SIEVING FOR THE PRIMES TO PROVE THEIR INFINITUDE

HUNDE EBA

Abstract. We prove the infinitude of prime numbers by the principle of contradiction, that is different from Euclid's proof in a way that it uses an explicit property of prime numbers. A sieve method that applies the inclusion-exclusion principle is used to give the property of the prime numbers in terms of the prime counting function.

1. Introduction

In number theory a prime number is a natural number that has exactly two divisors, 1 and itself. Though this definition of prime numbers is easy to comprehend, working with the prime numbers is one of the most difficult things in mathematics. Consequently there are more conjectures than there are theorems on related topics. In ancient Greece around 300 BC, Eratosthenes introduced the first sieve idea that generates primes up to a positive integer x knowing the list of primes up to $\sqrt{x}$. Legendre developed this same idea to a more mathematical explicit formula in the early 19th century. Most of the basic theorems on prime numbers were given by Euclid. From the following two theorems; although Euclid and others had an indirect contribution to the first theorem, Carl Friedrich Gauss was the first to develop the theorem as a systematic science in 1801 and the second theorem was stated by Euclid himself.

Notations.

- $\mathbb{W}$ is the set of nonnegative integers.
- $\mathbb{P}$ is the set of prime numbers.
- $\mathbb{Z}_+$ is the set of positive integers.

Theorem 1.1 (Fundamental Theorem of Arithmetic). *For $n \in \mathbb{Z}_+$, $\epsilon_i \in \mathbb{W}$, and $p_i \in \mathbb{P}$*

$$n = p_1^{\epsilon_1} p_2^{\epsilon_2} p_3^{\epsilon_3} \cdots p_n^{\epsilon_n} \tag{1.1}$$

and every n has a unique factorization.

Research supported by Addis Ababa Science and Technology University.

Proof. The proof can be given by either mathematical induction or contradiction. See [3, pp. 20–21] and [2, p. 15]. □

Theorem 1.2. *There are infinitely many prime numbers.*

Proof. Suppose p_n is the largest element of the finite set of prime numbers and $q = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot p_n + 1$; that is, one more than the product of all the positive integers from 1 through p_n. The integer q is larger than p_n and is not divisible by any positive integer from 2 through p_n. Thus, the positive divisors of q other than 1 must be greater than p_n, which is a contradiction. Therefore, there must be an infinitely many prime numbers. □

Definition 1.3. *Floor function:* $\lfloor x \rfloor$ *is the greatest integer less than or equal to* x*.*

Definition 1.4. *Möbius function is denoted by* $\mu(n)$*, where* $n \in \mathbb{Z}_+$ *and is*

$$\mu(n) = \begin{cases} 1, & if\, n = 1 \\ 0, & if\, p^2 | n\ for\ some\ prime\ p \\ (-1)^r, & if\, n = p_1 p_2 \cdots p_r,\ where\ p_i\ are\ distinct\ primes. \end{cases}$$

Definition 1.5. *Prime counting function:* $\pi(x)$ *is the function counting the number of primes less than or equal to some real number* x*.*

2. An Identity on Sets

Identity (2.1), proved in this section, is an original result. Hence we have used this identity to derive other identities and well-known theorems. The interested reader can try to obtain the results listed in the Appendix from (2.1).

Lemma 2.1. *Let* s_i *be the ith number defined by the set* S*, which is either a finite or infinite set,* $n \in \mathbb{Z}_+$*,*

$$S = \{s_1, s_2, s_3, \ldots, s_n\}.$$

Then

$$1 = s_1 + \sum_{i=1}^{n-1} s_{i+1} \prod_{j=1}^{i} (1 - s_j) + \prod_{i=1}^{n} (1 - s_i). \tag{2.1}$$

Proof. We prove (2.1) by induction on n.
Base case: (2.1) is true for $n = 1$.
Inductive hypothesis: Suppose (2.1) is true for $n = k$.

$$1 = s_1 + \sum_{i=1}^{k-1} s_{i+1} \prod_{j=1}^{i} (1 - s_j) + \prod_{i=1}^{k} (1 - s_i). \tag{2.2}$$

Inductive step: From the inductive hypothesis we intend to imply (2.1) is also true for $n = k+1$. Separating the sum of products in (2.2) into two and combining one of them to the product term, we have

$$1 = s_1 + \sum_{i=1}^{k} s_{i+1} \prod_{j=1}^{i} (1 - s_j) - s_{k+1} \prod_{j=1}^{k} (1 - s_j) + \prod_{i=1}^{k} (1 - s_i);$$

$$1 = s_1 + \sum_{i=1}^{k} s_{i+1} \prod_{j=1}^{i} (1 - s_j) + \prod_{i=1}^{k+1} (1 - s_i). \tag{2.3}$$

Thus, identity (2.1) holds for $n = k+1$.
Conclusion: By the principle of mathematical induction, (2.1) is true for all $n \in \mathbb{Z}_+$. □

Corollary 2.2. *Let x be any number and $p_i \in \mathbb{P}$. Then*

$$x = \frac{x}{p_1} + \sum_{i=1}^{n-1} \frac{x}{p_{i+1}} \prod_{j=1}^{i} \left(1 - \frac{1}{p_j}\right) + x \prod_{i=1}^{n} \left(1 - \frac{1}{p_i}\right). \tag{2.4}$$

Proof. Let R be the set of all reciprocals of primes; $R = \left\{\frac{1}{p_1}, \frac{1}{p_2}, \frac{1}{p_3}, \ldots, \frac{1}{p_n}\right\}$, letting $R = S$ and multiplying x to both sides of equation (2.1), from Lemma 2.1, (2.4) directly follows. □

3. Sieving for the Prime Numbers

We sieve for the primes by the use of inclusion-exclusion principle.

Lemma 3.1 (Inclusion-Exclusion Principle)**.** *Let A_i be the ith set from the sets: $A_1, A_2, \cdots, A_n$ for $n \in \mathbb{Z}_+$ and $|A_i|$ be the cardinality of the set A_i. Then*

$$\left|A_1 \cup A_2 \cup \cdots \cup A_n\right| = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \left|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right|$$

or

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{J \neq \emptyset, J \subseteq \{1,2,\cdots,n\}} (-1)^{|J|-1} \left|\bigcap_{i \in J} A_i\right| \tag{3.1}$$

Proof. The proof uses the principle of mathematical induction on n and is shown in [6] and [1]. Equation (3.1) is known as the inclusion-exclusion principle. □

Proposition 3.2. *Considering all the assumptions taken in Lemma 3.1, let* $A = \{1, 2, 3, \ldots, x\}$, $P = \{p_1, p_2, \ldots, p_n\}$ *for* P *the set of prime numbers up to* x, $x \in \mathbb{Z}_+$, $p_n \leq x$, *and* $A_i = \{a \in A : a \bmod p_i = 0\}$, *where* $p_i \in P$ *and* $A_i \subset A \subseteq \mathbb{Z}_+$. *Then*

$$\frac{|A|}{\prod_{i\in J} p_i} - \frac{|A| \bmod \prod_{i\in J} p_i}{\prod_{i\in J} p_i} = \left|\bigcap_{i\in J} A_i\right| \tag{3.2}$$

where $J \neq \emptyset, J \subseteq \{1, 2, \ldots, n\}$

Proof. From the assumptions taken in the proposition we can directly imply: $A_i = \{p_i, 2p_i, \ldots, |A_i|p_i\}$, $0 \leq x - |A_i|p_i < p_i$ and $|A| = x$. Knowing that $\bigcap_{i\in J} A_i = \{a \in A : a \bmod \prod_{i\in J} p_i = 0\}$, we can yet imply

$$\bigcap_{i\in J} A_i = \left\{ \prod_{i\in J} p_i,\, 2\prod_{i\in J} p_i, \ldots, \left|\bigcap_{i\in J} A_i\right| \prod_{i\in J} p_i \right\} \quad \text{and}$$

$$0 \leq x - \left|\bigcap_{i\in J} A_i\right| \prod_{i\in J} p_i < \prod_{i\in J} p_i \quad \text{where} \quad J \neq \emptyset, J \subseteq \{1, 2, \ldots, m\}.$$

For $B = \{b, 2b, \ldots, |B|b\}$, $B \subset A$, $b \in A$, and $0 \leq x - |B|b < b$, from a simple division rule we have

$$\frac{x}{b} - \frac{x \bmod b}{b} = |B|. \tag{3.3}$$

If $B \equiv \bigcap_{i\in J} A_i$ then $b = \prod_{i\in J} p_i$. Substituting this condition into (3.3) we have (3.2), which completes the proof. □

Identities (3.4) and (3.9) are well-known to many mathematicians. The result (3.4) appears as an exercise or observation in several number theory books and is an application of Hardy and Wright's Theorem 268 (taking the constant function $F(x) = 1$ so that $G(x) = \lfloor x \rfloor$ [3, p. 237, formula (3.9)], in a different form, is due to Legendre [4]. But the way we prove them is new and uses an interesting application of the inclusion-exclusion principle.

Lemma 3.3. *Let* x *be any number greater than or equal to* 1, $\lfloor x \rfloor$ *be the floor function and* $\mu(j)$ *be the Möbius function. Then*

$$\sum_{n=1}^{\infty} \left\lfloor \frac{x}{j} \right\rfloor \mu(j) = 1. \tag{3.4}$$

Proof. Combining (3.1) and (3.2), we have

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{J\neq\emptyset, J\subseteq\{1,2,\ldots,n\}} (-1)^{|J|-1} \left(\frac{|A|}{\prod_{i\in J} p_i} - \frac{|A| \bmod \prod_{i\in J} p_i}{\prod_{i\in J} p_i} \right). \tag{3.5}$$

It can be seen that $A = \{1\} \cup \bigcup_{i=1}^{n} A_i$, $|A| = x \Rightarrow \left|\bigcup_{i=1}^{n} A_i\right| = x - 1$ and

$$\frac{|A|}{\prod_{i\in J} p_i} - \frac{|A| \bmod \prod_{i\in J} p_i}{\prod_{i\in J} p_i} = \left\lfloor \frac{x}{\prod_{i\in J} p_i} \right\rfloor, \tag{3.6}$$

where $J \neq \emptyset, J \subseteq \{1, 2, \ldots, n\}$
Substituting (3.6) into (3.5) we have

$$x - 1 = \sum_{J\neq\emptyset, J\subseteq\{1,2,\ldots,n\}} (-1)^{|J|-1} \left\lfloor \frac{x}{\prod_{i\in J} p_i} \right\rfloor. \tag{3.7}$$

As J runs through all subsets, $\prod_{i\in J} p_i$ runs through all integers 2 through x as well as some additional larger integers. For $x < \prod_{i\in J} p_i$ we have $\left\lfloor \frac{x}{\prod_{i\in J} p_i} \right\rfloor = 0$ and hence, we can view $\prod_{i\in J} p_i$ as $\prod_{i\in J} p_i = \{2, 3, 4, \ldots, x\} \cup \{a \in \prod_{i\in J} p_i : a > x\}$ to deduce

$$x - 1 = \sum_{1<j\leq x} -\mu(j) \left\lfloor \frac{x}{j} \right\rfloor. \tag{3.8}$$

By rearranging (3.8) we have $\sum\limits_{j\leq x} \left\lfloor \frac{x}{j} \right\rfloor \mu(j) = 1$. Considering the fact that $\left\lfloor \frac{x}{j} \right\rfloor$ is always zero for $j > x$ we get the result (3.4). □

Lemma 3.4. *Let $\pi(x)$ be the prime counting function, $p_m \leq \sqrt{x}$ for $m \in \mathbb{Z}_+$ so that $m = \pi(\sqrt{x})$ and $\{y\}$ be the fractional part of y for any number y. Also consider H; $H \neq \emptyset, H \subseteq \{1, 2, \ldots, m\}$. Then we have*

$$\pi(x) = \pi(\sqrt{x}) + x \prod_{p\leq\sqrt{x}} \left(1 - \frac{1}{p}\right) + \sum_{H} (-1)^{|H|-1} \left\{ \frac{x}{\prod_{i\in H} p_i} \right\} - 1. \tag{3.9}$$

Proof. $\frac{|A| \bmod \prod_{i\in J} p_i}{\prod_{i\in J} p_i}$ is the fractional part of $\frac{x}{\prod_{i\in J} p_i}$ and hence,

$$\frac{|A| \bmod \prod_{i\in J} p_i}{\prod_{i\in J} p_i} = \left\{ \frac{x}{\prod_{i\in J} p_i} \right\}. \tag{3.10}$$

Substituting (3.10) in to (3.7), we have

$$x - 1 = \sum_{J\neq\emptyset, J\subseteq\{1,2,\ldots,n\}} (-1)^{|J|-1} \left(\frac{x}{\prod_{i\in J} p_i} - \left\{ \frac{x}{\prod_{i\in J} p_i} \right\} \right). \tag{3.11}$$

Consider the nonempty sets H, K, and M to be used as indexes, for which all are subsets of index J. Let $H \subseteq \{1, 2, \ldots, m\}$, $K \subseteq \{m+1, m+2, \ldots, n\}$ and let $L \subseteq H$ (including empty set). And also M is a set of all possible combinations of L and K, so $M = (L \cup K)$. Consequently, the sum $\sum_M (-1)^{|M|-1} \left(\frac{x}{\prod_{i\in M} p_i} - \left\{ \frac{x}{\prod_{i\in M} p_i} \right\} \right) = \sum_{\sqrt{x}<p\le x} 1$, means the positive integers $\le x$ that are divisible by p_j, $j \in K$ are only the prime numbers found between $\sqrt{x}$ and x. Since $J = H \cup M$, we have $\sum_J f(J) = \sum_H f(H) + \sum_M f(M)$, where $f(J)$ is the summand found in (3.11). Hence, considering these facts to (3.11), we have

$$x - 1 = \sum_H (-1)^{|H|-1} \left(\frac{x}{\prod_{i\in H} p_i} - \left\{ \frac{x}{\prod_{i\in H} p_i} \right\} \right) + \sum_{\sqrt{x}<p\le x} 1. \tag{3.12}$$

From (3.12) we can say that the sum (over the primes less than or equal to x) of numbers of positive integers less than or equal to x that are divisible by the ith prime but not by the preceding primes is exactly equal to $x-1$ and the sum (over the primes between $\sqrt{x}$ and x) of numbers of positive integers that are divisible by ith prime but not by the preceding primes is exactly equal to the numbers of primes between $\sqrt{x}$ and x.
It can be seen that, $\sum_{\sqrt{x}<p\le x} 1 = \pi(x) - \pi(\sqrt{x})$ and

$$\begin{aligned}
\sum_H (-1)^{|H|-1} \frac{x}{\prod_{i\in H} p_i} &= \frac{x}{p_1} + \left(\frac{x}{p_2} - \frac{x}{p_1 p_2} \right) \\
&\quad + \left(\frac{x}{p_3} - \frac{x}{p_1 p_3} - \frac{x}{p_2 p_3} + \frac{x}{p_1 p_2 p_3} \right) \\
&\quad + \cdots + \left(\frac{x}{p_m} - \frac{x}{p_1 p_m} - \frac{x}{p_2 p_m} + \cdots + \frac{(-1)^{m-1} x}{p_1 p_2 \cdots p_m} \right) \\
&= \frac{x}{p_1} + \sum_{i=1}^{m-1} \frac{x}{p_{i+1}} \prod_{j=1}^{i} \left(1 - \frac{1}{p_j} \right).
\end{aligned} \tag{3.13}$$

Combining (2.4) from Corollary 2.2 and (3.13), we obtain

$$\sum_H (-1)^{|H|-1} \frac{x}{\prod_{i\in H} p_i} = x - x \prod_{i=1}^{m} \left(1 - \frac{1}{p_i} \right). \tag{3.14}$$

Combining (3.12) with (3.14), we obtain the result

$$\pi(x) = \pi(\sqrt{x}) + x \prod_{p\le\sqrt{x}} \left(1 - \frac{1}{p} \right) + \sum_H (-1)^{|H|-1} \left\{ \frac{x}{\prod_{i\in H} p_i} \right\} - 1.$$

□

4. An Alternative Proof of the Infinitude of Prime Numbers

Though there are many proofs of the infinitude of prime numbers in the literature, the way we prove it here is new and different in that it uses the explicit formula from the previous section which is gained by the application of the inclusion-exclusion principle.

Theorem 4.1. *Let P be the set of prime numbers; $P = \{p_1, p_2, \ldots, p_n\}$. Then there are infinitely many elements of P (there are infinitely many prime numbers).*

Proof. Suppose there are finitely many prime numbers so that n is a finite number. From Lemma 3.4 we have

$$\pi(x) = \pi(\sqrt{x}) + x \prod_{p \le \sqrt{x}} \left(1 - \frac{1}{p}\right) + \sum_{\substack{H \neq \emptyset \\ H \subseteq \{1,2,\cdots,m\}}} (-1)^{|H|-1} \left\{ \frac{x}{\prod_{i \in H} p_i} \right\} - 1.$$

If there are only finitely many primes, then the number of terms in the last sum is bounded by 2^m and each term in the sum is bounded in absolute value by 1, so the last sum is bounded. Also we know $\pi(x)$ and $\pi(\sqrt{x})$ are bounded. But the term $x \prod_{p \le \sqrt{x}} \left(1 - \frac{1}{p}\right)$ is not bounded, because the product is constant for large x. This is a contradiction to our premise that there are finitely many prime numbers. Hence, there must be infinitely many prime numbers. □

5. Appendix: Some Identities Derived From Identity (2.1)

- Geometric series formula: $\sum\limits_{i=1}^{n} r^i = \frac{r - r^{n+1}}{1-r}$.
- An identity that R. Apéry used in his famous proof of irrationality of $\zeta(3)$: $\sum\limits_{i=1}^{n} \frac{a_1 a_2 \cdots a_{i-1}}{(x+a_1)(x+a_2)\cdots(x+a_i)} = \frac{1}{x} - \frac{a_1 a_2 \cdots a_n}{x(x+a_1)(x+a_2)\cdots(x+a_n)}$ (see [5]).
- The factorial: $n! - \frac{1}{n!} = \sum\limits_{i=1}^{n-1} i \left(i! + \frac{1}{(i+1)!}\right)$.
- The limit: $e^x - 1 = \lim\limits_{n\to\infty} \frac{1}{n} \sum\limits_{i=0}^{n} \left(1 + \frac{1}{n}\right)^{xi}$.
- Identity on Riemann Zeta function; for $p_i \in \mathbb{P}$ and $\Re s > 1$:
  $\zeta(s) = \frac{p_1^s}{p_1^s - 1} + \sum\limits_{i=1}^{\infty} \frac{1}{p_{i+1}^s - 1} \prod\limits_{j=1}^{i} \frac{p_j^s}{p_j^s - 1} = \left(\frac{p_1^s - 1}{p_1^s} - \sum\limits_{i=1}^{\infty} \frac{1}{p_{i+1}^s} \prod\limits_{j=1}^{i} \frac{p_j^s - 1}{p_j^s} \right)^{-1}$

and possibly other identities could be derived.

References

[1] M. Balazs and B. Toth, *Inclusion-exclusion principle*, Oct. 2014. `http://people.maths.bris.ac.uk/~mb13434/incl_excl_n.pdf`.

[2] W. E. Clark, revised by J. Hefferon, *Elementary Number Theory*, St Michael's College, 2003. `http://joshua.smcvt.edu/numbertheory/book.pdf`.

[3] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Fourth Edition, Oxford University Press, 1975.

[4] D. H. Lehmer, *On the exact number of primes less than a given limit*, Illinois J. Math., **3** (1959), 381–388. `https://projecteuclid.org/euclid.ijm/1255455259`.

[5] A. van der Poorten, *A proof that Euler missed, An informal report*, Math. Intelligencer, **1.4** (1978/79), 195–203. `http://pracownicy.uksw.edu.pl/mwolf/Poorten_MI_195_0.pdf`.

[6] S. Weiss, *The inclusion exclusion principle and its more general version*, June 28, 2009. `http://compsci.hunter.cuny.edu/~sweiss/resources/inclusion_exclusion.pdf`.

MSC2010: 11A41, 11N35

Key words and phrases: prime numbers, sieves, inclusion-exclusion principle

Department of Electro-Mechanical Engineering, Addis Ababa Science and Technology University, Addis Ababa, Ethiopia
E-mail address: hunde.eba@aastu.edu.et