ON THE GAP SEQUENCE AND GILBREATH'S CONJECTURE

T. AGAMA

Abstract. Motivated by Gilbreath's conjecture, we develop the notion of the gap sequence induced by any sequence of numbers. We introduce the notion of the path and associated circuits induced by an originator and study the conjecture via the notion of the trace and length of a path.

1. Introduction

Let $\mathbb{P}$ denotes the set of all prime numbers and $\{p_i\}_{i=1}^{n}$ be sequence of consecutive prime numbers. Then Gilbreath's conjecture - a not-well studied conjecture in additive number theory - is a conjecture about the distribution of the sequences generated by applying the forward difference operator on consecutive prime numbers leaving the result unassigned and repeating the process. The conjecture named after Norman L. Gilbreath - who presented it to the mathematical community in 1958 after discovering the pattern working arithmetic on a napkin - is the assertion that the first term in each of the sequences generated in this way must always be a unit. More formally the conjecture can be stated in the following manner

Conjecture 1.1 (Gilbreath). Let $\{p_n\}$ be the ordered sequence of prime numbers p_n and define each term in the sequence $\{d_n^1\}$ by

$$d_n^1 = p_{n+1} - p_n$$

where n is positive. Also for any integer $k \geq 2$ let the terms in $\{d_n^k\}$ be given by

$$d_n^k = |d_{n+1}^{k-1} - d_n^{k-1}|.$$

Then Gilbreath's conjecture is the assertion that $d_1^k = 1$ for all $k \geq 1$.

The conjecture was studied long before Gilbreath's observation by Francois Proth who allegedly had obtained a proof which was invalidated [1]. The conjecture remains unresolved as of now but it had been verified computationally to be true by Andrew Odlyzko, that $d_1^k = 1$ for all $k \leq n = 3.4 \times 10^{11}$ in 1993 [3]. There has also been spates of attempts generalizing Gilbreath' conjecture by many authors to other non-prime sequences obeying similar distribution of prime numbers with certain specifications on their gaps [4] but various counter examples have now been found. Nonetheless a careful study by Andrew Odlyzko confirms the generalization to sequences starting with 1 and others of even parity with not-too-large gaps and sufficiently random [2].
In this paper we introduce and develop the notion of a **path** and **circuit** induced by a sequence. We study associated statistics of paths such as the **trace** and

Date: April 12, 2021.
2000 *Mathematics Subject Classification.* Primary 54C40, 14E20; Secondary 46E25, 20C20.
Key words and phrases. path; order; step.

formulate a finite version of Gilbreath's conjecture in the language of trace, so that establishing the finite version would immediately imply the actual version of the conjecture. Our studies and further studies of this conjecture will be studied in the following language:

Conjecture 1.2 (Gilbreath)**.** Let $\mathbb{P}$ denotes the set of all prime numbers and $\{d_j^k\}_{j\geq 1}$ for all $1\leq k\leq n-1$ be the circuit induced by the originator $\{p_i\}_{i=1}^n$ where each $p_i\in \mathbb{P}$. Then $d_1^k>0$ for all $1\leq k\leq n-1$ and $\tau_{n,1}=n-1$ for all $n\geq 2$.

2. **The notion of a path induced by a sequence**

In this section we introduce and study the notion of a path induced by an originator. We study some few statistics in this direction and establish some inequalities for our studies in the sequel.

Definition 2.1. Let $\{a_i\}_{i=1}^n$ be any finite sequence. Then by the **path** of **order** 1 with **steps** $l\geq 1$ induced by the sequence, we mean the sequence $\{d_j^1\}_{j=1}^l$ such that

$$d_1^1=|a_2-a_1|,d_2^1=|a_3-a_2|,\ldots,d_l^1=|a_{l+1}-a_l|.$$

Similarly by the **path** of **order** $k\geq 2$ with t $(t<l)$ **steps** induced by the sequence $\{a_i\}_{i=1}^n$, we mean the sequence $\{d_j^k\}_{j=1}^t$ such that

$$d_1^k=|d_2^{k-1}-d_1^{k-1}|,\ldots,d_t^k=|d_{t+1}^{k-1}-d_t^{k-1}|$$

and we call each d_j^k for $1\leq j\leq t$ a **segment** of the path induced. We call d_1^k the **prime** segment of the path. We call the sequence $\{a_i\}_{i=1}^n$ the **originator** of the paths and we denote with $a_i=d_i^0$ for $1\leq i\leq n$. Similarly we call the originator the **trivial** path induced with $\{a_i\}_{i=1}^n=\{d_i^0\}_{i=1}^n$.

Proposition 2.1. *Let $\{d_j^k\}_{j=1}^t$ be a path of order $k\geq 1$ with step t with originator $\{a_i\}_{i=1}^n$. Then the path $\{d_i^{k+1}\}_{i\geq 1}$ has exactly $t-1$ steps.*

Proof. Suppose $\{d_j^k\}_{j=1}^t$ is a path of order $k\geq 1$ with step t with originator $\{a_i\}_{i=1}^n$. Then $d_i^{k+1}=|d_{j+1}^k-d_j^k|$ for $t-1\geq j\geq 1$ is a segment of the path $\{d_i^{k+1}\}_{i\geq 1}$ and each such segment is **uniquely** determined by $t-1$ segments of the path $\{d_j^k\}_{j=1}^t$. It follows that the path $\{d_i^{k+1}\}_{i\geq 1}$ must have exactly $t-1$ steps. $\square$

It follows that the number of steps of paths induced by any sequence must experience some amount of drop with an increase in the order of the path. In particular, the number of steps in a path produced by some originator of order l must be a unit more step than the path of order $l+1$ with the same originator.

Proposition 2.2. *Let $\{a_i\}_{i=1}^n$ be an originator of paths, then the total number of steps in all induced paths must be*

$$\frac{n(n-1)}{2}.$$

Proof. Suppose $\{a_i\}_{i=1}^{n}$ is an originator of paths, then appealing to Proposition 2.1 the path of order 1 must have exactly $(n-1)$ steps. The path of order 2 must have exactly $(n-2)$ steps. By induction the path of order $k \geq 2$ must have $(n-k)$ steps. By iterating downwards we generate the steps of all such induced path by the originator terminating to 1. Thus the total number of such steps of all induced paths is given by

$$1+2+\cdots+(n-2)+(n-1)=\frac{n(n-1)}{2}.$$

□

Remark 2.2. We relate the notion of the step and the order of a path to the number of terms in an originator. This is an easy consequence of Proposition 2.2.

Proposition 2.3 (Step-order equation)**.** *Let $\{a_i\}_{i=1}^{n}$ be an originator of the path $\{d_j^k\}_{j=1}^{t}$. Then we have*

$$n=k+t.$$

Proof. Let $\{d_j^k\}_{j=1}^{t}$ be the path induced by the originator $\{a_i\}_{i=1}^{n}$. Then by appealing to Proposition 2.2 the number of steps t in the path must satisfy

$$t=n-k.$$

□

3. The length of a path

In this section we introduce and study the notion of the **length** of a path.

Definition 3.1. Let $\{d_j^k\}_{j=1}^{t}$ be a path of order $k \geq 1$ with step t induced by the sequence $\{a_i\}_{i=1}^{n}$. Then by the length of the path, denoted $\iota(d_j^k)$, we mean the finite sum

$$\iota_{t,k}=\sum_{j=1}^{t} d_j^k.$$

Remark 3.2. Next we establish a somewhat crude inequality that relates the length of each path to the worst segment of the previous consecutive path. This relationship will turn out to be useful to our further studies in the sequel.

Proposition 3.1. *Let $\{d_j^k\}_{j=1}^{t}$ be a path with originator $\{a_i\}_{i=1}^{n}$. Then for all $k \geq 1$ the inequality holds*

$$|d_{n-k}^{k-1}-d_1^{k-1}| \leq \iota_{n-k,k} \leq (n-k)\max\{|d_{j+1}^{k-1}-d_j^{k-1}|\}_{j=1}^{n-k}.$$

Proof. Appealing to Definition 3.1 and Proposition 2.3, we can write

$$\begin{aligned}
\iota_{n-k,k} &= \sum_{j=1}^{t} d_j^k \\
&= \sum_{j=1}^{n-k} |d_{j+1}^{k-1}-d_j^{k-1}| \\
&\leq \max\{|d_{j+1}^{k-1}-d_j^{k-1}|\}_{j=1}^{n-k} \sum_{j=1}^{n-k} 1
\end{aligned}$$

thereby establishing the upper bound. The lower bound however follows by adding and deleting of the segments of the path of order $(k-1)$ and subsequent appeal to the triangle inequality. $\square$

It turns out that a good knowledge of the largest value of a segment in a given path provides at least an information about at least one segment in the closest previous path. We leverage the inequality devised in Proposition 3.1 to make this assertion more formal.

Proposition 3.2. *Let $\{d_j^k\}_{j=1}^t$ be a path with originator $\{a_i\}_{i=1}^n$. If $\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k} \leq c$ for some $c>0$, then there exists at least some $1 \leq m \leq (n-k)$ such that $d_m^k \leq c$.*

Proof. Suppose $\{d_j^k\}_{j=1}^t$ is a path with originator $\{a_i\}_{i=1}^n$. Then by appealing to Proposition 3.1, we have the inequality

$$\iota_{n-k,k} \leq (n-k)\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}.$$

Under the requirement $\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k} \leq c$ for some $c>0$, then it follows that

$$\begin{aligned}\iota_{n-k,k} &\leq (n-k)\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k} \\ &\leq c(n-k)\end{aligned}$$

so that the average value of segments in the path with $(n-k)$ steps is given by

$$\frac{\iota_{n-k,k}}{(n-k)} = \frac{1}{(n-k)}\sum_{j=1}^{n-k}|d_{j+1}^{k-1} - d_j^{k-1}| \leq c.$$

It follows that there must exists some $1 \leq m \leq (n-k)$ such that $d_m^k \leq c$ for $c>0$. Suppose for all such $1 \leq m \leq (n-k)$ then $d_m^k > c$, then $\iota_{n-k,k} > c(n-k)$. It follows that

$$c(n-k) < \iota_{n-k,k} \leq (n-k)\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}$$

so that $c < \max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}$, which is a contradiction. $\square$

Proposition 3.3. *Let $\{d_j^k\}_{j=1}^t$ and $\{d_j^{k+1}\}_{j=1}^{t-1}$ be any two paths of the same originator such that $|d_{j+1}^k - d_j^k| \leq d_{j+1}^k$ for all $1 \leq j \leq t$. Then the inequality holds*

$$\iota_{t-1,k+1} < \iota_{t,k}$$

for all $k \geq 1$.

Proof. Appealing to Definition 3.1 we can write

$$\begin{aligned}\iota_{t-1,k+1} &= \sum_{j=1}^{t-1} d_j^{k+1} \\ &= \sum_{j=1}^{t-1}|d_{j+1}^k - d_j^k|.\end{aligned}$$

Under the requirement that $|d_{j+1}^k - d_j^k| \leq d_{j+1}^k$ for all $1 \leq j \leq t$, we have the inequality

$$\sum_{j=1}^{t-1} |d_{j+1}^k - d_j^k| \leq \sum_{j=1}^{t-1} d_{j+1}^k$$
$$= \sum_{j=1}^{t} d_j^k = \iota_{t,k}.$$

$\square$

Remark 3.3. It suggests very clearly that for all the paths induced by the originator $\{a_i\}_{i=1}^n$ the worst order and the least step attainable are $n-1$ and 1, respectively. Next we introduce and study the notion of a circuit and associated statistics.

4. **The notion of a circuit**

In this section we introduce and study the notion of a circuit generated by paths induced by a certain originator.

Definition 4.1. Let $\{a_i\}_{i=1}^n$ be a generator of the paths $\{d_j^k\}_{j\geq 1}$. Then we call the collection of all such paths for all $1 \leq k \leq n-1$ the circuit induced by the originator.

Definition 4.2. Let $\{d_j^k\}_{j\geq 1}$ for all $1 \leq k \leq n-1$ be the circuit induced by the originator $\{a_i\}_{i=1}^n$. Then we denote the length of the circuit with

$$\kappa(n) := \sum_{k=1}^{n-1} \iota_{n-k,k}.$$

Proposition 4.1. *Let $\{d_j^k\}_{j\geq 1}$ for all $1 \leq k \leq n-1$ be the circuit induced by the originator $\{a_i\}_{i=1}^n$. Then the inequality holds*

$$(n-2)\mathrm{min}\{|d_{n-k}^{k-1} - d_1^{k-1}|\}_{k=1}^{n-2} \leq \kappa(n) \leq \sum_{k=1}^{n-1} \mathrm{max}\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}$$
$$+ \int_1^{n-1} \left(\sum_{s=1}^{t} \mathrm{max}\{|d_{j+1}^{s-1} - d_j^{s-1}|\}_{j=1}^{n-s} \right) dt.$$

Proof. The lower bound follows by an appeal to the lower bound in Proposition 3.1. The upper bound follows by an application of partial summation to the sum

$$\kappa(n) := \sum_{k=1}^{n-1} \iota_{n-k,k}$$
$$\leq \sum_{k=1}^{n-1} (n-k)\mathrm{max}\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}.$$

$\square$

Definition 4.3. Let $\{d_j^k\}_{j\geq 1}$ for all $1\leq k\leq n-1$ be the circuit induced by the originator $\{a_i\}_{i=1}^n$. Then by the **trace** of the sth segment of paths in a circuit, denoted $\tau_{n,s}$, we mean the finite sum

$$\tau_{n,s}:=\sum_{k=1}^{n-s}d_s^k.$$

Proposition 4.2. *Let $\{d_j^k\}_{j\geq 1}$ for all $1\leq k\leq n-1$ be the circuit induced by the originator $\{a_i\}_{i=1}^n$, then the inequality holds*

$$2\tau_{n,s}\geq (a_{s+1}-a_s)+d_s^{n-s}+\tau_{n,s+1}.$$

Proof. First we note that we can write

$$\begin{aligned}\tau_{n,s}:&=\sum_{k=1}^{n-s}d_s^k\\ &=\sum_{k=1}^{n-s}|d_{s+1}^{k-1}-d_s^{k-1}|\\ &\geq\sum_{k=1}^{n-s}(d_{s+1}^{k-1}-d_s^{k-1})\\ &=\sum_{k=1}^{n-s}d_{s+1}^{k-1}-\sum_{k=1}^{n-s}d_s^{k-1}\\ &=\sum_{i=0}^{n-s-1}d_{s+1}^{i}-\sum_{i=0}^{n-s-1}d_s^{i}\\ &=d_{s+1}^0+\sum_{i=1}^{n-(s+1)}d_{s+1}^i-\sum_{i=1}^{n-s}d_s^i-d_s^0+d_s^{n-s}\\ &=(a_{s+1}-a_s)+d_s^{n-s}+\tau_{n,s+1}-\tau_{n,s}\end{aligned}$$

thereby establishing the desired inequality. ∎

It follows that we can write the length of a circuit $\kappa(n)$ with originator $\{a_i\}_{i=1}^n$ as the sum of the trace of segments of each kind within paths in the circuit. To that end, we can write

$$\begin{aligned}\kappa(n)&=\sum_{k=1}^{n-1}\iota_{n-k,k}\\ &=\sum_{k=1}^{n-1}\sum_{s=1}^{n-k}d_s^k\end{aligned}$$

so that by interchanging the order of summation we have

$$\begin{aligned}
\kappa(n) &= \sum_{k=1}^{n-1}\sum_{s=1}^{n-k} d_s^k \\
&= \sum_{s=1}^{n-1} d_s^1 + \sum_{s=1}^{n-2} d_s^2 + \cdots + \sum_{s=1}^{n-(n-2)} d_s^{n-2} + d_s^{n-1} \\
&= \left(d_1^1 + d_1^2 + \cdots + d_1^{n-1}\right) + \left(d_2^1 + d_2^2 + \cdots + d_2^{n-2}\right) + \cdots + d_1^{n-1} \\
&= \sum_{k=1}^{n-1} d_1^k + \sum_{k=1}^{n-2} d_2^k + \cdots + \sum_{k=1}^{n-(n-1)} d_{n-1}^k \\
&= \sum_{s=1}^{n-1}\sum_{k=1}^{n-s} d_s^k \\
&= \sum_{s=1}^{n-1} \tau_{n,s}.
\end{aligned}$$

It follows that the total length of any given circuit can also be obtained by summing the trace of each segment in a circuit, so that we can upper and lower bound the average trace in a circuit by an appeal to Proposition 4.1 as

Proposition 4.3. *Let* $\{d_j^k\}_{j\geq 1}$ *for all* $1 \leq k \leq n-1$ *be the circuit induced by the originator* $\{a_i\}_{i=1}^n$*. Then the inequality holds*

$$(n-2)\min\{|d_{n-k}^{k-1} - d_1^{k-1}|\}_{k=1}^{n-2} \leq \sum_{s=1}^{n-1} \tau_{n,s} \leq (n-1)\max_{1\leq k\leq n-1}\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}$$
$$+ \int_1^{n-1} \left(\sum_{s=1}^{t} \max\{|d_{j+1}^{s-1} - d_j^{s-1}|\}_{j=1}^{n-s}\right) dt.$$

Proof. The lower bound follows from the lower bound in Proposition 4.1. The upper bound follows by appealing to the upper bound in Proposition 4.1 and noting that

$$\sum_{k=1}^{n-1} \max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k} \leq (n-1)\max_{1\leq k\leq n}\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}.$$

□

The upper bound in Proposition 4.3 does suggests on average the trace of segments in a circuit must be at most

$$\leq \max_{1\leq k\leq n}\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}$$

so that there must exists some $1 \leq m \leq n-1$ such that $\tau_{n,m} \leq \max_{1\leq k\leq n}\max\{|d_{j+1}^{k-1} - d_j^{k-1}|\}_{j=1}^{n-k}$. Next we leverage the inequality in Proposition 4.2 to establish an inequality relating the length of a circuit to the terms of the originator and the trace of the first segment in each path in the circuit.

Theorem 4.4. *Let $\{d_j^k\}_{j\geq 1}$ for all $1 \leq k \leq n-1$ be the circuit induced by the originator $\{a_i\}_{i=1}^n$. Then the inequality holds*

$$\kappa(n) + \tau_{n,1} \geq (2a_n - a_{n-1} - a_1) + \sum_{j=1}^{n-2} d_j^{n-j}.$$

Proof. By iterating the inequality in Proposition 4.2, we obtain the following chains of inequalities

$$2\tau_{n,1} \geq (a_2 - a_1) + d_1^{n-1} + \tau_{n,2}$$

$$2\tau_{n,2} \geq (a_3 - a_2) + d_2^{n-2} + \tau_{n,3}$$

$$\vdots$$

$$\vdots$$

$$2\tau_{n,n-2} \geq (a_{n-1} - a_{n-2}) + d_{n-2}^2 + \tau_{n,n-1}.$$

Adding the left hand-sides and the right-hand sides of the chain, we obtain further the inequality

$$2\sum_{s=1}^{n-2} \tau_{n,s} \geq (a_{n-1} - a_1) + \sum_{j=1}^{n-2} d_j^{n-j} + \sum_{s=2}^{n-1} \tau_{n,s}.$$

By adding and deleting the term $2\tau_{n,n-1}$ on the left-hand side of the inequality and $\tau_{n,1}$ on the right-hand side, we obtain the refined inequality

$$2\sum_{s=1}^{n-1} \tau_{n,s} \geq \sum_{s=1}^{n-1} \tau_{n,s} + \sum_{j=1}^{n-2} d_j^{n-j} + (a_{n-1} - a_1) + 2\tau_{n,n-1} - \tau_{n,1}.$$

It follows that we can write

$$\begin{aligned}\sum_{s=1}^{n-1} \tau_{n,s} &\geq \sum_{j=1}^{n-2} d_j^{n-j} + (a_{n-1} - a_1) + 2\tau_{n,n-1} - \tau_{n,1} \\ &= (2a_n - a_{n-1} - a_1) + \sum_{j=1}^{n-2} d_j^{n-j} - \tau_{n,1}\end{aligned}$$

by exploiting the relation $\tau_{n,n-1} = d_{n-1}^1 = a_n - a_{n-1}$, thereby establishing the inequality. □

Proposition 4.4. *Let $\{d_j^k\}_{j\geq 1}$ for all $1 \leq k \leq n-1$ be the circuit induced by the originator $\{a_i\}_{i=1}^n$ with each $a_i \in \mathbb{Z}$. If $\tau_{n,s} < n-s$ then there exists at least some t such that $d_s^t = 0$ for $1 \leq t \leq n-s$.*

Proof. Under the assumption that $\{d_j^k\}_{j\geq 1}$ for all $1 \leq k \leq n-1$ is the circuit induced by the originator $\{a_i\}_{i=1}^n$, then we obtain the lower bound

$$\begin{aligned}\tau_{n,s} &:= \sum_{k=1}^{n-s} d_s^k \\ &\geq \min\{d_s^k\}_{k=1}^{n-s} \sum_{k=1}^{n-s} 1 = (n-s)\min\{d_s^k\}_{k=1}^{n-s}\end{aligned}$$

and under the requirement $\tau_{n,s} < n-s$ with $\min\{d_s^k\}_{k=1}^{n-s} \in \mathbb{Z}^+ \cup \{0\}$, we can take $\min\{d_s^k\}_{k=1}^{n-s} = 0$, thereby ending the proof. □

Proposition 4.5. *Let $\{d_j^k\}_{j\geq 1}$ for all $1 \leq k \leq n-1$ be the circuit induced by the originator $\{a_i\}_{i=1}^n$. If $d_1^k > 0$ for all $1 \leq k \leq n-1$ and $\tau_{n,1} = n-1$ for all $n \geq 2$ then $d_1^k = 1$ for all $1 \leq k \leq n-1$.*

Proof. Under the assumption that for the circuit $\{d_j^k\}_{j\geq 1}$ for all $1 \leq k \leq n-1$ induced by the originator $\{a_i\}_{i=1}^n$ with $\tau_{n,1} = n-1$, then it follows that

$$\tau_{n,1} = \sum_{k=1}^{n-1} d_1^k = n-1.$$

Since there are $n-1$ prime segments in the sum and each prime segment $d_1^k > 0$ for all $1 \leq k \leq n-1$, then $d_1^k = 1$ for $1 \leq k \leq n-1$. □

Remark 4.5. It turns out we can restate Gilbreath's conjecture in the language of the trace, so that proving this version of the conjecture would certainly imply the actual version of Gilbreath's conjecture.

Conjecture 4.1 (Gilbreath)**.** Let $\mathbb{P}$ denotes the set of all prime numbers and $\{d_j^k\}_{j\geq 1}$ for all $1 \leq k \leq n-1$ be the circuit induced by the originator $\{p_i\}_{i=1}^n$ where each $p_i \in \mathbb{P}$. Then $d_1^k > 0$ for all $1 \leq k \leq n-1$ and $\tau_{n,1} = n-1$ for all $n \geq 2$.

[1].

References

1. Proth, Francois *Théoremes sur les nombres premiers*, CR Acad. Sci. Paris, vol. 87(2), Elsevier, 1878, 926.
2. David, Darling, *The Universal Book of Mathematics*, by John Wiley & Sons, Inc., Hoboken, 2004.
3. Odlyzko, Andrew M., *Iterated absolute values of differences of consecutive primes*, Mathematics of computation, vol.61(203), 1993, 373–380.
4. Guy, Richard, *Unsolved problems in number theory*, Springer Science & Business Media, vol.1, Taylor & Francis, 2004.

Department of Mathematics, African Institute for Mathematical science, Ghana
E-mail address: theophilus@aims.edu.gh/emperordagama@yahoo.com

[1] .