A Study of Primes

Amolak Singh

Spring 2017

Abstract

For the purposes of graduation, this paper was written for *MATH 4997W: Senior Project* under the guidance of Professor Paul Garrett. This paper discusses important prime number theory, such as some facts, theorems, and relevant proofs about properties of prime numbers. Also included is a brief discussion of the Riemann Hypothesis and the relevance of primes in nature.

Contents

1 Introduction

Definition 1. *A prime number, n, is a natural number with exactly two positive divisors:* 1 *and* n.

The study of primes throughout the past two millenia seems to follow a certain pattern: mathematicians seek to order and understand these special natural numbers, which often seem to follow no order or structure. With different applications such as RSA ciphers in cryptography and playing an important role in number theory, prime numbers and their properties have been the subject of extensive inquiry since the times of the ancient Greeks. Seemingly random, they have confounded mathematicians in their attempts to understand the properties of primes [Rockmore 2005].

Interested in studying "numbers for their mystical and numerological properties," mathematicians from Pythagoras' school made several discoveries about primes [O'Connor & Robertson 2009]. These results culminated in Euclid's *Elements*, in which he proved two significant theorems: the Fundamental Theorem of Arithmetic and the infinitude of primes [Weisstein 2017].

2 Some Prime Properties

To understand primes, we must investigate their relationship to other numbers and understand their basic properties. First, we will state and prove three propositions necessary to prove the Fundamental Theorem of Arithmetic, give a proof of the Fundamental Theorem of Arithmetic, and give a relatively simple argument about the infinitude of prime numbers.

First, we will prove the division algorithm by showing that the quotient and remainder not only exist, but are also unique [Garrett 2016].

Proposition 2.1. *Let m be a nonzero integer and let N be any integer. Then, there exists unique integers q and r such that $0 \leq r \leq |m|$ and $N = qm + r$.*

Proof. Fix N and m. Let S be the set of all non-negative integers that can be written as $N - sm$ for some $s \in \mathbb{Z}$. By the well-ordering property, S has a least element $r = N - qm$.

We will first show that $0 \leq r < |m|$. Suppose, for contradiction's sake, that $r \geq |m|$. It follows that $r - |m| \geq 0$, and so $r - |m| = n - (q+1)|m| \geq 0$ implying $r - |m| \in S$. However, $r - |m| \leq r$ contradicting our statement

above that r is the smallest element of S, Thus, $r < |m|$. Now, suppose $m > 0$. Then, $N = q|m| + r = qm + r$.

Now, let us show that q and r are unique. Let

$$N = qm + r = q^{'}m + r'$$

such that $0 \leq r \leq m$ and $0 \leq r^{'} \leq m$. Without loss of generality, suppose that $r \leq r^{'}$. Then, $(q - q^{'})m = r^{'} - r \geq 0$. If $r^{'} - r \neq 0$, then $q - q^{'} \neq 0$ and we have

$$r^{'} - r = |r^{'} - r| = |q - q^{'}| * |m| \geq 1|m|.$$

From earlier, we know that $r^{'} - r \leq r^{'} < |m|$ and so

$$|m| \leq r^{'} - r < |m|,$$

implying, by contradiction, that $r^{'} = r$. Therefore, $(q - q^{'})m = r^{'} - r = 0$ and $q^{'} = q$ [Garrett 2016]. □

Now, we will show the existence of the greatest common divisor, represented as $gcd()$. For example, the $gcd(m, n)$ is the smallest positive integer of the form $xm + yn$ such that $x, y \in \mathbb{Z}$ [Garrett 2016].

Proposition 2.2. *Let $m, n \in \mathbb{Z} - \{0\}$. Among all common divisors of m, n there exists a unique $d > 0$ such that for every other common divisor e of m, n we have $e|d$. This d is the greatest common divisor of m, n, such that $d = gcd(m, n)$.*

Proof. Let $D = x_0m + y_0n$ be the least positive integer expressible in the form $xm+yn$. Also let $m = m^{'}d$ and $n = n^{'}d$ with $n^{'}, m^{'} \in \mathbb{Z}$. First, we will show that any divisor d of both m and n divides D. We have the following equalities:

$$D = x_0m + y_0n = x_0(m^{'}d) + yo(n^{'}d) = (x_0m^{'} + y_0n^{'}) * d,$$

implying d divides D.

Now, let $m = qD + r$ with $0 \leq r < D$. We have the following equalities:

$$0 \leq r = m - qD = m - q(x_0m + y_0n) = (1 - qx_0)m + (-y_0)n.$$

So, $r = m + yn$ where $x = 1 - qx_0$ and $y = -y_0$. It follows that since $r < D$, and D is the smallest positive integer, $r = 0$. Thus, $D|m$ and $D|n$ [Garrett 2007]. □

Proposition 2.3. *Let p be prime and let $x, y \in \mathbb{N}$. If p divides ab, then p divides a or p divides b.*

Proof. If p divides a, then we are done. Else, p does not divide a and consequently $gcd(p, a) = 1$, since p is prime. We know that $gcd(ab, pb) = b$. Since p divides pb and, p divides ab, it follows that p divides $gcd(pb, ab) = b * gcd(p, a) = b * 1 = b$ [Stein 2017]. □

These propositions play a key role in the proof of the Fundamental Theorem of Arithmetic, showing that not only does every natural number have a prime factorization, but in fact that the prime factorization is unique. The Fundamental Theorem of Arithmetic plays a key role in the Euclidean Algorithm and finding greatest common denominators which is of utmost importance to several ciphers in cryptography [Apostol 1976].

Theorem 2.1. (Fundamental Theorem of Arithmetic) *A number n, for each $n \in \mathbb{N}$, can be written as a product of primes uniquely, barring rearrangement.*

Proof. First, we will prove that any natural can be written as a product of primes. Suppose, for contradiction's sake, that a given natural number $a > 1$ is not a product of finitely many primes. Then, by the well-ordering principle, there exists a smallest a that cannot be written as a product of primes. Since a is composite by our assumption, we can factor a such that $a = x * y$, where $x, y \in \mathbb{N}$, and $1 < x, y < a$.

It follows that we can write x and y as products of prime factors because a is the smallest positive integer than cannot be written as a product of primes. However, $a = x * y$, so a must be a product of prime numbers, contradicting our assumption.

Now, we will show that this factorization is unique. Suppose, for contradiction's sake, that a given natural number $a > 1$ is the smallest number that does not have a unique prime factorization such that

$$a = p_1 \cdots p_m = q_1 \cdots q_n,$$

where p_i and q_j are the prime factors. Since q_1 and $q_2 \cdots q_n$ are less than a, both q_1 and $q_2 \cdots q_n$ must have unique prime factorizations. By Proposition 2.3, either p_1 divides q_1, or p_1 divides $q_2 \cdots q_n$. Therefore, $p_1 = q_j, 1 < j < n$.

Let a' be a smaller integer with p_1 and q_j removed from the initial equivalence such that $a' = p_2 \cdots p_m = q_1 \cdots q_{n-1}$. We now have a smaller integer factorizable in two non-unique ways, contradicting our initial assumption. Consequently, a number such as a cannot exist, and all natural numbers must have a unique prime factorization [Coppel 2009]. □

Euclid's proof of the infinitude of prime numbers is quite simple; yet, it is often considered one of the most elegant proofs in mathematics. This proof has allowed mathematicians to address different questions about primes, such as their distribution.

Theorem 2.2. *There exist infinitely many prime numbers.*

Proof. Suppose that $p_1, p_2, \ldots, p_n$ are n distinct primes. We will construct a prime p_{n+1} not equal to any of $p_1, \ldots, p_n$. Let a number

$$N = p_1 * p_2 * \cdots * p_n + 1.$$

Then, we can factor N such that

$$N = q_1 * q_2 * \cdots * q_m$$

with each q_i being prime and $m > 1$. If $q_1 = p_i$ for $1 \leq i \leq n$, then p_i must divide N. Additionally, p_i must divide $N - 1$.

So, p_i must divide $1 = N - (N - 1)$, which contradicts our assumption about p_i. Thus, we have constructed our new prime such that $p_{n+1} = q_1$ and $p_{n+1} \notin \{p_1, p_2, \ldots, p_n\}$ [Stein 2017]. □

Centuries later, Leonard Euler strengthened the argument of the primes' infinitude by proving that the series of the reciprocal of primes diverges.

Proposition 2.4. *The infinite series* $\sum_{n=1}^{\infty} \frac{1}{p_n} = 1$, *where* p_i *are primes, diverges.*

To prove this proposition, Euler used his product formula

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{p_n} = \prod_{p} \frac{1}{1 - p^{-}s} \tag{1}$$

For $x \geq 2$,

$$\sum_{n \leq x} \frac{1}{n} < \prod_{p \leq x} \frac{1}{1 - p^{-s}} < \prod_{p \leq x} e^{\frac{2}{p}} = exp\Big\{\sum_{p \leq x} \frac{1}{p}\Big\}.$$

By the divergence of the harmonic series, he concluded that the series of the reciprocal primes must converge [Bateman & Diamond 2004]. Similarly, Euler showed that since $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ has an infinite limit, there must be infinitely many primes.

3 The Distribution of Primes

While Euclid's theorem proves to us that there are infinitely many primes, it does not tell us how big an infinity [Caldwell 2017]. Looking at just a few primes, there seems to be no obvious pattern. However, upon closer inspection, the seemingly random and irregular distribution of prime numbers lends itself to a specific pattern.

One way to understand the distribution of primes is to ask the following question: how many primes are less than a given natural number? It is quite clear that there is no reasonable closed form formula that can answer this question. However, it is reasonable to ask whether there exists some relatively simple formula which can *approximate* the number of primes less than a given natural number [Jameson 2003].

In the late eighteenth century, Legendre and Gauss independently began investigating this exact question; while they were unable to prove anything rigorously, a century later, Hadamard and de la Valle-Poussin proved an important theorem about the distribution of primes, as described below [Garrett 2016].

Theorem 3.1. Prime Number Theorem *The number of primes, represented by the notation $\pi(x)$, not exceeding x is asymptotic to x divided by $log(x)$ such that*

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

I will not give a proof of the Prime Number Theorem in this paper. However, the intuition behind this theorem is worth understanding. In 1798, Legendre conjectured that

$$\pi(x) \sim \frac{x}{log(x) - c}$$

where he calculated c to equal 1.08366, based on a table of values up to $\pi(40000)$ [Caldwell 2017]. However, choosing $c = 1$ will give the better approximation of $\pi(x)$ [Garrett 2017].

Gauss proposed an even better estimator of $\pi(x)$:

$$li(x) = \int_2^x \frac{1}{log(t)}\, dt$$

that actually slightly overestimates $\pi(x)$ [Jameson 2003]. See Table 1 below for a comparison.

It is quite clear from Table 1 that all three approximations are very good. Legendre's conjecture is quite accurate when $c = 1$, but it does

x	$\pi(x)$	$\frac{x}{log(x)}$	$\frac{x}{log(x)-1}$	$li(x)$
1000	168	145	169	177
10000	1229	1086	1218	1246
50000	5133	4621	5092	5166
100000	9592	8686	9512	9630
500000	41538	38103	41246	41607
1000000	78498	72382	78030	78628
10000000	664579	620420	661459	664918

Table 1: Comparing estimates of $\pi(x)$

underestimate $\pi(x)$. In fact, it has been shown that $\pi(x) > x/log(x)$ for $x \geq 17$ [Coppel 2009]. It is known that $li(x)$ is a better estimate. Additionally, while it may seem like $\pi(x) < li(x)\ \forall x$, it is known that $\pi(x) > li(x)$ for some $x \in [1.398201 * 10^{316}, 1.398244 * 10^{316}]$ [Coppel 2009].

There are even better estimates for $\pi(x)$. Riemann used his zeta function to give an exact formula for $\pi(x)$ by summing over the zeta function's nontrivial zeros, increasing from the smallest zero [Rockmore 2005]. His formula is as follows:

$$\pi(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} li(x^{1/n}) - \sum_{n=1}^{\infty} \frac{\mu(n)}{n} li(x^{\rho/n}) +$$

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} \int_{x^{1/n}}^{\infty} \frac{dt}{(t^2-1)t\ log(t)} - log2 \sum_{n=1}^{\infty} \frac{\mu(n)}{n}$$

where $\mu(n)$ is the Möbius function, ρ runs over the nontrivial zeros of the zeta function and $li(x)$ is the logarithmic integral as discussed earlier [Caldwell 2017]. The first term of the above function, the Riemann function $R(x)$ approximates the number of primes, while the remaining terms remove the errors that $R(x)$ would have produced on its own by tracking the difference between the exact number of primes and asymptotic estimate of the Prime Number Theorem [du Savoy 2003, Rockmore 2005].

3.1 Proof of Euler's Product Formula

We will give a non-rigorous proof connecting Euler's product formula to the Riemann Zeta function. This proof is unique due to its simplicity and use of basic algebra to connect the zeta function with prime numbers.

Proof. We will start with the following two equations:

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + \dots \tag{2}$$

$$\frac{1}{2^s} * \zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + \frac{1}{6^s} + \frac{1}{10^s} + \dots \tag{3}$$

Subtracting (3) from (2), we can eliminate all terms divisible by 2:

$$\left(1 - \frac{1}{2^s}\right) * \zeta(s) = 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{9^s} + \frac{1}{11^s} + \dots \tag{4}$$

We then repeat the process for the next term as follows:

$$\frac{1}{3^s} * \left(1 - \frac{1}{2^s}\right) * \zeta(s) = 1 + \frac{1}{3^s} + \frac{1}{9^s} + \frac{1}{15^s} + \frac{1}{21^s} + \frac{1}{27^s} + \dots \tag{5}$$

Subtracting as before, we get:

$$\left(1 - \frac{1}{3^s}\right) * \left(1 - \frac{1}{2^s}\right) * \zeta(s) = 1 + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{11^s} + \frac{1}{13^s} + \frac{1}{17^s} + \dots \tag{6}$$

Now, we have removed all the elements divisible by two and three. We can repeat this sieving process infinitely, until we get the following equation:

$$\cdots \left(1 - \frac{1}{11^s}\right) * \left(1 - \frac{1}{7^s}\right) * \left(1 - \frac{1}{5^s}\right) * \left(1 - \frac{1}{3^s}\right) * \left(1 - \frac{1}{2^s}\right) * \zeta(s) = 1 \tag{7}$$

Dividing (7) by all of the factors except $\zeta(s)$, we get:

$$\zeta(s) = \frac{1}{\cdots \left(1 - \frac{1}{11^s}\right) * \left(1 - \frac{1}{7^s}\right) * \left(1 - \frac{1}{5^s}\right) * \left(1 - \frac{1}{3^s}\right) * \left(1 - \frac{1}{2^s}\right)} \tag{8}$$

Thus, in its general form, $\zeta(s)$ can be written as

$$\prod_p \frac{1}{1 - p^- s}$$

[Dittrich 2016]. □

4 The Riemann Hypothesis

The most famous (and arguably the most important) unsolved problem in mathematics is probably the Riemann Hypothesis [Stein 2017]. In the mid-nineteenth century, Riemann conjectured that all nontrivial solutions of $\zeta(s) = 0$ lie on a certain straight line: $Re(s) = 1/2$ [Goodman & Weisstein 2017]. He also showed that the distribution of primes is related to the location of the zeros of zeta function. First, he noticed that the trivial solutions were symmetric to $Re(s) = 1/2$, and after calculating a few nontrivial solutions, he found them to be lie directly on $Re(s) = 1/2$.

In 1900, David Hilbert proclaimed the problem of proving or disproving the Riemann hypothesis as one of the single most important problems for mathematicians to solve in the twentieth century [Apostol 1976]. To this day it remains unsolved. Of note, though, is that Riemann's conjecture has been shown to hold true for the first ten trillion non-trivial zeros of the zeta function [Gourdon 2004].

5 A Biological Relevance of Primes

Throughout this paper, we have tried to understand the theory behind prime numbers as developed over the past two millenia. However, it is also worth understanding some applications. The applications of prime numbers are quite extensive, ranging from public key cryptography to random number generation to even biology. While many sources will talk about the computing applications of prime numbers, some occurrences of prime numbers in nature are particularly interesting.

A famous example is the life cycle of two species of cicada, *Magicicada septendecim* and *Magicicada tredecim*, who, as their species names imply have life cycles of seventeen and thirteen years, respectively [du Savoy 2003]. Cicadas spend all but one year of their life underground, after which they become adults, emerge from the ground, sing, mate and die. This cycle continues every thirteen or seventeen years depending on the species.

This brings us to the following question: why are the lengths of their life cycles prime numbers of years? Scientists have proposed a few possible evolutionary reasons. The first is that the two species survive on similar resources in similar regions; with prime number life cycles, they will only emerge simultaneously once every 221 years (since $lcm(13, 17) = 221$) thus maximizing the natural resources available to each other [du Savoy 2003].

Contrarily, if their life cycles were only slightly different, say twelve and

sixteen years, they would both emerge after 48, 96, 144, and 192 years in the same period of time.

In a similar vein, the second possible reason is that, early on in their evolutionary history, a certain fungus developed that preyed on the cicadas. To avoid the deadly effects of the fungus, the cicadas developed a life cycle that would minimize the number of times they would interact with the fungus thus improving the survival of their species.

6 Acknowledgment

This paper was written as my undergraduate senior project at the University of Minnesota, Twin Cities, thanks to the invaluable guidance of Professor Paul Garrett, both in terms of mathematical content and the writing itself. I would also like to thank Professor Adil Ali for introducing me to number theoretic concepts through his cryptology and number theory course and recommending me to write this paper under Professor Garrett's guidance.

7 References

[Bateman & Diamond 2004] P. Bateman & H. Diamond, *Analytic Number Theory: An Introductory Course*, New Jersey: World Scientific (2004).

[Caldwell 2017] C. Caldwell, *The Prime Pages: Prime Number Research, Records, and Resources*, `https://primes.utm.edu/`, retrieved 6 Feb 2016.

[Coppel 2006] W. Coppel, *Number theory: An Introduction to Mathematics*, (2nd ed.), New York: Springer (2009).

[Dittrich 2016] W. Dittrich, *On Riemann's Paper, "On the Number of Primes Less Than a Given Magnitude"*, `https://arxiv.org/pdf/1609.02301.pdf`, retrieved 11 Feb 2016.

[Du Sautoy 2003] M. Du Sautoy, *The Music of the Primes: Searching to Solve the Greatest Mystery in Mathematics*, (1st ed.), New York: HarperCollins (2003).

[Garrett 2007] P. Garrett, *Abstract Algebra*, `http://www-users.math.umn.edu/~garrett/m/algebra/Whole_with_TOC.pdf`, retrieved 11 Mar 2016.

[Garrett 2016] P. Garrett, *Cryptology and Number Theory (MATH 5248)*, University of Minnesota, Alpha Print (2016).

[Goodman & Weisstein 2017] L. Goodman & E. Weisstein, *Riemann Hypothesis*, MathWorld: A Wolfram Web Resource, `http://mathworld.wolfram.com/RiemannHypothesis.html`, retrieved 6 Feb 2016.

[Gourdon 2004] X. Gourdon, *The* 10^{13} *First Zeros of the Riemann Zeta Function & Zeros Computation at Very Large Height*, `http://numbers.computation.free.fr/Constants/Miscellaneous/zetazeros1e13-1e24.pdf`, retrieved 1 Feb 2016.

[Jameson 2003] G. Jameson, it The Prime Number Theorem. London Mathematical Society Student Texts **34**, Cambridge: Cambridge University Press (2003).

[O'Connor & Robertson 2009] J. O'Connor & E. Robertson, *Prime Numbers*, `http://www-groups.dcs.st-and.ac.uk/history/HistTopics/Prime_numbers.html`, retrieved 6 Feb 2016.

[Rockmore 2005] D. Rockmore, *Stalking the Riemann Hypothesis: The Quest to Find the Hidden Law of Prime Numbers*, (1st ed.), New York: Pantheon Books (2005).

[Stein 2009] W. Stein, *Elementary Number Theory: Primes, Congruences, and Secrets*, Undergraduate Texts in Mathematics, London: Springer (2009).

[Weisstein 2017] E. Weisstein, *Fundamental Theorem of Arithmetic*, MathWorld: A Wolfram Web Resource, `http://mathworld.wolfram.com/FundamentalTheoremofArithmetic.html`, retrieved 11 Feb 2016.