\\ \left(\frac{p-1}{2}\right)^r \log\frac{p-1}{2} + \left(\frac{p+1}{2}\right)^r \log\frac{p+1}{2} &< 2p^r\log\frac{p}{2}. \end{aligned} \tag{154}$$

Thus, for (148),

$$\begin{aligned} g_n'(r) &> 2\cdot p^r\log p - 2\cdot 2^r\log 2 - 2\cdot p^r\log\frac{p}{2}\\ &= 2\log 2(p^r-2^r)\\ &> 0. \end{aligned} \tag{155}$$

Therefore, $g_n(r)$ increases on $r>0$. Meanwhile, because $g_n(0)<0$ and $g_n(1)>0$, there exists one and only r with $f_r(n+1)=f_r(n)$, and $0<r<1$. As mentioned, we want to show that for any n_0 that is an r_0-th power *Ruth-Aaron* number, we can always find $n>n_0$ that is an $r-$th power *Ruth-Aaron* number, where $0<r<r_0$. Now, because $g_n(r)$ increases on $r>0$, it suffices to show that we can always find $n>n_0$ with $f_{r_0}(n+1)>f_{r_0}(n)$. We consider the function $x^r/\log x$. If we take our n to be greater than e^{1/r_0}, then the function $x^{r_0}/\log x$ will be increasing. We thus have

$$\frac{P(n)^{r_0}}{\log P(n)} < \frac{n^{r_0}}{\log n} \tag{156}$$

holds. Meanwhile, because $\frac{p-1}{2}$ and $\frac{p+1}{2}$ are consecutive integers, when they are both greater than 8, at most one of them is of the form $2^a\cdot 3^b$, and the number of $n\le x$ with $n=2^a\cdot 3^b$ is at most $O((\log x)^2)$ (Lemma 5.2), which means we can apply Lemma 2.7 to most n. Thus,

$$\begin{aligned} f_{r_0}(n+1)-f_{r_0}(n) &= 2\cdot p^{r_0}-2\cdot 2^{r_0}-f_{r_0}\left(\frac{p-1}{2}\right)-f_{r_0}\left(\frac{p+1}{2}\right)\\ &> 2\cdot p^{r_0}-2\cdot 2^{r_0}-P\left(\frac{p-1}{2}\right)^{r_0}\log\left(\frac{p-1}{2}\right)/\log P\left(\frac{p-1}{2}\right)\\ &- P\left(\frac{p+1}{2}\right)^{r_0}\log\left(\frac{p+1}{2}\right)/\log P\left(\frac{p+1}{2}\right) \text{ (Lemma 2.7)}\\ &> 2\cdot p^{r_0}-2\cdot 2^{r_0}-\left(\frac{p-1}{2}\right)^{r_0}-\left(\frac{p+1}{2}\right)^{r_0}\\ &> 2\cdot\left(p^{r_0}-2^{r_0}-\left(\frac{p+1}{2}\right)^{r_0}\right)\\ &> 0 \end{aligned} \tag{157}$$

when p is sufficiently large. Therefore, when $n=p^2-1$ is sufficiently large, $f_{r_0}(n+1)>f_{r_0}(n)$, which means the *Ruth-Aaron* r with $f_r(n+1)=f_r(n)$ is less than r_0 and that we can always construct an infinite decreasing series of r, completing our proof. □

6 Future Work

There are a few directions for future work on the *Ruth-Aaron* function. First of all, although we improved Erdős and Pomerance's result [11] by a factor of $(\log\log x)^3 \cdot (\log\log\log x)/\log x$, it is still far from the actual number of $r-$th power *Ruth-Aaron* numbers up to x, as they appear to be extremely rare. Therefore, further research might work towards tightening the existing bounds, as in many cases our estimation is relatively loose. Second, future work might consider placing the *Ruth-Aaron* function in a linear equation. Inspired by Luca and Stănică's result [18] on the number of integers up to x with $\varphi(n) = \varphi(n-1) + \varphi(n-2)$ (See Appendix A), we decide to apply the Fibonacci equation to *Ruth-Aaron* numbers.

Definition 6.1 *A Rabonacci number n is an integer which satisfies*

$$f(n) \;=\; f(n-1) + f(n-2). \tag{158}$$

Similarly, a $r-$th power Rabonacci number n satisfies

$$f_r(n) \;=\; f_r(n-1) + f_r(n-2). \tag{159}$$

It's obvious that $\mathcal{R}_{(1,1,-1)}(x)+2$ (where $+2$ indicates adding 2 to every element in the set $\mathcal{R}_{(1,1,-1)}(x)$) is the same as the set of *Rabonacci* numbers not exceeding x. Meanwhile, it seems that *Rabonacci* numbers might be even rarer than *Ruth-Aaron* numbers, as there are 42 *Rabonacci* numbers below 10^6, in contrast to 149 *Ruth-Aaron* numbers within the same range. Using an approach similar to the proof of Theorem 4.3, we are able to present a partial result on the upper bound on the Rabonacci numbers.

Proposition 6.2 *When $P(n) \leq x^{1/3}$ and $f(n) > f(n-1) \geq f(n-2)$, the number of Rabonacci numbers up to x is at most*

$$\#\mathcal{R}_{(1,1,-1)}(x) \;=\; O\left(\frac{x(\log\log x \log\log\log x)^2}{(\log x)^2}\right). \tag{160}$$

Moreover, Nelson, Penney, and Pomerance proved the following under the hypothesis of Schinzel's Conjecture H [24].

Conjecture 6.3 (Erdős) There are infinitely many n with

$$f(n) \;=\; f(n+1). \tag{161}$$

Meanwhile, inspired by his conjecture on the Euler Totient Function [10], we have

Conjecture 6.4 There exists, for every $k \geq 1$, consecutive integers $n, n+1, \ldots, n+k$ such that

$$f(n) \;=\; f(n+1) \;=\; \cdots \;=\; f(n+k). \tag{162}$$

So far, we are able to confirm that there is at least one n for $k = 1$ and $k = 2$. In fact, as of now, two integers are found to be *Ruth-Aaron* numbers when $k = 2$. Due to the rarity of *Ruth-Aaron* triples, we conjecture that

Conjecture 6.5 There are finitely many integers n with

$$f(n) \;=\; f(n+1) \;=\; f(n+2). \tag{163}$$

It's obvious that in (162), as k increases, the number of n satisfying the equation decreases significantly. Therefore, if we could prove the conjecture for $k = 2$, then for all $k \geq 2$, the equation $f(n) = f(n+1) = \cdots = f(n+k)$ has only finitely many solutions.

Clearly we are unable to prove the conjecture at this point; however, considering the scarcity of *Ruth-Aaron* pairs, we can show that the conjecture hold true if we allow a better result of $\#\mathcal{R}_{(1,-1)}(x)$ for the estimation of the number of integers up to x with $f(n) = f(n+1) = f(n+2)$. In fact, if $\#\mathcal{R}_{(1,-1)}(x) \leq O\left(x^{(1-\epsilon)/2}\right)$, for any $0 < \epsilon < 1$, and *Ruth-Aaron* numbers are uniformly distributed, then standard probabilistic models predict that the number of *Ruth-Aaron* triples (and hence quadruples and higher) is finite. We could create a more sophisticated model that takes into account the decay in the density of *Ruth-Aaron* numbers, which will also give a finite bound on the number of triples; we prefer to do the simple model below to highlight the idea.

Proposition 6.6 *If $\#\mathcal{R}_{(1,-1)}(x) \leq O\left(x^{(1-\epsilon)/2}\right)$, for any $0 < \epsilon < 1$, and Ruth-Aaron numbers are uniformly distributed, then the number of Ruth-Aaron triples is $O(1)$.*

Proof. First, we consider the probability that integer n is a *Ruth-Aaron* number when $2^x \leq n < 2^{x+1}$. Because we are assuming a uniform distribution, the probability is

$$O\left(\frac{2^{x(1-\epsilon)/2}}{2^x}\right) \;=\; O\left(2^{-(1+\epsilon)x/2}\right). \tag{164}$$

Then the probability of two *Ruth-Aaron* numbers being consecutive is $O\left(2^{-(1+\epsilon x)}\right)$. Since there are $2^x - 2$ consecutive integer triples in $[2^x, 2^{x+1})$, the expected number of *Ruth-Aaron* triples between 2^x and 2^{x+1} is at most

$$\begin{aligned} O\left((2^x - 2)\cdot 2^{-(1+\epsilon)x}\right) \;&\ll\; O\left(2^x \cdot 2^{-(1+\epsilon)x}\right) \\ &=\; O\left(\frac{1}{2^{\epsilon x}}\right). \end{aligned} \tag{165}$$

Therefore, the total number of integers with $f(n) = f(n+1) = f(n+2)$ is at most

$$O\left(\sum_{x=1}^{\infty} \frac{1}{2^{\epsilon x}}\right) \;\ll\; O(1), \tag{166}$$

which means there are finitely many *Ruth-Aaron* triples if $\#\mathcal{R}_{(1,-1)}(x)$ has a rate of growth of $O\left(x^{(1-\epsilon)/2}\right)$ or lower, and *Ruth-Aaron* numbers are uniformly distributed. □

7 Acknowledgments

We thank Carl Pomerance, Serban Raianu, and the referee for many helpful comments which improved the paper.

A List of Other Arithmetic Functions with Similar Results

A.1 Euler Totient Function

Erdős, Pomerance, and Sárközy [12] proved that the number of n up to x with

$$\varphi(n) \;=\; \varphi(n+1), \tag{167}$$

where $\varphi(x)$ is the Euler Totient function, or the number of integers up to x that are relatively prime to x, is at most $x/\exp\{(\log x)^{1/3}\}$ [12]. A similar result holds for the Sigma function $\sigma(x) = \sum_{d|n} d$. In the meantime, they conjectured that for every $\epsilon > 0$ and $x > x_0(\epsilon)$, the equations $\varphi(n) = \varphi(n+1)$ and $\sigma(n) = \sigma(n+1)$ each have at least $x^{1-\epsilon}$ solutions up to x [12].

Meanwhile, Luca and Stănică [18] proved that the number of n up to x with $\phi(n) = \phi(n-1) + \phi(n-2)$ is

$$O\left(C(t) \cdot \frac{x \log\log\log x}{\sqrt{\log\log x}}\right), \tag{168}$$

where $1 \leq t < \exp(\log x/\log\log x)$ and $C(t)$ is a constant dependent upon t. Such n are referred to as *Phibonacci* numbers [2].

A.2 Prime Omega Function

Let the unique prime factorization of n be $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{n_k}$, then the Prime Omega function $\omega(n) = k$. Erdős et al. conjectured that the number of $n \leq x$ with

$$\omega(n) \;=\; \omega(n+1) \tag{169}$$

is $O\left(\frac{x}{\sqrt{\log\log x}}\right)$ [12].

B Important Theorems and Conjectures

B.1 The Prime Number Theorem

The Prime Number Theorem gives an asymptotic form for the prime counting function $\pi(x)$. In 1795, the 15-year-old Gauss conjectured that

$$\pi(x) \;\sim\; \frac{x}{\log x}. \tag{170}$$

He later refined his estimate to

$$\pi(x) \sim \text{Li}(x), \tag{171}$$

where $\text{Li}(x) = \int_2^x \frac{x}{\ln x} dx$ is the logarithmic integral. This result has then been refined numerous times up to this day, involving other arithmetic functions as well as modification on the error term [32]. In this paper, we apply the weaker version of the Prime Number Theorem:

$$\begin{aligned} \pi(x) &= \frac{x}{\log x} + O\left(\frac{x}{(\log x)^2}\right) \\ &= O\left(\frac{x}{\log x}\right) \end{aligned} \tag{172}$$

in order to compute the number of n up to x under restrictions on the largest prime factor of n, $n+1$, and $n/P(n)$.

B.2 On the Largest Prime Factor

De Bruijn [5] introduces the upper bound on the number of $n \leq x$ with prime factors no greater than y, denoted as $\Psi(x, y)$:

$$\Psi(x, y) < x \exp\left\{-\frac{\log\log\log y}{\log y}\log x + \log\log y + \log\log y + O\left(\frac{\log\log y}{\log\log\log y}\right)\right\}. \tag{173}$$

He then refined the result to, for $2 < y \leq x$,

$$\log\Psi(x, y) = Z\left\{1 + O\left(\frac{1}{\log x}\right) + O\left(\frac{1}{\log\log x}\right) + O\left(\frac{1}{u+1}\right)\right\}, \tag{174}$$

where $u = \log x / \log y$ and

$$Z = \left(\log\left(1 + \frac{\log y}{\log x}\right)\right) \cdot \frac{\log x}{\log y} + \left(\log\left(1 + \frac{\log x}{y}\right)\right) \cdot \frac{y}{\log y}. \tag{175}$$

As stated in Lemma 2.8, De Bruijn's result provides us with a crucial lower bound on the largest prime factor of $P(n)$, which we then used extensively throughout the paper.

B.3 The Chinese Remainder Theorem

Theorem B.1 (The Chinese Remainder Theorem) *If $m_1, m_2, \ldots, m_k$ are pairwise relatively prime positive integers, and if $a_1, a_2, \ldots, a_k$ are any integers, then the simultaneous congruences*

$$\begin{aligned} x &\equiv a_1 \ (mod\ m_1) \\ x &\equiv a_2 \ (mod\ m_2) \\ &\cdots \\ x &\equiv a_k \ (mod\ m_k) \end{aligned} \tag{176}$$

have solution that is unique modulo the product $m_1 m_2 \cdots m_k$ [8]. In our case, for $p = P(n), q = P(n+1)$, we have

$$\begin{aligned} n &\equiv 0 \ (mod\ p) \\ n &\equiv -1 \ (mod\ q). \end{aligned} \tag{177}$$

Applying the Chinese Remainder Theorem, we have the number of n up to x is at most $1 + \left\lfloor \frac{x}{pq} \right\rfloor$. This result is used extensively in the paper.

B.4 Mersenne Primes

We first considered using the conjectured infinitude of Mersenne primes, or primes of the form $2^n - 1$, to show the infinitude of $r-$th power *Ruth-Aaron* numbers for real r. In fact, Lenstra, Pomerance, and Wagstaff [31] have conjectured the infinitude of Mersenne primes, and they also proposed that the number of Mersenne primes less than x is approximately $e^\gamma \log_2 \log_2(x)$, where γ is the Euler-Mascheroni constant. Mathematicians have been in search of Mersenne primes both manually and computationally. As of now, the first 51 Mersenne primes have been found, with the largest being $2^{8258933} - 1$, which was discovered in December 2018 and is the largest publicly known prime.

References

[1] T. Albu, The irrationality of sums of radicals via Cogalois theory, *An. Ştiinţ. Univ. "Ovidius" Constanţa, Ser. Mat.*, **19** (2011), 15–36.

[2] A. Bager, Proposed Problem AMME 2833, *Amer. Math. Monthly*, **87** (1980), 404; with a solution in *ibid.*, **88** (1981), 622.

[3] É. Bézout, *Théorie générale des équations algébriques*, Ph.-D. Pierres, 1779.

[4] Harvard College, *The Harvard College mathematics review*, Cambridge, MA: Harvard College, 2007, available online at the URL: `http://abel.harvard.edu/hcmr/issues/2.pdf`.

[5] N.G. de Bruijn, On the number of positive integers $\leq x$ and free of prime factors $> y$ II, *Koninklijke Nederl. Akad. Wetensch. Proc. Ser. A : Mathematical Science*, **69** (1966), 239–247.

[6] P. Cohen, K. Cordwell, A. Epstein, C.H. Kwan, A. Lott, S.J. Miller, On Near Perfect Numbers, *Acta Arith.*, **194** (2020), 341–366.

[7] K. Dickman, On the density of the abundant numbers containing prime factors of a certain relative magnitude, *Arkiv Mat., Astron Fys.*, **22** (1930).

[8] C. Ding, D. Pei, and A. Salomaa, *Chinese Remainder Theorem: Applications in Computing, Coding, Cryptography*, World Scientific Publishing Co., Inc., 1996.

[9] P.D.T.A. Elliott *Probabilistic number theory (Vol. II)*, New York: Springer, 1980.

[10] P. Erdős, Some remarks on Euler's ϕ function and some related problems, *Bull. Amer. Math. Soc.*, **51** (1945), 540–544.

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

[12] P. Erdős, C. Pomerance, and A. Sárközy, On locally repeated values of certain arithmetic functions. II, *Acta Math, Hungar.*, **49** (1987), 251–259.

[13] K. Ford and S. Konyagin, On two conjectures of Sierpinski concerning the arithmetic functions σ and ϕ, (2019).

[14] K. Ford, Solutions of $\phi(n) = \phi(n+k)$ and $\sigma(n) = \sigma(n+k)$, *International Mathematics Research Notices*, (2020).

[15] S.W. Graham, J.J. Holt, and C. Pomerance, On the solutions to $\phi(n) = \phi(n+k)$, *Number Theory in Progress*, **2** (1999), 867–882.

[16] H. Halberstam, H.E. Richert, *Sieve Methods*, Academic Press, 1974.

[17] E. Landau, Über die zahlentheoretische Funktion $\phi(n)$ und ihre Beziehung zum Goldbachschen Satz, *Nachr. Köni. Ges. Wiss. Göttingen, math.-phys. KI.*, (1900), 177–186.

[18] F. Luca and P. Stănică, Linear equations with the Euler totient function, *Acta Arith.*, **128** (2007), 135–147.

[19] X. Lü and Z. Wang, On the largest prime factors of consecutive integers, 2018.

[20] J.M. De Koninck, A. Ivić, The distribution of the average prime divisor of an integer, *Arch. Math.*, **43** (1984), 37–43.

[21] P. Mihăilescu, Primary cyclotomic units and a proof of Catalans conjecture, *Journal für die reine und angewandte Mathematik*, **572** (2004), 167–195.

[22] S.J. Miller, *The Probability Lifesaver*, Princeton University Press, 2018.

[23] H.L. Montgomery, R.C. Vaughan, The large sieve. *Mathematika*, **20** (1973), 119–134.

[24] C. Nelson, D.E. Penney, and C. Pomerance, 714 and 715. *J. Recreational Math.*, **7** (1974), 87–89.

[25] Nombres - curiosités, théorie et usages, available online at URL: `http://villemin.gerard.free.fr/aNombre/TYPDIVIS/RuthAaro.htm`

[26] OEIS, Sequence A001366, available online at URL: https://oeis.org/A001366

[27] C. Pomerance, Ruth-Aaron numbers revisited, *Paul Erdős and His Mathematics, I (Series: Bolyai Soc. Math. Stud.*, **11** (2002), 567–579.

[28] R.A. Rankin, The Difference Between Consecutive Prime Numbers. IV, *Proc. Amer. Math. Soc.*, **1** (1950), 143–150.

[29] P. Ribenboim, *Catalan Conjecture*, Academic Press, 1994.

[30] C.R. Shalizi, *Advanced Data Analysis from an Elementary Point of View*, Carnegie Mellon University, Appendix A.

[31] S.S. Wagstaff, Jr. Divisors of Mersenne numbers, *Math. Comp.*, **40** (1983), 385–397.

[32] E.W. Weisstein, Cauchy's Inequality, From MathWorld—A Wolfram Web Resource, available online at the URL: https://mathworld.wolfram.com/CauchysInequality.html
— — —, Jensen's Inequality. From MathWorld—A Wolfram Web Resource, available online at URL: https://mathworld.wolfram.com/JensensInequality.html
— — —, Linearly Independent, From MathWorld—A Wolfram Web Resource, available online at URL: https://mathworld.wolfram.com/LinearlyIndependent.html
— — —, Prime Number Theorem, From MathWorld–A Wolfram Web Resource, available online at URL: https://mathworld.wolfram.com/PrimeNumberTheorem.html

Yanan Jiang
Milton Academy
170 Centre St
Milton, MA 02186
E-mail: yanan_jiang22@milton.edu

Steven J. Miller
Williams College
880 Main St
Williamstown, MA 01267
E-mail: sjm1@williams.edu

Received: April 31, 2017 **Accepted**: June 31, 2017
Communicated by Some Editor