A family of sequences generated by eliminating some prime factors from the sequence of the naturals

Zhihe Li

Ningbo Xiaoshi High School

Adviser: Cindy Boyd & David Ochoro

Table of the Paper

Abstract..**03**

Section 1..**05**

Section 2..**08**

Section 3..**15**

Section 4..**26**

Summary and Prospect...**40**

Reference..**42**

ABSTRACT

This paper starts with an interesting property of the sequence of the naturals $\{A_n\}$, i.e. $\{A_n\}:=\{1,2,3,...\}$. Considering the generated sequence $\{L_n\}$ by eliminating all the factor 2 in A_n if it has, we show that the partial sum of the new sequence $\{L_n\}$ from 2^{n-1}th term to (2^n-1) th term is square number 2^{2n-2}.

This property is further extended to the more general case: $\{H(n)\}$ is the sequence with each term $H(n)$ generated according to the following transforming rule:

ER $(\{p_1,p_2,...,p_m\})$: Given m primes $p_i, i=1,2,...,m$, for each $i\in\{1,2,...,m\}$, if A_n has factor p_i, then substitute A_n with A_n/p_i until it does not have factor p_i.

We focus on the partial sum of $\{H(n)\}$. In **section one**, we research on Question 1: Given one prime p, we eliminate all the factor p in $\{A_n\}$ according to the transforming rule ER$(\{p\})$. Then the new sequence $\{L_n\}$ obtained has the property as follows:

$$\sum_{i=kp^{n-1}}^{kp^n-1} L_i=\frac{(p-1)}{2}k^2p^{2n-1}.$$

The main results in this section are Theorem 1, 2 and the relative corollaries.

In **section two,** we mainly study Question 2: Given two prime p,q, we eliminate all factors p and q in $\{A_n\}$ according to the transforming rule ER$(\{p,q\})$. Then the new sequence $\{H(n)\}$ obtained has the property as follows:

$$\sum_{i=1}^{kp^rq^s} H(i)=\frac{k^2pq(p^{2r}-1)(q^{2s}-1)}{2(p+1)(q+1)}+\sum_{i=1}^{kp^r}H(i)+\sum_{i=1}^{kq^s}H(i)-\sum_{i=1}^{k}H(i)$$

(or $\sum_{i=kp^r}^{kp^rq^s} H(i)=\frac{k^2pq(p^{2r}-1)(q^{2s}-1)}{2(p+1)(q+1)}+\sum_{i=k}^{kq^s}H(i)$).

We also focus on the number of valued “1” point of $\{H(n)\}$ from this section, where the term in $\{A_n\}$ which becomes the number 1 in $\{H(n)\}$ is called a valued “1” point of

$\{H(n)\}$.Let $Y(n)=\{i \mid H(i)=1, i\in N, i\le n\}$, we have

$$|Y(kp^r q^s)|=|Y(kp^r)|+|Y(kq^s)|-|Y(k)|+rs,$$

The main results in this section are Theorem 3, 4, 5, 6 and the relative corollaries.

In **section three,** we extend to the case of given m primes $p_i, i=1,2,...,m$, i.e. Question 3, and obtain similar results in Theorem 7,8,9,10 and the relative corollaries.

In **section four,** we propose Question 4 and obtain several concise inequalities to estimate the partial sum of the first n terms and the number of valued "1" point in the first n terms. Furthermore, we derive a new proof for the noted fact that the set of primes is infinite.

For given two primes case, we have (Theorem 11,12)

$$\frac{pq}{2(p+1)(q+1)}x^2-2x<\sum_{i\le x}H(i)<\frac{pq}{2(p+1)(q+1)}x^2+2x$$

$$\frac{1}{2}\frac{\log x}{\log p}\left(\frac{\log x}{\log q}+1\right)<|Y(x)|<\frac{1}{2}\left(\frac{\log x}{\log p}+1\right)\left(\frac{\log x}{\log q}+2\right) \qquad \text{(where } p<q\text{)}$$

For given m primes case, we have (Theorem 13,14)

$$\frac{1}{2}\prod_{i=1}^{m}\frac{p_i}{p_i+1}x^2-2^{m-1}x<\sum_{i\le x}H(i)<\frac{1}{2}\prod_{i=1}^{m}\frac{p_i}{p_i+1}x^2+2^{m-1}x.$$

$$|Y(x)|=\frac{1}{m!}\frac{\log^m x}{\prod_{i=1}^{m}\log p_i}+O(\log^{m-1}x).$$

Finally, we point out that the multiplicative functions $H(n), H'(n)$ have further research value.

Key words: sequence of the naturals, multiplicative function, partial sum, valued "1" point.

Section 1

The set of naturals is buried with numerous secrets. Among those, I find an interesting rule. Let { A_n } be the sequence of the naturals i.e. $\{A_n\}=\{1,2,3,...,n,...\}$.We obtain a new sequence { L_n } by eliminating all the factor 2 for every term of $\{A_n\}$ according to the following rule:

(1)If A_n has factor 2,then we substitute A_n with $A_n/2$ until it does not have factor 2;

(2) If A_n does not have factor 2,then keep it unchanged.

Thus, we obtain a new sequence $\{L_n\}$ as follows:

1,1,3,1,5,3,7,1,9,5,11,3,13,7,15,1,17,9,19,5,21,11,23,3,25,13,27,7,29,15,31,1,33,17,35,9,37,19,39,5,41,… （1）

We can dig out some secrets if we divide it into groups with 1 being the first term of each group as follows:

(1),(1,3),(1,5,3,7),(1,9,5,11,3,13,7,15),(1,17,9,19,5,21,11,23,3,25,13,27,7,29,15,31),(1,33,17,35,9,37,19,39,5,41,……

Summing up each group, we derive that the sum of these groups are 1,4,16,64,… respectively. Thus we have the following conjecture：

Conjecture 1. The sum of the new sequence { L_n } from 2^{n-1} th term to (2^n-1) th term is 2^{2n-2}.

Since { L_n } contains lots of 3,5,…, we also can divide them by setting 3,5…be the first term of each group as follows:

(1,1),(3,1,5),(3,7,1,9,5,11),(3,13,7,15,1,17,9,19,5,21,11,23),(3,25,13,27,7,29,15,31,1,33,17,35,9,37,19,39,5,41,……

(1,1,3,1),(5,3,7,1,9),(5,11,3,13,7,,15,1,17,9,19),(5,21,11,23,3,25,13,27,7,29,15,31,1,33,17,35,9,37,19,39),(5,41,……

We find that if we divide {L_n} by 3, the sums of second, third, fourth,…groups are 9,36,144,… respectively; if we divide it by 5, the sums of second, third, fourth,…groups are 25, 100, 400,… respectively;…

Hence, we put forward the second conjecture.

Conjecture 2. The sum of the new sequence {L_n} from $k\times 2^{n-1}$th term to $(k\times 2^n-1)$ th term is $k^2\times 2^{2n-2}$,where $k\in N$.

The above transformation is eliminating all the factor 2 for every term of $\{A_n\}$. But what property does the new sequence have if we eliminate all the factor 3 for every term of $\{A_n\}$? Or even more generally, we consider the following question:

Question 1: Let p be a prime. We obtain a new sequence {L_n} by transforming $\{A_n\}$ according to the rule ER$(\{p\})$ as follows:

(1)If A_n has factor p, then we substitute A_n with A_n/p until it does not have factor p;

(2) If A_n does not have factor p, then keep it unchanged.

What properties does {L_n} have?

The rule ER$(\{p\})$above can also be written as follows:

If $A_n=p^r\times X$, where $r\ge 0, X\ge 1$are integers, $(X,p)=1$, then $L_n=X$.

For this question, we derive the following theorem.

Theorem 1. Let p be a prime. Let {L_n} be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p\})$of Question 1. Then for $k\in N, n\in N$,the following equality holds:

$$\sum_{i=kp^{n-1}}^{kp^n-1} L_i=\frac{(p-1)}{2}k^2p^{2n-1} \qquad (2)$$

Proof. The sum $\sum_{i=kp^{n-1}}^{kp^n-1} L_i$ is from kp^{n-1}th term to (kp^n-1)th term, which means the number of the terms is $kp^n - 1 - kp^{n-1} + 1 = k(p-1)p^{n-1}$.

Now, we claim that the $k(p-1)p^{n-1}$ terms are different from each other.

Indeed, assume that there are positive integers i and j such that $kp^{n-1} \le i < j \le kp^n - 1$, with $L_i = L_j$. Then by the definition of sequence{ L_n }, we have that $i = p^r \times L_i$, $j = p^s \times L_j$ where r and s are nonnegative integers. Under the assumption, we have $j = i \times p^{s-r}, s - r \ge 1$, which contradicts with $kp^{n-1} \le i < j \le kp^n - 1$.

Furthermore, those $k(p-1)p^{n-1}$ terms are just the positive integers no larger than kp^n except those multiples of p. In fact, the set of positive integers which are no larger than kp^n contains kp^{n-1} multiples of p. Thus, there are $k(p-1)p^{n-1}$ terms left exactly, after these kp^{n-1} multiples of p removed.

Note that the sum of these kp^{n-1} multiples of p removed is $p \times \sum_{i=1}^{kp^{n-1}} i$, so we have

$$\sum_{i=kp^{n-1}}^{kp^n-1} L_i = \sum_{i=1}^{kp^n} i - p \times \sum_{i=1}^{kp^{n-1}} i = \frac{1}{2}\left(1 + kp^n\right)kp^n - p \times \frac{1}{2}\left(1 + kp^{n-1}\right)kp^{n-1}$$

$$= \frac{1}{2}\left(kp^n\right)^2 - \frac{1}{2}kp^n kp^{n-1} = \frac{(p-1)}{2}k^2 p^{2n-1} \qquad \square$$

Theorem 2 is derived from Theorem 1.

Theorem 2. Let p be a prime. Let { L_n } be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p\})$of Question 1. Then for $k, n \in N$,the following equality holds:

$$\sum_{i=1}^{kp^n} L_i = \sum_{i=1}^{k} L_i + \frac{k^2 p(p^{2n} - 1)}{2(p+1)} \tag{3}$$

Proof. $\sum_{i=1}^{kp^n} L_i = \sum_{i=1}^{k} L_i + \sum_{j=1}^{n} \sum_{i=kp^{j-1}+1}^{kp^j} L_i$

$$=\sum_{i=1}^{k}L_i+\sum_{j=1}^{n}\frac{(p-1)}{2}k^2p^{2j-1}$$

$$=\sum_{i=1}^{k}L_i+\sum_{j=1}^{n}\frac{1}{2}\left(1-\frac{1}{p}\right)k^2p^{2j}$$

$$=\sum_{i=1}^{k}L_i+\frac{1}{2}\left(1-\frac{1}{p}\right)k^2\frac{p^2(1-p^{2n})}{1-p^2}$$

$$=\sum_{i=1}^{k}L_i+\frac{k^2p(p^{2n}-1)}{2(p+1)}. \qquad \square$$

For the case k＝1，the following corollary holds：

Corollary 1. Let p be a prime. Let $\{L_n\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p\})$ of Question 1. Then for $n\in N$,the following equality holds:

（1） $\sum_{i=p^{n-1}}^{p^n-1}L_i=\frac{(p-1)}{2}p^{2n-1}$

（2） $\sum_{i=1}^{p^n-1}L_i=\frac{p(p^{2n}-1)}{2(p+1)}$

For the case p＝2，both conjecture 1 and conjecture 2 are correct.

Section 2

In first section, we study the properties of the new sequence generated by eliminating all the factor p in $\{A_n\}$. Naturally, we wish to generalize it for the case of eliminating more primes. In this section, we consider the case of given two primes.

Question 2. Let p and q be primes. We obtain a new sequence $\{H(n)\}$ by transforming $\{A_n\}$ according to the rule ER$(\{p,q\})$ as follows:

(1)If A_n has factor p, then we substitute A_n with A_n/p until it does not have factor p; If A_n has factor q, then we substitute A_n with A_n / q until it does not have factor q;

(2) If A_n does not have factors p and q, then keep it unchanged.

What properties does $\{H(n)\}$ have?

The rule ER$(\{p,q\})$ above can also be written as follows:

If $A_n = p^r q^s \times X$, where $r \geq 0, s \geq 0, X \geq 1$ are integers, $(X, pq) = 1$, then $H(n) = X$.

For instance, let p=2, q=3. Then the generated sequence $\{H(n)\}$ is:

1,1,1,1,5,1,7,1,1,5,11,1,13,7,5,1,17,1,19,5,7,11,23,1,25,13,1,7,29,5,31,1,11,17,35,1,37,19,13,5,41,7,43,11,…… (4)

This sequence becomes much more complicated. What can be seen is that more number 1 occurs in this new sequence. In the following part, we study the partial sum of this new sequence and the frequency of number 1.

Firstly, the number-theoretic function $H(n): N \to R$ is a completely multiplicative function (cf. [1], p77-79) where R is the set of real numbers. And the function has following properties.

(1) $H(p) = H(q) = H(1) = 1$;

$H(p') = p'$, when p' is a prime, $p' \neq p, p' \neq q$.

(2) $H(mn) = H(m)H(n), H(pn) = H(n), H(qn) = H(n)$;

(3) $H(n) = n$, when n does not have factor p or q.

Theorem 3. Let p and q be primes. Let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p,q\})$ of Question 2. Then for $k \in N$, the following equality holds:

$$\sum_{i=1}^{kpq} H(i) = \frac{k^2}{2} pq(p-1)(q-1) + \sum_{i=1}^{kp} H(i) + \sum_{i=1}^{kq} H(i) - \sum_{i=1}^{k} H(i) \tag{5}$$

Note that this equality can be written as the following briefer forms:

$$\sum_{i=kp}^{kpq} H(i) = \frac{k^2}{2} pq(p-1)(q-1) + \sum_{i=k}^{kq} H(i)$$

or
$$\sum_{i=kq}^{kpq} H(i) = \frac{k^2}{2} pq(p-1)(q-1) + \sum_{i=k}^{kp} H(i) \qquad (5')$$

Proof. we divide $\{H(i), i=1,2,...,kpq\}$ into two groups and sum up for each group.

In the first group, i is neither a multiple of p nor a multiple of q. Then $H(i)=$i. By Principle of cross-classification (cf. [2], p123) , the sum of this group is

$$S_1 = \sum_{i=1}^{kpq} i - p\sum_{i=1}^{kq} i - q\sum_{i=1}^{kp} i + pq\sum_{i=1}^{k} i$$

In the second group, i is a multiple of p or q. By Principle of cross-classification as well, we derive that this group sums is

$$S_2 = \sum_{i=1}^{kq} H(pi) + \sum_{i=1}^{kp} H(qi) - \sum_{i=1}^{k} H(pqi).$$

Hence

$$\sum_{i=1}^{kpq} H(i) = S_1 + S_2$$

$$= \sum_{i=1}^{kpq} i - p\sum_{i=1}^{kq} i - q\sum_{i=1}^{kp} i + pq\sum_{i=1}^{k} i + \sum_{i=1}^{kq} H(pi) + \sum_{i=1}^{kp} H(qi) - \sum_{i=1}^{k} H(pqi)$$

$$= \frac{1}{2}kpq(1+kpq) - \frac{1}{2}p\times kq(1+kq) - \frac{1}{2}q\times kp(1+kp) + \frac{1}{2}pq\times k(1+k)$$

$$+ \sum_{i=1}^{kq} H(i) + \sum_{i=1}^{kp} H(i) - \sum_{i=1}^{k} H(i)$$

$$= \frac{kpq}{2}(1+kpq-1-kq-1-kp+1+k) + \sum_{i=1}^{kq} H(i) + \sum_{i=1}^{kp} H(i) - \sum_{i=1}^{k} H(i)$$

$$= \frac{k^2}{2} pq(p-1)(q-1) + \sum_{i=1}^{kp} H(i) + \sum_{i=1}^{kq} H(i) - \sum_{i=1}^{k} H(i)$$ □

Theorem 4. Let p and q be primes. Let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p,q\})$ of Question 2. Then for $k,r,s \in N$, the following equality holds:

$$\sum_{i=1}^{kp^r q^s} H(i) = \frac{k^2 pq(p^{2r}-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=1}^{kp^r} H(i) + \sum_{i=1}^{kq^s} H(i) - \sum_{i=1}^{k} H(i). \qquad (6)$$

Similarly, this equality can be written as the following briefer forms:

$$\sum_{i=kp^r}^{kp^r q^s} H(i) = \frac{k^2 pq(p^{2r}-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=k}^{kq^s} H(i)$$

or

$$\sum_{i=kq^s}^{kp^r q^s} H(i) = \frac{k^2 pq(p^{2r}-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=k}^{kp^r} H(i). \qquad (6')$$

Proof. We prove (6) by induction on r and s. It follows from Theorem 3 that (6) holds if r=1, s=1.

Assume that the equality hold for all r≤u, s≤v, then if r=u+1, s≤v, by assumption we have that

$$\sum_{i=1}^{kp^{u+1} q^s} H(i) = \frac{(kp)^2 pq(p^{2u}-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=1}^{(kp)p^u} H(i) + \sum_{i=1}^{(kp)q^s} H(i) - \sum_{i=1}^{(kp)} H(i)$$

$$= \frac{(kp)^2 pq(p^{2u}-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=1}^{kp^{u+1}} H(i) + \sum_{i=1}^{kpq^s} H(i) - \sum_{i=1}^{kp} H(i)$$

$$= \frac{(kp)^2 pq(p^{2u}-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=1}^{kp^{u+1}} H(i)$$

$$+ \left(\frac{k^2 pq(p^2-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=1}^{kp} H(i) + \sum_{i=1}^{kq^s} H(i) - \sum_{i=1}^{k} H(i) \right) - \sum_{i=1}^{kp} H(i)$$

(unfold $\sum_{i=1}^{kpq^s} H(i)$ by the induction hypothesis)

$$= \frac{k^2 pq(q^{2s}-1)}{2(p+1)(q+1)} \times (p^2 \times p^{2u} - p^2 + p^2 - 1) + \sum_{i=1}^{kp^{u+1}} H(i) + \sum_{i=1}^{kq^s} H(i) - \sum_{i=1}^{k} H(i)$$

$$= \frac{k^2 pq(q^{2s}-1)}{2(p+1)(q+1)} \times (p^{2(u+1)} - 1) + \sum_{i=1}^{kp^{u+1}} H(i) + \sum_{i=1}^{kq^s} H(i) - \sum_{i=1}^{k} H(i),$$

then (6) holds for r=u+1,s≤v.

Similarly one can prove that (6) holds if r≤u+1, s=v+1.

Hence the conclusion holds. □

In the case k＝1，the following corollary holds：

Corollary2. Let p and q be primes. Let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER $(\{p,q\})$ of Question 2. Then for $r,s \in N$,the following equalities hold:

(1) $\sum_{i=p}^{pq} H(i) = \frac{pq(p-1)(q-1)}{2} + \sum_{i=1}^{q} H(i)$

（ or $\sum_{i=q}^{pq} H(i) = \frac{pq(p-1)(q-1)}{2} + \sum_{i=1}^{p} H(i)$ ）

(2) $\sum_{i=p^r}^{p^r q^s} H(i) = \frac{pq(p^{2r}-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=1}^{q^s} H(i)$

（ or $\sum_{i=q^s}^{p^r q^s} H(i) = \frac{pq(p^{2r}-1)(q^{2s}-1)}{2(p+1)(q+1)} + \sum_{i=1}^{p^r} H(i)$ ）.

Now we work on the times that number 1 occurs in $\{H(n)\}$. Similar to zero point, we call the term in $\{A_n\}$ which becomes to the number 1 in the new sequence $\{H(n)\}$ a valued "1" point of $\{H(n)\}$. We define $Y(n) = \{i \mid H(i) = 1, i \in N, i \le n\}$, and the cardinality of Y(n) is denoted by $|Y(n)|$.

Our idea is to transform calculating number of valued "1" points to the partial sum of a sequence. Set a number-theoretic function $H'(n): N \to R$ as follows:

$H'(n) = H(n) = 1$, when $H(n) = 1$; $H'(n) = 0$, when $H(n) \ne 1$. （7）

It is clear that $H'(n)$ is a completely multiplicative function as well. And $H'(p) = H'(q) = 1$. If n has prime factors other than p, q, then $H'(n) = 0$.

Clearly, $|Y(n)| = \sum_{i=1}^{n} H'(i)$, that is one can transform the number of valued "1" point of $\{H(n)\}$ to the partial sum of $H'(n)$.

Theorem 5. Let p and q be primes. Let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p,q\})$ of Question 2. Then for $k \in N$, the following equality holds:

$$|Y(kpq)| = |Y(kp)| + |Y(kq)| - |Y(k)| + 1 \tag{8}$$

Proof. Let $n \in Y(kpq)$ satisfy $H(n)=1$. Then n can be divided by p or q if n does not equal to 1.

Case 1: if $p \mid n$, since $H(p)=1$, it follows that $H(\frac{n}{p}) = H(\frac{n}{p})H(p) = H(n) = 1$.

Hence $\frac{n}{p} \in Y(kq), n \in pY(kq)$;

Case 2: If $q \mid n$, since $H(q)=1$, it follows that $H(\frac{n}{q}) = 1$, $\frac{n}{q} \in Y(kp), n \in qY(kp)$.

So, $Y(kpq) \subset pY(kq) \bigcup qY(kp) \bigcup \{1\}$, and $Y(kpq) \supset pY(kq) \bigcup qY(kp) \bigcup \{1\}$ is clear.

Hence $Y(kpq) = pY(kq) \bigcup qY(kp) \bigcup \{1\}$.

The above $pY(kq)$ and $qY(kp)$ stand for the sets $\{pn \mid n \in Y(kq)\}$ and $\{qn \mid n \in Y(kq)\}$ respectively. It is clear that

$pY(kp) \bigcap qY(kp) = \{pqn \mid n \in Y(k)\} = pqY(k)$,

$|pY(kq)| = |Y(kq)|$, $|qY(kp)| = |Y(kp)|$, $|pqY(k)| = |Y(k)|$.

By Principle of cross-classification, we get that

$$|Y(kpq)| = |pY(kq)| + |qY(kp)| - |pY(kq) \bigcap qY(kp)| + 1$$

$$= |Y(kq)| + |Y(kp)| - |Y(k)| + 1.$$ □

Theorem 6. Let p and q be primes. Let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p,q\})$ of Question 2. Then for $k \in N$, the following equality holds:

$$|Y(kp^r q^s)| = |Y(kp^r)| + |Y(kq^s)| - |Y(k)| + rs \qquad (9)$$

Proof. Let number-theoretic function $H'(n)$ be defined by（7）, and we have $|Y(n)| = \sum_{i=1}^{n} H'(i)$,（9）can be written as follows:

$$\sum_{i=1}^{kp^r q^s} H'(i) = \sum_{i=1}^{kp^r} H'(i) + \sum_{i=1}^{kq^s} H'(i) - \sum_{i=1}^{k} H'(i) + rs. \qquad (9')$$

We prove $(9')$ by induction in the following part. It follows from Theorem 5 that $(9')$ holds if r＝1，s＝1.

Assume that $(9')$ holds for all r≤u, s≤v . If r＝u+1, s≤v, we have

$$\sum_{i=1}^{kp^{u+1}q^s} H'(i) = \sum_{i=1}^{(kp)p^u} H'(i) + \sum_{i=1}^{(kp)q^s} H'(i) - \sum_{i=1}^{kp} H'(i) + us$$

$$= \sum_{i=1}^{kp^{u+1}} H'(i) + \left(\sum_{i=1}^{kp} H'(i) + \sum_{i=1}^{kq^s} H'(i) - \sum_{i=1}^{k} H'(i) + 1 \times s \right) - \sum_{i=1}^{kp} H'(i) + us$$

（unfold $\sum_{i=1}^{(kp)q^s} H'(i)$ by induction hypothesis）

$$= \sum_{i=1}^{kp^{u+1}} H'(i) + \sum_{i=1}^{kq^s} H'(i) - \sum_{i=1}^{k} H'(i) + (u+1) \times s.$$

So, if r＝u+1, s≤v , $(9')$ holds. Similarly one can prove that if r≤u+1, s＝v+1, $(9')$ holds as well. Hence the result follows from induction principle.

For the case k＝1，the following corollary holds：

Corollary3. Let p and q be primes. Let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER $(\{p,q\})$ of Question 2. Then for $r, s \in N$, the following equality holds:

$$|Y(p^r q^s)| = |Y(p^r)| + |Y(q^s)| + rs - 1$$

For the case p, q equal to 2, 3 respectively, the following interesting corollary holds：

Corollary4. Let p=2, q=3, $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p,q\})$ of question two. Then for $k,r,s\in N$, the following equalities hold:

(1) $$\sum_{i=1}^{k2^r3^s} H(i) = \frac{k^2(2^{2r}-1)(3^{2s}-1)}{4} + \sum_{i=1}^{k2^r} H(i) + \sum_{i=1}^{k3^s} H(i) - \sum_{i=1}^{k} H(i)$$

（or $$\sum_{i=k2^r}^{k2^r3^s} H(i) = \frac{k^2(2^{2r}-1)(3^{2s}-1)}{4} + \sum_{i=k}^{k3^s} H(i)$$ ）

（or $$\sum_{i=k3^s}^{k2^r3^s} H(i) = \frac{k^2(2^{2r}-1)(3^{2s}-1)}{4} + \sum_{i=k}^{k2^r} H(i)$$ ）

(2) $$\sum_{i=2^r}^{2^r3^s} H(i) = \frac{(2^{2r}-1)(3^{2s}-1)}{4} + \sum_{i=1}^{3^s} H(i)$$

（or $$\sum_{i=3^s}^{2^r3^s} H(i) = \frac{(2^{2r}-1)(3^{2s}-1)}{4} + \sum_{i=1}^{2^r} H(i)$$ ）

(3) $$\sum_{i=3^s+1}^{2\times3^s} H(i) = \frac{3^{2s+1}+1}{4}, \qquad \sum_{i=2^r+1}^{3\times2^r} H(i) = 2^{2r+1}$$

(4) $$\sum_{i=2k}^{6k} H(i) - \sum_{i=k}^{3k} H(i) = 6k^2, \qquad \sum_{i=2k}^{18k} H(i) - \sum_{i=k}^{9k} H(i) = 60k^2$$

(5) $$|Y(k2^r3^s)| = |Y(k2^r)| + |Y(k3^s)| - |Y(k)| + rs$$

$$|Y(2^r3^s)| = |Y(2^r)| + |Y(3^s)| + rs - 1$$

Section 3

In this section, we extend the problem to the situation including m arbitrary primes as follows.

Question 3: Given m primes p_i, $i=1,2,\ldots,m$, we obtain a new sequence $\{H(n)\}$ by transforming $\{A_n\}$ according to the rule ER$(\{p_1, p_2, ..., p_m\})$ as follows:

(1)For each $i \in \{1, 2, ..., m\}$, if A_n has factor p_i, then we substitute A_n with A_n / p_i until it does not have factor p_i

(2) If A_n does not have factor $p_i, i = 1, 2, ..., m$, then keep it unchanged.

What properties does $\{H(n)\}$ have?

The rule ER$(\{p_1, p_2, ..., p_m\})$ above can also be written as follows:

If $A_n = X \times \prod_{i=1}^{m} p_i^{r_i}$, where $(X, \prod_{i=1}^{m} p_i) = 1, r_i \geq 0, i = 1, 2, ..., m, X \geq 1$ are integers, then $H(n) = X$.

For instance, let the given primes be 2, 3,5, then the generated sequence $\{H(n)\}$ is:

1,1,1,1,1,1,7,1,1,1,11,1,13,7,1,1,17,1,19,1,7,11,23,1,1,13,1,7,29,1,31,1,11,17,7,1,37,19, 13,1,41,7,43,11,1,23,47,1,49,…… (10)

The new sequence is more complicated. As in the previous section, we know $H(n)$ is a totally multiplicative function.

Let $\alpha = \prod_{i=1}^{m} p_i$, $\alpha_l = \alpha / p_l$, $\alpha_{l,j} = \alpha / p_l p_j, \alpha_{l,j,t} = \alpha / p_l p_j p_t$,…; $\beta = \prod_{i=1}^{m} p_i^{r_i}$,

$\beta_l = \beta / p_l^{r_l}, \beta_{l,j} = \beta / p_l^{r_l} p_j^{r_j}, \beta_{l,j,t} = \beta / p_l^{r_l} p_j^{r_j} p_t^{r_t}, \ldots 1 \leq l \leq m, 1 \leq l < j \leq m, \ldots$ (11)

For the sum problem of $\{H(n)\}$, we propose the following Theorem 7 and Theorem 8.

Theorem 7. Given m primes $p_1, p_2, ..., p_m$, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p_1, p_2, ..., p_m\})$ of Question 3. Let $\alpha, \alpha_l, \alpha_{l,j}, \cdots$ be described as (11), then for $k \in N$, the following equality holds:

$$\sum_{i=1}^{k\alpha} H(i) = \frac{k^2}{2} \prod_{i=1}^{m} p_i (p_i - 1) + \sum_{l=1}^{m} \sum_{i=1}^{k\alpha_l} H(i) - \sum_{1 \leq l < j \leq m} \sum_{i=1}^{k\alpha_{l,j}} H(i) + \cdots$$

$$+(-1)^{m-2}\sum_{l=1}^{m}\sum_{i=1}^{kp_l} H(i)+(-1)^{m-1}\sum_{i=1}^{k} H(i)\,. \qquad (12)$$

Proof. Similar to the proof of Theorem 3, we divide $\{H(i), i=1,2,...,k\alpha\}$ into two groups and sum up for each group.

In the first group, i is not any multiples of $p_i, i=1,2,......,m$. Then $H(i)=i$ .By Principle of cross-classification，the sum of this group is:

$$S_1=\sum_{i=1}^{k\alpha} i-\sum_{l=1}^{m} p_l\sum_{i=1}^{k\alpha_l} i+\sum_{1\le l<j\le m} p_l p_j\sum_{i=1}^{k\alpha_{l,j}} i-......+(-1)^{m} p_1 p_2 ... p_m\sum_{i=1}^{k} i \qquad (13)$$

Now we want to show the following expression by induction

$$S_1=\frac{k^2}{2}\prod_{i=1}^{m} p_i(p_i-1) \qquad (14)$$

For $m\le 2$, (14) can be easily proved by Theorem 2 and Theorem 3. Now assume that (14) holds for each $m\le w$. Then for $m=w+1$ we conclude that

$$S_1=\sum_{i=1}^{k\alpha'} i-\sum_{l=1}^{w+1} p_l\sum_{i=1}^{k\alpha'_l} i+\sum_{1\le l<j\le w+1} p_l p_j\sum_{i=1}^{k\alpha'_{l,j}} i-......+(-1)^{w+1} p_1 p_2 ... p_{w+1}\sum_{i=1}^{k} i$$

$$=\sum_{i=1}^{(kp_{w+1})\alpha} i-\left(\sum_{l=1}^{w} p_l\sum_{i=1}^{k\alpha'_l} i+p_{w+1}\sum_{i=1}^{k\alpha} i\right) \qquad (k\alpha=k\alpha'_{w+1})$$

$$+\left(\sum_{1\le l<j\le w} p_l p_j\sum_{i=1}^{k\alpha'_{l,j}} i+\sum_{1\le l<j=w+1} p_l p_{w+1}\sum_{i=1}^{k\alpha_l} i\right)-......+(-1)^{w+1} p_1 p_2 ... p_{w+1}\sum_{i=1}^{k} i$$

$$(k\alpha_l=k\alpha'_{l,w+1})$$

$$=\sum_{i=1}^{(kp_{w+1})\alpha} i+\left(-\sum_{l=1}^{w} p_l\sum_{i=1}^{(kp_{w+1})\alpha_l} i+\sum_{1\le l<j\le w} p_l p_j\sum_{i=1}^{(kp_{w+1})\alpha_{l,j}} i-......+(-1)^{w} p_1 p_2 ... p_w\sum_{i=1}^{(kp_{w+1})} i\right)$$

$$(kp_{w+1}\alpha_l=k\alpha'_l, kp_{w+1}\alpha_{l,j}=k\alpha'_{l,j})$$

$$-\left(p_{w+1}\sum_{i=1}^{k\alpha} i-p_{w+1}\sum_{l=1}^{w} p_l\sum_{i=1}^{k\alpha_l} i+......+(-1)^{w} p_{w+1}\times p_1 p_2 ... p_w\sum_{i=1}^{k} i\right)$$

$$=\left(\sum_{i=1}^{(kp_{w+1})\alpha} i-\sum_{l=1}^{w} p_l\sum_{i=1}^{(kp_{w+1})\alpha_l} i+\sum_{1\le l<j\le w} p_l p_j\sum_{i=1}^{(kp_{w+1})\alpha_{l,j}} i-......+(-1)^{w} p_1 p_2 ... p_w\sum_{i=1}^{(kp_{w+1})} i\right)$$

$$-p_{w+1}\left(\sum_{i=1}^{k\alpha} i-\sum_{l=1}^{w} p_l\sum_{i=1}^{k\alpha_l} i+\ldots+(-1)^w p_1 p_2 \ldots p_w\sum_{i=1}^{k} i\right).$$

(Applying the hypothesis to two parts of the above expression respectively)

$$=\frac{(kp_{w+1})^2}{2}\prod_{i=1}^{w} p_i(p_i-1)-p_{w+1}\times\frac{k^2}{2}\prod_{i=1}^{w} p_i(p_i-1)$$

$$=\frac{k^2}{2}\prod_{i=1}^{w+1} p_i(p_i-1)$$

So, the situation of $m=w+1$ is proved and (14) holds.

In the second group, i is at least one of the multiples of $p_i, i=1,2,......,m$. Also due to Principle of cross-classification, the sum of this group is:

$$S_2=\sum_{l=1}^{m}\sum_{i=1}^{k\alpha_l} H(p_l\times i)-\sum_{1\le l<j\le m}\sum_{i=1}^{k\alpha_{l,j}} H(p_l p_j\times i)+......+(-1)^{m-1}\sum_{i=1}^{k} H(\alpha\times i)$$

$$=\sum_{l=1}^{m}\sum_{i=1}^{k\alpha_l} H(i)-\sum_{1\le l<j\le m}\sum_{i=1}^{k\alpha_{l,j}} H(i)+\cdots\cdots+(-1)^{m-2}\sum_{l=1}^{m}\sum_{i=1}^{kp_l} H(i)+(-1)^{m-1}\sum_{i=1}^{k} H(i)$$

Hence,

$$\sum_{i=1}^{k\alpha} H(i)=S_1+S_2=\frac{k^2}{2}\prod_{i=1}^{m} p_i(p_i-1)$$

$$+\sum_{l=1}^{m}\sum_{i=1}^{k\alpha_l} H(i)-\sum_{1\le l<j\le m}\sum_{i=1}^{k\alpha_{l,j}} H(i)+\cdots\cdots+(-1)^{m-2}\sum_{l=1}^{m}\sum_{i=1}^{kp_l} H(i)+(-1)^{m-1}\sum_{i=1}^{k} H(i). \qquad \square$$

Theorem 8. Given m primes $p_1, p_2, ..., p_m$, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p_1, p_2, ..., p_m\})$ of Question 3. Let $\beta, \beta_l, \beta_{l,j}, \cdots$ be described as (11), then for $k\in N$, the following equality holds:

$$\sum_{i=1}^{k\beta} H(i)=\frac{k^2}{2}\prod_{i=1}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}+\sum_{l=1}^{m}\sum_{i=1}^{k\beta_l} H(i)-\sum_{1\le l<j\le m}\sum_{i=1}^{k\beta_{l,j}} H(i)+\cdots$$

$$+(-1)^{m-2}\sum_{l=1}^{m}\sum_{i=1}^{kp_l^{r_l}} H(i)+(-1)^{m-1}\sum_{i=1}^{k} H(i) \qquad (15)$$

Proof. Similar to the proof of Theorem 4. We prove (15) by induction on the power of p_i. It follows from Theorem 7 that (15) holds with the power of p_i being $r_i = 1, i = 1, 2, ..., m$.

Assume that the power of p_i is no more than $r_i, i = 1, 2, ..., m$, the equality holds. Without loss of generality, we first consider the power of p_1 changing from r_1 to $r_1 + 1$. Let $\beta = \prod_{i=2}^{m} p_i^{r_i} \times p_1^{r_1}$, $\beta' = \prod_{i=2}^{m} p_i^{r_i} \times p_1^{r_1+1}$, $\beta_l, \beta_{l,j}, ..., \beta'_l, \beta'_{l,j} ......$ be described as (11), then we have

$$\beta' = p_1\beta, \beta_1 = \beta'_1, \beta_{1,j} = \beta'_{1,j}, \beta_{1,j,t} = \beta'_{1,j,t}, ...,$$

$$p_1\beta_l = \beta'_l, p_1\beta_{l,j} = \beta'_{l,j}, p_1\beta_{l,j,t} = \beta'_{l,j,t}, ..., 1 < l < j < t < ... \le m.$$

From the induction hypothesis we have

$$\sum_{i=1}^{k\beta'} H(i) = \sum_{i=1}^{(kp_1)\beta} H(i) = \frac{(kp_1)^2}{2} \prod_{i=1}^{m} \frac{p_i(p_i^{2r_i} - 1)}{p_i + 1} + \sum_{l=1}^{m} \sum_{i=1}^{(kp_1)\beta_l} H(i) - \sum_{1 \le l < j \le m} \sum_{i=1}^{(kp_1)\beta_{l,j}} H(i) + \cdots\cdots$$
$$+ (-1)^{m-2} \sum_{l=1}^{m} \sum_{i=1}^{(kp_1)p_l^{r_l}} H(i) + (-1)^{m-1} \sum_{i=1}^{(kp_1)} H(i). \qquad (16)$$

Now we analyze right part of (16) one by one.

$$\sum_{l=1}^{m} \sum_{i=1}^{(kp_1)\beta_l} H(i) = \sum_{l=2}^{m} \sum_{i=1}^{(kp_1)\beta_l} H(i) + \sum_{i=1}^{(kp_1)\beta_1} H(i)$$

$$= \sum_{l=2}^{m} \sum_{i=1}^{k\beta'_l} H(i) + \frac{k^2}{2} \prod_{i=2}^{m} \frac{p_i(p_i^{2r_i} - 1)}{p_i + 1} \times \frac{p_1(p_1^2 - 1)}{p_1 + 1} + (\sum_{i=1}^{k\beta_1} H(i) + \sum_{l=2}^{m} \sum_{i=1}^{kp_1\beta_{1,l}} H(i))$$

$$(kp_1\beta_l = k\beta'_l)$$

$$-(\sum_{1=l<j\le m} \sum_{i=1}^{k\beta_{1,j}} H(i) + \sum_{1<l<j\le m} \sum_{i=1}^{kp_1\beta_{1,l,j}} H(i)) + ... + (-1)^{m-1} \sum_{i=1}^{k} H(i)$$

$$= \frac{k^2}{2} \prod_{i=2}^{m} \frac{p_i(p_i^{2r_i} - 1)}{p_i + 1} \times \frac{p_1(p_1^2 - 1)}{p_1 + 1} + (\sum_{l=2}^{m} \sum_{i=1}^{k\beta'_l} H(i) + \sum_{i=1}^{k\beta'_1} H(i)) \qquad (k\beta_1 = k\beta'_1)$$

$$+ (\sum_{l=2}^{m} \sum_{i=1}^{kp_1\beta_{1,l}} H(i) - \sum_{1<l<j\le m} \sum_{i=1}^{kp_1\beta_{1,l,j}} H(i) + \sum_{1<l<j<t\le m} \sum_{i=1}^{kp_1\beta_{1,l,j,t}} H(i) - ...... + (-1)^{m-2} \sum_{i=1}^{kp_1} H(i))$$

$$+ (- \sum_{1=l<j\le m} \sum_{i=1}^{k\beta_{1,j}} H(i) + \sum_{1=l<j<t\le m} \sum_{i=1}^{k\beta_{1,j,t}} H(i) - ...... (-1)^{m-1} \sum_{i=1}^{k} H(i))$$

$$=\frac{k^2}{2}\prod_{i=2}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}\times\frac{p_1(p_1^2-1)}{p_1+1}+\sum_{l=1}^{m}\ \sum_{i=1}^{k\beta'_l}H(i)$$

$$+\Big(\sum_{l=2}^{m}\ \sum_{i=1}^{kp_1\beta_{1,l}}H(i)-\sum_{1<l<j\le m}\ \sum_{i=1}^{kp_1\beta_{1,l,j}}H(i)+\sum_{1<l<j<t\le m}\ \sum_{i=1}^{kp_1\beta_{1,l,j,t}}H(i)-\ldots\ldots+(-1)^{m-2}\sum_{i=1}^{kp_1}H(i)\Big)$$

$$+\Big(-\sum_{1<j\le m}\ \sum_{i=1}^{k\beta'_{1,j}}H(i)+\sum_{1<j<t\le m}\ \sum_{i=1}^{k\beta'_{1,j,t}}H(i)-\ldots\ldots\ (-1)^{m-1}\sum_{i=1}^{k}H(i)\Big) \tag{17}$$

$$\left(k\beta_{1,j}=k\beta'_{1,j}\ ,\quad k\beta_{1,j,t}=k\beta'_{1,j,t}\right)$$

$$\sum_{1\le l<j\le m}\ \sum_{i=1}^{(kp_1)\beta_{l,j}}H(i)=\sum_{1=l<j\le m}\ \sum_{i=1}^{kp_1\beta_{1,j}}H(i)+\sum_{1<l<j\le m}\ \sum_{i=1}^{kp_1\beta_{l,j}}H(i)$$

$$=\sum_{j=2}^{m}\ \sum_{i=1}^{kp_1\beta_{1,j}}H(i)+\sum_{1<l<j\le m}\ \sum_{i=1}^{k\beta'_{l,j}}H(i) \qquad \left(kp_1\beta_{l,j}=k\beta'_{l,j}\right) \tag{18}$$

$$\sum_{1\le l<j<t\le m}\ \sum_{i=1}^{(kp_1)\beta_{l,j,t}}H(i)=\sum_{1=l<j<t\le m}\ \sum_{i=1}^{kp_1\beta_{1,j,t}}H(i)+\sum_{1<l<j<t\le m}\ \sum_{i=1}^{kp_1\beta_{l,j,t}}H(i)$$

$$=\sum_{1<j<t\le m}\ \sum_{i=1}^{kp_1\beta_{1,j,t}}H(i)+\sum_{1<l<j<t\le m}\ \sum_{i=1}^{k\beta'_{l,j,t}}H(i) \qquad \left(kp_1\beta_{l,j,t}=k\beta'_{l,j,t}\right) \tag{19}$$

……

$$\frac{(kp_1)^2}{2}\prod_{i=1}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}+\frac{k^2}{2}\prod_{i=2}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}\times\frac{p_1(p_1^2-1)}{p_1+1}$$

$$=\frac{k^2}{2}\prod_{i=2}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}\times\frac{p_1}{p_1+1}(p_1^{2r_1+2}-p_1^2+p_1^2-1)$$

$$=\frac{k^2}{2}\prod_{i=2}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}\times\frac{p_1(p_1^{2(r_1+1)}-1)}{p_1+1} \tag{20}$$

Connecting (16) with (20), we obtain

$$\sum_{i=1}^{k\beta'}H(i)=\frac{k^2}{2}\prod_{i=2}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}\times\frac{p_1(p_1^{2(r_1+1)}-1)}{p_1+1}+\Bigg(\sum_{l=1}^{m}\ \sum_{i=1}^{k\beta'_l}H(i)$$

$$+\Big(\sum_{l=2}^{m}\sum_{i=1}^{kp_1\beta_{1,l}}H(i)-\sum_{1<l<j\le m}\ \sum_{i=1}^{kp_1\beta_{1,l,j}}H(i)+\sum_{1<l<j<t\le m}\ \sum_{i=1}^{kp_1\beta_{1,l,j,t}}H(i)-\ldots\ldots+(-1)^{m-2}\sum_{i=1}^{kp_1}H(i)\Big)$$

$$+\Big(-\sum_{1<j\le m}\ \sum_{i=1}^{k\beta'_{1,j}}H(i)+\sum_{1<j<t\le m}\ \sum_{i=1}^{k\beta'_{1,j,t}}H(i)-\ldots\ldots\ (-1)^{m-1}\sum_{i=1}^{k}H(i)\Big)\Bigg)$$

$$-\left(\sum_{j=2}^{m}\sum_{i=1}^{kp_1\beta_{1,j}}H(i)+\sum_{1<l<j\le m}\sum_{i=1}^{k\beta'_{l,j}}H(i)\right)$$

$$+\left(\sum_{1<j<t\le m}\sum_{i=1}^{kp_1\beta_{1,j,t}}H(i)+\sum_{1<l<j<t\le m}\sum_{i=1}^{k\beta'_{l,j,t}}H(i)\right)+\cdots\cdots$$

$$=\frac{k^2}{2}\prod_{i=2}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}\times\frac{p_1(p_1^{2(r_1+1)}-1)}{p_1+1}+\left(\sum_{l=1}^{m}\sum_{i=1}^{k\beta'_l}H(i)\right.$$

$$+\left(-\sum_{1<j\le m}\sum_{i=1}^{k\beta'_{1,j}}H(i)-\sum_{1<l<j\le m}\sum_{i=1}^{k\beta'_{l,j}}H(i)\right)$$

$$\left.+\left(\sum_{1<j<t\le m}\sum_{i=1}^{k\beta'_{1,j,t}}H(i)+\sum_{1<l<j<t\le m}\sum_{i=1}^{k\beta'_{l,j,t}}H(i)\right)+\ldots\ldots+(-1)^{m-1}\sum_{i=1}^{k}H(i))\right)$$

$$=\frac{k^2}{2}\prod_{i=2}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}\times\frac{p_1(p_1^{2(r_1+1)}-1)}{p_1+1}$$

$$+\sum_{l=1}^{m}\sum_{i=1}^{k\beta'_l}H(i)\text{ - }\sum_{1\le l<j\le m}\sum_{i=1}^{k\beta'_{l,j}}H(i)+\cdots\cdots+(-1)^{m-1}\sum_{i=1}^{k}H(i)$$

So, when the power of p_1 changing from r_1 to r_1+1, (15) holds. We can use the same method to prove the conclusion for $p_i, i=2,3,...,n$. Hence the conclusion holds. □

When $k=1$, we have the following corollary:

Corollary 5. Given m primes $p_1, p_2, ..., p_m$, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p_1, p_2, ..., p_m\})$ of Question 3. Let $\beta, \beta_l, \beta_{l,j}, \cdots$ be described as (11), then, the following equality holds:

$$\sum_{i=1}^{\beta}H(i)=\frac{1}{2}\prod_{i=1}^{m}\frac{p_i(p_i^{2r_i}-1)}{p_i+1}+\sum_{l=1}^{m}\sum_{i=1}^{\beta_l}H(i)\text{ - }\sum_{1\le l<j\le m}\sum_{i=1}^{\beta_{l,j}}H(i)+\cdots$$

$$+(-1)^{m-2}\sum_{l=1}^{m}\sum_{i=1}^{p_l^{r_l}}H(i)+(-1)^{m-1}.$$

Now we focus on the times that number 1 occurs in $\{H(n)\}$. Similarly, we call the term in $\{A_n\}$ which becomes to the number 1 in the new sequence $\{H(n)\}$ a valued "1" point of $\{H(n)\}$, and define $Y(n)=\{i\,|\,H(i)=1, i\in N, i\le n\}$.

Set a number-theoretic function $H'(n): N\to R$ as follows:

$H'(n)=H(n)=1$, when $H(n)=1$; $H'(n)=0$, when $H(n)\ne 1$.

It is clear that $H'(n)$ is a completely multiplicative function as well. It follows that $|Y(n)|=\sum_{i=1}^{n}H'(i)$, and we propose the Theorem 9 and Theorem 10.

Theorem 9. Given m primes $p_1, p_2, ..., p_m$, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER $(\{p_1, p_2, ..., p_m\})$ of Question 3. Let $\alpha, \alpha_l, \alpha_{l,j}, \alpha_{l,j,t}, \cdots$ be described as (11), Then for $k\in N$, the following equality holds:

$$|Y(k\alpha)|=\sum_{i=1}^{m}|Y(k\alpha_i)|-\sum_{1\le i<j\le m}|Y(k\alpha_{i,j})|+...+\ (-1)^{m-2}\sum_{i=1}^{m}|Y(kp_i)|+(-1)^{m-1}|Y(k)|+1 \qquad (21)$$

Proof. Similar to the proof of Theorem 5, let $n\in Y(k\alpha)$ satisfy $H(n)=1$, then either $n=1$ or n can be divided by p_i. If $p_i\,|\,n$, then $H(\frac{n}{p_i})=H(\frac{n}{p_i})H(p_i)=H(n)=1$. It follows that

$$\frac{n}{p_i}\in Y(k\alpha_i),\quad n\in p_iY(k\alpha_i),$$

So, $Y(k\alpha)\subset\bigcup_{i=1}^{m} p_iY(k\alpha_i)\cup\{1\}$, and $Y(k\alpha)\supset\bigcup_{i=1}^{m} p_iY(k\alpha_i)\cup\{1\}$ is obvious.

Hence $Y(k\alpha)=\bigcup_{i=1}^{m} p_iY(k\alpha_i)\cup\{1\}$.

Noting that $p_iY(k\alpha_i)\cap p_jY(k\alpha_j)=p_ip_jY(k\alpha_{i,j})$,

$p_iY(k\alpha_i)\cap p_jY(k\alpha_j)\cap p_tY(k\alpha_t)=p_ip_jp_tY(k\alpha_{i,j,t}),......,$,

$|p_iY(k\alpha_i)|=|Y(k\alpha_i)|$, $|p_ip_jY(k\alpha_{i,j})|=|Y(k\alpha_{i,j})|$, $......, 1\le i<j<t<...\le m$,

we can get the conclusion by Principle of cross-classification. □

Theorem 10. Given m primes $p_1, p_2, ..., p_m$, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p_1, p_2, ..., p_m\})$ of Question 3. Let $\beta, \beta_l, \beta_{l,j}, ...$ be described as (11), then for $k \in N$, the following equality holds:

$$|Y(k\beta)| = \sum_{i=1}^{m} |Y(k\beta_i)| - \sum_{1\le i<j\le m} |Y(k\beta_{i,j})| + ... + (-1)^{m-2} \sum_{i=1}^{m} |Y(kp_i^{r_i})| + (-1)^{m-1} |Y(k)| + \prod_{i=1}^{m} r_i$$

(22)

Proof. Similar to the proof of Theorem 6, we introduce the completely multiplicative function $H': N \to R, H'(p_i) = 1, i = 1, 2, ..., m; H'(p) = 0, p$ is any other primes.

We also have $|Y(n)| = \sum_{i=1}^{n} H'(i)$, and (22) can be written as follows:

$$\sum_{i=1}^{k\beta} H'(i) = \sum_{l=1}^{m} \sum_{i=1}^{k\beta_l} H'(i) - \sum_{1\le l<j\le m} \sum_{i=1}^{k\beta_{l,j}} H'(i) + ... + (-1)^{m-2} \sum_{l=1}^{m} \sum_{i=1}^{kp_l^{r_l}} H'(i) + (-1)^{m-1} \sum_{i=1}^{k} H'(i) + \prod_{i=1}^{m} r_i$$

(22′)

We prove $(22')$ by induction. It follows from Theorem 9 that $(22')$ holds if the power of p_i is $r_i = 1, i = 1, 2, ..., m$.

Assume that the power of p_i is no more than $r_i, i = 1, 2, ..., m$, $(22')$ holds. Without loss of generality, we first consider the power of p_1 changing from r_1 to $r_1 + 1$. Let $\beta = \prod_{i=2}^{m} p_i^{r_i} \times p_1^{r_1}$, $\beta' = \prod_{i=2}^{m} p_i^{r_i} \times p_1^{r_1+1}$, $\beta_l, \beta_{l,j}, ..., \beta'_l, \beta'_{l,j} ......$ be described as (11), then we have

$$\beta' = p_1\beta, \beta_1 = \beta'_1, \beta_{1,j} = \beta'_{1,j}, \beta_{1,j,t} = \beta'_{1,j,t}, ...,$$

$$p_1\beta_l = \beta'_l, p_1\beta_{l,j} = \beta'_{l,j}, p_1\beta_{l,j,t} = \beta'_{l,j,t}, ..., 1 < l < j < t < ... \le m.$$

From the hypothesis we have

$$\sum_{i=1}^{k\beta'} H'(i) = \sum_{i=1}^{(kp_1)\beta} H'(i) = \sum_{l=1}^{m} \sum_{i=1}^{(kp_1)\beta_l} H'(i) - \sum_{1\le l<j\le m} \sum_{i=1}^{(kp_1)\beta_{l,j}} H'(i) + \cdots\cdots$$

$$+(-1)^{m-2}\sum_{l=1}^{m}\sum_{i=1}^{(kp_1)p_l^{\eta}}H'(i)+(-1)^{m-1}\sum_{i=1}^{(kp_1)}H'(i)+\prod_{i=1}^{m}r_i \qquad (23)$$

Similar to Theorem 8, we analyze right part of (23) one by one.

$$\sum_{l=1}^{m}\sum_{i=1}^{(kp_1)\beta_l}H'(i)=\sum_{l=2}^{m}\sum_{i=1}^{(kp_1)\beta_l}H'(i)+\sum_{i=1}^{(kp_1)\beta_1}H'(i)$$

$$=\sum_{l=2}^{m}\sum_{i=1}^{k\beta_l'}H'(i)+(\sum_{i=1}^{k\beta_1}H'(i)+\sum_{l=2}^{m}\sum_{i=1}^{kp_1\beta_{1,l}}H'(i)) \qquad (kp_1\beta_l=k\beta_l')$$

$$-(\sum_{1=l<j\le m}\sum_{i=1}^{k\beta_{1,j}}H'(i)+\sum_{1<l<j\le m}\sum_{i=1}^{kp_1\beta_{1,l,j}}H'(i))+......+(-1)^{m-1}\sum_{i=1}^{k}H'(i)+1\times\prod_{i=2}^{m}r_i$$

$$=(\sum_{l=2}^{m}\sum_{i=1}^{k\beta_l'}H'(i)+\sum_{i=1}^{k\beta_1'}H'(i)) \qquad (k\beta_1=k\beta_1')$$

$$+(\sum_{j=2}^{m}\sum_{i=1}^{kp_1\beta_{1,j}}H'(i)-\sum_{1<l<j\le m}\sum_{i=1}^{kp_1\beta_{1,l,j}}H'(i)+\sum_{1<l<j<t\le m}\sum_{i=1}^{kp_1\beta_{1,l,j,t}}H'(i)-......+(-1)^{m-2}\sum_{i=1}^{kp_1}H'(i))$$

$$+(-\sum_{1<j\le m}\sum_{i=1}^{k\beta_{1,j}}H'(i)+\sum_{1<j<t\le m}\sum_{i=1}^{k\beta_{1,j,t}}H'(i)-......(-1)^{m-1}\sum_{i=1}^{k}H'(i))+1\times\prod_{i=2}^{m}r_i$$

$$=\sum_{l=1}^{m}\sum_{i=1}^{k\beta_l'}H'(i)$$

$$+(\sum_{j=2}^{m}\sum_{i=1}^{kp_1\beta_{1,j}}H'(i)-\sum_{1<l<j\le m}\sum_{i=1}^{kp_1\beta_{1,l,j}}H'(i)+\sum_{1<l<j<t\le m}\sum_{i=1}^{kp_1\beta_{1,l,j,t}}H'(i)-......+(-1)^{m-2}\sum_{i=1}^{kp_1}H'(i))$$

$$+(-\sum_{1<j\le m}\sum_{i=1}^{k\beta_{1,j}'}H'(i)+\sum_{1<j<t\le m}\sum_{i=1}^{k\beta_{1,j,t}'}H'(i)-......(-1)^{m-1}\sum_{i=1}^{k}H'(i))+1\times\prod_{i=2}^{m}r_i \qquad (24)$$

$$(k\beta_{1,j}=k\beta_{1,j}', \quad k\beta_{1,j,t}=k\beta_{1,j,t}')$$

$$\sum_{1\le l<j\le m}\sum_{i=1}^{(kp_1)\beta_{l,j}}H'(i)=\sum_{j=2}^{m}\sum_{i=1}^{kp_1\beta_{1,j}}H'(i)+\sum_{1<l<j\le m}\sum_{i=1}^{kp_1\beta_{l,j}}H'(i)$$

$$=\sum_{j=2}^{m}\sum_{i=1}^{kp_1\beta_{1,j}}H'(i)+\sum_{1<l<j\le m}\sum_{i=1}^{k\beta_{l,j}'}H'(i) \qquad (kp_1\beta_{l,j}=k\beta_{l,j}') \qquad (25)$$

$$\sum_{1\le l<j<t\le m}\sum_{i=1}^{(kp_1)\beta_{l,j,t}}H'(i)=\sum_{1=l<j<t\le m}\sum_{i=1}^{kp_1\beta_{1,j,t}}H'(i)+\sum_{1<l<j<t\le m}\sum_{i=1}^{kp_1\beta_{l,j,t}}H'(i)$$

$$= \sum_{1<j<t\le m}^{m} \sum_{i=1}^{kp_1\beta_{1,j,t}} H'(i) + \sum_{1<l<j<t\le m} \sum_{i=1}^{k\beta'_{l,j,t}} H'(i) \qquad (kp_1\beta_{l,j,t} = k\beta'_{l,j,t}) \qquad (26)$$

……

Connecting (23) with (26), we obtain that

$$\sum_{i=1}^{k\beta'} H'(i) = \Bigg(\sum_{l=1}^{m} \sum_{i=1}^{k\beta'_l} H'(i)$$

$$+ \Big(\sum_{j=2}^{m} \sum_{i=1}^{kp_1\beta_{1,j}} H'(i) - \sum_{1<l<j\le m} \sum_{i=1}^{kp_1\beta_{1,l,j}} H'(i) + \sum_{1<l<j<t\le m} \sum_{i=1}^{kp_1\beta_{1,l,j,t}} H'(i) - \ldots\ldots + (-1)^{m-2} \sum_{i=1}^{kp_1} H'(i) \Big)$$

$$+ \Big(- \sum_{1<j\le m} \sum_{i=1}^{k\beta'_{1,j}} H'(i) + \sum_{1<j<t\le m} \sum_{i=1}^{k\beta'_{1,j,t}} H'(i) - \ldots\ldots + (-1)^{m-1} \sum_{i=1}^{k} H'(i) \Big) + 1 \times \prod_{i=2}^{m} r_i \Bigg)$$

$$- \left(\sum_{j=2}^{m} \sum_{i=1}^{kp_1\beta_{1,j}} H'(i) + \sum_{1<l<j\le m} \sum_{i=1}^{k\beta'_{l,j}} H'(i) \right) + \left(\sum_{1<j<t\le m}^{m} \sum_{i=1}^{kp_1\beta_{1,j,t}} H'(i) + \sum_{1<l<j<t\le m} \sum_{i=1}^{k\beta'_{l,j,t}} H'(i) \right) - \cdots\cdots$$

$$+ (-1)^{m-1} \sum_{i=1}^{(kp_1)} H'(i) + \prod_{i=1}^{m} r_i$$

$$= \sum_{l=1}^{m} \sum_{i=1}^{k\beta'_l} H'(i)$$

$$+ \left(- \sum_{1<j\le m} \sum_{i=1}^{k\beta'_{1,j}} H'(i) - \sum_{1<l<j\le m} \sum_{i=1}^{k\beta'_{l,j}} H'(i) \right) + \left(\sum_{1<j<t\le m} \sum_{i=1}^{k\beta'_{1,j,t}} H'(i) + \sum_{1<l<j<t\le m} \sum_{i=1}^{k\beta'_{l,j,t}} H'(i) \right)$$

$$+ \cdots\cdots + (-1)^{m-1} \sum_{i=1}^{k} H'(i) + 1 \times \prod_{i=2}^{m} r_i + \prod_{i=1}^{m} r_i$$

$$= \sum_{j=1}^{m} \sum_{i=1}^{k\beta'_j} H'(i) - \sum_{1\le l<j\le m} \sum_{i=1}^{k\beta'_{l,j}} H'(i) + \sum_{1\le l<j<t\le m} \sum_{i=1}^{k\beta'_{l,j,t}} H'(i) - \cdots\cdots$$

$$+ (-1)^{m-1} \sum_{i=1}^{k} H'(i) + (r_1 + 1) \times \prod_{i=2}^{m} r_i$$

So, when the power of p_1 changing from r_1 to r_1+1, $(22')$ holds. We can use the same method to prove that for the situation of the power of $p_i, i = 2,3,...,n$, $(22')$ holds as well. Hence the theorem is proved by induction. □

When $k = 1$, we have the following corollary:

Corollary 6. Given m primes $p_1, p_2, ..., p_m$, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p_1, p_2, ..., p_m\})$ of Question 3. Let $\alpha, \alpha_l, \alpha_{l,j}, ...$ $\beta, \beta_l, \beta_{l,j}, ...$ be described as (11), then, the following equality holds:

$$(1) |Y(\alpha)| = \sum_{i=1}^{m} |Y(\alpha_i)| - \sum_{1 \le i < j \le m} |Y(\alpha_{i,j})| + ... + (-1)^{m-2} \sum_{i=1}^{m} |Y(p_i)| + (-1)^{m-1} + 1$$

$$(2) |Y(\beta)| = \sum_{i=1}^{m} |Y(\beta_i)| - \sum_{1 \le i < j \le m} |Y(\beta_{i,j})| + ... + (-1)^{m-2} \sum_{i=1}^{m} |Y(p_i^{r_i})| + (-1)^{m-1} + \prod_{i=1}^{m} r_i$$

Section 4

In previous sections, we studied the properties of the sequences generated by transforming $\{A_n\}$ according to the rules of Question1,2,3, and established equalities about the partial sum and the number of valued "1" point, such as the sum from first term to $kp^r q^s$ th term, and the number of valued "1" point in the first $kp^r q^s$ terms, and so on. But it is usually difficult to establish the equality about the sum of the first n terms and the number of valued "1" point in the first n terms. So we propose the following question.

Question 4. Given some primes, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rules of Question 2 and 3, $|Y(n)|$ be denoted as the number of valued "1" point in the first n terms of $\{H(n)\}$. Can we establish some inequalities to estimate the sum of the first n terms and the number of valued "1" point in the first n terms?

Firstly, we studied the case that the number of given primes is two, and we proposed the following theorem.

Theorem 11. Let p and q be primes. Let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p, q\})$ of Question 2. Then for any $x > 0$, the following inequality holds:

$$\frac{pq}{2(p+1)(q+1)} x^2 - 2x < \sum_{i \le x} H(i) < \frac{pq}{2(p+1)(q+1)} x^2 + 2x. \qquad (27)$$

Proof. Let $G:[0,+\infty)\to R$ be a function defined by

$$G(x)=(1+2+...+[x])-p(1+2+...+\left[\frac{x}{p}\right])-q(1+2+...+\left[\frac{x}{q}\right])+pq(1+2+...+\left[\frac{x}{pq}\right]),$$

and we set $p(1+2+......\left[\frac{x}{p}\right])=0$ for $\left[\frac{x}{p}\right]=0$ (similarly, for $\left[\frac{x}{q}\right]=0$ and $\left[\frac{x}{pq}\right]=0$).

According to Principle of cross-classification, for a fixed $x>0$, $G(x)$ denotes the sum of the first $[x]$ terms of the sequence of the naturals, with all multiples of p and multiples of q removed.

We establish a class of sets to classify and analyze every term of $\sum_{i\le x}H(i)$. Let

$t=\left[\frac{\log x}{\log p}\right], s=\left[\frac{\log x}{\log q}\right]$, then we have the following results.

Denote $A(1)=\{i\le x\,|\,(i,p^tq^s)=1\}$, then $\sum_{i\in A(1)}H(i)=G(x)$;

Denote $A(p)=\{i\le x\,|\,(i,p^tq^s)=p\}$, then $\sum_{i\in A(p)}H(i)=G(\frac{x}{p})$;

Denote $A(q)=\{i\le x\,|\,(i,p^tq^s)=q\}$, then $\sum_{i\in A(q)}H(i)=G(\frac{x}{q})$;

Denote $A(pq)=\{i\le x\,|\,(i,p^tq^s)=pq\}$, then $\sum_{i\in A(pq)}H(i)=G(\frac{x}{pq})$;

Denote $A(p^2)=\{i\le x\,|\,(i,p^tq^s)=p^2\}$, then $\sum_{i\in A(p^2)}H(i)=G(\frac{x}{p^2})$;

...

Noting that the number of the sets $A(p),A(q),...$ is finite, the intersection of any two different sets is empty, and the union of them is the set of the first $[x]$ terms of $\{A_n\}$.Thus,

$$\sum_{i\le x}H(i)=\sum_{i\in A(1)}H(i)+\sum_{i\in A(p)}H(i)+\sum_{i\in A(q)}H(i)+\sum_{i\in A(pq)}H(i)+\sum_{i\in A(p^2)}H(i)+......$$

$$=G(x)+G(\frac{x}{p})+G(\frac{x}{q})+G(\frac{x}{pq})+G(\frac{x}{p^2})+......$$

$$=(1+2+......+[x])-p(1+2+......+\left[\frac{x}{p}\right])-q(1+......+\left[\frac{x}{q}\right])+pq(1+......+\left[\frac{x}{pq}\right])$$

$$+(1+2+......+\left[\frac{x}{p}\right])-p(1+2+......+\left[\frac{x}{p^2}\right])-q(1+......+\left[\frac{x}{pq}\right])+pq(1+......+\left[\frac{x}{p^2q}\right])$$

$$+(1+2+......+\left[\frac{x}{q}\right])-p(1+2+......+\left[\frac{x}{pq}\right])-q(1+......+\left[\frac{x}{q^2}\right])+pq(1+......+\left[\frac{x}{pq^2}\right])$$

$$+(1+2+......+\left[\frac{x}{pq}\right])-p(1+2+......+\left[\frac{x}{p^2q}\right])-q(1+......+\left[\frac{x}{pq^2}\right])+pq(1+......+\left[\frac{x}{p^2q^2}\right])$$

$$+(1+2+......+\left[\frac{x}{p^2}\right])-p(1+2+......+\left[\frac{x}{p^3}\right])-q(1+......+\left[\frac{x}{p^2q}\right])+pq(1+......+\left[\frac{x}{p^3q}\right])$$

$$+......$$

Analyzing the above equality, we note that every term of the form $(1+2+...+\left[\frac{x}{p^u}\right])$ only appears in $G(\frac{x}{p^{u-1}})$ and $G(\frac{x}{p^u})$ which are $-p(1+2+...+\left[\frac{x}{p^u}\right])$ and $(1+2+...+\left[\frac{x}{p^u}\right])$.

Every term of the form $(1+2+...+\left[\frac{x}{q^v}\right])$ appears only in $G(\frac{x}{q^{v-1}})$ and $G(\frac{x}{q^v})$ which are $-q(1+2+...+\left[\frac{x}{q^v}\right])$ and $(1+2+...+\left[\frac{x}{q^v}\right])$.

Every term of the form $(1+2+...+\left[\frac{x}{p^uq^v}\right])$ appears only in $G(\frac{x}{p^{u-1}q^{v-1}})$, $G(\frac{x}{p^{u-1}q^v})$, $G(\frac{x}{p^uq^{v-1}})$ and $G(\frac{x}{p^uq^v})$ which are $pq(1+2+...+\left[\frac{x}{p^uq^v}\right])$, $-p(1+2+...+\left[\frac{x}{p^uq^v}\right])$, $-q(1+2+...+\left[\frac{x}{p^uq^v}\right])$ and $(1+2+...+\left[\frac{x}{p^uq^v}\right])$, where u and v are positive integers.

Then we have that

$$\sum_{i\le x}H(i)=\frac{[x]([x]+1)}{2}-\frac{p-1}{2}\{\left[\frac{x}{p}\right](\left[\frac{x}{p}\right]+1)+\left[\frac{x}{p^2}\right](\left[\frac{x}{p^2}\right]+1)+\left[\frac{x}{p^3}\right](\left[\frac{x}{p^3}\right]+1)+......\}$$

$$-\frac{q-1}{2}\{\left[\frac{x}{q}\right](\left[\frac{x}{q}\right]+1)+\left[\frac{x}{q^2}\right](\left[\frac{x}{q^2}\right]+1)+\left[\frac{x}{q^3}\right](\left[\frac{x}{q^3}\right]+1)+......\}$$

$$+\frac{(p-1)(q-1)}{2}\{\left[\frac{x}{pq}\right](\left[\frac{x}{pq}\right]+1)+\left[\frac{x}{p^2q}\right](\left[\frac{x}{p^2q}\right]+1)+\left[\frac{x}{pq^2}\right](\left[\frac{x}{pq^2}\right]+1)+......\}$$

(28)

$$<\frac{x(x+1)}{2}-\frac{p-1}{2}\{(\frac{x}{p}-1)\frac{x}{p}+(\frac{x}{p^2}-1)\frac{x}{p^2}+(\frac{x}{p^3}-1)\frac{x}{p^3}+......\}$$

$$-\frac{q-1}{2}\{(\frac{x}{q}-1)\frac{x}{q}+(\frac{x}{q^2}-1)\frac{x}{q^2}+(\frac{x}{q^3}-1)\frac{x}{q^3}+......\}$$

$$+\frac{(p-1)(q-1)}{2}\{\frac{x}{pq}(\frac{x}{pq}+1)+\frac{x}{p^2q}(\frac{x}{p^2q}+1)+\frac{x}{pq^2}(\frac{x}{pq^2}+1)+......\}$$

$$=\frac{x(x+1)}{2}-\frac{p-1}{2}\{-(\frac{x}{p}+\frac{x}{p^2}+\frac{x}{p^3}+......)+(\frac{x^2}{p^2}+\frac{x^2}{p^4}+\frac{x^2}{p^6}+......)\}$$

$$-\frac{q-1}{2}\{-(\frac{x}{q}+\frac{x}{q^2}+\frac{x}{q^3}+......)+(\frac{x^2}{q^2}+\frac{x^2}{q^4}+\frac{x^2}{q^6}+......)\}$$

$$+\frac{(p-1)(q-1)}{2}\{x(\frac{1}{p}+\frac{1}{p^2}+......)(\frac{1}{q}+\frac{1}{q^2}+......)+x^2(\frac{1}{p^2}+\frac{1}{p^4}+......)(\frac{1}{q^2}+\frac{1}{q^4}+......)\}$$

$$=\frac{x(x+1)}{2}-\frac{p-1}{2}(\frac{-x}{p-1}+\frac{x^2}{p^2-1})-\frac{q-1}{2}(\frac{-x}{q-1}+\frac{x^2}{q^2-1})$$

$$+\frac{(p-1)(q-1)}{2}(x\times\frac{1}{p-1}\times\frac{1}{q-1}+x^2\times\frac{1}{p^2-1}\times\frac{1}{q^2-1})$$

$$=\frac{x^2+4x}{2}+\frac{-(q+1)-(p+1)+1}{2(p+1)(q+1)}\times x^2$$

$$=\frac{pq}{2(p+1)(q+1)}x^2+2x\ ,$$

which implies the right side of (27).

In above inequality, we used the inequality $y-1<[y]\le y$ for any y.

Noting that (28) is an equality, similarly, one obtains that by (28)

$$\sum_{i\le x}H(i)>\frac{x(x-1)}{2}-\frac{p-1}{2}\{\frac{x}{p}(\frac{x}{p}+1)+\frac{x}{p^2}(\frac{x}{p^2}+1)+\frac{x}{p^3}(\frac{x}{p^3}+1)+......\}$$

$$-\frac{q-1}{2}\{\frac{x}{q}(\frac{x}{q}+1)+\frac{x}{q^2}(\frac{x}{q^2}+1)+\frac{x}{q^3}(\frac{x}{q^3}+1)+......\}$$

$$+\frac{(p-1)(q-1)}{2}\{(\frac{x}{pq}-1)\frac{x}{pq}+(\frac{x}{p^2q}-1)\frac{x}{p^2q}+(\frac{x}{pq^2}-1)\frac{x}{pq^2}+......\}$$

$$=\frac{x(x-1)}{2}-\frac{p-1}{2}\{(\frac{x}{p}+\frac{x}{p^2}+\frac{x}{p^3}+......)+(\frac{x^2}{p^2}+\frac{x^2}{p^4}+\frac{x^2}{p^6}+......)\}$$

$$-\frac{q-1}{2}\{(\frac{x}{q}+\frac{x}{q^2}+\frac{x}{q^3}+......)+(\frac{x^2}{q^2}+\frac{x^2}{q^4}+\frac{x^2}{q^6}+......)\}$$

$$+\frac{(p-1)(q-1)}{2}\{-x(\frac{1}{p}+\frac{1}{p^2}+......)(\frac{1}{q}+\frac{1}{q^2}+......)+x^2(\frac{1}{p^2}+\frac{1}{p^4}+......)(\frac{1}{q^2}+\frac{1}{q^4}+......)\}$$

$$=\frac{x(x-1)}{2}-\frac{p-1}{2}(\frac{x}{p-1}+\frac{x^2}{p^2-1})-\frac{q-1}{2}(\frac{x}{q-1}+\frac{x^2}{q^2-1})$$

$$+\frac{(p-1)(q-1)}{2}(-x\times\frac{1}{p-1}\times\frac{1}{q-1}+x^2\times\frac{1}{p^2-1}\times\frac{1}{q^2-1})$$

$$=\frac{pq}{2(p+1)(q+1)}x^2-2x,$$

which means that the left side of (27) holds. □

Following is the analysis of valued "1" point of $\{H(n)\}$.

One obtains by $H(n)=1$ that $n=p^uq^v$, where $u\ge 0, v\ge 0$ are integers.

If $p^uq^v\le x$, we have $u\log p+v\log q\le \log x$.

Let $u_0=\left[\frac{\log x}{\log p}\right], v_0=\left[\frac{\log x}{\log q}\right]$ and let $v=0,1,2,...,v_0$, then we obtain that

$$|Y(x)|=\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log\frac{x}{q}}{\log p}\right]+\left[\frac{\log\frac{x}{q^2}}{\log p}\right]+...+\left[\frac{\log\frac{x}{q^{v_0}}}{\log p}\right]+1, \tag{29}$$

and we can estimate $|Y(x)|$ by (29).

Following we obtain the estimation of $|Y(x)|$ by using another method (Best to my knowledge, it is difficult to get the same result by (29)).

Theorem 12. Let p and q be primes such that $p<q$. Let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule ER$(\{p,q\})$ of Question 2, $|Y(x)|$ be denoted as

the number of valued "1" point in the first $[x]$ term of $\{H(n)\}$. Then for any $x>0$ large enough, the following inequality holds:

$$\frac{1}{2}\frac{\log x}{\log p}\left(\frac{\log x}{\log q}+1\right)<|Y(x)|<\frac{1}{2}\left(\frac{\log x}{\log p}+1\right)\left(\frac{\log x}{\log q}+2\right). \qquad (30)$$

Proof. Since valued "1" point in the first $[x]$ term of $\{H(n)\}$ satisfies $n=p^u q^v$, one has $u\log p+v\log q\le\log x$ which implies that the point (u,v) is the grid point in the triangle $\triangle OAB$ (containing three lines) surrounded by U axis, V axis and the line $u\log p+v\log q=\log x$ in the $U-V$ plane. So the number of grid points equals to $|Y(x)|$.We use the Pick Theorem (see [4, P41-48] for details) to estimate the number of grid points.

Firstly, we prove the inequality on the left side of (30). We fix a grid polygon ω which containing $\triangle OAB$ (Figure 1). And ω is formed by the two straight and one fold lines. One is the segment OW_0 on U axis, such that the coordinate of W_0 is $\left(\left[\frac{\log x}{\log p}\right]+1,0\right)$. One is the segment OW on V axis, such that the coordinate of W is $\left(0,\left[\frac{\log x}{\log q}\right]+1\right)$. The other is the fold line from W_0 to W, such that the horizontal coordinate of the intersection point W_k of the fold line and the line $v=k$ is

$$\left[\frac{\log\frac{x}{q^k}}{\log p}\right]+1=\left[\frac{\log x-k\log q}{\log p}\right]+1, k=1,2,3,......,v_0=\left[\frac{\log x}{\log q}\right],$$

We get the fold line by linking the points W_0, W_k, $k=1,2,...,v_0$ and W, in turn. Noting that the fold line has no intersection points with segment AB, it follows that

$$S_\omega>S_{\triangle OAB}=\frac{\log^2 x}{2\log p\log q}.$$

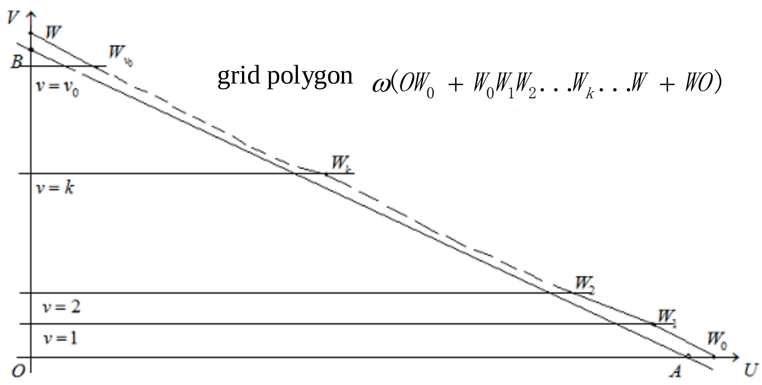

Figure 1

We denote E as the number of grid points inside the polygon ω, and F as the number on the border of ω. It is easy to see that, F is equal to the sum of the numbers of grid points on the line OW_0, line OW, and the fold line W_0W minus the number of repetitive grid points, which means that

$$F=\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log x}{\log q}\right]+3.$$

By Pick Theorem we have that

$$S_\omega = \mathrm{E}+\frac{F}{2}-1>\frac{\log^2 x}{2\log p\log q},$$

then it follows that

$$|Y(x)| = \text{the number of grid points inside } \triangle OAB \text{ (including the border)}$$

$$= \mathrm{E}+(1+\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right])>\frac{\log^2 x}{2\log p\log q}+\frac{1}{2}(1+\left[\frac{\log x}{\log p}\right])$$

$$>\frac{\log^2 x}{2\log p\log q}+\frac{1}{2}\frac{\log x}{\log p}=\frac{1}{2}\frac{\log x}{\log p}\left(\frac{\log x}{\log q}+1\right),$$

and the left side of (30) holds. And it is obvious that we still obtain the inequality when we exchange p and q.

Next we prove the inequality on the right side of (30). We fix the biggest grid polygon δ (Figure 2) contained in the triangle $\triangle OAB$. Similarly, δ is formed by the two straight and one fold lines. One is the segment OW_0' on U axis, such that the coordinate of the point W_0' is $\left(\left[\frac{\log x}{\log p}\right],0\right)$. One is the segment OW' on V axis, such that the coordinate of the point W' is $\left(0, \left[\dfrac{\log x}{\log q}\right]\right)$. The other is the fold line from the point W_0' to W', such that the horizontal coordinate of the intersection point W_k' of the fold line and the line $v = k$ is

$$\left[\frac{\log \dfrac{x}{q^k}}{\log p}\right] = \left[\frac{\log x - k\log q}{\log p}\right], k = 1, 2, 3, \ldots\ldots, v_0 .$$

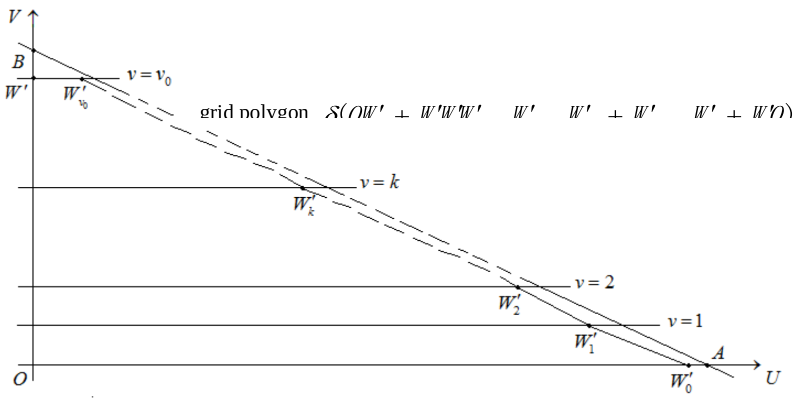

Figure 2

It is worthy to note that points W_{v_0}' and W' are both on the line $v = v_0 = \left[\dfrac{\log x}{\log q}\right]$. They may coincide or be different, or there may exist some other grid points between them. No matter what kind of situation, the number of grid points on the segment $W_{v_0}'W'$ is equal to $\left[\dfrac{\log \dfrac{x}{q^{v_0}}}{\log p}\right] + 1$ (including end points, if the two ends coincide, then $\left[\dfrac{\log \dfrac{x}{q^{v_0}}}{\log p}\right] = 0$).

One has by $p < q$ that the point W_k' is not on the V axis and it does not coincide with W' when $k < v_0$. Similarly, we get the fold line by linking the points W_0', W_k', $k = 1, 2, \ldots, v_0$ and W' in turn.

Since the points A and B cannot be grid points at the same time, we have that

$$S_\delta < S_{\triangle OAB} = \frac{\log^2 x}{2\log p \log q} .$$

We denote E_δ as the number of grid points inside the grid polygon δ, and F_δ as the number on the border of δ. It is easy to see that, F_δ is equal to the sum of the numbers of grid points on the segment OW_0', segment OW', and the fold line $W_0'W'$ minus the number of repetitive grid points which means that

$$F_\delta = (\left[\frac{\log x}{\log p}\right]+1)+(\left[\frac{\log x}{\log q}\right]+1)+(v_0+\left[\frac{\log\frac{x}{q^{v_0}}}{\log p}\right]+1)-3$$

$$=\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log\frac{x}{q^{v_0}}}{\log p}\right].$$

One has by Pick Theorem that

$$S_\delta = E_\delta+\frac{1}{2}\left(\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log\frac{x}{q^{v_0}}}{\log p}\right]\right)-1<\frac{\log^2 x}{2\log p\log q}.$$

Then it follows that

$|Y(x)|$ = the number of grid points inside $\triangle OAB$ (including the border)

$$=E_\delta+\left(\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log\frac{x}{q^{v_0}}}{\log p}\right]\right)$$

$$<\frac{\log^2 x}{2\log p\log q}+\frac{1}{2}\left(\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log\frac{x}{q^{v_0}}}{\log p}\right]\right)+1$$

$$=\frac{\log^2 x}{2\log p\log q}+\frac{1}{2}\left(\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log x-\left[\frac{\log x}{\log q}\right]\log q}{\log p}\right]\right)+1$$

$$=\frac{\log^2 x}{2\log p\log q}+\frac{1}{2}\left(\left[\frac{\log x}{\log p}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{\log x}{\log q}\right]+\left[\frac{(\frac{\log x}{\log q}-\left[\frac{\log x}{\log q}\right])\log q}{\log p}\right]\right)+1$$

$$< \frac{\log^2 x}{2\log p \log q} + \frac{1}{2}\left(\frac{\log x}{\log p} + \frac{\log x}{\log q} + \frac{\log x}{\log q}\right) + 1 + \frac{1}{2}\left[\frac{\log q}{\log p}\right]$$

$$= \frac{1}{2}(\frac{\log x}{\log p} + 2)(\frac{\log x}{\log q} + 1) + \frac{1}{2}\left[\frac{\log q}{\log p}\right].$$

Since $p < q, \frac{1}{\log p} > \frac{1}{\log q}$, for large enough x, we obtain that

$$\frac{1}{2}(\frac{\log x}{\log p} + 2)(\frac{\log x}{\log q} + 1) + \frac{1}{2}\left[\frac{\log q}{\log p}\right] < \frac{1}{2}(\frac{\log x}{\log p} + 1)(\frac{\log x}{\log q} + 2)$$

(as long as $\log x \times (\frac{1}{\log p} - \frac{1}{\log q}) > \left[\frac{\log q}{\log p}\right]$ is satisfied), and the right side of (30) holds. □

Now, we study the case that the number of primes is m. We obtain Theorem 13 and Theorem 14 by extending Theorem 11 and Theorem 12.

Theorem 13. Given m primes $p_1, p_2, ..., p_m$, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule $\mathrm{ER}(\{p_1, p_2, ..., p_m\})$ of Question 3. Then for any $x > 0$, the following inequality holds:

$$\frac{1}{2}\prod_{i=1}^{m}\frac{p_i}{p_i+1}x^2 - 2^{m-1}x < \sum_{i \le x} H(i) < \frac{1}{2}\prod_{i=1}^{m}\frac{p_i}{p_i+1}x^2 + 2^{m-1}x. \qquad (31)$$

Proof. We use induction to prove the theorem. When $m = 1, 2$, we obtain (31) by Theorem 11 and the proof of Theorem 11 (The proof is same as Theorem 11 for $m = 1$).

Assume that (31) holds for $m = r$. Following is the proof for $m = r + 1$.

Let $\{H_r(n)\}$ be the sequence generated by eliminating all $p_i (i = 1, 2, ..., r)$ factors for every term in $\{A_n\}$ according to ER($p_1, p_2, ..., p_r$) and $\{H_{r+1}(n)\}$ be the sequence generated by eliminating all $p_i (i = 1, 2, ..., r+1)$ factors according to ER($p_1, p_2, ..., p_{r+1}$). One has by assumption that

$$\frac{1}{2}\prod_{i=1}^{r}\frac{p_i}{p_i+1}x^2 - 2^{r-1}x < \sum_{i \le x} H_r(i) < \frac{1}{2}\prod_{i=1}^{r}\frac{p_i}{p_i+1}x^2 + 2^{r-1}x. \qquad (32)$$

It is easy to see that $H_{r+1}(i) = H_r(i)$ holds when $i \le x$ and i is not the multiple of p_{r+1}.

Define a function $G:[0,+\infty)\to R$ by

$$G(x)=\sum_{i\le x}H_r(i)-p_{r+1}\sum_{i\le x/p_{r+1}}H_r(i).$$

That is, $G(x)$ equals to the sum of the first $[x]$ terms of $\{H_r(n)\}$ with the all multiples of p_{r+1} in $\{H_r(n)\}$ removed.

Similar to the proof of Theorem 11, in order to classify and analyze every term of $\sum_{i\le x}H_{r+1}(i)$ we establish a class of set which satisfies that the intersection of any two different sets is empty. Let $t=\left[\frac{\log x}{\log p}\right], s=\left[\frac{\log x}{\log q}\right]$, then we have the following results.

Denote $A(0)=\{i\le x\mid (i,p_{r+1}^t)=1=p_{r+1}^0\}$, then $\sum_{i\in A(0)}H_{r+1}(i)=G(x)$;

Denote $A(1)=\{i\le x\mid (i,p_{r+1}^t)=p_{r+1}\}$, then $\sum_{i\in A(1)}H_{r+1}(i)=G(\frac{x}{p_{r+1}})$;

Denote $A(2)=\{i\le x\mid (i,p_{r+1}^t)=p_{r+1}^2\}$, then $\sum_{i\in A(2)}H_{r+1}(i)=G(\frac{x}{p_{r+1}^2})$;

Denote $A(3)=\{i\le x\mid (i,p_{r+1}^t)=p_{r+1}^3\}$, then $\sum_{i\in A(3)}H_{r+1}(i)=G(\frac{x}{p_{r+1}^3})$; …

We notice that the number of the sets $A(0),A(1),A(2),A(3),\ldots$ is $t+1$, the intersection of any two different sets is empty, and the union of them is the set of the first $[x]$ terms of the sequence of naturals. Thus

$$\sum_{i\le x}H_{r+1}(i)=\sum_{i\in A(0)}H_{r+1}(i)+\sum_{i\in A(1)}H_{r+1}(i)+\sum_{i\in A(2)}H_{r+1}(i)+\sum_{i\in A(3)}H_{r+1}(i)+\ldots+\sum_{i\in A(t)}H_{r+1}(i)$$

$$=G(x)+G(\frac{x}{p_{r+1}})+G(\frac{x}{p_{r+1}^2})+G(\frac{x}{p_{r+1}^3})+\ldots+G(\frac{x}{p_{r+1}^t})$$

$$=\left(\sum_{i\le x}H_r(i)-p_{r+1}\sum_{i\le\frac{x}{p_{r+1}}}H_r(i)\right)+\left(\sum_{i\le\frac{x}{p_{r+1}}}H_r(i)-p_{r+1}\sum_{i\le\frac{x}{p_{r+1}^2}}H_r(i)\right)$$

$$+\left(\sum_{i\le\frac{x}{p_{r+1}^2}}H_r(i)-p_{r+1}\sum_{i\le\frac{x}{p_{r+1}^3}}H_r(i)\right)+\ldots+\left(\sum_{i\le\frac{x}{p_{r+1}^t}}H_r(i)-p_{r+1}\sum_{i\le\frac{x}{p_{r+1}^{t+1}}}H_r(i)\right)$$

(Noting that $p_{r+1}\sum_{i\le \frac{x}{p_{r+1}^{t+1}}} H_r(i)=0$)

$$=\sum_{i\le x}H_r(i)-(p_{r+1}-1)\left(\sum_{i\le \frac{x}{p_{r+1}}}H_r(i)+\sum_{i\le \frac{x}{p_{r+1}^2}}H_r(i)+\ldots+\sum_{i\le \frac{x}{p_{r+1}^t}}H_r(i)\right). \qquad (33)$$

Noting that (33) is an equality, then by assumption (32) we obtain that

$$\sum_{i\le x}H_{r+1}(i)<(\frac{1}{2}\prod_{i=1}^{r}\frac{p_i}{p_i+1}x^2+2^{r-1}x)$$

$$-(p_{r+1}-1)\left(\frac{1}{2}\prod_{i=1}^{r}\frac{p_i}{p_i+1}(\frac{x}{p_{r+1}})^2-2^{r-1}(\frac{x}{p_{r+1}})+\frac{1}{2}\prod_{i=1}^{r}\frac{p_i}{p_i+1}(\frac{x}{p_{r+1}^2})^2-2^{r-1}(\frac{x}{p_{r+1}^2})+\ldots\right)$$

$$=\left(\frac{1}{2}\prod_{i=1}^{r}\frac{p_i}{p_i+1}x^2\left(1-(p_{r+1}-1)\left(\frac{1}{p_{r+1}^2}+\frac{1}{p_{r+1}^4}+\frac{1}{p_{r+1}^6}+\ldots\right)\right)\right)$$

$$+\left(x\left(2^{r-1}+2^{r-1}(p_{r+1}-1)\left(\frac{1}{p_{r+1}}+\frac{1}{p_{r+1}^2}+\frac{1}{p_{r+1}^3}+\ldots\right)\right)\right)$$

$$=\frac{1}{2}\prod_{i=1}^{r}\frac{p_i}{p_i+1}x^2\left(1-\frac{1}{p_{r+1}+1}\right)+x\left(2^{r-1}+2^{r-1}(p_{r+1}-1)\frac{1}{p_{r+1}-1}\right)$$

$$=\frac{1}{2}\prod_{i=1}^{r+1}\frac{p_i}{p_i+1}x^2+2^r x,$$

and the right side of (31) is established for $m=r+1$. The left side of (31) can be established by using the same method. Then the proof is complete by induction. □

By using the notation "O" (see [2, P147]) we have a more concise conclusion of Theorem 13:

$$\sum_{i\le x}H(i)=\frac{1}{2}\prod_{i=1}^{m}\frac{p_i}{p_i+1}x^2+O(x). \qquad (34)$$

We derive a new proof for the noted fact that **the set of primes is infinite** by Theorem 13 (see [3] for the four kinds of proofs of the proposition). Here is a brief proof.

Suppose that the primes are $p_1,p_2,\ldots,p_m$, where m is a positive integer. Then the terms of the positive integers series generated from $\{A_n\}$ become "1" after eliminating all the prime factors according to the rule ER $\{p_1,p_2,\ldots,p_m\}$, and one has that

$\sum_{i\le x} H(i)=\sum_{i\le x} 1=[x]$. It is easy to see that the equality contradicts with (31) when $x\to\infty$, which implies that there cannot be only a finite number of primes. □

We apply the notation $"O"$ to Theorem 12, and extend it to the following theorem.

Theorem 14. Given m primes $p_1, p_2, \dots, p_m$, let $\{H(n)\}$ be the sequence generated by transforming $\{A_n\}$ according to the rule $\mathrm{ER}(\{p_1, p_2, \dots, p_m\})$ of Question 3, and $|Y(x)|$ be denoted as the number of valued "1" point in the first $[x]$ term of $\{H(n)\}$. Then for any $x>0$, the following inequality holds:

$$|Y(x)|=\frac{1}{m!}\frac{\log^m x}{\prod_{i=1}^{m}\log p_i}+O(\log^{m-1} x). \tag{35}$$

Proof. As the beginning of the proof, we list two useful conclusions:

(1) $$\frac{C_m^0}{1}-\frac{C_m^1}{2}+\frac{C_m^2}{3}-\dots+(-1)^m\frac{C_m^m}{m+1}=\frac{1}{m+1} \tag{36}$$

multiplying $m+1$ on the left side, one obtains by $\frac{C_m^i}{i+1}\times(m+1)=C_{m+1}^{i+1}$ that

$$\left(\frac{C_m^0}{1}-\frac{C_m^1}{2}+\frac{C_m^2}{3}-\dots+(-1)^m\frac{C_m^m}{m+1}\right)\times(m+1)$$

$$=C_{m+1}^1-C_{m+1}^2+C_{m+1}^3-\dots+(-1)^m C_{m+1}^{m+1}=-(1-1)^{m+1}+1=1,$$

(36) is established.

(2) $$\sum_{n\le x} n^s=\frac{x^{s+1}}{s+1}+O(x^s),\ \text{for}\ \ s\ge 0 \qquad \text{(see [2,p49])}. \tag{37}$$

We use induction to prove the theorem. It is obvious that (35) holds for $m=1,2$. Assume that (35) holds for $m=r$. Following is the proof for $m=r+1$.

We denote $\{H_r(n)\}$ (resp. $\{H_{r+1}(n)\}$) as the sequence generated by eliminating all $p_i(i=1,2,\dots,r)$ (resp. $p_i(i=1,2,\dots,r+1)$) factors for every term of $\{A_n\}$ according to rule of Question 3 and $Y_r(x)$ (resp. $Y_{r+1}(x)$) as set of valued "1" point in the first $[x]$ terms of $\{H_r(n)\}$ (resp. $\{H_{r+1}(n)\}$).

Let $s=\left[\dfrac{\log x}{\log p_{r+1}}\right]$. By analyzing elements in $Y_{r+1}(x)$ we have that

$$Y_{r+1}(x)=Y_r(x)\cup p_{r+1}Y_r(\frac{x}{p_{r+1}})\cup p_{r+1}^2Y_r(\frac{x}{p_{r+1}^2})\cup p_{r+1}^3Y_r(\frac{x}{p_{r+1}^3})\cup\ldots p_{r+1}^sY_r(\frac{x}{p_{r+1}^s}).$$

Noting that there are $s+1$ sets on the right side and the intersection of any two different sets is empty. This equality together with the assumption in the induction implies that

$$|Y_{r+1}(x)|=|Y_r(x)|+|Y_r(\frac{x}{p_{r+1}})|+|Y_r(\frac{x}{p_{r+1}^2})|+|Y_r(\frac{x}{p_{r+1}^3})|+\ldots+|Y_r(\frac{x}{p_{r+1}^s})|$$

$$=\left(\frac{1}{r!}\frac{\log^r x}{\prod_{i=1}^r\log p_i}+O(\log^{r-1}x)\right)+\left(\frac{1}{r!}\frac{\log^r(x/p_{r+1})}{\prod_{i=1}^r\log p_i}+O(\log^{r-1}(x/p_{r+1}))\right)$$

$$+\left(\frac{1}{r!}\frac{\log^r(x/p_{r+1}^2)}{\prod_{i=1}^r\log p_i}+O(\log^{r-1}(x/p_{r+1}^2)\right)+\ldots+\left(\frac{1}{r!}\frac{\log^r(x/p_{r+1}^s)}{\prod_{i=1}^r\log p_i}+O(\log^{r-1}(x/p_{r+1}^s)\right)$$

$$=\frac{1}{r!}\frac{1}{\prod_{i=1}^r\log p_i}\left(\log^r x+\log^r(x/p_{r+1})+\log^r(x/p_{r+1}^2)+\ldots+\log^r(x/p_{r+1}^s)\right)$$

$$+O(\log^{r-1}(x)+O(\log^{r-1}(x/p_{r+1})+O(\log^{r-1}(x/p_{r+1}^2)+\ldots+O(\log^{r-1}(x/p_{r+1}^s).\quad (38)$$

By denoting U as part with $"O"$, we have that

$$U=O(\log^{r-1}(x)\times(\left[\frac{\log x}{\log p_{r+1}}\right]+1))=O(\log^r(x)).$$

Let $W=\left(\log^r x+\log^r(x/p_{r+1})+\log^r(x/p_{r+1}^2)+\ldots+\log^r(x/p_{r+1}^s)\right)$. Following is the analysis of W, we choose the $(t+1)$th term

$$\log^r(x/p_{r+1}^t)=(\log x-t\log p_{r+1})^r$$

$$=C_r^0\log^r x-C_r^1\log^{r-1}x(t\log p_{r+1})+C_r^2\log^{r-2}x(t\log p_{r+1})^2-\ldots+(-1)^rC_r^r(t\log p_{r+1})^r.$$

We set $t=0,1,2,\ldots,s$, and sum the corresponding terms up in the expansion. It follows from (37) that

$$\sum_{t=0}^{s}(-1)^l C_r^l \log^{r-l} x(t\log p_{r+1})^l = (-1)^l C_r^l \log^{r-l} x \log^l p_{r+1}\sum_{t=0}^{s} t^l$$

$$=(-1)^l C_r^l \log^{r-l} x \log^l p_{r+1}(\frac{1}{l+1}s^{l+1}+O(s^l))$$

$$=(-1)^l C_r^l \log^{r-l} x \log^l p_{r+1}\left(\frac{1}{l+1}(\frac{\log x}{\log p_{r+1}})^{l+1}+O(\frac{\log x}{\log p_{r+1}})^l\right)$$

$$=(-1)^l C_r^l \times\frac{1}{l+1}\times\frac{\log^{r+1} x}{\log p_{r+1}}+O(\log^r x)\,. \qquad (39)$$

Let $l=0,1,2,...,r$ in (39), and we sum them up. One has by (36) that

$$W=\sum_{l=0}^{r}\left((-1)^l C_r^l\times\frac{1}{l+1}\times\frac{\log^{r+1} x}{\log p_{r+1}}+O(\log^r x)\right)$$

$$=\sum_{l=0}^{r}(-1)^l C_r^l\times\frac{1}{l+1}\times\frac{\log^{r+1} x}{\log p_{r+1}}+O(\log^r x)$$

$$=\frac{\log^{r+1} x}{\log p_{r+1}}\times\sum_{l=0}^{r}((-1)^l C_r^l\times\frac{1}{l+1})+O(\log^r x)$$

$$=\frac{1}{r+1}\times\frac{\log^{r+1} x}{\log p_{r+1}}+O(\log^r x)\,. \qquad (40)$$

Substituting (40) into (39), we can obtain that

$$|Y_{r+1}(x)|=\frac{1}{r!}\frac{1}{\prod_{i=1}^{r}\log p_i}\times\left(\frac{1}{r+1}\times\frac{\log^{r+1} x}{\log p_{r+1}}+O(\log^r x)\right)+O(\log^r x)$$

$$=\frac{1}{(r+1)!}\frac{\log^{r+1} x}{\prod_{i=1}^{r+1}\log p_i}+O(\log^r x)\,,$$

(35) holds for $m=r+1$. The proof is complete by induction.

Summary and prospect

1.In this paper, we start from an interesting property of the sequence generated by eliminating all the factor 2 for each term of the sequence of the naturals. Eliminating all the prime factor p, or p, q, or $p_i (i = 1, 2, ..., m)$ according to the relative transforming rules in the different cases, we obtain the new sequence $\{H(n)\}$ respectively. We studied the partial sum and the number of valued "1" points of the new sequence $\{H(n)\}$. Furthermore, we get several concise inequalities which estimate the partial sum of the first n terms and the numbers of valued "1"points in the first n terms of $\{H(n)\}$.

However, these results are still relatively preliminary. The conclusions are complicated in the case of given m primes. Though the inequality in the fourth section is relatively concise and beautiful, there exists deficiencies in (32) and (33) that the coefficient "O" is in connection with the number of the given primes.

2.On the problem of the sum related to the sequence of the naturals, there are some known equalities as follows:

（1） $\sum_{i=1}^{n} i = \frac{1}{2}(n^2 + n)$;

（2） $\sum_{n \le x} n^s = \frac{x^{s+1}}{s+1} + O(x^s)$, where $s \ge 0$

Compare (34)

$$\sum_{i \le x} H(i) = \frac{1}{2} \prod_{i=1}^{m} \frac{p_i}{p_i + 1} x^2 + O(x)$$

with the equalities above, it's meaningful to study the sum of the form like $\sum_{n \le x} H^s(n)$, and the following equality may be correct:

$$\sum_{n \le x} H^s(n) = \frac{1}{s+1} \prod_{i=1}^{m} \frac{p_i}{p_i + 1} x^{s+1} + O(x^s), \qquad \text{where } s \ge 0.$$

When $s < 0$ there are some known equalities about $\sum_{n \le x} n^s$ as well, and we can also study the partial sum such as $\sum_{n \le x} H^s(n)$ (where $s < 0$), maybe there are some interesting connections between them.

3. $H(n)$ is a new sequence generated by transforming the sequence of naturals according to the eliminating rule. Maybe, we can consider to apply the transforming rule on the

sequence consisted of integers, the new sequence generated may have some interesting connection with the sequence. For example, given m primes, we have transforming rule ER $(\{p_1, p_2, ..., p_m\})$. Let p be a prime, $\{a_n\}$ be an arithmetic sequence, $a_n = pn + l$, $(p, l) = 1, l \in N$,Then what properties does the new sequence generated by transforming $\{a_n\}$ according the eliminating rule ER $(\{p_1, p_2, ..., p_m\})$ have?

4.Furthermore, we can consider the more general transforming as follows:

Let Q be a set of some primes, finite or infinite, define a transforming rule ER(Q) as: For every $n \in N$ and every $p \in Q$, if some term A_n of $\{A_n\}$ has factor $p \in Q$, then substitute A_n with A_n / p , until it does not have factor p. Thus we obtain a new sequence $\{H_Q(n)\}$. Maybe we can write as

$$ER(Q)(\{A_n\}) = \{H_Q(n)\},$$

Also, we can apply ER(Q) on the sequence consisted of integers, for example, on the arithmetic sequence $\{a_n\}$ above, what properties does the new sequence $ER(Q)(\{a_n\})$ have?

I think these questions are interesting and meaningful, but it's difficult for me to answer as I'm just a high school student now. It is my dream to learn more knowledge about mathematics and do further research on them in future.

Finally, I would like to express my gratitude to my two foreign tutors, to Mr. Yau for providing this platform for students who love mathematics, to all teachers for your review and guidance for this paper.

Reference：

[1]《简明数论》 潘承洞 潘承彪 著 2016 年
[2]《Introduction to Analytic Number Theory》 Tom M. Apostol 2012
[3]《素数定理的初等证明》 潘承彪 潘承洞 著 2017 年
[4]《毕克定理》 佩 捷 主编 2014 年