Towards The Riemman Zeta Hypothesis

Tariq Rashid

Contents

1. Introduction

sssss

Approach

no mathematical proof not a texbook

but a journey to show the magic of primes - and seeing some of the beuty and surprising reults from proofs and some anlaysis

proptitise understanding and intuition over textbook rgiour

not terse, not ultra concie - but elabroate and repeat and exxplain

Why – beacuse i found it hard, an it doesn't need to be

a tour guide

Why Primes

ss so simp,e a chile can underatand them

taight in school nut not made aware of their mystery and power eg rncrption

Riemann Hypothesis millenium challenge - millin dollars

resists anlaysis

mystery

no simple foruma

Millenium Problem

sss

mysterious...

Praesent pulvinar, nisl quis interdum efficitur, risus metus convallis eros, quis congue elit sapien non nunc. Nam bibendum bibendum nunc, quis sagittis augue tincidunt consectetur. Curabitur fringilla at nibh sit amet auctor. Maecenas sit amet orci venenatis, mattis enim non, mollis massa. Quisque orci velit, auctor at neque molestie, vestibulum convallis mi. Sed rhoncus metus elit, in tincidunt mi pellentesque non. Fusce nec turpis nec neque posuere iaculis in nec sapien. Aenean quis lectus mauris. Etiam commodo maximus est, id molestie nulla hendrerit a. Proin fermentum fermentum velit, sollicitudin accumsan nulla porta ac. Nulla vitae felis at metus volutpat commodo ut at nunc. Ut in dictum leo. Ut imperdiet quis elit et accumsan. Integer eget neque vehicula, suscipit ligula rhoncus, consectetur risus. Suspendisse bibendum purus lectus, nec rhoncus erat hendrerit id.

Part I.

Exploring Prime Numbers

2. What Are Prime Numbers?

Let's start by looking at the most ordinary numbers we know, the **counting numbers**.

$$1, 2, 3, 4, 5, 6, 7, 8, \ldots$$

We became familiar with these numbers when we were just toddlers, counting apples in a bowl, for example.

Multiplication

We soon learned to add and multiply these numbers. Many of us learned our times tables by heart. Almost without thinking we could recite multiplications like $2 \times 4 = 8$, and $5 \times 5 = 25$.

When we multiply 3 by 4, the answer is 12. This 12 is called a **product**, and the 3 and 4 are called **factors**.

If we pick any two numbers a and b and multiply them, the result is another number, which we can call c.

$$a \times b = c$$

Because a and b are whole numbers, so is c.

An Innocent Question

Those factors a and b can be any counting number we feel like choosing. Does this freedom apply to c as well?

Surely some combination of a and b can give us any number c that we desire. Let's try a couple of examples.

- If we want c to be 12, we could choose $a = 3$ and $b = 4$. We could have chosen $a = 2$ and $b = 6$, and that would work too.
- If we want c to be 100, we could choose $a = 2$ and $b = 50$. Another combination that works is $a = 10$ and $b = 10$.

What if we want c to be 7?

If we try for a short while, we'll find there doesn't seem to be a combination of factors a and b that gives 7 as a product. In fact, if we try all the numbers in the range $2 \ldots 6$ we'll see for ourselves there really is no combination that gives $a \times b = 7$.

What if we want c to be 11? Again we'll find no combination of whole number factors gives 11 as a product.

So the answer to our innocent-looking question is no, c can't be any whole number.

Numbers like 7 and 11 that don't have whole number factors, are called **prime numbers**. Here are the first few.

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, \ldots$$

In short, if we multiply two counting numbers, the answer is never a prime number.

What About 1?

You might have spotted that when we were trying to find factors of 7 we didn't consider combinations like $a = 1$ and $b = 7$. That's because we exclude 1 as a legitimate factor. Why? Because every number has 1 as a factor, and that's not particularly interesting.

If we didn't exclude 1, there would be no prime numbers because every number c would have factors $a = 1$ and $b = c$.

Even worse, a number could have lots of factors as 1, which is also rather unhelpful. The number 12 could have an infinite number of factors.

$$12 = 4 \times 3 \times 1 \times 1 \times 1 \times 1 \times 1 \times 1 \times \dots$$

Negative Numbers?

Prime numbers were known about and discussed in ancient times, well before the idea of a negative number was accepted.

Over the hundreds of years since then, new ideas and insights were developed about prime numbers, and they were built on the original assumption that prime numbers could only be **positive** whole numbers.

Today almost all exploration of prime numbers continues under the same constraint that products, factors and primes are positive whole numbers greater than 1. This constraint really doesn't limit the mysteries and surprises that prime numbers hold.

Apparent Randomness

Looking back at the list of prime numbers, there doesn't seem to be a pattern to them. Apart from never being even numbers, with the exception of 2, they seem to be fairly randomly located along the number line.

For hundreds of years, mathematicians puzzled over the primes, attacking them with all sorts of exotic tools, trying to crack them open to reveal any elusive rules that govern their location. That endeavour continues to this day.

3. How Many Primes Are There?

At first thought it might seem obvious that there is an unending supply of prime numbers.

If we think a little longer, a bit of doubt might intrude on our certainty. A small number like 6 has factors 2 and 3. Every multiple of 2 is not a prime number, every multiple of 3 is not a prime number, every multiple of 4 is not a prime number, and so on. All these multiples are reducing the probability that a large number is prime.

We might be tempted to think that eventually prime numbers just fizzle out. Instead of relying on intuition, let's decide the matter with rigorous mathematical proof.

Proof There Are Infinitely Many Primes

A proof is not an intuition, nor is it a set of convincing examples. A proof is a watertight logical argument that leads to a conclusion we can't argue with.

The proof that there is no limit to the number of primes is ancient and rather elegant, due to Euclid around 300 BC, and a nice one to have as our first example.

Let's start by assuming the number of primes is not endless but finite. If there are n primes, we can list them.

$$p_1, p_2, p_3, p_4 \dots p_n$$

We can create a new number x by multiplying all these primes together.

$$x = p_1 \cdot p_2 \cdot p_3 \cdot p_4 \cdot \ldots \cdot p_n$$

This x is clearly not a prime number. It's full of factors like p_1 , p_3 and p_n.

Let's make another number y in the same way, but this time we'll also add 1.

$$y = p_1 \cdot p_2 \cdot p_3 \cdot p_4 \cdot \ldots \cdot p_n + 1$$

Now y could be a prime number, or it could not be a prime number. These are the only two options for any positive whole number.

If y is prime then we have a problem because we've just found a new prime number which isn't part of the original finite set $p_1, p_2 \dots p_n$. How do we know it's not part of the original set? Well y is bigger than any of the primes in the list because we created it by multiplying them all together, and adding 1 for good measure.

So perhaps y is not a prime. In this case, it must have factors. And the factors must be one or more of the known primes $p_1, p_2 \dots p_n$. That means y can be divided by one of those primes p_i exactly, leaving no remainder. Let's write this out.

$$\frac{y}{p_i} = \frac{p_1 \cdot p_2 \cdot p_3 \cdot p_4 \cdot \ldots \cdot p_n}{p_i} + \frac{1}{p_i}$$

The first part divides neatly without a remainder because p_i is one of the primes $p_1, p_2 \dots p_n$. The second part doesn't divide neatly at all.

That means y can't be divided by any of known primes. Which again suggests it is a new prime, not in the original list.

Both of these options point to the original list of primes being incomplete.

And that's the proof. No finite list of primes can be a complete list of primes. So there are infinitely many primes.

A Common Misunderstanding

It is easy to think that $p_1 \cdot p_2 \cdot p_3 \cdot p_4 \cdot \ldots \cdot p_n + 1$ is a way of generating prime numbers. This is not correct. The proof only asks what the consequences are if $p_1 \cdot p_2 \cdot p_3 \cdot p_4 \cdot \ldots \cdot p_n + 1$ is prime, under the assumption that we have a limited list of primes $p_1, p_2, p_3, p_4 \ldots p_n$.

We can prove that $p_1 \cdot p_2 \cdot p_3 \cdot p_4 \cdot \ldots \cdot p_n + 1$ is not always prime by finding just one counter-example. If we use prime numbers 2, 3, 5, 7, 11 and 13, we can see that $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 30031$ which is not prime because $30031 = 59 \cdot 509$.

4. Primes Are The Building Blocks Of Numbers

We saw earlier that positive whole numbers have factors if they're not a prime number. Let's explore this a little further.

Breaking A Number Into Its Factors

Let's think about the number 12 and its factors. We can think of two combinations straight away.

$$12 = 2 \times 6$$
$$12 = 3 \times 4$$

Looking again at those factors we can see that 6 itself can be broken down into smaller factors 3 and 2. That 4 can also be broken down into factors 2 and 2.

$$12 = 2 \times (3 \times 2)$$
$$12 = 3 \times (2 \times 2)$$

We can't break these smaller factors down any further, which means they're prime numbers. Both combinations now look very similar. If we put those factors in order of size, we can see they are in fact exactly the same.

$$12 = 2 \times 2 \times 3$$
$$12 = 2 \times 2 \times 3$$

Perhaps every number can be broken down into a list of prime factors that is unique to that number, much like DNA is unique to people. Let's prove it.

Fundamental Theorem Of Arithmetic

We'll split this proof into two steps.

- First we'll show that any positive whole number can be broken down into a list of factors that are all prime.
- Second we'll show this list of primes is unique to that number.

Let's imagine a number N and write it out as a product of factors.

$$N = f_1 \cdot f_2 \cdot f_3 \cdot \ldots \cdot f_m$$

We can look at each of these factors f_i in turn. If a factor is not prime, we can break it down into smaller factors. For example, the factor f_1 might be broken down as $f_1 = g_1 \cdot g_2$. If a factor is prime, $f_2 = p_1$ for example, we leave it because we can't break it into smaller factors.

$$N = (g_1 \cdot g_2) \cdot p_1 \cdot (g_3 \cdot g_4 \cdot g_5) \cdot \ldots \cdot (g_x \cdot g_y)$$

If we keep repeating this process, all the factors will eventually be prime. How can we be so sure? Well, if any number in the list isn't prime, we can apply the process again, breaking that number down into smaller factors. The only thing that stops us applying the process again is when all the factors are eventually prime.

Figure 4.1 shows an example of this iterative process applied to the number 720.

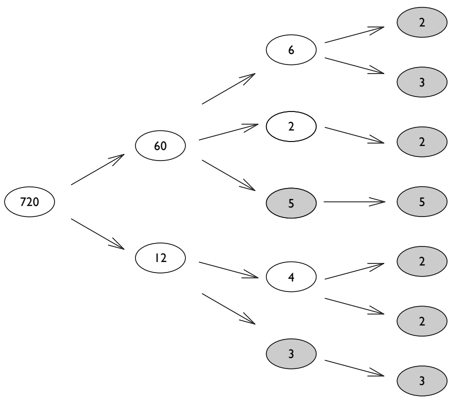


Figure 4.1.: Breaking 720 into factors until only primes remain.

We can now write N as a product of these primes.

$$N = p_2 \cdot p_3 \cdot p_1 \cdot p_5 \cdot p_4 \cdot p_6 \cdot p_7 \cdot \ldots \cdot p_n$$

These primes won't necessarily be in order of size. They may also repeat, for example p_1 might be the same as p_7. It doesn't matter. We've shown that any positive whole number can be written as a product of primes.

Let's now show that this list of primes is unique to that number N. For the moment, imagine this isn't true and a number N can be written as a product of two different lists of primes.

$$N = p_1 \cdot p_2 \cdot p_3 \cdot \ldots \cdot p_a$$
$$N = q_1 \cdot q_2 \cdot q_3 \cdot \ldots \cdot q_a \cdot q_b \cdot q_c \cdot q_d$$

These primes are not necessarily in order of size, and some might be repeated, so p_2 could be the same as p_4. Again, we won't let that bother us. To keep our argument general, we'll assume that the number of primes in the second list, d, is larger than the number of primes in the first list, a.

Now, we can see that p_1 is a factor of N. That means it must also be a factor of the second list. That means p_1 is one of the factors q_i. Because we we didn't assume any order in these primes, let's say it is q_1. That means we can divide both lists by $p_1 = q_1$.

$$\cancel{p_1} \cdot p_2 \cdot p_3 \cdot \ldots \cdot p_a = \cancel{q_1} \cdot q_2 \cdot q_3 \cdot \ldots \cdot q_a \cdot q_b \cdot q_c \cdot q_d$$

We can apply the same logic again. The first list has a factor p_2 which means it must also be a factor of the second list. We can say that $p_2 = q_2$, and divide both lists by this factor.

$$\cancel{p_1} \cdot \cancel{p_2} \cdot p_3 \cdot \ldots \cdot p_a = \cancel{q_1} \cdot \cancel{q_2} \cdot q_3 \cdot \ldots \cdot q_a \cdot q_b \cdot q_c \cdot q_d$$

We can keep doing this until all the factors in the first list have been matched up with factors in the second list. It doesn't matter if a prime repeats, for example if p_1 is the same as p_3, the factors will still be matched correctly, in this case $p_1 = q_1$ and $p_3 = q_3$.

$$\cancel{p_1} \cdot \cancel{p_2} \cdot \cancel{p_3} \cdot \ldots \cdot \cancel{p_a} = \cancel{q_1} \cdot \cancel{q_2} \cdot \cancel{q_3} \cdot \ldots \cdot \cancel{q_a} \cdot q_b \cdot q_c \cdot q_d$$

Let's simplify the algebra.

$$1 = q_b \cdot q_c \cdot q_d$$

What we've just shown is that if a number N can be written as two separate lists of prime factors, their factors can be paired up as being equal, and if any are left over, they must equal 1. That is, the two lists are identical.

We've shown that any whole number N can be decomposed into a list of prime factors, and this list of primes is unique to that number. This is rather profound, and is called the Fundamental Theorem of Arithmetic.

5. Primes Are Rather Elusive

If we listed all the counting numbers $1, 2, 3, 4, 5, \ldots$, excluded 1, and then crossed out all the multiples of $2, 3, 4, \ldots$ we'd be left with the primes. This sieving process emphasises that primes are defined more by what they are not, than by what they are.

If there was a simple pattern in the primes, we'd be able to encode it into a simple formula for generating them. For example, the triangle numbers 1,3,6,10,15,... can be generated by the simple expression $\frac{1}{2}n(n + 1)$. The prime numbers, however, have resisted attempts by mathematicians over hundreds of years to find precise and simple patterns in them.

One of the first questions anyone enthusiastic about prime numbers asks is whether a polynomial can generate the n^{th} prime. Polynomials are both simple and rather flexible, and it would be quite pleasing if one could generate primes.

Let's prove that prime numbers are so elusive that no simple polynomial in n can generate the n^{th} prime.

No Simple Polynomial Generates Only Primes

A **polynomial** in n has the following general form, simple yet flexible.

$$P(n) = a + bn + cn^2 + dn^3 + \ldots + \alpha n^\beta$$

By simple polynomial we mean the coefficients $a, b, c \ldots \alpha$ are whole numbers. Let's also say that $b, c, d, \ldots \alpha$ are not all zero. This way we exclude trivial polynomials like $P(n) = 7$ that only generate a single value no matter what n is.

Let's start our proof by assuming there is indeed a $P(n)$ that generates only primes, given a counting number n. When $n = 1$, it generates a prime, which we can call p_1.

$$p_1 = P(1) = a + b + c + d + \ldots + \alpha$$

Now let's try $n = (1 + p_1)$.

$$P(1 + p_1) = a + b(1 + p_1) + c(1 + p_1)^2 + d(1 + p_1)^3 + \ldots$$

That looks complicated, but all we need to notice is that if we expand out all the terms, we'll have two kinds, those with p_1 as a factor, and those without. We can collect together all those terms with factor p_1 and call them $p_1 \cdot X$.

$$P(1 + p_1) = (a + b + c + d + e + \ldots \alpha) + p_1 \cdot X$$

We then notice that $(a + b + c + d + e + \ldots \alpha)$ is actually p_1.

$$\begin{aligned} P(1 + p_1) &= p_1 + p_1 \cdot X \\ &= p_1(1 + X) \end{aligned}$$

Since X is a whole number, this is divisible by p_1. It shouldn't be because $P(1+p_1)$ is supposed to be a prime. This contradiction means the starting assumption that there is a simple polynomial $P(n)$ that generates only primes is wrong.

We've actually proved a stronger statement than we intended. We intended to prove that there is no simple polynomial $P(n)$ that generates the n^{th} prime. We ended up proving that no simple polynomial $P(n)$ can generate only primes.

Polynomials With Rational Coefficients

Insisting on integer coefficients for polynomials might seem overly restrictive. Let's broaden our definition to allow **rational** coefficients of the form $\frac{s}{t}$ where s and t are integers.

We again assume $P(n)$ does indeed generate only primes, and so $p_1 = P(1)$ is prime. This time we'll consider $n = (1 + k \cdot p_1)$.

$$\begin{aligned} P(1 + k \cdot p_1) &= a + b(1 + k \cdot p_1) + c(1 + k \cdot p_1)^2 + \dots \\ &= p_1 + k \cdot p_1 \cdot X \\ &= p_1(1 + k \cdot X) \end{aligned}$$

Here X contains terms that are combinations of the rational coefficients $a, b, c \dots \alpha$ multiplied together. We can choose a k which cancels all the denominators of the rational coefficients leaving $k \cdot X$ as an integer. The lowest common multiple of all the denominators is one way to do this.

Our proof by contradiction then continues as before because we've found an example of $P(n)$ that is not prime.

The primes really are rather elusive if even polynomials with rational coefficients can't generate only primes.

6. Primes Aren't That Spread Out

We've seen there is no limit to the supply of primes. A good question to ask next is how frequently they occur.

One way to explore this is by looking at the sum of their inverses, or **reciprocals**.

Infinite Sum Of Reciprocals

The counting numbers $1, 2, 3, 4, \ldots$ are spaced 1 apart. The sum of their inverses is called the **harmonic series**.

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \ldots$$

This series is known to diverge, that is, the sum is infinitely large. Appendix A has an easy short proof.

The square numbers $1, 4, 9, 16, \ldots$ are spaced further apart than the counting numbers.

$$1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \ldots$$

The sum of their inverses converges. Appendix B walks through Euler's historic and rather adventurous proof showing it converges to $\frac{\pi^2}{6}$.

We can interpret this to mean the squares n^2 are so spread out that the terms in the series become small quickly enough for the sum not to become infinitely large.

It is natural to then ask the same question about the primes. Are they so spread out that the infinite sum of their inverses converges too?

Infinite Sum Of Prime Reciprocals

Let's start by assuming the infinite series of prime reciprocals does in fact converge to a finite sum S.

$$S = \sum_{n=1} \frac{1}{p_n}$$

Because S is finite, and each term is smaller than the previous one, there must be a value of k such that the infinite series after $\frac{1}{p_k}$ sums to less than 1. We can call this sum x.

$$x = \sum_{n=k+1} \frac{1}{p_n} < 1$$

Let's build an infinite geometric series based on this x.

$$G = x + x^2 + x^3 + x^4 + \dots$$

This new series G converges because the ratio between terms x is less than 1.

Let's think a little more carefully about the terms in G. Any term in G will be of the form $\frac{1}{N}$ where N has prime factors p_{k+1} or larger. This is because x was intentionally constructed with primes p_{k+1} and larger.

Now consider a second series F where, in contrast to G, the terms are constructed from all the primes p_k and smaller.

$$F = \sum_{j=1} \frac{1}{1 + j \cdot (p_1 \cdot p_2 \cdot p_3 \ldots p_k)}$$

Between each term, only j changes. Now let's look more closely at the expression $1 + j \cdot (p_1 \cdot p_2 \cdot p_2 \ldots p_k)$. This has no prime factors from the range p_1 to p_k. Since all whole numbers have prime factors, its prime factors must be from the set p_{k+1} and larger.

That means F is a subseries of G. That is, the terms of F appear in the terms of G.

Now, if we compare the terms of F to the harmonic series, we can test whether F diverges.

We do this with the limit comparison test, which tests what happens to the ratio of terms from each series as they extend to infinity. If the ratio is finite, the series either both converge, or both diverge.

$$\lim_{j \to \infty} \frac{1 + j \cdot (p_1 \cdot p_2 \cdot p_3 \ldots p_k)}{j} = p_1 \cdot p_2 \cdot p_3 \ldots p_k$$

The ratio is finite, and since the harmonic series diverges, so does F.

Since F diverges, and is a subseries of G, then G must also diverge. But we constructed G to converge. This contradiction proves the initial assumption that the infinite series of prime reciprocals converges was wrong.

That $\sum 1/p_n$ diverges is a little surprising because our intuition was that primes thin out rather rapidly.

Legendre's Conjecture

The fact that $\sum 1/n^2$ converges suggests the primes are not as sparse as the squares. This leads us to an interesting proposal attributed to Legendre, but actually first published by Desboves in 1855, that there is at least one prime number between two consecutive squares.

$$n^2 < p < (n+1)^2$$

This remains a deep mystery of mathematics. Nobody has been able to prove or disprove it.

7. Distribution Of Primes

Given primes are so resistant to encoding into a simple generating formula, let's take a detour and try a different approach, **experimental mathematics**.

Number of Primes Up To A Number

We showed that primes don't run out as we explore larger and larger numbers. We also showed they don't thin out as quickly as the squares. So how quickly do they thin out?

One way to explore this is to keep a count of the number of primes as we progress along the whole numbers.

The expression $\pi(n)$ has become an abbreviation for 'the number of primes up to, and including, n'. For example, $\pi(5) = 3$ because there are 3 primes up to, and including, 5. The next number 6 is not prime, so $\pi(6)$ remains 3. The use of the symbol π can be confusing at first.

Figure 7.1 shows $\pi(n)$ for n up to 100. A fairly smooth curve seems to be emerging. This is slightly unexpected because the primes appear to be randomly placed amongst the numbers. The curve suggests the primes are governed by some kind of constraint. It wouldn't be too adventurous to say the curve looks logarithmic, like $\ln(n)$.

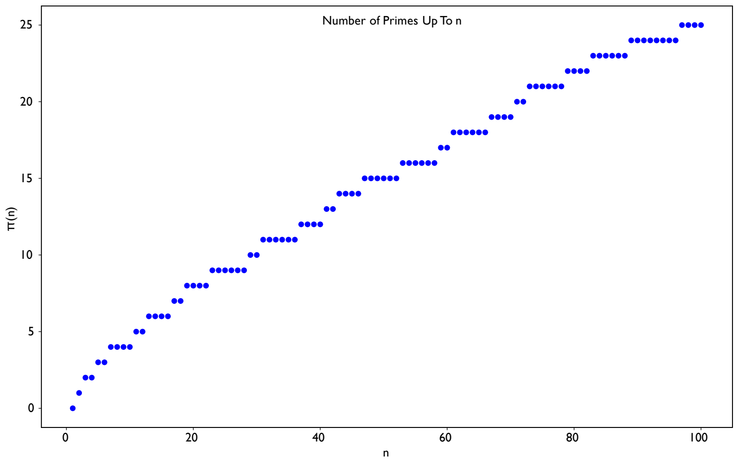


Figure 7.1.: $\pi(n)$ for n from 1 to 100.

Rather Good Approximations for $\pi(n)$

Gauss, one the most prolific mathematicians in history, was the first to find an expression that approximates $\pi(n)$ fairly well. He was aged about 15 at the time.

$$\pi(n) \approx \frac{n}{\ln(n)}$$

The expression is surprisingly simple. It is worth pondering on what hidden pattern in the primes is captured by the natural logarithm $\ln(n)$.

Just a year later, Gauss developed a different expression that approximates $\pi(n)$ even more closely.

$$\pi(n) \approx \int_0^n \frac{1}{\ln(x)} dx$$

At first glance, this logarithmic integral function, shortened to $\mathrm{li}(n)$, appears to be a continuous form of the first approximation.

Figure 7.2 shows a comparison of these approximations with the actual $\pi(n)$ for n all the way up to 10,000. It's clear the logarithmic integral function is much closer to the actual prime counts.

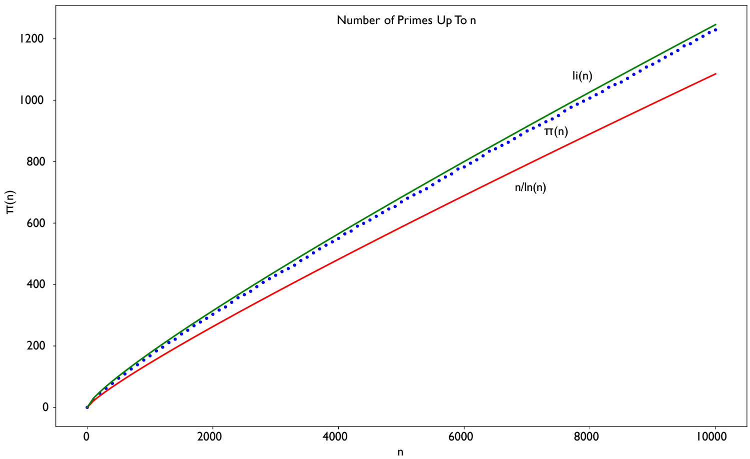


Figure 7.2.: Comparing $\mathrm{li}(n)$ and $n/\ln(n)$ with $\pi(n)$.

Proportional Error

Looking again at the previous chart, the prime counting approximation $n/\ln(n)$ appears to be diverging away from the true prime count $\pi(n)$ as n gets larger. That is, the error appears to be getting ever larger. If we looked at the numbers, we'd also see $\mathrm{li}(n)$ diverging away from

$\pi(n)$ too. Does this mean the approximations become useless as n gets larger?

Figure 7.3 paints a different picture. It shows the error as a proportion of $\pi(n)$. We can see this proportional error becomes smaller as n grows to 10,000. It's also clear that $\text{li}(n)$ has a distinctly smaller proportional error than $n/\ln(n)$.

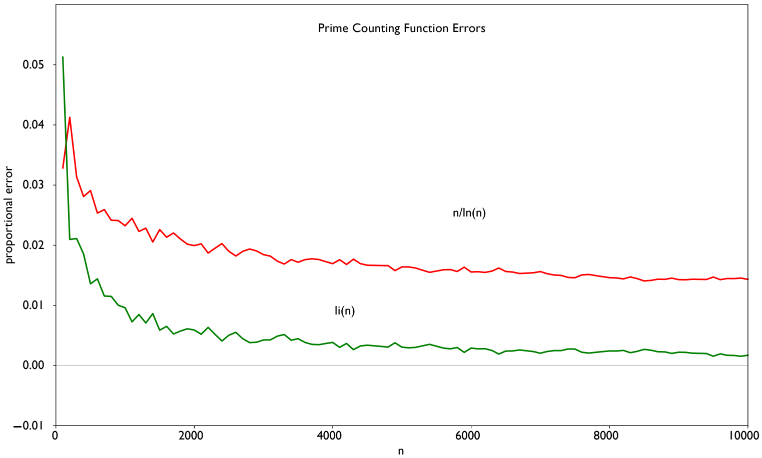


Figure 7.3.: Proportional errors for $\text{li}(n)$ and $n/\ln(n)$.

There are 1229 primes amongst the first 10,000 whole numbers. The logarithmic integral gives us $\text{li}(10,000) = 1246$. The error is just 17, and as a proportion of 1229, an impressively small 0.0138.

If we extended n to even larger values, we'd find the proportional error would fall further towards zero. Perhaps these approximations are correct in the limit $n \to \infty$?

Prime Density

Let's look again at those approximations and see if we can interpret their form. The following compares Gauss' first approximation with a general expression for calculating the mass of a volume of stuff with a given average density.

$$\text{mass} = \text{density} \times \text{volume}$$

$$\pi(n) \approx \frac{1}{\ln(n)} \times n$$

The comparison suggests that $1/\ln(n)$ is the average density of primes. If true, this would be a remarkable insight into the primes.

We can apply a similar analogy to Gauss' second approximation too. This time we compare it with another general expression for calculating mass where the density is not assumed to be constant throughout its volume.

$$\text{mass} = \int (\text{density}) dv$$

$$\pi(n) \approx \int_0^n \frac{1}{\ln(x)} dx$$

Again, $1/\ln(x)$ emerges as a more locally accurate density of primes around a number x.

It was this density of primes around a number that the young Gauss first noticed as he studied the number of primes in successive ranges of whole numbers, 1-1000, 1001-2000, 2001-3000, and so on.

Nth Prime?

If the density of primes is $1/\ln(n)$ then we can say average distance between primes is $\ln(n)$. We can then make the short leap to say the n^{th} prime is approximately $n\ln(n)$.

Before we get too excited about having found a simple function for generating the n^{th} prime, this expression is an approximation, based on an average, itself based on Gauss' approximation for $\pi(n)$.

It's still interesting to see how well this expression for the n^{th} prime performs. Figure 7.4 shows the error in $n\ln(n)$ as a proportion of the actual n^{th} prime. As n increases to 10,000,000, the proportional error falls to approximately 0.1017.

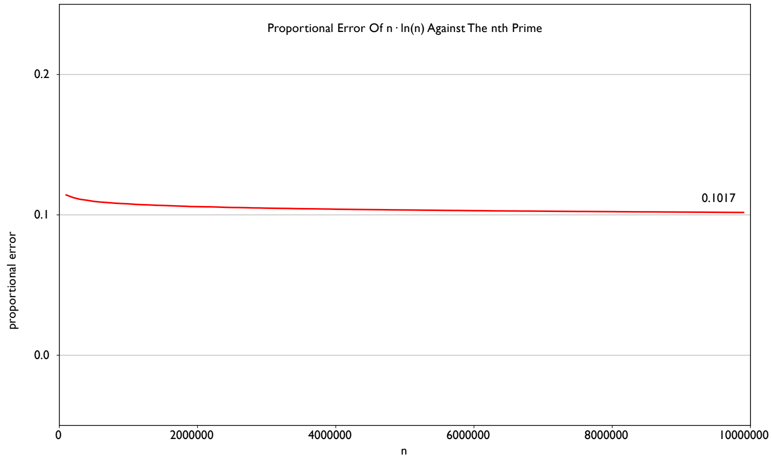


Figure 7.4.: Proportional error of $n\ln(n)$ against the n^{th} prime.

Looking at the graph, it is tempting to conclude the proportional error in $n\ln(n)$ approaches 0.1 as $n \to \infty$. We should be cautious because the error could be falling to zero very very slowly.

Anecdotal evidence is not mathematical proof, no matter how compelling it looks.

For example, our experiments show $\mathrm{li}(n)$ is always a bit higher than the true $\pi(n)$. If we extended n from 10,000 to 10,000,000 we'd still find $\mathrm{li}(n)$ was always higher. But in 1914 Littlewood proved that $\mathrm{li}(n)$ can become lower than $\pi(n)$ infinitely many times. More recent proofs show this starts to happen somewhere between 10^{19} and a mind-blowingly large 10^{316}.

Imperfect History

The question of who first developed an approximation for $\pi(n)$ is not perfectly clear. Gauss didn't always publish his work, leaving us to reconstruct history from notes and letters.

In his 1797 book on number theory, Legendre first published a form $n/(A\ln(n) + B)$, which he updated in his 1808 second edition to $n/(\ln(n) - 1.08366)$.

However, in 1849 Gauss wrote a letter to astronomer, and former student, Encke telling him that he had, in '1792 or 1793', developed the logarithmic integral approximation, which he wrote as $\int \frac{dn}{\mathrm{logn}}$. His collected works also reveal that in 1791 he had written about the simpler approximation, $\frac{a}{la}$ as he wrote it.

Appendix C presents reproductions of the relevant parts of these historical works.

8. The Prime Number Theorem

We've just seen experimental evidence that $n/\ln(n)$ approximates $\pi(n)$ fairly well. Although the error itself grows as $n \to \infty$, the proportional error gets ever smaller.

Let's write that out.

$$\lim_{n\to\infty} \frac{\pi(n) - n/\ln(n)}{\pi(n)} = 0$$

Rearranging this gives us the following.

$$\lim_{n\to\infty} \frac{\pi(n)}{n/\ln(n)} = 1$$

This says the ratio of $\pi(n)$ and the approximation $n/\ln(n)$ tends to 1 as $n \to \infty$. And this is precisely what the **prime number theorem** says.

$$\pi(n) \sim n/\ln(n)$$

The symbol $\sim$ says that both sides are **asymptotically equivalent**. For example, $f(n) \sim g(n)$ means $f(n)/g(n) = 1$ as $n \to \infty$.

Asymptotic Equivalence

Some examples of asymptotic equivalence will help clarify its meaning.

If $f(x) = x^2 + x$ and $g(x) = x^2$, then $f \sim g$. Both f and g have the same dominant term x^2.

$$\lim_{x \to \infty} \frac{x^2 + x}{x^2} = \lim_{x \to \infty} 1 + \frac{1}{x} = 1$$

Swapping f and g doesn't break asymptotic equivalence, $g \sim f$.

$$\lim_{x \to \infty} \frac{x^2}{x^2 + x} = \lim_{x \to \infty} \frac{1}{1 + \frac{1}{x}} = 1$$

However, if $f(x) = x^3$ and $g(x) = x^2$, then f and g are not asymptotically equivalent because the ratio f/g tends to x, not 1.

If we know that $f \sim g$ and $g \sim h$, then we can also say $f \sim h$. This property, called **transitivity**, is familiar from normal equality.

What About li(n)?

Gauss' second approximation $\text{li}(n)$ appeared to be a better approximation for $\pi(n)$. You'll find the prime number theorem is sometimes expressed using the logarithmic integral.

$$\pi(n) \sim \text{li}(n)$$

Surely the prime number theorem must be about one of the approximations, not both? The only solution is for both approximations to be asymptotically equivalent. Let's see that this is indeed the case.

Let's set $f(n) = \frac{n}{\ln(n)}$ and $g(n) = \int_0^n \frac{1}{\ln(x)} dx$.

To show $f \sim g$ we need to find the limit of $f(n)/g(n)$ as $n \to \infty$ and confirm it is 1. Sadly, both $f(n)$ and $g(n)$ become infinitely large as $n \to \infty$, so the ratio is undefined.

When this happens, we usually try l'Hopital's rule as an alternative way to find the limit.

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} = \lim_{n\to\infty} \frac{f'(n)}{g'(n)}$$

It's fairly easy to work out $f'(n) = \frac{\ln(n)-1}{\ln^2(n)}$, and $g'(n) = \frac{1}{\ln(n)}$ pops out of the definition of $\mathrm{li}(n)$.

$$\begin{aligned}
\lim_{n\to\infty} \frac{f'(n)}{g'(n)} &= \lim_{n\to\infty} \frac{(\ln(n)-1)\ln(n)}{\ln^2(n)} \\
&= \lim_{n\to\infty} 1 - \frac{1}{\ln(n)} \\
&= 1
\end{aligned}$$

So the prime number theorem can refer to either of the two approximations, $n/ln(n)$ and $\mathrm{li}(n)$, because they are asymptotically equivalent.

What Does The Prime Number Theorem Really Say?

The prime number theorem says that $\pi(n)$ grows in a way that is asymptotically equivalent to functions like $n/\ln(n)$ and $\mathrm{li}(n)$.

It doesn't say that these are the only or best functions for approximating $\pi(n)$, which leaves open the intriguing possibility of other functions that are even better than $\mathrm{li}(n)$.

Bertrand's Postulate

The prime number theorem, even if it looks imprecise, is rather useful. Let's look at a particularly simple example.

In 1845 Bertrand proposed that there is at least one prime between a counting number and its double, $n < p < 2n$.

Let's compare the number of primes up to $2x$, with the number of primes up to x.

$$\frac{\pi(2x)}{\pi(x)} \sim \frac{2x}{\ln(2x)} \cdot \frac{\ln(x)}{x} \sim 2$$

This shows that for sufficiently large x, between x and $2x$, there are $x/\ln(x)$ primes. This is a stronger statement than Bertrand's postulate which merely suggests there is at least one prime.

There are proofs that don't require x to be sufficiently large, but they're nowhere near as simple as this deduction from the prime number theorem.

9. Euler's Golden Bridge

Euler was the first to find a connection between the world of primes and the world of ordinary counting numbers. This 'golden bridge' has become a path through which many new insights about the primes have been revealed.

Let's recreate Euler's discovery and experience some of his genius ourselves.

A Simple Series

Let's start with a simple and familiar series, valid for $|x| < 1$.

$$\frac{1}{(1-x)} = 1 + x + x^2 + x^3 + \ldots$$

Every power of x is in this series, a fact we'll be making good use of.

In our enthusiasm to involve the primes, we might be tempted to say that x is a prime p. This won't work because $|x|$ needs to be < 1, and the primes are all larger than 1.

One idea is to set x to $\frac{1}{p}$ which is always < 1.

$$\frac{1}{(1-\frac{1}{p})} = 1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \ldots$$

Again, every power of that prime p can be found in this series.

Multiplying Two Series

Now imagine taking that expression for one prime p_1, and multiplying it with the expression for a different prime p_2.

$$\frac{1}{(1-\frac{1}{p_1})} \cdot \frac{1}{(1-\frac{1}{p_2})} = \left(1+\frac{1}{p_1}+\frac{1}{p_1^2}+\dots\right) \cdot \left(1+\frac{1}{p_2}+\frac{1}{p_2^2}+\dots\right)$$

If we multiplied out the product on the right, we would get every combination of powers for p_1 and p_2. For example, somewhere in that series will be denominators $p_1^7 \cdot p_2^3$, as well as p_1^8 and p_2^{99}.

Also, each combination will appear only once. We wouldn't get p_1^2 or $p_1^7 \cdot p_2^3$ appearing twice, for example.

All The Primes

Let's now extend the product, from two primes p_1 and p_2, to all primes p_i. The symbol $\prod$ means product, just like $\sum$ means sum.

$$\prod_{p_i} \frac{1}{(1-\frac{1}{p_i})} = \prod_{p_i} \left(1+\frac{1}{p_i}+\frac{1}{p_i^2}+\dots\right)$$

If we multiplied out the product, again, we would get every combination of powers for every combination of primes p_i. For example, somewhere in that series will be denominators $p_1^2 \cdot p_2^{10} \cdot p_5^2$, as well as p_{23}^{43} and p_{76}.

And again, each combination would appear only once.

Unique Prime Factors

We saw earlier the fundamental theorem of arithmetic tell us that every positive integer n can be expressed as a unique product of prime factors.

The series we just multiplied out contains every combination of powers for every combination of primes. That is, it contains all the unique prime factorisations of every positive integer.

For example, the number 15750 has prime factors $2 \cdot 3^2 \cdot 5^3 \cdot 7$, and that series will have a term $(2 \cdot 3^2 \cdot 5^3 \cdot 7)^{-1}$ in it.

Euler's Product Formula

If that series contains the unique prime factorisation of every whole number n, and only contains it once, then we can finally make the magic leap that leads to **Euler's product formula**.

$$\boxed{\prod_p \frac{1}{(1-\frac{1}{p})} = \sum_n \frac{1}{n}}$$

The product of $\frac{1}{(1-\frac{1}{p})}$ over all primes p, is the sum of $\frac{1}{n}$ over all positive integers.

To say this result is amazing would not be an exaggeration. It reveals a previously well-hidden connection between the primes and the ordinary counting numbers.

Even better, that connection is beautifully simple. And that simplicity looks ripe for revealing more insights into the primes.

10. Walking Euler's Golden Bridge

Euler's product formula connects prime numbers to ordinary counting numbers.

$$\sum_n \frac{1}{n} = \prod_p \frac{1}{(1 - \frac{1}{p})}$$

Its simplicity almost demands we apply mathematical tools to reveal new insights. Let's try a few ourselves.

Another Proof Of Infinite Primes

The product in Euler's formula involves the all primes p. If the number of primes were finite, then that product would also be finite. We can say this because each expression is of the form $(1 - \frac{1}{p})^{-1}$ which is always finite, and never zero.

However, the sum in Euler's product formula is the harmonic series $\sum \frac{1}{n}$, which we've already seen becomes infinitely large. That means the number of primes can't be finite.

Euler's golden bridge has gifted us a really simple proof that primes are endless without too much effort at all.

This might look too easy. Perhaps this proof is a circular argument because the construction of the Euler product formula itself assumed an infinity of primes? If we look back, we'll see that there was no such assumption. The proof really is that simple and elegant.

Growth Of Prime Reciprocals

Euler was gifted with a deep intuition that led him to groundbreaking results, even if his routes didn't always meet modern standards of rigour. Here we'll follow his famous intuition to find the rate at which the inverse primes diverge.

Turning a product into a sum, by taking the natural logarithm, is often the start of a fruitful journey.

$$\ln\left(\sum_n \frac{1}{n}\right) = \ln\left(\prod_p \frac{1}{(1-\frac{1}{p})}\right)$$

Let's focus on the right hand side.

$$\begin{aligned}\ln\left(\prod_p \frac{1}{(1-\frac{1}{p})}\right) &= \sum_p \ln\frac{1}{(1-\frac{1}{p})} \\ &= -\sum_p \ln(1-\frac{1}{p})\end{aligned}$$

We can use the well known series $\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \ldots$ to expand $\ln(1-\frac{1}{p})$.

$$\begin{aligned}-\sum_p \ln(1-\frac{1}{p}) &= \sum_p(\frac{1}{p} + \frac{1}{2p^2} + \frac{1}{3p^3} + \frac{1}{4p^4}\ldots) \\ &= \sum_p \frac{1}{p} + C\end{aligned}$$

The expansion gives us the sum over the reciprocals of primes, which we now know is divergent. The rest of the expansion C converges.

How do we know C converges? Well, we know the sum of the reciprocals of squares $\sum \frac{1}{n^2}$, and indeed higher powers $\sum \frac{1}{n^x}$, converges. Any series $\sum \frac{1}{p^x}$, summming only over primes p and not all integers n, also converges because it is a subseries of $\sum \frac{1}{n^x}$.

Let's now focus on the left hand side, and look more closely at a finite form of the harmonic series $\sum_1^n \frac{1}{x}$.

$$\begin{aligned}\ln\left(\sum_1^n \frac{1}{x}\right) &> \ln \int_1^{n+1} \frac{1}{x} dx \\ &= \ln\left(\ln(n+1)\right)\end{aligned}$$

Appendix D explains how the inequality emerges by comparing the area under the curve $y = \frac{1}{x}$ with discrete sums of $\frac{1}{n}$. This is a simple but powerful technique used frequently in number theory, and worth becoming familiar with.

Examining these two results, Euler made a rather brave leap to say the sum of prime reciprocals less than n grows like $\ln(\ln(n))$.

$$\sum_{p<n} \frac{1}{p} \sim \ln\left(\ln(n)\right)$$

His conclusion was correct, but his argument wasn't quite watertight. We can't naively take the logarithm of infinity. Even so, this is a deep insight about primes prompted by Euler's product formula.

Density of Primes

We can try Eulerian brave leaps ourselves. The following expressions are all equivelent, with the last one derived using $u = \ln(x)$ and $du =$

$\frac{1}{x}dx$.

$$\ln(\ln(n)) = \int_1^{\ln(n)} \frac{1}{u}du = \int_e^n \frac{1}{x}\frac{1}{\ln(x)}dx$$

Now let's write the sum of prime reciprocals as an integral.

$$\sum_{p<n} \frac{1}{p} = \int_1^n \frac{1}{x}P(x)dx$$

Here $P(x)$ is a selector for primes, and is 1 if x is prime, 0 if not. We can make a small leap to say that in the neighbourhood of x, $P(x)$ is like a prime density.

Let's compare the two integrals.

$$\sum_{p<n} \frac{1}{p} \sim \int_e^n \frac{1}{x}\frac{1}{\ln(x)}dx$$
$$\sum_{p<n} \frac{1}{p} = \int_1^n \frac{1}{x}P(x)dx$$

This suggests that $1/\ln(x)$ is the density of primes around x. We saw experimental evidence for this earlier, and now we've drawn the same conclusion, albeit not rigorously, from Euler's golden bridge.

11. Prime Reciprocals Grow Like LogLog

We're going to end Part 1 by being ambitious and proving that sum of prime reciprocals grows like $\ln(\ln(x))$. This is ambitious because we going to go much further than simpler proofs which only prove a lower bound to the sum.

Part II.

Prime Number Theorem

Part III.

Appendices

A. $\sum 1/n$ Diverges

Infinite Series

Have a look at the following infinite series.

$$1 + 1 + 1 + 1 + \ldots$$

We can easily see this sum is infinitely large. The series **diverges**.

The following shows a different infinite series. Each term is half the size of the previous one.

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$$

We can intuitively see this series gets ever closer to 2. Many would simply say the sum is in fact 2. The series **converges**.

Harmonic Series

Now let's look at this infinite series, called the **harmonic series**.

$$S = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \ldots$$

Each term is smaller than the previous one, and so contributes an ever smaller amount to the sum. Perhaps surprisingly, the harmonic series doesn't converge. The sum is infinitely large.

The following, rather fun, proof is based on Oresme's which dates back to the early 1300s.

We start by grouping the terms in the series as follows.

$$S = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots$$

The brackets will have $2, 4, 8, 16 \dots$ terms inside them. Replacing each term in a group by its smallest member gives us the following new series.

$$\begin{aligned} T &= 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \dots \\ &= 1 + \frac{1}{2} + \left(\frac{1}{2}\right) + \left(\frac{1}{2}\right) + \dots \end{aligned}$$

We can see straight away this series diverges.

Because we replaced terms in S by smaller ones to make T, we can say $S > T$.

And because T diverges, so must the harmonic series S.

B. $\sum 1/n^2$ Converges

The sum of the reciprocals of the square numbers was a particularly difficult challenge, first posed around 1650, and later named the Basel problem.

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots$$

Although there are more modern proofs, we will follow Euler's original proof from 1734 because his methods were pretty audacious, and later influenced Riemann's work on the prime number theorem.

Taylor Series For sin(x)

We start with the familiar Taylor series for $\sin(x)$.

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \ldots$$

Euler's New Series For sin(x)

The polynomial $f(x) = (1 - \frac{x}{a})(1 + \frac{x}{a})$ has factors $(1 - \frac{x}{a})$ and $(1 + \frac{x}{a})$, and zeros at $+a$ and $-a$. We can shorten it to $f(x) = (1 - \frac{x^2}{a^2})$.

Euler's novel idea was to write $\sin(x)$ as a product of similar linear factors, which would lead him to a different series.

The zeros of $\sin(x)$ are at $0, \pm\pi, \pm 2\pi, \pm 3\pi, \ldots$ so the product of factors looks like the following.

$$\sin(x) = A \cdot x \cdot \left(1 - \frac{x^2}{\pi^2}\right) \cdot \left(1 - \frac{x^2}{(2\pi)^2}\right) \cdot \left(1 - \frac{x^2}{(3\pi)^2}\right) \cdot \ldots$$

The constant A is 1 because we know $\frac{\sin(x)}{x} \to 1$ as $x \to 0$. Alternatively, taking the first derivative of both sides gives $A = 1$ when $x = 0$.

The second factor is x and not x^2 because the zero of $\sin(x)$ at $x = 0$ has multiplicity 1.

Euler then expanded out the series.

$$\sin(x) = x \cdot \left[1 + \frac{x^2}{\pi^2}\left(\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots\right) + X\right]$$

Inside the square brackets, the terms with powers of x higher than 2 are contained in X.

Comparing The Two Series

The terms in Euler's new series and the Taylor series must be equivalent because they both represent $\sin(x)$. Let's pick out the x^3 terms from both series.

$$\frac{x^3}{3!} = \frac{x^3}{\pi^2}\left(\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots\right)$$

We can easily rearrange this to give us the desired infinite sum.

$$\frac{\pi^2}{6} = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots$$

Euler, aged 28, had solved the long standing Basel problem, not only proving the infinite series of squared reciprocals converged, but giving it an exact value.

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

Rigour

Euler's original proof was adventurous in expressing $\sin(x)$ as an infinite product of simple linear factors. It made intuitive sense, but at the time was not rigorously justified.

It was almost 100 years later when Weierstrass developed and proved a factorisation theorem that confirmed Euler's leap was legitimate.

C. Historical References For $\pi(n)$

Gauss, 1791

Gauss' 1791 'Some Asymptotic Laws Of Number Theory' can be found in volume 10 of his collected works. In it he presents his approximation for $\pi(n)$.

$$\frac{a}{la}$$

Today, this would be written as $n/\ln(n)$.

Source: `http://resolver.sub.uni-goettingen.de/purl?PPN236018647`

NACHLASS.

EINIGE ASYMPTOTISCHE GESETZE DER ZAHLENTHEORIE.

[I.]

[Handschriftliche Eintragung in dem Buche:] JOHANN CARL SCHULZE, Neue und erweiterte Sammlung logarithmischer .... Tafeln. I, Berlin 1778; [von GAUSS' Hand] Gauß. 1791.

[Auf der Rückseite des letzten Blattes.]

[1.]

Primzahlen unter a $(=\infty)$

$$\frac{a}{la}.$$

[2.]

Zahlen aus zwei Factoren

$$\frac{lla \, . \, a}{la},$$

(wahrsch.) aus 3 Factoren

$$\frac{\frac{1}{2}(lla)^2 a}{la}, \ldots$$

et sic in inf.

2*

Figure C.1.: Gauss' 1971 Some Asymptotic Laws Of Number Theory.

Legendre, 1797

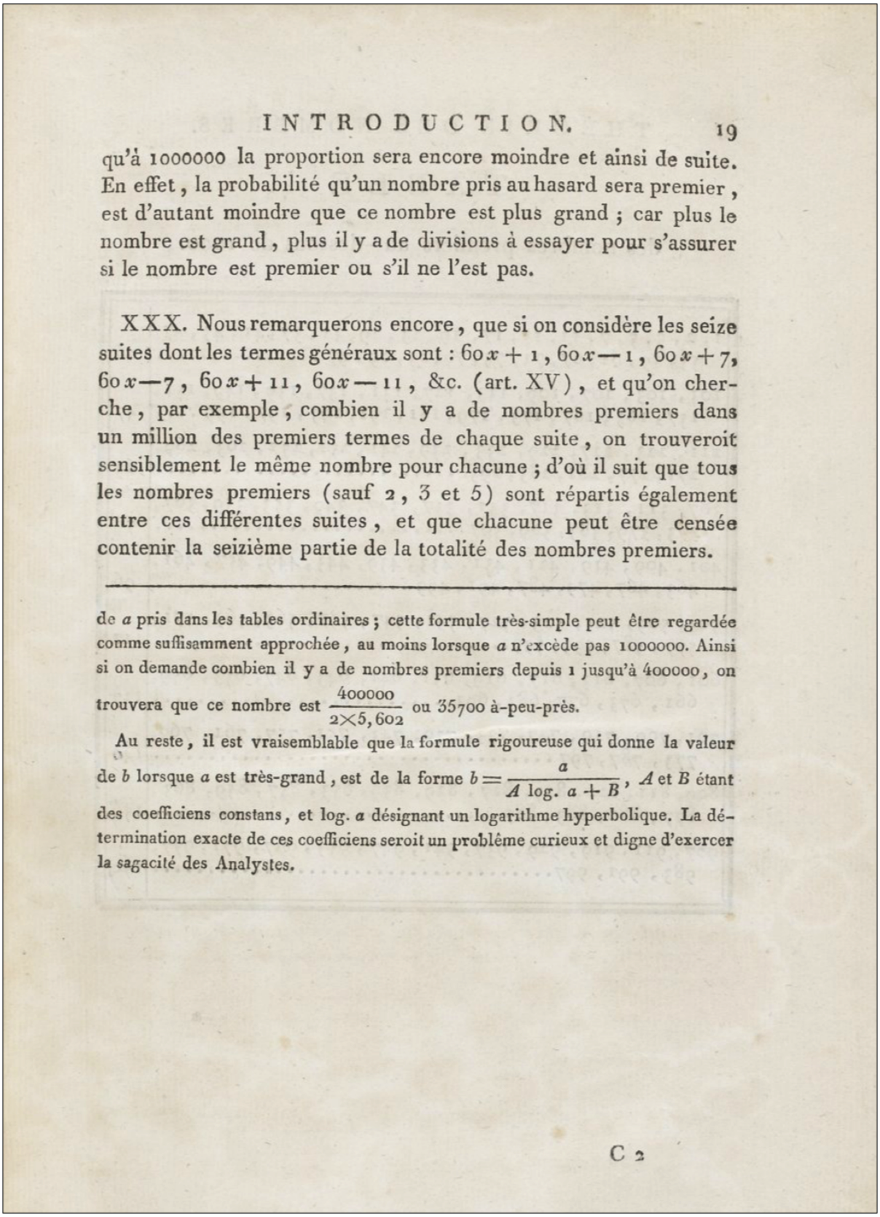
Legendre in his first edition of 'Essai Sur La Theorie Des Nombres' presented his approximation.

$$\frac{a}{A\log(a)+B}$$

The logarithm is the natural $\ln(a)$. In his 1808 second edition he quantifies the constants.

$$\frac{x}{\log(x)-1.08366}$$

Source: `https://gallica.bnf.fr/ark:/12148/btv1b8626880r/f55.image`



INTRODUCTION. 19

qu'à 1000000 la proportion sera encore moindre et ainsi de suite. En effet, la probabilité qu'un nombre pris au hasard sera premier, est d'autant moindre que ce nombre est plus grand; car plus le nombre est grand, plus il y a de divisions à essayer pour s'assurer si le nombre est premier ou s'il ne l'est pas.

XXX. Nous remarquerons encore, que si on considère les seize suites dont les termes généraux sont : $60x+1$, $60x-1$, $60x+7$, $60x-7$, $60x+11$, $60x-11$, &c. (art. XV), et qu'on cherche, par exemple, combien il y a de nombres premiers dans un million des premiers termes de chaque suite, on trouveroit sensiblement le même nombre pour chacune; d'où il suit que tous les nombres premiers (sauf 2, 3 et 5) sont répartis également entre ces différentes suites, et que chacune peut être censée contenir la seizième partie de la totalité des nombres premiers.

---

de a pris dans les tables ordinaires; cette formule très-simple peut être regardée comme suffisamment approchée, au moins lorsque a n'excède pas 1000000. Ainsi si on demande combien il y a de nombres premiers depuis 1 jusqu'à 400000, on trouvera que ce nombre est $\frac{400000}{2\times 5,602}$ ou 35700 à-peu-près.

Au reste, il est vraisemblable que la formule rigoureuse qui donne la valeur de b lorsque a est très-grand, est de la forme $b=\frac{a}{A\ \log.\ a+B}$, A et B étant des coefficiens constans, et log. a désignant un logarithme hyperbolique. La détermination exacte de ces coefficiens seroit un problême curieux et digne d'exercer la sagacité des Analystes.

C 2

Figure C.2.: Legendre's 1797 Essai Sur La Theorie Des Nombres.

Gauss, 1849

Gauss wrote a letter to astronomer Encke dated Decemer 24th 1849, in which he first presents an integral form of a prime counting function. He states this is based on work he started in 1792 or 1793.

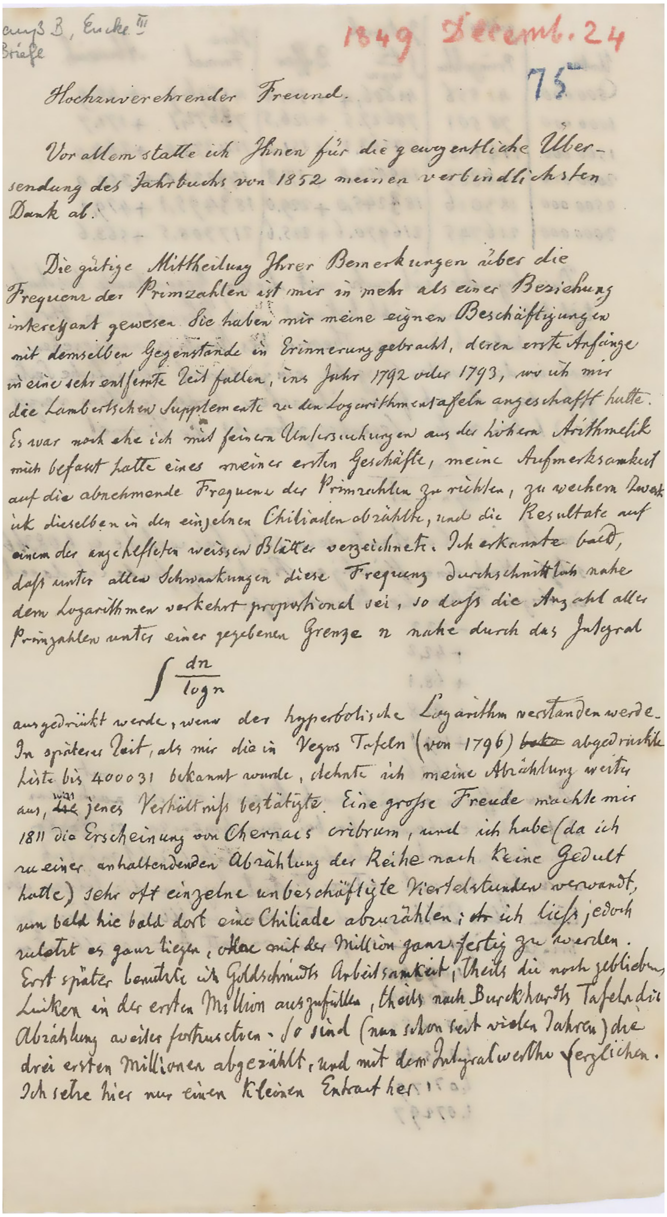
Gauss uses the following expression.

$$\int \frac{dn}{\mathrm{logn}}$$

Today this would be written as the logarithmic integral function.

$$\int_0^n \frac{1}{ln(x)} dx$$

Source: `https://gauss.adw-goe.de/handle/gauss/199`



aus B, Encke III
Briefe

1849 Decemb. 24

75

Hochzuverehrender Freund.

Vor allem statte ich Jhnen für die gewogentliche Übersendung des Jahrbuchs von 1852 meinen verbindlichsten Dank ab.

Die gütige Mittheilung Jhrer Bemerkungen über die Frequenz der Primzahlen ist mir in mehr als einer Beziehung interessant gewesen. Sie haben mir meine eignen Beschäftigungen mit demselben Gegenstande in Erinnerung gebracht, deren erste Anfänge in eine sehr entfernte Zeit fallen, ins Jahr 1792 oder 1793, wo ich mir die Lambertschen Supplemente zu den Logarithmentafeln angeschafft hatte. Es war noch ehe ich mit feinern Untersuchungen aus der höhern Arithmetik mich befasst hatte eines meiner ersten Geschäfte, meine Aufmerksamkeit auf die abnehmende Frequenz der Primzahlen zu richten, zu welchem Zweck ich dieselben in den einzelnen Chiliaden abzählte, und die Resultate auf einem der angehefteten weissen Blätter verzeichnete. Ich erkannte bald, dass unter allen Schwankungen diese Frequenz durchschnittlich nahe dem Logarithmen verkehrt proportional sei, so dass die Anzahl aller Primzahlen unter einer gegebenen Grenze n nahe durch das Integral

$$\int \frac{dn}{\log n}$$

ausgedrückt werde, wenn der hyperbolische Logarithm verstanden werde. In späterer Zeit, als mir die in Vegas Tafeln (von 1796) ~~bek~~ abgedruckte Liste bis 400031 bekannt wurde, dehnte ich meine Abzählung weiter aus, ~~die~~ was jenes Verhältniss bestätigte. Eine grosse Freude machte mir 1811 die Erscheinung von Chernacs cribrum, und ich habe (da ich zu einer anhaltenden Abzählung der Reihe nach keine Geduld hatte) sehr oft einzelne unbeschäftigte Viertelstunden verwandt, um bald hie bald dort eine Chiliade abzuzählen; ich liess jedoch zuletzt es ganz liegen, ohne mit der Million ganz fertig zu werden. Erst später benutzte ich Goldschmidts Arbeitsamkeit, theils die noch gebliebenen Lücken in der ersten Million auszufüllen, theils nach Burckhardts Tafeln die Abzählung weiter fortzusetzen. So sind (nun schon seit vielen Jahren) die drei ersten Millionen abgezählt, und mit dem Integralwerthe verglichen. Ich setze hier nur einen kleinen Extract her:

Figure C.3.: First page of Gauss' 1849 letter to Encke.

D. Bounds For $\sum \frac{1}{x}$

Understanding the behaviuor of continuous functions is often easier than discrete functions like $\sum_1^n \frac{1}{x}$ where x takes whole number values from 1 to n. A good strategy is to find continuous functions that are upper and lower bounds for the discrete function.

Figure D.1 shows a graph of $y = \frac{1}{x}$, together with rectangles representing the fractions $\frac{1}{n}$.

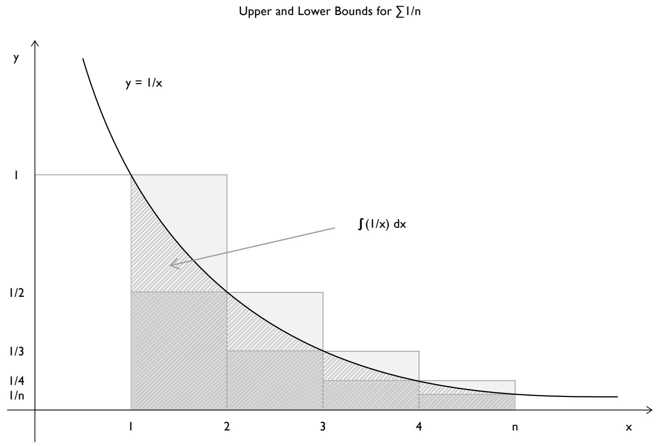


Figure D.1.: Comparing discrete $1/n$ with continuous $1/x$.

Lower Bound

If we consider the range $1 \leq x \leq 4$ we can see the area of the three taller rectangles $1 + \frac{1}{2} + \frac{1}{3}$ is greater than the area under the curve $\int_1^4 \frac{1}{x} dx$. By extending the range to n we can make a general observation.

$$\sum_1^n \frac{1}{x} > \int_1^{n+1} \frac{1}{x} dx$$

The integral has an upper limit of $n+1$ because the width of the last rectangle extends from $x = n$ to $x = n+1$. We can perform the integral to simplify the expression.

$$\boxed{\sum_1^n \frac{1}{x} > \ln(n+1)}$$

This is a rather nice lower bound on the growth of the harmonic series.

Upper Bound

Let's now look at the shorter rectanges. In the range $1 \leq x \leq 4$ we can see the area of the three shorter rectangles $\frac{1}{2} + \frac{1}{3} + \frac{1}{4}$ is less than the area under the curve $\int_1^4 \frac{1}{x} dx$. Again, by extending the range to n we can make a general observation.

$$\sum_2^n \frac{1}{x} < \int_1^n \frac{1}{x} dx$$

The harmonic sum starts at 2 because this time we're looking at rectanges extending to the left of a given x. We can easily fix the limits of the sum using $\sum_1^n \frac{1}{x} = 1 + \sum_2^n \frac{1}{x}$.

$$\sum_1^n \frac{1}{x} - 1 < \int_1^n \frac{1}{x} dx$$

Again, we can perform the integral.

$$\boxed{\sum_1^n \frac{1}{x} < \ln(n) + 1}$$

This is a nice upper bound to the growth of the harmonic series.

Comparison Tests

This method of comparing a discrete series with a continuous function to obtain nice functions for lower or upper bounds is fairly powerful, and very common in the pratice of number theory.