EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 17, No. 1, 2024, 42-58
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Application of the Inclusion-Exclusion Principle to Prime Number Subsequences

Michael P. May

Department of Mechanical & Aerospace Engineering, Missouri University of Science and Technology, Rolla, Missouri, USA

Abstract. We apply the Inclusion-Exclusion Principle to a unique pair of prime number subsequences to determine whether these subsequences form a small set or a large set and thus whether the infinite sum of the inverse of their terms converges or diverges. In this paper, we analyze the complementary prime number subsequences $\mathbb{P}'$ and $\mathbb{P}''$ as well as revisit the twin prime subsequence $\mathbb{P}_2$.

2020 Mathematics Subject Classifications: 11A41, 11L20, 11B05, 11K31, 11N36

Key Words and Phrases: Prime numbers, Higher-order prime number sequences, Inclusion-Exclusion Principle, Sums over prime reciprocals, Prime density, Sieves

1. The Prime Subsequences $\mathbb{P}'$ and $\mathbb{P}''$

The prime number subsequence [4]

$$\mathbb{P}^{'} = \{p'\} = \{2, 5, 7, 13, 19, 23, 29, 31, 37, 43, 47, 53, 59, 61, 71, ...\}$$

can be generated via an alternating sum of the prime number subsequences of increasing order [7], i.e.,

$$\mathbb{P}^{'} = \left\{(-1)^{n-1}\left\{p^{(n)}\right\}\right\}_{n=1}^{\infty} \tag{1}$$

where the right-hand side of Eq. 1 is an expression of the alternating sum

$$\left\{p^{(1)}\right\} - \left\{p^{(2)}\right\} + \left\{p^{(3)}\right\} - \left\{p^{(4)}\right\} + \left\{p^{(5)}\right\} - \dots . \tag{2}$$

The prime number subsequences of increasing order [1] in Expression 2 are defined as

DOI: https://doi.org/10.29020/nybg.ejpam.v17i1.4979

Email address: `mike.may.bbi@gmail.com` (M. P. May)

 © 2024 EJPAM All rights reserved.

$$\left\{p^{(1)}\right\} = \{p_n\}_{n=1}^{\infty} = \{2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, ...\}$$

$$\left\{p^{(2)}\right\} = \{p_{p_n}\}_{n=1}^{\infty} = \{3, 5, 11, 17, 31, 41, 59, 67, 83, 109, 127, ...\}$$

$$\left\{p^{(3)}\right\} = \{p_{p_{p_n}}\}_{n=1}^{\infty} = \{5, 11, 31, 59, 127, 179, 277, 331, ...\}$$

$$\left\{p^{(4)}\right\} = \left\{p_{p_{p_{p_n}}}\right\}_{n=1}^{\infty} = \{11, 31, 127, 277, 709, ...\}$$

$$\left\{p^{(5)}\right\} = \left\{p_{p_{p_{p_{p_n}}}}\right\}_{n=1}^{\infty} = \{31, 127, 709, ...\}$$

and so on and so forth. Thus, the operation performed on the right-hand side of Eq. 1 denotes an infinite alternating sum of the sets of prime number subsequences of increasing order.

The prime number subsequence $\mathbb{P}^{'}$ can also be generated by performing a structured alternating summation of the individual elements across the sets denoted on the right-hand side of Eq. 1. To illustrate this, we arrange the subsequences of increasing order [1] in Expression 2 side-by-side and sum elements laterally across the rows to create the $\mathbb{P}^{'}$ subsequence term-by-term as follows:

Table 1: Alternating Sum of $p^{(n)}$

(row)	$+p^{(1)}$	$-p^{(2)}$	$+p^{(3)}$	$-p^{(4)}$	$+p^{(5)}$	$-p^{(6)}$	...	p'
(1)	2	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	2
(2)	3	3	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	0
(3)	5	5	5	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	5
(4)	7	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	7
(5)	11	11	11	11	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	0
(6)	13	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	13
(7)	17	17	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	0
(8)	19	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	19
(9)	23	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	23
(10)	29	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	$\longrightarrow$	29
(11)	31	31	31	31	31	$\longrightarrow$	$\longrightarrow$	31
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Thus, the infinite prime number subsequence $\mathbb{P}^{'}$ of higher order [4] that emerges in the rightmost column of Table 1 is

$$\mathbb{P}^{'} = \left\{p'\right\} = \{2, 5, 7, 13, 19, 23, 29, 31, 37, 43, 47, 53, 59, 61, 71, ...\}.$$

The prime number subsequence $\mathbb{P}^{'}$ can also be generated by performing a sieving operation on the natural numbers $\mathbb{N}$ [7]. Starting with $n = 1$, choose the prime number with subscript 1 (i.e., $p_1 = 2$) as the first term of the subsequence and eliminate that prime number from the natural number line. Next, proceed forward on $\mathbb{N}$ from 1 to the next available natural number. Since 2 was eliminated from the natural number line in the previous step, one moves forward to the next available natural number that has not been eliminated, which is 3. The prime number 3 then becomes the subscript for the next $\mathbb{P}^{'}$ term which is $p_3 = 5$, and 5 is then eliminated from the natural number line, and so on and so forth. Such a sieving operation has been carried out in Table 2 for the natural numbers 1 to 100:

Table 2: Sieving $\mathbb{N}$ to generate $\mathbb{P}'$

1	**2**	3	4	**5**	6	**7**	8	9	10
11	12	**13**	14	15	16	17	18	**19**	20
21	22	**23**	24	25	26	27	28	**29**	30
31	32	33	34	35	36	**37**	38	39	40
41	42	**43**	44	45	46	**47**	48	49	50
51	52	**53**	54	55	56	57	58	**59**	60
61	62	63	64	65	66	67	68	69	70
71	72	**73**	74	75	76	77	78	**79**	80
81	82	83	84	85	86	87	88	**89**	90
91	92	93	94	95	96	**97**	98	99	100

Thus, we may also designate $\mathbb{P}'$, which has been alternately created via the sieving operation in Table 2, by the following notation [7] to indicate that the natural numbers $\mathbb{N}$ have been sieved to produce this prime number subsequence:

$$\lfloor \underline{\mathbb{N}} \rfloor = \mathbb{P}' = \{2, 5, 7, 13, 19, 23, 29, 31, 37, 43, 47, 53, 59, 61, 71, ...\}.$$

Regardless of which one of these three methods is used to generate $\mathbb{P}'$, when the prime numbers in this unique subsequence are applied as indexes to the set of all prime numbers $\mathbb{P}$, one obtains the next higher-order prime number subsequence $\mathbb{P}''$ [3]:

$$\mathbb{P}^{''} = \left\{p''\right\} = \{3, 11, 17, 41, 67, 83, 109, 127, 157, 191, 211, 241, ...\}.$$

By definition, the sequence $\mathbb{P}''$ can also be generated via the expression [7]

$$\mathbb{P}^{''} = \left\{(-1)^n \left\{p^{(n)}\right\}\right\}_{n=2}^{\infty} \tag{3}$$

where an expansion of the right-hand side of Eq. 3 reveals the alternating sum

$$\left\{p^{(2)}\right\} - \left\{p^{(3)}\right\} + \left\{p^{(4)}\right\} - \left\{p^{(5)}\right\} + \left\{p^{(6)}\right\} - ... .$$

The prime number subsequence of higher order $\mathbb{P}^{''}$ can also be generated by performing the aforementioned sieving operation on the set of all prime numbers $\mathbb{P}$, similar to how the primes $\mathbb{P}^{'}$ were sifted from the set of all natural numbers $\mathbb{N}$. Furthermore, it has been shown [7] that the subsequences $\mathbb{P}^{'}$ and $\mathbb{P}^{''}$, when added together, form the entire set of prime numbers $\mathbb{P}$:

$$\mathbb{P} = \mathbb{P}' + \mathbb{P}''. \tag{4}$$

We sketch a proof of Eq. 4 here:

Proof.
It has been shown [7] that

$$\mathbb{P}' = \left\{(-1)^{n-1}\left\{p^{(n)}\right\}\right\}_{n=1}^{\infty} = \left\{p^{(1)}\right\} - \left\{p^{(2)}\right\} + \left\{p^{(3)}\right\} - ...$$

and

$$\mathbb{P}'' = \left\{(-1)^{n}\left\{p^{(n)}\right\}\right\}_{n=2}^{\infty} = \left\{p^{(2)}\right\} - \left\{p^{(3)}\right\} + \left\{p^{(4)}\right\} - ... \, .$$

Therefore,

$$\begin{aligned}\mathbb{P}' + \mathbb{P}'' = &\left\{p^{(1)}\right\} - \left\{p^{(2)}\right\} + \left\{p^{(3)}\right\} - ... \\ &\qquad\qquad + \\ &\left\{p^{(2)}\right\} - \left\{p^{(3)}\right\} + \left\{p^{(4)}\right\} - ... = \left\{p^{(1)}\right\} = \mathbb{P}.\end{aligned}$$

An interesting property was observed in the relationship between the set of all prime numbers $\mathbb{P}$ and the complementary prime number sets $\mathbb{P}'$ and $\mathbb{P}''$. Since $\mathbb{P}'' = \mathbb{P}_{\mathbb{P}'}$, Eq. 4 can be rewritten as

$$\begin{aligned}\mathbb{P}'' &= \mathbb{P} - \{2, 5, 7, 13, 19, 23, 29, ...\} \\ &= \left\{p_2, p_5, p_7, p_{13}, p_{19}, p_{23}, p_{29}, ...\right\} = \mathbb{P}_{\mathbb{P}'}\end{aligned}$$

where the prime numbers of the subsequence $\mathbb{P}'$ form the indexes for the complement set of primes $\mathbb{P}''$ such that

$$\mathbb{P}'' = \mathbb{P}_{\mathbb{P}'} = \{p_k \mid k \in \mathbb{P}'\}.$$

2. Asymptotic Densities of $\mathbb{P}'$ and $\mathbb{P}''$

We will now derive the asymptotic densities for the prime number subsequences $\mathbb{P}'$ and $\mathbb{P}''$ assuming that $1/\ln n$ is the asymptotic density of the set of all prime numbers $\mathbb{P}$ as $n \to \infty$. We approach this task by alternately adding and subtracting the prime number densities (or "probabilities") of the prime number subsequences of increasing order, also known as "superprimes" [1], to arrive at values for the asymptotic densities for $\mathbb{P}'$ and $\mathbb{P}''$. We begin by recalling [7] that the prime number subsequence $\mathbb{P}'$ is formed by the alternating series

$$\mathbb{P}' = \left\{(-1)^{n-1}\left\{p^{(n)}\right\}\right\}_{n=1}^{\infty} = \left\{p^{(1)}\right\} - \left\{p^{(2)}\right\} + \left\{p^{(3)}\right\} - ...$$

where

$$\left\{p^{(k)}\right\} = \left\{p_{p_{\cdot_{\cdot_{\cdot p_n}}}}\right\} \text{ (p "k" times).}$$

Broughan and Barnett have shown [1] that for the general case of higher-order "superprimes" $p_{p_{\cdot_{\cdot_{\cdot p_k}}}}$, the asymptotic density is approximately

$$\frac{n}{p_{p_{\cdot_{\cdot_{\cdot p_n}}}}} \sim \frac{n}{n\,(\ln n)^k} \sim \frac{1}{(\ln n)^k}$$

for large $n \in \mathbb{N}$. Now, assuming that $1/\ln n$ is the asymptotic density for the set of all prime numbers $\mathbb{P}$, we derive an expression for the density d' for the prime number subsequence $\mathbb{P}'$ at ∞. We begin with the geometric series

$$S = 1 - x + x^2 - x^3 + x^4 - x^5 + ... = \frac{1}{1+x} \quad (|x| < 1).$$

Then let

$$T' = 1 - S$$

so that

$$\begin{aligned} T' &= x - x^2 + x^3 - x^4 + x^5 + ... \\ &= \frac{-1}{1+x} + 1 \\ &= \frac{x}{1+x}. \end{aligned}$$

Now substitute $\dfrac{1}{\ln n}$ for x to get

$$\frac{\frac{1}{\ln n}}{1+\frac{1}{\ln n}} = \frac{1}{\ln n + 1}$$

so that we have

$$\begin{aligned} d' &\approx \frac{1}{\ln n} - \frac{1}{(\ln n)^2} + \frac{1}{(\ln n)^3} - \frac{1}{(\ln n)^4} + ... && (5)\\ &= \frac{1}{\ln n + 1}. && (6)\end{aligned}$$

Similarly, we derive the asymptotic density for the prime number subsequence $\mathbb{P}''$. When we set

$$T'' = S - (1-x)$$

we have

$$\begin{aligned} T'' &= S - (1-x)\\ &= x^2 - x^3 + x^4 - x^5 + ...\\ &= \frac{1}{1+x} - 1 + x\\ &= \frac{x^2}{1+x}.\end{aligned}$$

Now substitute $\frac{1}{\ln n}$ for x to get

$$\frac{\left(\frac{1}{\ln n}\right)^2}{1+\frac{1}{\ln n}} = \frac{1}{\ln n(\ln n + 1)}$$

so that

$$\begin{aligned} d'' &\approx \frac{1}{(\ln n)^2} - \frac{1}{(\ln n)^3} + \frac{1}{(\ln n)^4} - \frac{1}{(\ln n)^5} + ... && (7)\\ &= \frac{1}{\ln n(\ln n + 1)}. && (8)\end{aligned}$$

Based on our assumption that $1/\ln n$ is the asymptotic density of the set of all prime numbers $\mathbb{P}$ as $n \to \infty$, Eqs. 6 and 8 provide us with the densities (or probabilities of occurrence) of the primes in the complementary sets $\mathbb{P}'$ and $\mathbb{P}''$, respectively, as n approaches ∞. Thus, the average gap size g' between prime numbers in the subsequence $\mathbb{P}'$ on the natural number line as $n \to \infty$ is the inverse of the density d' of $\mathbb{P}'$ such that

$$\begin{aligned} g' = \frac{1}{d'} &\approx \frac{1}{\frac{1}{\ln n} - \frac{1}{(\ln n)^2} + \frac{1}{(\ln n)^3} - \frac{1}{(\ln n)^4} + ...} \\ &= \ln n + 1. \end{aligned}$$

Similarly, the average gap size g'' between prime numbers in the subsequence $\mathbb{P}''$ on the natural number line as $n \to \infty$ is the inverse of the density d'' of $\mathbb{P}''$ such that

$$\begin{aligned} g'' = \frac{1}{d''} &\approx \frac{1}{\frac{1}{(\ln n)^2} - \frac{1}{(\ln n)^3} + \frac{1}{(\ln n)^4} - \frac{1}{(\ln n)^5} + ...} \\ &= \ln n(\ln n + 1). \end{aligned}$$

Since it has been shown via the sieving operation [7] that the prime number subsequence $\mathbb{P}'$ has fewer primes than the set of all prime numbers $\mathbb{P}$, it intuitively follows that the average gap size for $\mathbb{P}'$ will always be larger than the gap size for $\mathbb{P}$ and that the larger gap size for $\mathbb{P}'$ results from omitting the count of the prime numbers $\mathbb{P}''$ on $\mathbb{N}$.

3. $\pi'(x)$ and $\pi''(x)$

We have shown that when we remove the prime number subsequence $\mathbb{P}''$ from the set of all prime numbers $\mathbb{P}$, we create the prime number subsequence $\mathbb{P}'$ [7]. Thus, we define the prime number count for the sequences $\mathbb{P}'$ and $\mathbb{P}''$ up to x as

$$\pi'(x) = |\mathbb{P}'(x)|$$

and

$$\pi''(x) = |\mathbb{P}''(x)|$$

where $|\mathbb{P}'(x)|$ and $|\mathbb{P}''(x)|$ represent the cardinality of the prime number subsequences $\mathbb{P}'$ and $\mathbb{P}''$ up to x. However, since neither $\pi'(x)$ nor $\pi''(x)$ have been shown up to this point to be calculable without manually counting each term up to x, we will begin by generating an estimate of the count $\pi(x)$ of set of all primes $\mathbb{P}$ up to x via the Inclusion-Exclusion Principle and then perform an operation on that result to reduce the count of all primes down to $\pi'(x)$ and $\pi''(x)$.

4. $\pi(x)$ via the Inclusion-Exclusion Principle

To calculate $\pi(x)$, we invoke the Inclusion-Exclusion Principle [8] [6]. Let r represent the number of primes less than $\sqrt{x}$. Then let $P = \{n \in N | 1 < n \leq x\}$ such that n is not a multiple of $p_1, p_2, ..., p_r$. If $A(x, r)$ represents the cardinality of P, then it follows that the number of primes $\leq x$ is

$$\pi(x) \leq r + A(x, r).$$

Now, let M_i be the set of integers from 1 to n which are multiples of p_i, and let M_{ij} be the set of integers from 1 to n that are multiples of both p_i and p_j. Then,

$$M_{ij} = M_i \cap M_j$$

so that

$$|M_i| = \lfloor \frac{x}{p_i} \rfloor \quad \text{and} \quad |M_{ij}| = \lfloor \frac{x}{p_i p_j} \rfloor.$$

Then it follows by the Inclusion-Exclusion Principle that

$$A(x, r) = \lfloor x \rfloor - \sum_{i=1}^{r} \lfloor \frac{x}{p_i} \rfloor + \sum_{i<j\leq r}^{r} \lfloor \frac{x}{p_i p_j} \rfloor - \ldots + (-1)^r \lfloor \frac{x}{p_1 p_2 \cdots p_r} \rfloor. \tag{9}$$

If we approximate the RHS of Eq. 9 by ignoring the round-downs, then we have

$$x - \sum_{i=1}^{r} \frac{x}{p_i} + \sum_{i<j\leq r}^{r} \frac{x}{p_i p_j} - \ldots + (-1)^r \frac{x}{p_1 p_2 \cdots p_r}$$

with an error of at most

$$1 + \binom{r}{1} + \binom{r}{2} + \ldots + \binom{r}{r} = 2^r.$$

Thus, we now have for our estimate of the number of primes less than or equal to x as

$$\pi(x) \leq r + x \cdot \prod_{i=1}^{r} \left(1 - \frac{1}{p_i}\right) + 2^r. \tag{10}$$

We now want to choose r relatively small compared to x. In order to do so, we need a good estimate of the coefficient of x in the middle term on the RHS of 10 in terms of r.

Theorem 1. *If $x \geq 2$, then* $\prod_{p\leq x} \left(1 - \frac{1}{p}\right) < \frac{1}{\ln x}$.

Proof.

$$\prod_{p\le x}\frac{1}{1-\dfrac{1}{p}}=\prod_{p\le x}\left(1+\frac{1}{p}+\frac{1}{p^2}+\dots\right).$$

Now,

$$\prod_{p\le x}\frac{1}{1-\dfrac{1}{p}}>\sum_{k=1}^{n}\frac{1}{k}>\int_{1}^{\lceil x\rceil}\frac{du}{u}>\ln x.$$

$$\therefore\ \prod_{p\le x}\left(1-\frac{1}{p}\right)<\frac{1}{\ln x}\Rightarrow\prod_{i=1}^{r}\left(1-\frac{1}{p_i}\right)<\frac{1}{\ln p_r}.$$

We now have as our estimate of $\pi(x)$,

$$\pi(x)\le r+\frac{x}{\ln p_r}+2^r. \tag{11}$$

5. Estimating $\pi''(x)$

In order to calculate an estimate of $\pi''(x)$, we begin by taking a look at the coefficient of x in the middle term on the RHS of 10. We can write that coefficient as

$$\prod_{i=1}^{r}\left(1-\frac{1}{p_i}\right)=\prod_{i=1}^{r'}\left(1-\frac{1}{p'_i}\right)\cdot\prod_{i=1}^{r''}\left(1-\frac{1}{p''_i}\right) \tag{12}$$

where r' is the number of p' < the number of the first r primes $\le\sqrt{x}$, and r'' is the number of p'' < the number of the first r primes $\le\sqrt{x}$ such that $r'+r''=r$. We found that one cannot simply divide the product on the LHS of Eq. 12 by either product on the RHS of Eq. 12 and expect the quotient to represent a pure count of p' or $p''\le x$. Therefore, we must approach the problem from a different direction; i.e., we must find another way to reduce the coefficient of x on the RHS of 10 such that the estimate will leave the count of p'' only with no p' and no composites remaining when the coefficient is multiplied by x. Hence, we model the inequality in Theorem 1 as

$$\prod_{i=1}^{r}\left(1-\frac{1}{p_i}\right)<\frac{1}{\ln p_r}=\frac{1}{\ln^j p_r\cdot\ln^k p_r}. \tag{13}$$

It was found that if the last term on the RHS of Eq. 13 is multiplied by either $\ln^j p_r$ or $\ln^k p_r$, then the resultant value is greater than

$$\frac{1}{\ln p_r}$$

which is counterintuitive to our proof that the complementary prime number subsequences $\mathbb{P}'$ and $\mathbb{P}''$ add to form the complete set of prime numbers $\mathbb{P}$. Multiplying the last term of Eq. 13 by either $\ln^j p_r$ or $\ln^k p_r$ actually increases the cardinality of $\mathbb{P}'(x)$ or $\mathbb{P}''(x)$ to be greater than the cardinality of the entire set of prime numbers $\mathbb{P}(x)$ when that coefficient is multiplied by x on the RHS of 10. Hence our motivation to approach the solution from a different direction. In that light, it was found that if we let

$$\frac{1}{\ln^j p_r \cdot \ln^k p_r} = \frac{1}{\ln p_r} \cdot \left[j + k\right] \tag{14}$$

we can then subtract

$$\frac{j}{\ln p_r} \quad \text{or} \quad \frac{k}{\ln p_r}$$

from the LHS of Eq. 14 (or from the RHS of Eq. 13) and obtain the proper coefficient to multiply times x on the RHS of 11 to obtain the correct estimate of the quantity of $\mathbb{P}'(x)$ or $\mathbb{P}''(x)$ depending upon which of these quantities is subtracted from the RHS of Eq. 14. So the task at hand is to find a j and a k that will satisfy both sides of Eq. 14. To that end, it is seen in Eq. 14 that j and k must sum to unity on both sides of the equation to make this approach work. In order to do so, we recall the asymptotic densities that we derived earlier for $\pi'(x)$ and $\pi''(x)$ as

$$\pi'(x) \sim \frac{1}{\ln n + 1} \;\text{ and }\; \pi''(x) \sim \frac{1}{\ln n(\ln n + 1)}.$$

Now, since $j + k = 1$ must hold true to satisfy Eq. 14, we let

$$j = \frac{\frac{1}{\ln n(\ln n+1)}}{\frac{1}{\ln n}} = \frac{\ln n}{\ln n(\ln n + 1)}$$

and

$$k = \frac{\frac{1}{\ln n+1}}{\frac{1}{\ln n}} = \frac{\ln n}{\ln n + 1}$$

such that $j =$ the ratio of the asymptotic density of the prime subsequence $\mathbb{P}''$ divided by the asymptotic density of the set of all primes $\mathbb{P}$; and $k =$ the ratio of the asymptotic density of the prime subsequence $\mathbb{P}'$ to the asymptotic density of all primes $\mathbb{P}$. We now introduce a lemma:

Lemma 1. *For* $x > 1, j + k = 1$.

Proof.

$$\begin{aligned} j+k &= \frac{\ln x}{\ln x+1} + \frac{\ln x}{\ln x(\ln x+1)} \\ &= \frac{\ln x(\ln x+1) + \ln x \ln x(\ln x+1)}{\ln x(\ln x+1)^2} \\ &= \frac{\ln x(\ln x+1)(\ln x+1)}{\ln x(\ln x+1)^2} \\ &= 1. \end{aligned}$$

Since $j+k=1$ is valid in the lemma for all $x>1$, we have established that

$$\begin{aligned} \frac{1}{\ln^j p_r \cdot \ln^k p_r} &= \frac{1}{\ln p_r} \cdot \left[j+k \right] \\ &= \frac{1}{\ln p_r} \cdot \left[\frac{\ln p_r}{\ln p_r(\ln p_r+1)} + \frac{\ln p_r}{\ln p_r+1} \right]. \end{aligned}$$

Thus, in harmony with the lemma, we have

$$\pi''(x) \leq r'' + \frac{x}{\ln p_r(\ln p_r+1)} + 2^{r''} \tag{15}$$

where r'' = the number of $p'' \leq p_r$ and $2^{r''}$ = the maximum error resulting from the main term. We now proceed with our estimate of 15. We know that

$$r'' \leq \lfloor \frac{r}{2} \rfloor \quad \text{for} \quad p'' > 2$$

because it was proven that there are fewer primes in the subsequence $\mathbb{P}''$ than in the complementary prime subsequence $\mathbb{P}'$ (recall Eq. 4). Thus,

$$\begin{aligned} \pi''(x) &\leq r'' + \frac{x}{\ln p_r(\ln p_r+1)} + 2^{r''} && (15) \\ &< r'' + \frac{x}{\ln r(\ln r+1)} + 2^{r''} && (r < p^r) \\ &\leq \lfloor \frac{r}{2} \rfloor + \frac{x}{\ln r(\ln r+1)} + 2^{\lfloor \frac{r}{2} \rfloor} && \left(\lfloor \frac{r}{2} \rfloor \geq r'' \right) \\ &< \frac{x}{\ln r(\ln r+1)} + 2^{\lfloor \frac{r}{2} \rfloor + 1} && \left(2^{\lfloor \frac{r}{2} \rfloor} > \lfloor \frac{r}{2} \rfloor \right) \\ &< \frac{x}{\ln r(\ln r+1)} + 2^{\frac{r}{2}+1} && \left(\frac{r}{2} > \lfloor \frac{r}{2} \rfloor \right). \end{aligned}$$

Now, let

$$r = x^m \quad \text{such that} \quad m = \frac{1}{c \cdot \ln\ln x} \quad \text{for some positive constant } c.$$

We now have

$$\begin{aligned}
\pi''(x) &< \frac{x}{\ln x^m(\ln x^m + 1)} + 2^{\frac{1}{2}(x^m+2)} \qquad\qquad \left(m = \frac{1}{c \cdot \ln\ln x}\right) \\
&= \frac{x}{\left(\dfrac{\ln x}{c \cdot \ln\ln x}\right)^2 + \dfrac{\ln x}{c \cdot \ln\ln x}} + 2^{\frac{1}{2}(x^m+2)} \\
&= \frac{x}{\dfrac{c \cdot \ln\ln x \cdot \ln^2 x + (c \cdot \ln\ln x)^2 \cdot \ln x}{(c \cdot \ln\ln x)^3}} + 2^{\frac{1}{2}(x^m+2)} \\
&= x \cdot \frac{(c \cdot \ln\ln x)^2}{\ln^2 x + c \cdot \ln\ln x \cdot \ln x} + 2^{\frac{1}{2}(x^m+2)} \\
&< C \cdot x \cdot \frac{(\ln\ln x)^2}{\ln^2 x + c \cdot \ln\ln x \cdot \ln x} + 2^{\frac{1}{2}(x^m+2)}.
\end{aligned}$$

Since $\ln\ln x \cdot \ln x < c \cdot \ln\ln x \cdot \ln x$ for $c \geq 1$, and since $2^{\frac{1}{2}(x^m+2)} \ll$ than the main term when $c \geq 5$, we finally arrive at

$$\pi''(x) < C \cdot x \cdot \frac{(\ln\ln x)^2}{(\ln x)^2 + \ln\ln x \cdot \ln x}. \tag{16}$$

Thus, we see that for some positive constant $C < +\infty$, the sum of the reciprocals of the infinite subsequence of prime numbers $\mathbb{P}''$ converges, and this is confirmed when we compare 16 to the count of $p_2 \leq x$ (see 20).

6. Estimating $\pi'(x)$

In order to calculate an estimate of $\pi'(x)$, we invoke the lemma and begin with

$$\pi'(x) \leq r' + \frac{x}{\ln p_r + 1} + 2^{r'}. \tag{17}$$

Proceeding as before, we know that

$$r' \leq \lfloor r \rfloor \quad \text{for} \quad p' \geq 2$$

since there are fewer primes in the subsequence $\mathbb{P}'$ than in the set of all prime numbers $\mathbb{P}$. Thus,

$$\begin{aligned}
\pi'(x) &\leq r' + \frac{x}{\ln p_r + 1} + 2^{r'} && (17)\\
&< r' + \frac{x}{\ln r + 1} + 2^{r'} && (r < p^r)\\
&\leq \lfloor r \rfloor + \frac{x}{\ln r + 1} + 2^{\lfloor r \rfloor} && \left(\lfloor r \rfloor \geq r'\right)\\
&< \frac{x}{\ln r + 1} + 2^{\lfloor r \rfloor + 1} && \left(2^{\lfloor r \rfloor} > \lfloor r \rfloor\right)\\
&< \frac{x}{\ln r + 1} + 2^{r+1} && (r > \lfloor r \rfloor)\,.
\end{aligned}$$

Now, let

$$r = x^m \quad \text{such that} \quad m = \frac{1}{c \cdot \ln\ln x} \quad \text{for some positive constant } \ c.$$

We now have

$$\begin{aligned}
\pi'(x) &< \frac{x}{\ln x^m + 1} + 2^{x^m+1} && \left(m = \frac{1}{c \cdot \ln\ln x}\right)\\
&= \frac{x}{\dfrac{\ln x}{c \cdot \ln\ln x} + 1} + 2^{x^m+1}\\
&= \frac{x}{\dfrac{\ln x + c \cdot \ln\ln x}{c \cdot \ln\ln x}} + 2^{x^m+1}\\
&= x \cdot \frac{c \cdot \ln\ln x}{\ln x + c \cdot \ln\ln x} + 2^{x^m+1}\\
&< C \cdot x \cdot \frac{\ln\ln x}{\ln x + c \cdot \ln\ln x} + 2^{x^m+1}.
\end{aligned}$$

Since $\ln\ln x < c \cdot \ln\ln x$ for $c \geq 1$, and since $2^{x^m+1} \ll$ than the main term when $c \geq 5$, we finally arrive at

$$\pi'(x) < C \cdot x \cdot \frac{\ln\ln x}{\ln x + \ln\ln x}.$$

Since we've shown that $\mathbb{P}''$ is a small set in that the infinite sum of its reciprocals converges, and since it is known that the sum of the reciprocals of the set of all prime numbers $\mathbb{P}$ diverges, we can deduce from the relation

$$\mathbb{P} = \mathbb{P}' + \mathbb{P}''$$

that the prime number subsequence $\mathbb{P}'$ is a large set and that the infinite sum of its reciprocals diverges.

7. Estimating $\pi_2(x)$

We now take a look at how the twin prime count $\pi_2(x)$ can be estimated using the technique heretofore disclosed. If we assume that

$$\pi_2(x) \sim \frac{C}{\ln^2 x}$$

for some positive constant C [2], we can then model j found on the RHS of Eq. 14 as

$$j = \frac{\frac{C}{\ln^2 p_r}}{\frac{1}{\ln p_r}} = C \cdot \frac{1}{\ln p_r}$$

and we can model k found on the RHS of Eq. 14 as

$$k = 1 - C \cdot \frac{1}{\ln p_r}$$

such that $j =$ the ratio of the asymptotic density of the twin prime subsequence $\mathbb{P}_2$ divided by the asymptotic density of the set of all primes $\mathbb{P}$; and $k =$ the ratio of the asymptotic density of the set of remaining prime numbers $[\mathbb{P} - \mathbb{P}_2]$ to the asymptotic density of all primes $\mathbb{P}$ so that $j + k = 1$. We can then model $\pi_2(x)$ as

$$\begin{aligned}\pi_2(x) &\leq r_2 + x \cdot \frac{1}{\ln p_r} \cdot \left[C \cdot \frac{1}{\ln p_r}\right] + 2^{r_2} && (18)\\ &= r_2 + x \cdot C \cdot \frac{1}{\ln^2 p_r} + 2^{r_2} && (19)\end{aligned}$$

where $r_2 =$ the number of $p_2 \leq p_r$ and $2^{r_2} =$ the maximum error resulting from the main term. Similar to r'', we know that

$$r_2 \leq \lfloor \frac{r}{2} \rfloor \quad \text{for} \quad p_2 > 2$$

because there are fewer twin primes $\mathbb{P}_2$ than half the count of all prime numbers $\mathbb{P}$. Thus,

$$\begin{aligned}\pi_2(x) &\leq r_2 + x \cdot \frac{C}{\ln^2 p_r} + 2^{r_2} && (19)\\ &< r_2 + x \cdot \frac{C}{\ln^2 r} + 2^{r_2} && (r < p^r)\\ &\leq \lfloor \frac{r}{2} \rfloor + x \cdot \frac{C}{\ln^2 r} + 2^{\lfloor \frac{r}{2} \rfloor} && \left(\lfloor \frac{r}{2} \rfloor \geq r_2\right)\\ &< x \cdot \frac{C}{\ln^2 r} + 2^{\lfloor \frac{r}{2} \rfloor + 1} && \left(2^{\lfloor \frac{r}{2} \rfloor} > \lfloor \frac{r}{2} \rfloor\right)\end{aligned}$$

$$< x \cdot \frac{C}{\ln^2 r} + 2^{\frac{r}{2}+1} \qquad \left(\frac{r}{2} > \lfloor\frac{r}{2}\rfloor\right).$$

Now, let

$r = x^m$ such that $m = \dfrac{1}{c \cdot \ln\ln x}$ for some positive constant c. We now have

$$\begin{aligned}
\pi_2(x) &< x \cdot \frac{C}{\ln^2 x^m} + 2^{\frac{1}{2}(x^m+2)} \\
&= x \cdot \frac{C}{\left(\dfrac{\ln x}{c \cdot \ln\ln x}\right)^2} + 2^{\frac{1}{2}(x^m+2)} \\
&= x \cdot \frac{C \cdot (c \cdot \ln\ln x)^2}{(\ln x)^2} + 2^{\frac{1}{2}(x^m+2)} \\
&= C \cdot x \cdot \frac{(\ln\ln x)^2}{(\ln x)^2} + 2^{\frac{1}{2}(x^m+2)}.
\end{aligned}$$

And since $2^{\frac{1}{2}(x^m+2)} \ll$ than the main term for $c \geq 5$, we arrive at

$$\pi_2(x) < C \cdot x \cdot \frac{(\ln\ln x)^2}{(\ln x)^2}. \tag{20}$$

Thus, it is confirmed via this approach that for some positive constant $C < +\infty$, the sum of the reciprocals of the twin primes $\mathbb{P}_2$ converges. Further, when we compare the inequality 20 with the inequality for $\mathbb{P}''$ in 16, we see that the count of $p'' \leq x$, or $\pi''(x)$, is less than the count of twin primes $p_2 \leq x$, or $\pi_2(x)$.

8. Mathematica calculations

Mathematica [5] was programmed to calculate the sum of the reciprocals of $\mathbb{P}''$ and $\mathbb{P}_2$ for various ranges of x up to $10E6$, and a table of the computations appears below. Table 3 reveals that through the ranges of x calculated, the sum of the reciprocals of p'' is smaller than the sum of the reciprocals for the twin primes p_2, both of which converge at ∞.

Table 3: $p''(x)$ and $p_2(x)$ reciprocal sums

x	$\sum \frac{1}{p''(x)}$	$\sum \frac{1}{p_2(x)}$
1E02	0.534430	1.28989
1E03	0.606479	1.40995
1E04	0.644283	1.47370
1E05	0.668046	1.51443
1E06	0.683968	1.54268
2E06	0.687789	1.54950
3E06	0.689858	1.55321
4E06	0.691258	1.55573
5E06	0.692310	1.55763
6E06	0.693139	1.55915
7E06	0.693834	1.56040
8E06	0.694421	1.56148
9E06	0.694932	1.56240
10E6	0.695379	1.56322

9. Conclusion

In this paper, we applied the Inclusion-Exclusion Principle to the complementary prime number subsequences $\mathbb{P}'$ and $\mathbb{P}''$ to derive the respective prime counting functions $\pi'(x)$ and $\pi''(x)$ to determine whether these subsequences form a small set or a large set and thus whether the infinite sum of the inverse of their terms converges or diverges. In this study, we concluded that the sum of the reciprocals of the prime number subsequence $\mathbb{P}'$ diverges, similar to that for the set of all prime numbers $\mathbb{P}$, while the sum of the reciprocals of the prime number subsequences $\mathbb{P}''$ converges, similar to that for the set of all twin primes $\mathbb{P}_2$.

References

[1] KA Broughan and AR Barnett. On the Subsequence of Primes Having Prime Subscripts. *Journal of Integer Sequences*, 12(09.2.3), 2009.

[2] GH Hardy and JE Littlewood. Some Problems of 'Partitio Numerorum' III: On the Expression of a Number as a Sum of Primes. *Acta Math*, 44:1–70, 1923.

[3] OEIS Foundation Inc. Entry A262275 in The On-Line Encyclopedia of Integer Sequences. *http://oeis.org/A262275*, 2023.

[4] OEIS Foundation Inc. Entry A333242 in The On-Line Encyclopedia of Integer Sequences. *http://oeis.org/A333242*, 2023.

[5] Wolfram Research, Inc. Mathematica, Version 13.3. Champaign, IL, 2023.

[6] WJ Leveque. *Fundamentals of Number Theory.* Dover Publications, New York, New York, 1977.

[7] MP May. Properties of Higher-Order Prime Number Sequences. *Missouri Journal of Mathematical Sciences*, 32(2):158–170, 2020.

[8] J McKernan. Math 104B Lecture 5 Notes. *University of California-San Diego*, https://mathweb.ucsd.edu/~jmckerna/Teaching/17-18/Winter/104B/lectures.html, 2017.