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.

First Steps

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.

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

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.

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

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.

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

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.

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

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

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.

$$
\begin{aligned}
12 &= 2 \times 6 \\
12 &= 3 \times 4
\end{aligned}
$$

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.

$$
\begin{aligned}
12 &= 2 \times (3 \times 2) \\
12 &= 3 \times (2 \times 2)
\end{aligned}
$$

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.

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

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.

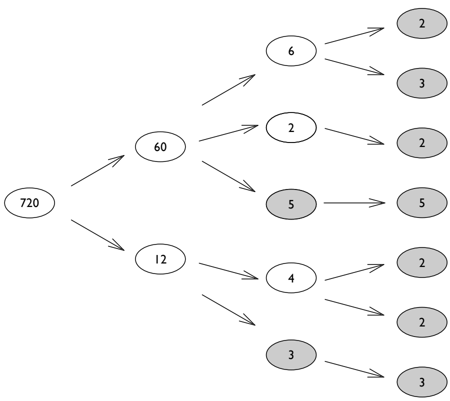


Figure 4.1.: Breaking 720 into factors until only primes remain.

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 greater than 1 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 how frequently particular numbers occur is to look 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 to avoid the sum becoming infinitely large.

It's natural to 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.

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

Between each term, only j changes. Now let's look more closely at the expression $1 + j \cdot (p_1 \cdot p_2 \cdot p_2 \dots p_k)$. 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.

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

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.

$$\sum \frac{1}{p} \to \infty$$

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.

Another Proof

Appendix C presents a different proof that $\sum 1/p_n$ diverges. Although it moves at a slightly faster pace, it is worth a look because it is short and rather fun.

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 $\ln(n)$.

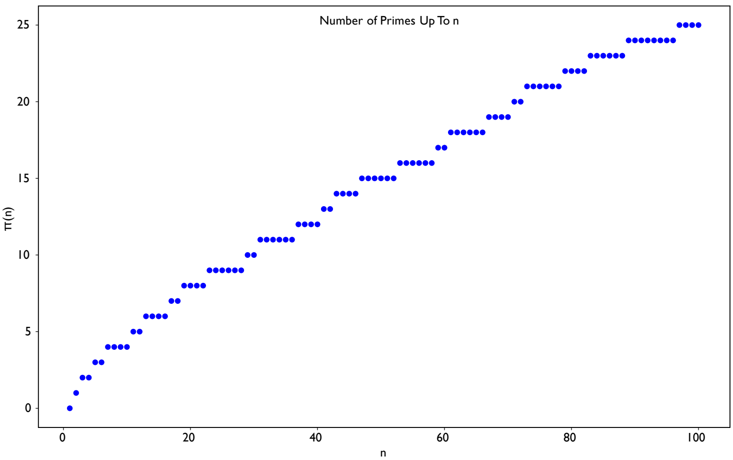


Figure 7.1.: $\pi(n)$ for n from 1 to 100.

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.

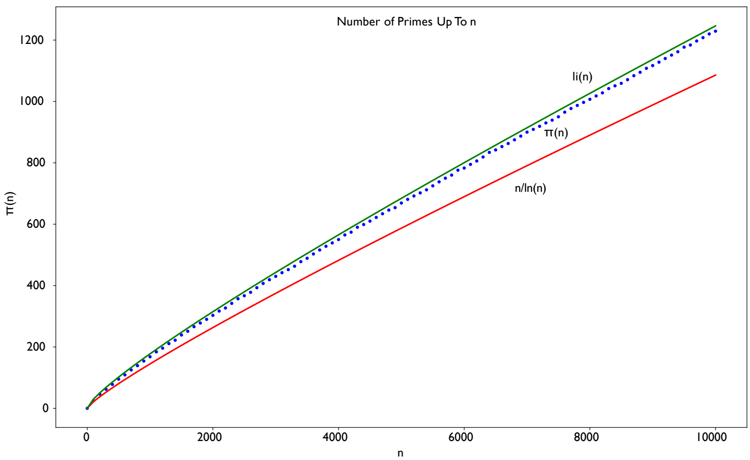


Figure 7.2.: Comparing $\mathrm{li}(n)$ and $n/\ln(n)$ with $\pi(n)$.

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 $\mathrm{li}(n)$ has a distinctly smaller proportional error than $n/\ln(n)$.

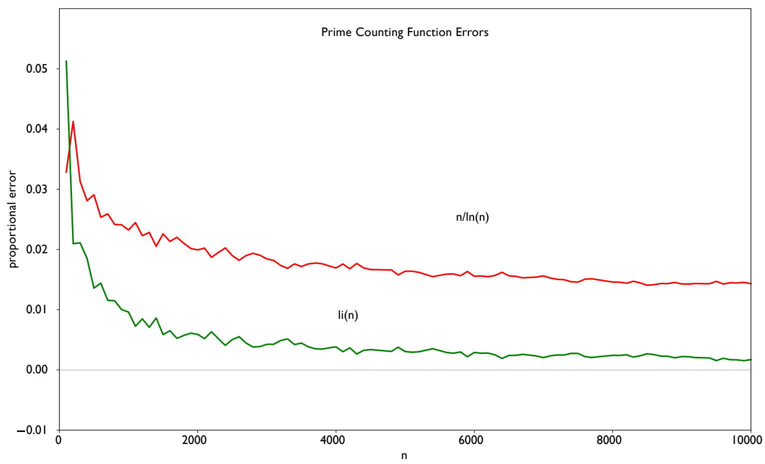


Figure 7.3.: Proportional errors for $\mathrm{li}(n)$ and $n/\ln(n)$.

There are 1229 primes amongst the first 10,000 whole numbers. The logarithmic integral gives us $\mathrm{li}(10,000) = 1246$. 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.

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 $\int \frac{dn}{\text{logn}}$. His collected works also reveal that in 1791 he had written about the simpler approximation, $\frac{a}{la}$ as he wrote it.

Appendix D 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, $g \sim f$. This is clear from its definition as a ratio.

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 $n \to \infty$, which is a little unhelpful.

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, can provide easy insights into questions about the primes.

In 1845 Bertrand proposed that there is at least one prime between a counting number and its double, $n < p < 2n$. A proof would take a few pages to walk through.

We can use the prime number theorem to asymptotically compare the number of primes up to $2n$, with the number of primes up to n.

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

With very little work, this tells us that between n and $2n$, there are approximately $n/\ln(n)$ primes, an approximation that becomes truer for larger n.

This is actually a stronger statement than Bertrand's postulate which merely suggests there is at least one prime.

9. Gaps Between Primes

We've seen the primes aren't so thinly spread out that $\sum 1/p$ converges, and we've just seen the primes occur with an average density of $1/\ln(x)$ around the number x.

This suggests the gaps between primes must be constrained. Let's find out.

Prime-Free Sequences

The factorial $n!$ of a whole number n is the product of all the whole numbers between 1 and n. For example, $4! = 4 \times 3 \times 2 \times 1 = 24$.

Let's look at the sequence of numbers from $(5!+2)$ to $(5!+5)$.

$$
\begin{aligned}
5! + 2 &= (5 \cdot 4 \cdot 3 \cdot 1 + 1) \cdot 2 \\
5! + 3 &= (5 \cdot 4 \cdot 2 \cdot 1 + 1) \cdot 3 \\
5! + 4 &= (5 \cdot 3 \cdot 2 \cdot 1 + 1) \cdot 4 \\
5! + 5 &= (4 \cdot 3 \cdot 2 \cdot 1 + 1) \cdot 5
\end{aligned}
$$

We can see 2 is a factor of $(5!+2)$, 3 is a factor of $(5!+3)$, 4 is a factor of $(5!+4)$, and 5 is a factor of $(5!+5)$. So the sequence $(5!+2)$ to $(5!+5)$ is entirely free of primes.

Using the same method, we can show any sequence that follows the same pattern $(n!+2)$ to $(n!+n)$ is prime-free. This is a sequence of length $n-1$.

Because n can be as large as desired, we've just proved there is no upper limit on prime gaps.

The primes continue to surprise us. We started with the fact that the primes can't be too sparse, and yet we have shown there is no upper limit on the gaps between primes. Both facts can be true if large prime gaps are rare. This leads us to explore the distribution of prime gaps.

Distribution of Prime Gaps

If we counted all the prime gaps in a moderately large range of numbers, we would expect very long gaps to be rare. It isn't obvious whether the smallest gaps would occur more or less often than medium sized gaps.

Figure 9.1 shows the logarithm of the count for each prime gap amongst the first 500 million numbers. For clarity, the definition of a **prime gap** is $(p_{n+1} - p_n)$, and not the length of the prime-free sequence, so the prime gap between 3 and 5 is 2, not 1.

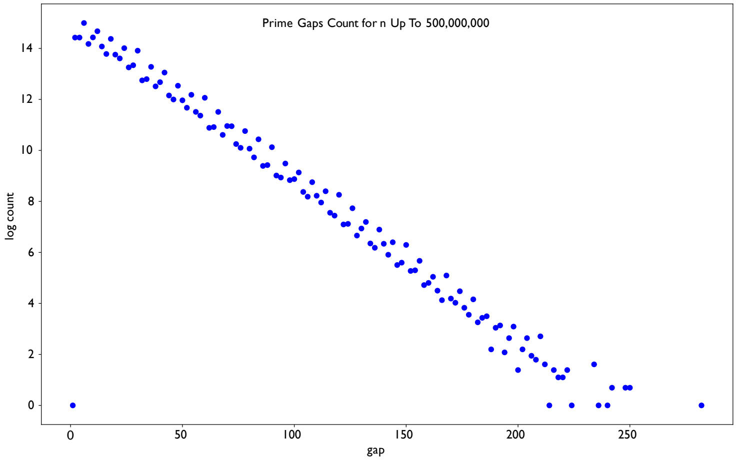


Figure 9.1.: Count of prime gaps in the first 500 million numbers.

The distribution is surprisingly constrained to a very linear band, with clear regions above and below where no prime gap counts seem to occur. It does seem to confirm that the larger the prime gap, the less frequently it occurs in a given range of numbers.

Just as with the distribution of prime numbers, encoding the exact distribution of prime gaps into a simple formula is beyond the reach of today's mathematics. Appendix E presents an alternative approach to modelling this distribution using a probabilistic model of prime numbers.

Twin Prime Conjecture

Looking again at the distribution, gaps of size 2 seem to occur more often than almost all other sizes. You may have noticed these **twin**

primes do pop up more often than expected along the number line:

$$(3,5),(5,7),(11,13),(17,19)\ldots(101,103),(107,109)\ldots$$

A natural question to ask is whether these twin primes ever run out. The **Twin Prime Conjecture** says they don't. This is yet another simple statement about primes that remains unproven.

10. Euler's Golden Bridge

Euler was the first to find a connection between the world of primes and the world of ordinary counting numbers. Many insights about the primes have been revealed by travelling over this 'golden bridge'.

Let's recreate Euler's discovery for ourselves.

Riemann Zeta Function

We know the harmonic series $\sum 1/n$ diverges. We also know the series $\sum 1/n^2$ converges.

It's natural to ask for which values of s the more general series, known as the **Riemann Zeta** function $\zeta(s)$, converges.

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \ldots$$

Series like $\sum 1/n^3$ and $\sum 1/n^4$ converge because each term is smaller than the corresponding one in $\sum 1/n^2$. Less obvious is when $s < 2$.

Appendix F presents a short proof that $\zeta(s)$ converges for $s > 1$.

The proof compares the discrete sum $\zeta(s)$ with a related continuous integral that is easier to analyse. This simple technique is used a lot in number theory, and worth becoming familiar with.

Sieving The Zeta Function

Let's write out the zeta function again, noting that $1^s = 1$.

$$\zeta(s) = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + \frac{1}{6^s} + \ldots$$

We can divide this series by 2^s.

$$\frac{1}{2^s}\zeta(s) = \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + \frac{1}{8^s} + \frac{1}{10^s} + \frac{1}{12^s} \ldots$$

These denominators are multiples of 2^s. By subtracting these terms from $\zeta(s)$, we sieve out terms with these multiples of 2^s.

$$(1 - \frac{1}{2^s}) \cdot \zeta(s) = 1 + \frac{1}{3^s} + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{9^s} + \frac{1}{11^s} + \ldots$$

This dividing and subtracting of infinite series is only valid because they are absolutely convergent for $s > 1$.

Let's now divide this series by 3^s.

$$\frac{1}{3^s} \cdot (1 - \frac{1}{2^s}) \cdot \zeta(s) = \frac{1}{3^s} + \frac{1}{9^s} + \frac{1}{15^s} + \frac{1}{21^s} + \frac{1}{27^s} + \ldots$$

These denominators are all multiples of 3^s, but not all multiples of 3^s are here. Some like 6^s and 12^s were removed in the previous step. Subtracting this series from the previous series leaves terms with denominators that are not multiples of 2 or 3.

$$(1 - \frac{1}{3^s}) \cdot (1 - \frac{1}{2^s}) \cdot \zeta(s) = 1 + \frac{1}{5^s} + \frac{1}{7^s} + \frac{1}{11^s} + \frac{1}{13^s} + \ldots$$

We can't remove terms with multiples of 4^s because they were sieved out when we removed multiples of 2^s. The next useful step is to remove multiples of 5^s. Doing this leaves terms with denominators that are not multiples of 2, 3 or 5.

$$(1-\frac{1}{5^s})\cdot(1-\frac{1}{3^s})\cdot(1-\frac{1}{2^s})\cdot\zeta(s) = 1+\frac{1}{7^s}+\frac{1}{11^s}+\frac{1}{13^s}+\ldots$$

Repeating this several times, we'll see that only multiples of successive primes are available to be removed. We'll also see that after each removal, the very first term 1 always survives.

If we kept going, we'd end up with an all the primes on the left, and only 1 on the right. This is because every n in $1/n^s$ is either a prime, or a multiple of a prime, which we know from the fundamental theroem of arithmetic.

$$\ldots\cdot(1-\frac{1}{11^s})\cdot(1-\frac{1}{7^s})(1-\frac{1}{5^s})\cdot(1-\frac{1}{3^s})\cdot(1-\frac{1}{2^s})\cdot\zeta(s) = 1$$

We can rearrange this to isolate $\zeta(s)$.

$$\zeta(s) = \prod_p (1-\frac{1}{p^s})^{-1}$$

The symbol $\prod$ means product, just like $\sum$ means sum.

Euler's Product Formula

We've arrived at **Euler's product formula**.

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

The product of $(1 - \frac{1}{p^s})^{-1}$ over all primes p is the sum of $\frac{1}{n^s}$ over all positive integers n, as long as we remember to keep $s > 1$.

To say this result is amazing would not be an exaggeration. It reveals a deep connection between the primes and the ordinary counting numbers, a connection that doesn't appear too complicated at first sight.

Another Proof Of Infinite Primes

The simplicity of Euler's product formula demands we explore its implications. Let's take our first steps.

We know the harmonic series $\zeta(1) = \sum 1/n$ diverges. We can use Euler's formula to write this sum as a product over primes.

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

Each of the factors $(1 - \frac{1}{p})^{-1}$ is always finite, and never zero. For the product to diverge, there must be an infinite number of these factors. This means there must be an infinite number of primes. Some consider this to be the first new proof of infinite primes since Euclid's from around 300BC.

Is this proof a circular argument? That is, does the derivation of the Euler product formula itself assume an infinity of primes? Looking back, the derivation sieves out all integers greater than 1 because they are either a prime, or a multiple of a prime. This is the fundamental theorem of arithmetic, which makes no claims about the number of primes.

Easy Proof $\sum 1/p$ Diverges

Turning a product into a sum by taking its logarithm is often fruitful, and we'll do it here for $\zeta(1)$ which we know diverges.

$$\ln\left(\sum \frac{1}{n}\right) = -\sum_p \ln(1 - \frac{1}{p})$$

We can use $\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} \ldots$ for $|x| < 1$ to expand the logarithm.

$$\begin{aligned}\ln\left(\sum \frac{1}{n}\right) &= \sum_p (\frac{1}{p} + \frac{1}{2p^2} + \frac{1}{2p^3} + \ldots) \\ &= \sum_p \frac{1}{p} + C\end{aligned}$$

The remainder C is finite, and showing it requires just a bit of algebra.

$$C = \sum_p \sum_{k=2} \frac{1}{kp^k} < \sum_p \sum_{k=2} \frac{1}{p^k} < \sum_{n=2} \sum_{k=2} \frac{1}{n^k}$$

The last inequality is because there are fewer primes p than normal whole numbers n. Noticing that $\sum_{k=2} 1/n^k$ is a geometric series, we can simplify the last expression.

$$
\begin{aligned}
C < \sum_{n=2} \frac{1}{n(n-1)} &= \sum_{n=2} \left(\frac{1}{n-1} - \frac{1}{n} \right) \\
&= \frac{1}{1} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \ldots = 1
\end{aligned}
$$

Now, because we know the harmonic series $\sum 1/n$ diverges, then so must $\sum 1/p + C$. But because we've shown that $C < 1$, then the prime harmonic series $\sum 1/p$ must diverge.

Part II.

Next Steps

11. Into The Complex Domain

The Riemann Zeta function encodes information about the primes. We saw at the end Part 1 that we can use it to easily prove the primes are endless, and that the prime harmonic series diverges, telling us the primes aren't so sparse.

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

We've thought about the zeta function with s as an integer, for example the harmonic series with $s = 1$, and the Basel problem with $s = 2$. We've also considered s as a real number, and asked whether the zeta function converges for values of s between 1 and 2. We proved the zeta function converges for $s > 1$.

Riemann was the first to consider s as a **complex number**. If we think the zeta function over the complex domain might reveal new insights into the primes, we need to understand how it behaves. Exploring where it converges is a good start.

Convergence For $\sigma > 1$

It has become tradition to write complex s as $s = \sigma + it$, where σ is the real part of s.

$$\sum \frac{1}{n^s} = \sum \frac{1}{n^{\sigma+it}} = \sum \frac{1}{n^\sigma} \frac{1}{n^{it}}$$

Let's look at the series with each term replaced by its magnitude.

$$\sum \left| \frac{1}{n^s} \right| = \sum \left| \frac{1}{n^\sigma} \frac{1}{n^{it}} \right|$$

Rewriting n^{it} as $e^{it\ln(n)}$ makes clear it has a magnitude of 1.

$$\sum \left| \frac{1}{n^s} \right| = \sum \frac{1}{n^\sigma}$$

We've already shown $\sum 1/n^\sigma$ converges for $\sigma > 1$ where σ is real. This tells us $\sum 1/n^s$ converges absolutely for for $\sigma > 1$. Since absolute convergence implies convergence, we can say $\sum 1/n^s$ converges whenever the real part of s is more than 1, that is $\sigma > 1$.

Divergence For $\sigma \leq 0$

Let's look again at the terms in the sum.

$$\left| \frac{1}{n^s} \right| = \left| \frac{1}{n^\sigma} \frac{1}{n^{it}} \right| = \frac{1}{n^\sigma}$$

If $\sigma < 0$, the magnitude of the terms grows larger than 1. If $\sigma = 0$, the magnitude of each term is exactly 1. For any series to converge, a necessary requirement is that the terms get smaller towards zero. This means $\sum 1/n^s$ diverges for $\sigma \leq 0$.

Divergence For $\sigma < 1$

Having found the Riemann Zeta function converges for $\sigma > 1$, and diverges for $\sigma \leq 0$, we're left with a gap $0 < \sigma \leq 1$. To fill this gap we need to understand more generally when series of the form $\sum a_n/n^s$, called **Dirichlet series**, converge or diverge.

Appendix G explains how Dirichlet series converge in half-planes to the right of an **abscissa of convergence** σ_c. That is, they converge at any $s = \sigma + it$ where $\sigma > \sigma_c$.

Because $\zeta(s)$ converges for $\sigma > 1$, and we know it diverges at $s = 1+0i$, the abscissa of absolute convergence is at $\sigma_c = 1$. We can now say $\zeta(s)$ diverges for $\sigma < 1$.

Visualising The Zeta Function For $s > 1$

Visualising a function helps us understand it. Figure 11.1 shows a surface plot of $\ln|\zeta(s)|$ for $\sigma > 1$. Because the values of $\zeta(s)$ are complex, it is easier to plot the magnitude $|\zeta(s)|$. Taking the logarithm scales down very large values to better show the function's features.

The spike around $s = 1 + 0i$ corresponds to the divergent harmonic series $\zeta(1)$.

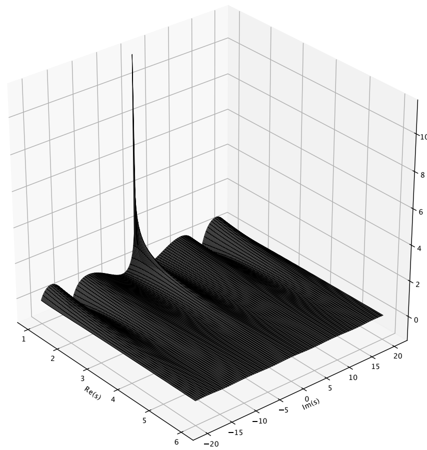


Figure 11.1.: Surface plot of $\ln|\zeta(s)|$ for $\sigma > 1$.

The surface seems to smooth out to the right as σ grows larger. It is natural to ask what value it approaches. Let's look again at the terms of the sum $\zeta(s)$.

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

As $\sigma \to \infty$, the magnitude of each term $|n^{-s}| = n^{-\sigma}$ tends to zero as $\sigma \to \infty$ for all n, except $n = 1$ where $1/1^s = 1$. That means $\zeta(s) \to 1$ as $\sigma \to \infty$.

We've assumed the limit of the sum is the sum of the limits, $\lim \sum 1/n^s = \sum \lim 1/n^s$. This isn't always true, but is true here, as explained in Appendix L.

No Zeros In $\sigma > 1$

Another natural question to ask is whether the function is ever zero. The product form of $\zeta(s)$ can help because we know products are only zero if one of the factors is zero.

$$\zeta(s) = \prod_p (1 - \frac{1}{p^s})^{-1}$$

For any factor $(1 - 1/p^s)^{-1}$ to be zero would require p^s to be zero. We can see this isn't possible by writing $|p^s| = \left|e^{s\ln(p)}\right| = e^{\sigma\ln(p)} > 0$.

Our logic isn't complete because infinite products can take a value of zero even if no individual factor is zero. This happens when each multiplication makes the product ever smaller, for example we can see $1/2 \times 1/3 \times 1/4 \times 1/5 \times \ldots$ gets ever smaller towards zero. Appendix L discusses infinite products, and shows that $\zeta(s)$ doesn't have any zeros in $\sigma > 1$.

Hints The Function Extends Into $\sigma \leq 1$

Looking again at Figure 11.1, we can see that aside from $s = 1 + 0i$, the function doesn't seem to diverge along the line $s = 1 + it$. It looks like the surface has been prematurely cut off, and would continue smoothly into $\sigma \leq 1$ if allowed.

Figure 11.2 shows a contour plot of $\ln|\zeta(s)|$. The contours do indeed appear artificially cut off, as if they should continue smoothly to the left of $\sigma = 1$.

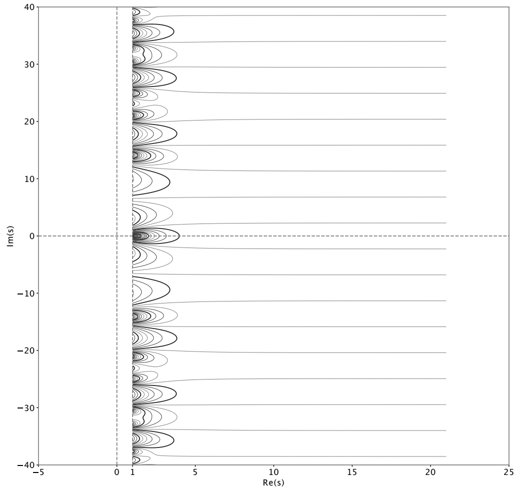


Figure 11.2.: Contour plot of $\ln|\zeta(s)|$ for $\sigma > 1$.

Almost Symmetric

The plots of the magnitude $|\zeta(s)|$ suggest the function is symmetric about the real axis. This isn't quite true.

Remembering the complex conjugate $\overline{s}$ is a reflection of s in the real axis, for example $\overline{3+2i} = 3-2i$, let's look again at the terms in the series.

$$n^{-\overline{s}} = e^{-\overline{s}\ln(n)} = \overline{e^{-sln(n)}} = \overline{n^{-s}}$$

This means $\zeta(\overline{s})$ is the complex conjugate of $\zeta(s)$. So, although the magnitude of $\zeta(s)$ is mirrored above and below the real axis, the sign of the imaginary part is inverted.

12. A New Series For $\zeta(s)$

Before we try to extend the Riemann Zeta series to the left of $\sigma = 1$ we should first understand how it might even be possible.

Series And Functions

Let's take a fresh look at the well known Taylor series expansion of $f(x) = (1-x)^{-1}$ developed around $x = 0$.

$$S_0 = 1 + x + x^2 + x^3 + \ldots$$

The series S_0 is only valid for $|x| < 1$, but the function $f(x)$ is defined for all x except $x = 1$. This apparent discrepancy requires some clarification.

That series S_0 is just one possible representation of the function. We can use the standard method for working out Taylor series to find a different representation of $f(x)$ valid outside $|x| < 1$. For example, the following series S_3 is developed around $x = 3$, and is valid for $1 < x < 5$.

$$S_3 = -\frac{1}{2} + \frac{1}{4}(x-3) - \frac{1}{8}(x-3)^2 + \frac{1}{16}(x-3)^3 - \ldots$$

So the series S_0 and S_3 both represent $f(x) = (1-x)^{-1}$ but over different parts of its domain. This clarifies the distinction between a function, and series which represents it in different parts of its domain.

Perhaps the series $\sum 1/n^s$ only gives us a partial view of a much richer function that encodes information about the primes. Could that function be represented by a different series over a different domain?

A New Series

Let's write out the familiar series for $\zeta(s)$.

$$\zeta(s) = \sum \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} + \ldots$$

An alternating version of the zeta function is called the **eta** function $\eta(s)$.

$$\eta(s) = \sum \frac{(-1)^{n+1}}{n^s} = \frac{1}{1^s} - \frac{1}{2^s} + \frac{1}{3^s} - \frac{1}{4^s} + \frac{1}{5^s} - \ldots$$

This is a Dirichlet series which, as explained in Appendix G, converges for $\sigma > 0$. If we could express $\zeta(s)$ in terms of $\eta(s)$, we would have a new series for the Riemann Zeta function that extends to the left of $\sigma = 1$, even if only as far as $\sigma > 0$.

Looking at the difference $\zeta(s) - \eta(s)$, we can see a pattern to exploit.

$$\zeta(s) - \eta(s) = \frac{2}{2^s} + \frac{2}{4^s} + \frac{2}{6^s} + \ldots$$

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

$$= 2^{1-s}\zeta(s)$$

Rearranging gives us a new series for $\zeta(s)$.

$$\boxed{\zeta(s) = \frac{1}{1-2^{1-s}} \sum \frac{(-1)^{n+1}}{n^s}}$$

We do need to check whether the factor $(1-2^{1-s})^{-1}$ diverges in the newly gained domain. The denominator $(1-2^{1-s})$ is zero at $s = 1+0i$, and provides $\zeta(s)$ with its familiar divergence at that point.

The denominator also has zeros at $s = 1 + 2\pi ik/\ln(2)$ for integers $k \neq 0$, but the resulting **poles** are cancelled by exactly matched zeros of the sum $\eta(s)$. This is explained in Appendix K.

We can finally say the new series for $\zeta(s)$ converges in the domain $\sigma > 0$, with the exception of just one pole at $s = 1$.

Finding a new representation of a complex function that agrees with a previously known representaion, but is valid over an extended or overlapping domain, is called **analytic continuation**.

Visualising The New Series

Any enthusiastic mathematician would be impatient to visualise this new series. Figure 12.1 shows a contour plot of $\ln|\zeta(s)|$ evaluated using the new series.

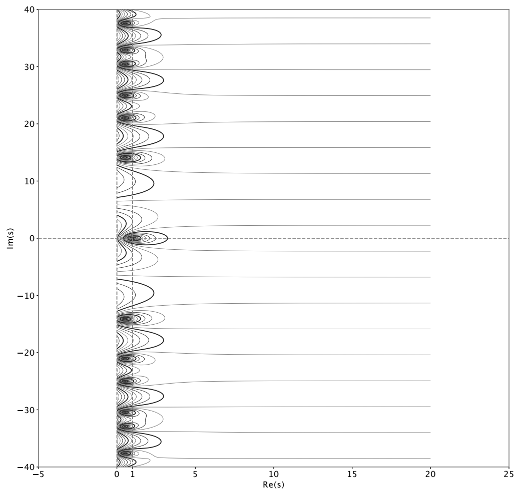


Figure 12.1.: Contour plot of $\ln|\zeta(s)|$ for $\sigma > 0$.

Our previous intuition was justified, the surface does continue smoothly to the left of $\sigma = 1$. In fact, the contours suggest the function should again continue smoothly even further into $\sigma < 0$.

Even more intriguing is the appearance of what look like zeros, all of which seem to be on the line $\sigma = 1/2$.

We'll have to wait until a little later to show these really are zeros.

13. The Riemann Hypothesis

We're ready to make a first statement of the **Riemann Hypothesis**.

Extending the Riemann Zeta function to $\sigma > 0$ revealed zeros, all of which seem to be on the line $\sigma = 1/2$. Figure 13.1 emphasises these zeros by excluding values of $|\zeta(s)| \geq 1$.

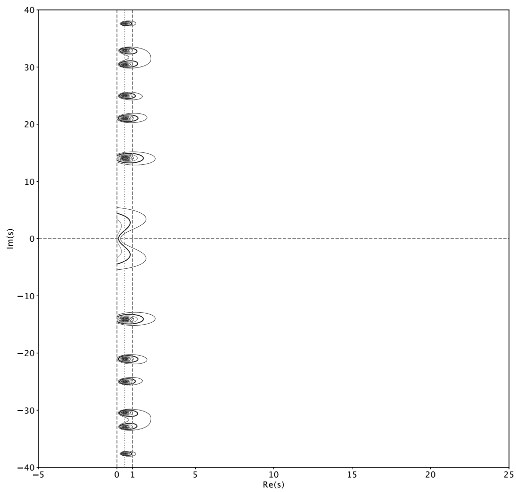


Figure 13.1.: Contour plot of $\ln|\zeta(s)|$ for $\sigma > 0$.

Proving those zeros exist only on the line $\sigma = 1/2$ is the central challenge of the Riemann Hypothesis.

The challenge looks suprisingly trivial. Proving the zeros of a complex function lie on a straight line is surely a textbook exercise?

In fact, the Riemann Hypothesis has resisted proof since it was first proposed by Bernhard Riemann in his seminal 1859 paper, "On the Number of Prime Numbers Less than a Given Quantity".

Like Fermat on his Last Theorem, *"I have discovered a truly marvelous proof of this, which this margin is too narrow to contain"*, Riemann appears not to have foreseen the years of frustrated effort ahead.

> *"... it is very probable that all roots [of a related function] are real. Of course one would wish for a rigorous proof here; I have for the time being, after some fleeting vain attempts, provisionally put aside the search for this ..."*

The Riemann Hypothesis is simple to state, but its importance isn't clear yet. As we continue our journey we'll see the connection between these zeros and the distribution of prime numbers.

14. Pairs of Symmetric Zeros

A common strategy for wrestling with the Riemann Hypothesis is to explore the properties of its zeros. Here we'll take our first steps along this path.

Let's start with the property $\zeta(\overline{s}) = \overline{\zeta(s)}$, which we showed holds for the series $\zeta(s) = \sum 1/n^s$, valid for $\sigma > 1$.

We can show the new series extending $\zeta(s)$ to $\sigma > 0$ also maintains this property. Instead of doing this with slightly laborious algebra, we'll take this opportunity to introduce the elegant and rather powerful principle of **analytic continuation**.

Intuition For Analytic Continuation

Let's first develop an intuition for analytic continuation.

The left of Figure 14.1 shows a complex function $f(z)$ defined on a domain A. It also shows a larger domain B which includes domain A. The complex function $g(z)$ is defined on domain B. Note that $f(x)$ is not valid on the larger domain B.

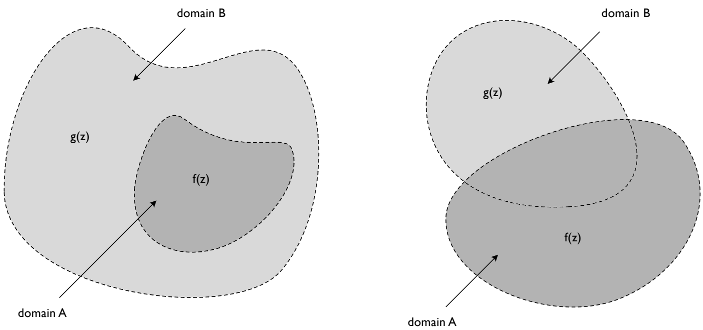


Figure 14.1.: Analytic continuation of $f(z)$.

If $g(z)$ agrees with $f(z)$ on the domain A, then we can see that $g(z)$ is more general than $f(z)$. We call $g(z)$ an **analytic continuation** of $f(z)$.

Let's go a little further. The right of Figure 14.1 shows a domain B which overlaps, but doesn't completely cover, domain A. If $f(z)$ agrees with $g(z)$ where A and B overlap, it isn't too difficult to see how $f(z)$ and $g(z)$ are representations of the same function. If we initially know $f(z)$ and then later find $g(z)$, we can say again $g(z)$ is an analytic continuation of $f(z)$.

So a useful intuition is that if two functions match over a shared domain, they must represent the same function.

For this intuition to be valid, the functions must be well behaved enough to only allow a single unique extension. **Analytic functions** have this good behaviour.

Appendix L defines analytic functions and explains analytic continuation more precisely.

Using Analytic Continuation

Let's construct a function $f(s) = \zeta(s) - \overline{\zeta(\overline{s})}$. Wherever $\zeta(s)$ is analytic, so is $f(s)$. This is because $\overline{\zeta(\overline{s})}$ is analytic too, as explained in Appendix 14.1.

We know $f(s) = 0$ along the real line where $\sigma > 1$. Using analytic continuation, $f(z)$ must also be zero in any domain that $f(s)$ is analytic, as long as that domain includes the real line $\sigma > 1$.

So $f(s) = 0$ in the complex half-plane $\sigma > 1$, but also $\sigma > 0$ because we extended $\zeta(s)$ to this larger domain, where it remains analytic except at $s = 1$.

But $f(s) = 0$ means $\zeta(\overline{s}) = \overline{\zeta(s)}$, which means this property holds in $\sigma > 0$. If later we are able to analytically continue $\zeta(s)$ into $\sigma < 0$, this property will continue to hold there too.

Symmetric Zeros

If $\zeta(s) = 0$ then the property $\zeta(\overline{s}) = \overline{\zeta(s)}$ tells us $\zeta(\overline{s}) = 0$.

$$\zeta(s) = 0 \implies \zeta(\overline{s}) = 0$$

This means the zeros exist in symmetric pairs $\sigma + it$ and $\sigma - it$. That is, they are mirrored above and below the real line, or lie on it.

For a first attempt, this is quite an enlightening insight into the zeros of the Riemann Zeta function $\zeta(s)$.

Zeros existing in symmetric pairs $\sigma \pm it$ is compatible with the Riemann Hypothesis, but sadly it doesn't mean they all lie on a single line $\sigma = a$, never mind the holy grail $\sigma = 1/2$.

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.

$$\boxed{\sum \frac{1}{n} \to \infty}$$

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 multiplied out his new formula.

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

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.

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

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.

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

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. $\sum 1/p$ Diverges

This additional proof that the infinite sum of inverse primes diverges is based on one by Ivan Niven, published in 1971.

You can read the original at `https://www.tandfonline.com/doi/abs/10.1080/00029890.1971.11992740`.

Square-Free Numbers

We can write any counting number m as a unique product of a square j^2 and **square-free** factor k.

$$m = k \cdot j^2$$

Remember that any integer is a unique product of primes. We can split these primes into two groups, one group with primes raised to an even power, which together can be written as a square, and the other group with primes not raised to any power.

For example, $360 = (2 \cdot 5) \cdot (2 \cdot 3)^2$ has a square-free factor of 10, and a square factor of 36. On the other hand, $30 = (2 \cdot 3 \cdot 5)$ is entirely square-free.

Infinite Sum Of Square-Free Reciprocals

Let's look at the following inequality, where k are the square-free integers less than n.

$$\left(\sum_{k<n}\frac{1}{k}\right)\left(\sum_{j<n}\frac{1}{j^2}\right) \geq \sum_{m<n}\frac{1}{m}$$

The inequality is true because multiplying out the two series would give us not just the terms $1/m$, but also many more. The two sides are only equal when $n = 2$.

As $n \to \infty$, the right hand side becomes the harmonic series which we know diverges. We also know the second sum converges to $\pi^2/6$. That means the sum over square-free integers $\sum 1/k$ must diverge, a neat result we'll use very soon.

Infinite Sum Of Prime Reciprocals

Let's assume, perhaps incorrectly, the sum of prime reciprocals $\sum 1/p$ converges to a finite β.

The partial sum is less than the full sum, $\sum_{p<n} 1/p < \beta$, so we can write the following.

$$\exp(\beta) > \exp\left(\sum_{p<n}\frac{1}{p}\right) = \prod_{p<n}\exp(\frac{1}{p})$$

We can also truncate the Taylor series for e^x to say $e^x > 1+x$. Applying this to $\exp(1/p)$ lets us write the following.

$$\prod_{p<n} \exp(\frac{1}{p}) > \prod_{p<n} (1 + \frac{1}{p})$$

Multiplying out that product would give a series with terms $1/k$ where each k is square-free. This is because each prime contributes to any k at most once. We'd also end up with more terms than are in $\sum_{p<n} 1/k$ for $n > 3$.

$$\prod_{p<n} (1 + \frac{1}{p}) \geq \sum_{k<n} \frac{1}{k}$$

Let's put all this together.

$$\exp(\beta) > \exp\left(\sum_{p<n} \frac{1}{p}\right) = \prod_{p<n} \exp(\frac{1}{p}) > \prod_{p<n} (1 + \frac{1}{p}) \geq \sum_{k<n} \frac{1}{k}$$

This suggests that as $n \to \infty$ the finite $\exp(\beta)$ is greater than $\sum 1/k$, which we showed was divergent. This is clearly a contradiction, so our assumption that $\sum 1/p$ converges was wrong.

$$\boxed{\sum \frac{1}{p} \to \infty}$$

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

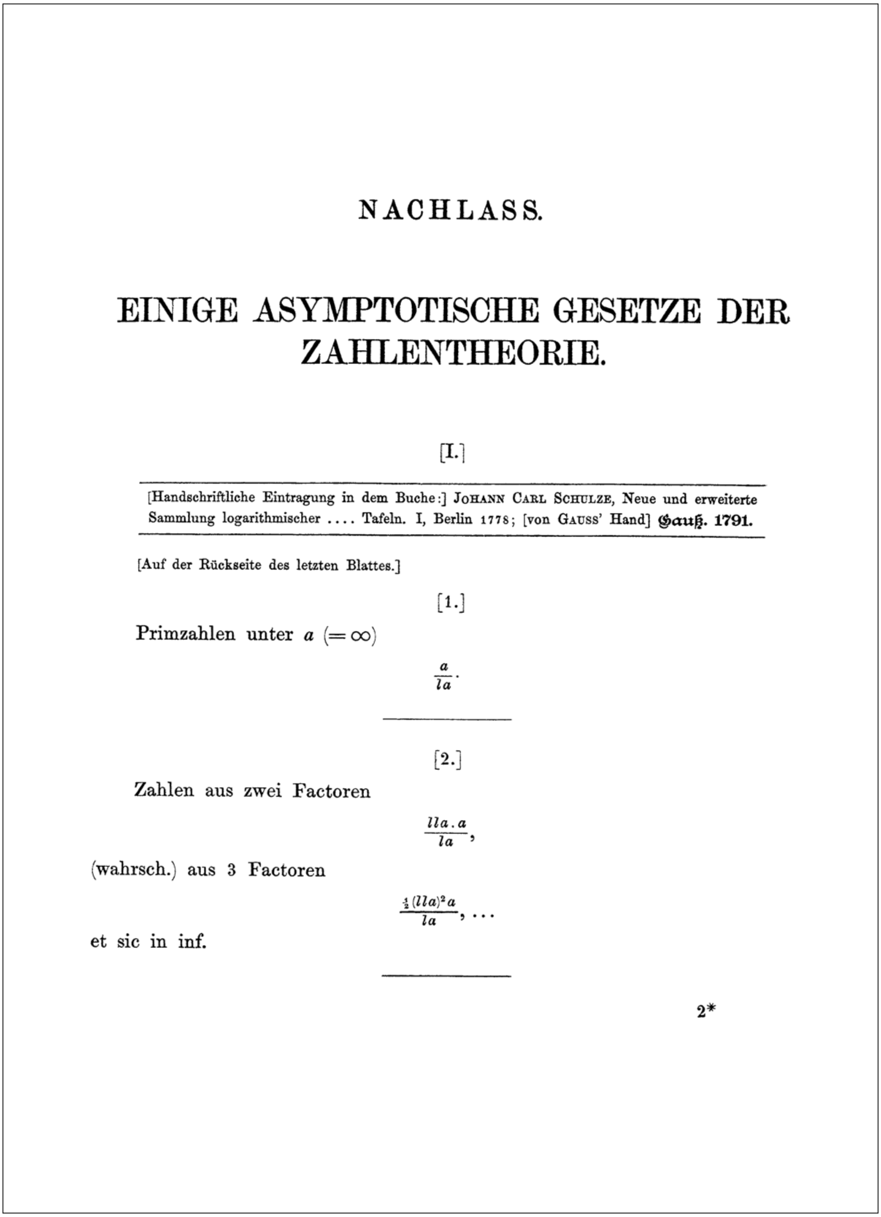
Gauss, 1791

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 D.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` 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.

Figure D.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.

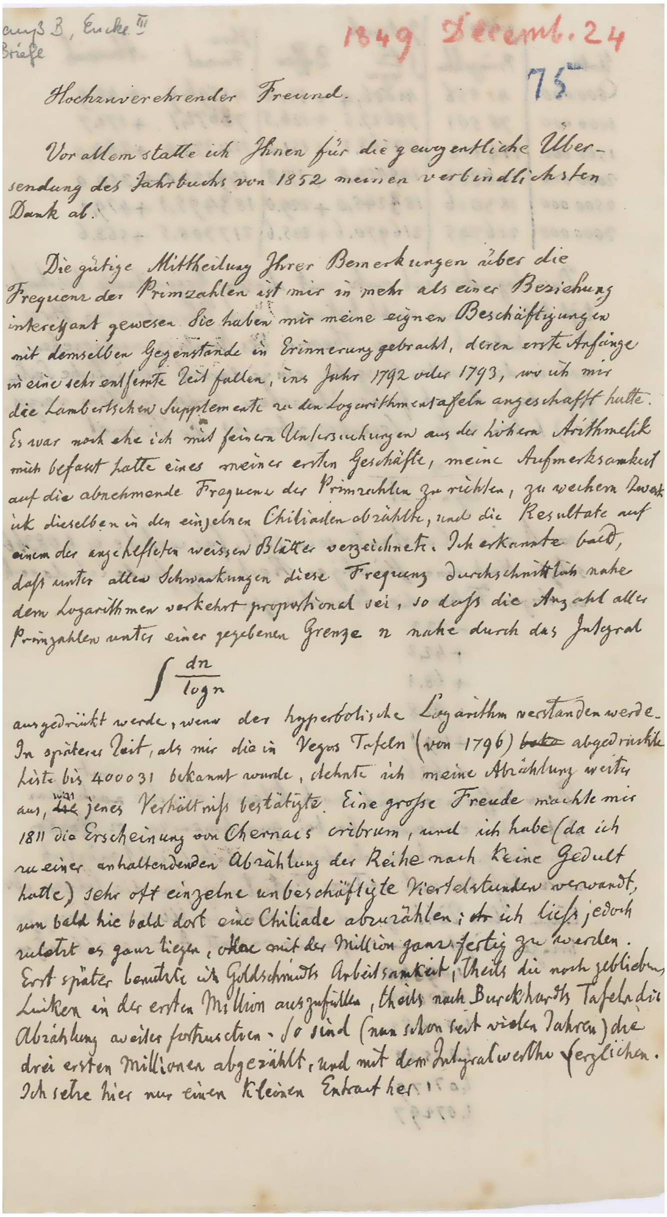
Gauss uses the following expression.

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

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`



Gauss B, Encke III
Briefe

1849 Decemb. 24

75

Hochzuverehrender Freund.

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

Die gütige Mittheilung Ihrer 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) ~~bate~~ 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; doch liess ich 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 D.3.: First page of Gauss' 1849 letter to Encke.

E. Probabilistic Primes

Probabilistic models of primes, despite being built on rather broad simplifications, can match numerical evidence fairly well. Predictions from these models can become worthy conjectures about the primes.

Here we build a particularly simple probabilistic model, and use it to predict the distribution of prime gaps.

Probability of a Prime

The Prime Number Theorem tells us the density of primes around x is approximately $1/\ln(x)$. It isn't a big leap to interpret this density as a probability. So the probability n is prime is simply $1/\ln(n)$.

$$\Pr(n \text{ prime}) = \frac{1}{\ln(n)}$$

Prime Gaps

A prime gap at n of size 4 is a sequence {prime, not prime, not prime, not prime, prime}. The probability of this sequence is

$$\frac{1}{\ln(n)} \cdot \left(1 - \frac{1}{\ln(n+1)}\right) \cdot \left(1 - \frac{1}{\ln(n+2)}\right) \cdot \left(1 - \frac{1}{\ln(n+3)}\right) \cdot \frac{1}{\ln(n+4)}$$

For n much larger than a, we can approximate $\ln(n+a) \approx \ln(n)$, simplifying the probability

$$\left(1 - \frac{1}{\ln(n)}\right)^3 \cdot \left(\frac{1}{\ln(n)}\right)^2$$

We can generalise to prime gaps of size k having a sequence {prime, $k-1$ not prime, prime}.

$$\left(1 - \frac{1}{\ln(n)}\right)^{k-1} \cdot \left(\frac{1}{\ln(n)}\right)^2$$

Prime Gap Counts

Now we know the probability of a gap of size k, we can say the expected number of gaps of that size amongst the first N numbers is approximately the probability summed over all the n up to N.

$$\sum_{n}^{N} \left(1 - \frac{1}{\ln(n)}\right)^{k-1} \cdot \left(\frac{1}{\ln(n)}\right)^2$$

Using another approximation that for most n between 1 and N, $\ln(n) \approx \ln(N)$, this simplifies further.

$$N \cdot \left(1 - \frac{1}{\ln(N)}\right)^{k-1} \cdot \left(\frac{1}{\ln(N)}\right)^2$$

We can take the logarithm of this count to give us a linear function of the gap size k.

$$(k-1) \cdot \ln\left(1 - \frac{1}{\ln(N)}\right) + 2\ln\left(\frac{1}{\ln(N)}\right) + \ln(N)$$

Plotting a graph of the logarithm of the counts for each prime gap k should give us a straight line. The gradient will be negative because it is the logarithm of $\left(1 - \frac{1}{\ln(N)}\right)$, which is always less than 1.

Figure E.1 shows this probabilistic model works surprisingly well given the approximations we applied.

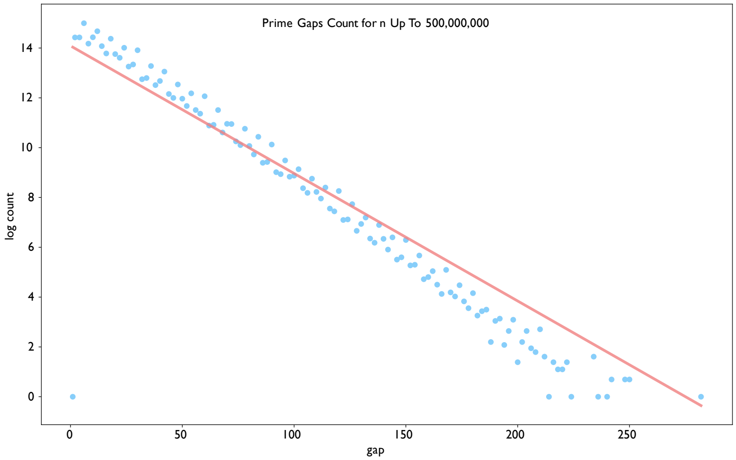


Figure E.1.: Modelling prime gaps in the first 500 million numbers.

Improved Model

Our simple model has many imperfections, and a significant one is that it doesn't take into account that every other number in a given range, the even numbers, are never prime.

If we want to assert that the probability of an odd number being prime is zero, but also preserve the density of primes being $1/\ln(x)$ in the neighbourhood of x, we can double the probability of the odd numbers being prime to $2/\ln(x)$.

The probability of a prime gap of size k now only includes odd numbers in the sequence {prime, $k/2-1$ not prime, prime }. Following the same steps, and noting that we're only applying the probability to half the numbers in the range 1 to N, we have a new estimate for the prime gap counts.

$$\frac{N}{2}\cdot\left(1-\frac{2}{\ln(N)}\right)^{\frac{1}{k}-1}\cdot\left(\frac{2}{\ln(N)}\right)^{2}$$

Figure E.2 shows that for smaller gaps this model is indeed an improvement.

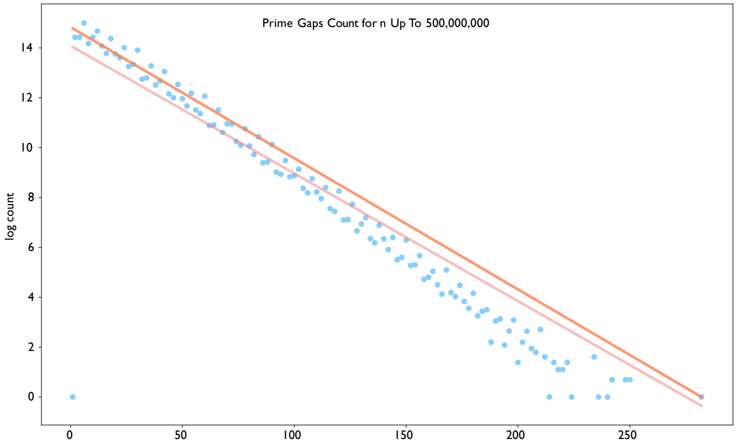


Figure E.2.: Improved model of prime gaps.

The model doesn't work as well for larger gaps, not just because the approximation $\ln(n + a) \approx \ln(n)$ is better for smaller gaps, but mostly because it ignores the fact that multiples of numbers larger than 2 can't be prime, and larger gaps are more likely than smaller gaps to contain these multiples.

F. Integral Comparison

Understanding the behaviour of continuous functions is often easier than discrete functions. We can gain insights into discrete sums like $\sum \frac{1}{x}$ by exploring the related continuous integral $\int \frac{1}{x} dx$.

Lower & Upper Bounds For The Growth Of $\sum 1/n$

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

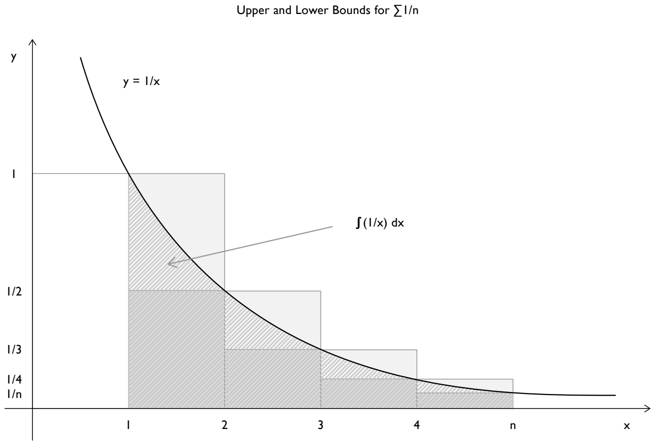


Figure F.1.: Comparing discrete $1/n$ with continuous $1/x$.

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.

Let's now look at the shorter rectangles. 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 rectangles extending to the left of a given x. We can adjust the limit 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.

Convergence Of $\zeta(s) = \sum 1/n^s$

Figure F.2 shows a graph of $y = \frac{1}{x^s}$, together with rectangles representing the fractions $\frac{1}{x^s}$.

The shape of the graph assumes $s > 0$. If s was ≤ 0 then it is easy to see $\sum 1/n^s$ would diverge because each term would be ≥ 1.

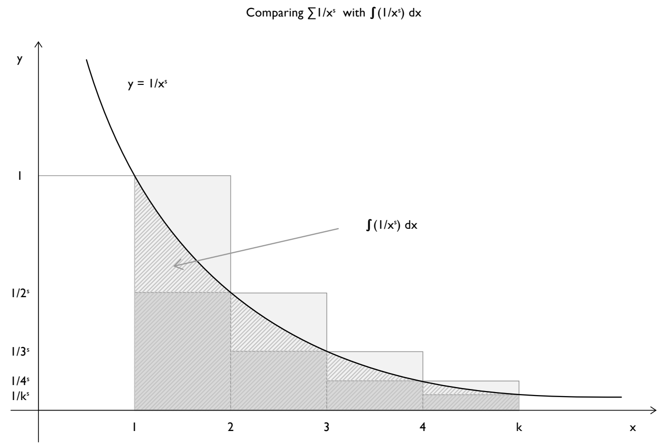


Figure F.2.: Comparing discrete $\sum 1/x^s$ with continuous $\int 1/x^s dx$.

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

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

The sum starts at 2 because we're looking at rectangles extending to the left of a given x. We can adjust the limit of the sum using $\sum_1^k \frac{1}{x^s} = 1 + \sum_2^k \frac{1}{x^s}$.

$$\sum_{1}^{k} \frac{1}{x^s} - 1 < \int_{1}^{k} \frac{1}{x^s} dx$$

The integral is easily evaluated.

$$\sum_{1}^{k} \frac{1}{x^s} < \frac{k^{1-s} - 1}{1 - s} + 1$$

As $k \to \infty$, the right hand side only **converges** when $s > 1$. Because it is less than the right hand side, the sum $\sum 1/x^s$ also converges when $s > 1$. We haven't yet ruled out the possibility the sum might also converge for some $s \leq 1$.

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

$$\sum_{1}^{k} \frac{1}{x^s} > \int_{1}^{k+1} \frac{1}{x^s} dx$$

The integral has an upper limit of $k + 1$ because we're looking at rectangles extending to the right of a given x. We can perform the integral to simplify the expression.

$$\sum_{1}^{k} \frac{1}{x^s} > \frac{(k+1)^{1-s} - 1}{1 - s}$$

As $k \to \infty$, the right hand side **diverges** when $s \leq 1$. Because it is greater than the right hand side, the sum $\sum 1/x^s$ also diverges when

$s \leq 1$. We have now ruled out the possibility the sum might converge for some $s \leq 1$.

$$\boxed{\zeta(s) = \sum 1/n^s \text{ only converges for } s > 1}$$

We can go further. The two inequalities together provide a lower and upper bound for the zeta function.

$$\boxed{\frac{1}{s-1} < \zeta(s) < \frac{1}{s-1} + 1}$$

G. Convergence of Dirichlet Series

The Riemann Zeta series is an example of a Dirichlet series.

$$\zeta(s) = \sum \frac{1}{n^s} = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \ldots$$

Dirichlet series have the general form $\sum a_n/n^s$, in contrast to the more familiar power series $\sum a_n z^n$.

In this appendix we'll explore when these series converge, first looking at absolute convergence, and then more general convergence.

Abscissa Of Absolute Convergence

A series **converges absolutely** even when all its terms are replaced by their magnitudes, sometimes called absolute values. This is quite a strong condition, and not all series that converge do so absolutely.

Let's assume a Dirichlet series converges absolutely at $s_1 = \sigma_1 + it_1$, and consider another point $s_2 = \sigma_2 + it_2$ where $\sigma_2 \geq \sigma_1$. On the complex plane, s_2 is to the right of s_1.

Now let's compare the magnitudes of the terms in this series at s_1 and s_2. Remember $n^{\sigma+it} = n^{\sigma}e^{it\ln n}$, and because the magnitude of any

$e^{i\theta}$ is 1, we can simplify $\left|n^{\sigma+it}\right| = n^{\sigma}$.

$$\sum\left|\frac{a_n}{n^{s_1}}\right| = \sum\frac{|a_n|}{n^{\sigma_1}} \geq \sum\frac{|a_n|}{n^{\sigma_2}} = \sum\left|\frac{a_n}{n^{s_2}}\right|$$

This is simply telling us that the magnitude of each term in the series at s_2 is less than or equal to the magnitude of the same term at s_1. So if the series converges at s_1, it must also converge at s_2. More generally, the series converges at any $s = \sigma + it$ where $\sigma \geq \sigma_1$.

If our series doesn't converge everywhere, the s for which it diverges must therefore have $\sigma < \sigma_1$. We can see there must be a minimum σ_a, called the **abscissa of absolute convergence**, such that the series converges absolutely for all $\sigma > \sigma_a$.

Notice how absolute convergence depends only on the real part of s. Working out the domain of convergence along the real line automatically gives us the domain of convergence in the complex plane.

For example, in Appendix F we showed the series $\sum 1/n^{\sigma}$ converges for real $\sigma > 1$. We also know the series diverges at $\sigma = 1$. These two facts allow us to say $\sigma_a = 1$, and so the series $\sum 1/n^s$ converges for all complex $s = \sigma + it$ where $\sigma > 1$.

It's interesting that the region of convergence for a Dirichlet series is a half-plane, whereas the region for the more familiar power series $\sum a_n z^n$ is a circle.

Abscissa Of Convergence

Absolute convergence is easier to explore as we don't need to consider the effect of complex terms which contribute a negative amount to the overall magnitude of the series. For example, a term $e^{i\pi} = -1$ can partially cancel the effect of a term $2e^{i2\pi} = +2$. This cancelling effect can mean some series do converge, even if not absolutely.

Our strategy, inspired by Apostol, will be to show that if a Dirichlet series is bounded at $s_0 = \sigma_0 + it_0$ then it is also bounded at $s = \sigma + it$, where $\sigma > \sigma_0$, and then push a little further to show it actually converges at that s.

Let's start with a Dirichlet series $\sum a_n/n^s$ that we know has bounded partial sums at a point $s_0 = \sigma_0 + it_0$ for all $x \geq 1$.

$$\left| \sum_{n \leq x} \frac{a_n}{n^{s_0}} \right| \leq M$$

We'll use Abel's partial summation formula, explained in Appendix J, which relates a discrete sum to a continuous integral.

$$\sum_{x_1 < n \leq x_2} b_n f(n) = B(x_2)f(x_2) - B(x_1)f(x_1) - \int_{x_1}^{x_2} B(t)f'(t)dt$$

Because we're comparing to s_0, we'll define $f(x) = x^{s_0 - s}$ and $b_n = a_n/n^{s_0}$. Here $B(x)$ is defined as $\sum_{n \leq x} b_n$, and so $|B(x)| \leq M$.

$$\begin{aligned} \sum_{x_1 < n \leq x_2} \frac{a_n}{n^s} &= \sum_{x_1 < n \leq x_2} b_n f(n) \\ &= \frac{B(x_2)}{x_2^{s-s_0}} - \frac{B(x_1)}{x_1^{s-s_0}} + (s - s_0) \int_{x_1}^{x_2} \frac{B(t)}{t^{s-s_0+1}} dt \end{aligned}$$

We now consider the magnitude of the series, which is never more than the sum of the magnitudes of its parts, and make use of $|B(x)| \leq M$.

$$\left|\sum_{x_1<n\leq x_2}\frac{a_n}{n^s}\right| \leq \left|\frac{B(x_2)}{x_2^{s-s_0}}\right| + \left|\frac{B(x_1)}{x_1^{s-s_0}}\right| + \left|(s-s_0)\int_{x_1}^{x_2}\frac{B(t)}{t^{s-s_0+1}}dt\right|$$
$$\leq Mx_2^{\sigma_0-\sigma} + Mx_1^{\sigma_0-\sigma} + |s-s_0|\,M\int_{x_1}^{x_2}t^{\sigma_0-\sigma-1}dt$$

Because $x_1 < x_2$, we can say $Mx_2^{\sigma_0-\sigma} + Mx_1^{\sigma_0-\sigma} < 2Mx_1^{\sigma_0-\sigma}$ for $\sigma > \sigma_0$. Despite appearances, evaluating the integral is easy.

$$\left|\sum_{x_1<n\leq x_2}\frac{a_n}{n^s}\right| \leq 2Mx_1^{\sigma_0-\sigma} + |s-s_0|\,M\left(\frac{x_2^{\sigma_0-\sigma}-x_1^{\sigma_0-\sigma}}{\sigma_0-\sigma}\right)$$
$$\leq 2Mx_1^{\sigma_0-\sigma}\left(1+\frac{|s-s_0|}{\sigma-\sigma_0}\right)$$

The last step uses $\left|x_2^{\sigma_0-\sigma} - x_1^{\sigma_0-\sigma}\right| = x_1^{\sigma_0-\sigma} - x_2^{\sigma_0-\sigma} < x_1^{\sigma_0-\sigma} < 2x_1^{\sigma_0-\sigma}$.

This tells us that $\sum_{x_1<n\leq x_2} a_n/n^s$ is bounded if $\sum_{n\leq x} a_n/n^{s_0}$ is bounded, where $\sigma > \sigma_0$.

Let's see if we can push this result about boundedness to convergence.

$$\left|\sum_{x_1<n\leq x_2}\frac{a_n}{n^s}\right| \leq 2Mx_1^{\sigma_0-\sigma}\left(1+\frac{|s-s_0|}{\sigma-\sigma_0}\right) = Kx_1^{\sigma_0-\sigma}$$

Here K doesn't depend on x_1. If we let $x_1 \to \infty$ then $Kx_1^{\sigma_0-\sigma} \to 0$, which means the magnitude of the tail of the infinite sum $\sum a_n/n^s$ diminishes to zero, and so the series is not just bounded, it also converges.

Let's summarise our results so far:

- If $\sum_{n \le x} a_n/n^{s_0}$ is bounded, then $\sum a_n/n^s$ converges for $\sigma > \sigma_0$.
- With the special case of $s_0 = 0$, if $\sum_{n \le x} a_n$ is bounded, then $\sum a_n/n^s$ converges for $\sigma > 0$.

The special case is particularly useful as we can sometimes say whether a series converges for $\sigma > 0$ just by looking at the coefficients a_n.

Following the same logic as for σ_a, it is clear there is an abscissa of convergence σ_c where a Dirichlet series converges for $\sigma > \sigma_c$, and diverges for $\sigma < \sigma_c$.

Maximum Difference Between σ_c And σ_a

We know that not all convergent series are absolutely convergent, so we can say $\sigma_a \ge \sigma_c$. We shouldn't have to increase σ by too much before a conditionally convergent series converges absolutely.

If a series converges at s_0, the magnitude of terms is bounded. We can call this bound C.

$$\sum \left| \frac{a_n}{n^s} \right| = \sum \left| \frac{a_n}{n^{s_0}} \cdot \frac{1}{n^{s-s_0}} \right| \le C \sum \frac{1}{n^{\sigma - \sigma_0}}$$

We know series of the form $\sum 1/n^{\sigma-\sigma_0}$ only converge for $\sigma - \sigma_0 > 1$, so we can say if σ is larger than σ_c by at least 1, the series converges absolutely.

$$\boxed{0 \le \sigma_a - \sigma_c \le 1}$$

Example: Alternating Zeta Function

Let's apply our results to the **alternating zeta function**, also called the **eta function**.

$$\eta(s) = \sum \frac{(-1)^{n+1}}{n^s} = \frac{1}{1^s} - \frac{1}{2^s} + \frac{1}{3^s} - \frac{1}{4^s} + \ldots$$

At $s_0 = 0$ we can see by simple inspection that the partial sum $\sum_{n \leq x} (-1)^{n+1}$ oscillates but is always bounded ≤ 1, and so $\eta(s)$ converges for $\sigma > 0$.

H. Swapping $\lim \sum$ For $\sum \lim$

Let's consider what happens to the Riemann Zeta function $\zeta(s)$ as $\sigma \to +\infty$.

$$\lim_{\sigma \to \infty} \sum_n \frac{1}{n^s} = \lim_{\sigma \to \infty} \left(\frac{1}{1^s} + \frac{1}{2^s} + \dots \right)$$

It's tempting to look at each term and notice that $|n^{-s}| = n^{-\sigma} \to 0$ as $\sigma \to \infty$ for all n except $n = 1$, then conclude $\zeta(s) \to 1$ as $\sigma \to \infty$. In effect, we've taken the limit inside the sum.

$$\sum_n \lim_{\sigma \to \infty} \frac{1}{n^s} = \lim_{\sigma \to \infty} \left(\frac{1}{1^s} \right) + \lim_{\sigma \to \infty} \left(\frac{1}{2^s} \right) + \dots$$

However, the limit of an infinite sum is not always the sum of the limits. Tannery's Theorem tells us when we can safely swap sum and limit operators.

Tannery's Theorem

The theorem has three requirements

- An infinite sum $S_j = \sum_k f_k(j)$ that converges
- The limit $\lim_{j \to \infty} f_k(j) = f_k$ exists

- An $M_k \geq |f_k(j)|$ independent of j, where $\sum_k M_k$ converges

If the requirements are met, we can take the limit inside the sum.

$$\lim_{j \to \infty} \sum_k f_k(j) = \sum_k \lim_{j \to \infty} f_k(j)$$

Proof

Let's first show the sum of limits actually exists.

By definition, $|f_k(j)| \leq M_k$, and $\sum_k M_k$ converges. Taking $j \to \infty$ gives us $|f_k| \leq M_k$, and so $\sum_k |f_k|$ converges, which in turn means $\sum_k f_k$ converges absolutely. This quantity is the sum of limits.

Now let's show the limit of the sum is the sum of the limits.

Since $\sum_k M_k$ converges there must be an N so that $\sum_{k=N} M_k < \epsilon$, where ϵ is as small as we require.

This gives us a useful inequality.

$$\left| \sum_{k=N} f_k(j) \right| \leq \sum_{k=N} |f_k(j)| \leq \sum_{k=N} M_k < \epsilon$$

Taking $j \to \infty$ gives us a similar inequality.

$$\left| \sum_{k=N} f_k \right| \leq \sum_{k=N} |f_k| \leq \sum_{k=N} M_k < \epsilon$$

Let's now consider the absolute difference between $\sum_k f_k(j)$ and $\sum_k f_k$. Although the following looks complicated, it is simply splitting the sums over $[0,\infty]$ into sums over $[0,N-1]$ and $[N,\infty]$.

$$\begin{aligned}
\left|\sum_k f_k(j) - \sum_k f_k\right| &= \left|\sum_k^{N-1} f_k(j) + \sum_{k=N} f_k(j) - \sum_k^{N-1} f_k - \sum_{k=N} f_k\right| \\
&\leq \left|\sum_{k=N} f_k(j)\right| + \left|\sum_{k=N} f_k\right| + \left|\sum_k^{N-1} f_k(j) - \sum_k^{N-1} f_k\right| \\
&< 2\epsilon + \left|\sum_k^{N-1} (f_k(j) - f_k)\right|
\end{aligned}$$

As $j \to \infty$, the finite sum $\sum_k^{N-1} (f_k(j) - f_k) \to 0$, which leaves a simpler inequality.

$$\lim_{j\to\infty} \left|\sum_k f_k(j) - \sum_k f_k\right| < 2\epsilon$$

Because ϵ can be as small as we require, we finally have $\lim_{j\to\infty} \sum_k f_k(j) = \sum_k f_k$, which proves the theorem.

$$\lim_{j\to\infty} \sum_k f_k(j) = \sum_k f_k = \sum_k \lim_{j\to\infty} f_k(j)$$

Application To $\zeta(s)$

Let's apply Tannery's Theorem to $\zeta(s)$. Here $f_k(j)$ is written as $f_n(s) = 1/n^s$.

We start with the convergent infinite sum.

$$\zeta(s) = \sum_n \frac{1}{n^s} \text{ converges for } \sigma > 1$$

We confirm $f_n(s)$ exists when $\sigma \to \infty$.

$$\lim_{\sigma \to \infty} \frac{1}{n^s} = f_n = \begin{cases} 1 & n = 1 \\ 0 & n > 1 \end{cases}$$

We also find an $M_n \geq |f_n(s)|$ independent of σ.

$$\left| \frac{1}{n^s} \right| = \frac{1}{n^\sigma} \leq M_n = \frac{1}{n^\alpha}$$

Here $1 < \alpha \leq \sigma$. The sum $\sum_n M_n$ converges because $\alpha > 1$.

The criteria have been met, so we can safely move the limit inside the sum.

$$\lim_{\sigma \to \infty} \sum_n \frac{1}{n^s} = \sum_n \lim_{\sigma \to \infty} \frac{1}{n^s} = 1 + 0 + 0 + \ldots$$

So $\zeta(s) \to 1$, as $\sigma \to +\infty$.

I. Infinite Products

Let's first develop an intuition for infinite products through some examples.

$$2 \times 3 \times 4 \times 5 \times \dots$$

It is easy to see the above product diverges. Each factor increases the size of the product.

$$2 \times 0 \times 4 \times 5 \times \dots$$

It is a fundamental idea that multiplying by zero causes a product to be zero. The product is zero because one of the factors is zero.

$$\frac{1}{2} \times \frac{1}{3} \times \frac{1}{4} \times \frac{1}{5} \times \dots$$

This product is more interesting. Each factor is a fraction that reduces the size of the product. As the number of these reducing factors grows, the product gets ever closer to zero. We can make the leap to say the value of the infinite product is zero.

We have found two different ways for the product to be zero. We'll need to keep both in mind as we work with infinite products.

Definition

Similar to **infinite series**, we say an **infinite product** converges if the limit of the partial products is a **finite** value.

$$\lim_{N\to\infty} \prod_{n=1}^{N} a_n = P$$

We'll see why it is conventional to insist the finite value is **non-zero**.

Examples

Does the infinite product $\prod_{n=1}^{\infty} (1 + 1/n)$ converge?

Each factor $(1 + 1/n)$ is larger than one, so we expect the product to keep growing. Let's be more rigorous and see how the partial products actually grow.

$$\begin{aligned}
\prod_{n=1}^{N} \left(1 + \frac{1}{n}\right) &= \prod_{n=1}^{N} \left(\frac{n+1}{n}\right) \\
&= \frac{\cancel{2}}{1} \times \frac{\cancel{3}}{\cancel{2}} \times \frac{\cancel{4}}{\cancel{3}} \times \ldots \times \frac{N+1}{\cancel{N}} \\
&= N + 1
\end{aligned}$$

As $N \to \infty$, the product **diverges**.

Lets now look at the the similar infinite product $\prod_{n=2}^{\infty} (1 - 1/n)$. Notice

n starts at 2 to ensure the first factor is not zero.

Each factor $(1 - 1/n)$ is smaller than one, so we expect the product to keep shrinking. Let's see if the partial products do indeed get smaller.

$$\prod_{n=2}^{N} \left(1 - \frac{1}{n}\right) = \prod_{n=2}^{N} \left(\frac{n-1}{n}\right)$$

$$= \frac{1}{\cancel{2}} \times \frac{\cancel{2}}{\cancel{3}} \times \frac{\cancel{3}}{\cancel{4}} \times \ldots \times \frac{\cancel{N+1}}{N}$$

$$= \frac{1}{N}$$

As $N \to \infty$, the product tends to zero. Remember that for convergence we insist the limit is non-zero. For this reason we say the product **diverges to zero**.

Convergence And a_n

We know that for an infinite series $\sum a_n$ to converge, the terms a_n must $\to 0$. For an infinite product $\prod a_n$ to converge, the terms $a_n \to 1$.

If each term a_n was larger than 1, the product would get ever larger. If each term a_n was smaller than 1, the product would get ever smaller towards zero. Negative a_n cause partial product to oscillate, meaning convergence only happens if $a_n \to 1$.

Removing Zero-Valued Factors

A single zero-valued factor collapses an entire product to zero. If an infinite product has a **finite** number of zero-valued factors, they can be removed to leave a potentially interesting different product.

For example, the following product is zero because the first factor is zero.

$$\prod_{n=1}\left(1-\frac{1}{n^2}\right)=0$$

Removing the first factor leaves a much more interesting product.

$$\prod_{n=2}\left(1-\frac{1}{n^2}\right)=\frac{1}{2}$$

Convergence Criterion 1

Since the terms in a convergent infinite product tend to 1, it is useful to write the factors as $(1+a_n)$.

$$P=\prod(1+a_n)$$

We can turn a product into a sum by taking the logarithm.

$$\ln(P)=\ln\prod(1+a_n)=\sum\ln(1+a_n)$$

Using $1+x\leq e^x$ we arrive at a nice inequality.

$$\ln(P) \leq \sum a_n$$

This tells us that if the sum is bounded, the product is bounded too. If the terms a_n are always positive, then the sum can only grow monotonically (without oscillation), so the boundedness is convergence. This is a useful result but we can strengthen it.

Expanding out the product $\prod(1 + a_n)$ gives us a sum which includes the terms 1, all the individual a_n, and also the combinations of different a_n multiplied together. This gives us an inequality for $\sum a_n$ in the other direction.

$$1 + \sum a_n \leq \prod (1 + a_n) = P$$

This tell us that if the product converges, so does the sum. The two results together give us our first convergence criterion.

$$\boxed{\sum a_n \text{ converges } \Leftrightarrow \prod (1 + a_n) \text{ converges, for } a_n > 0}$$

This allows us to say $\prod(1 + 1/n)$ diverges because $\sum 1/n$ diverges, and $\prod(1 + 1/n^2)$ converges because $\sum 1/n^2$ converges.

Convergence Criterion 2

A very similar argument that uses $1 - x \leq e^{-x}$ leads to a criterion for products of the form $\prod(1 - a_n)$.

$$\boxed{\sum a_n \text{ converges } \Leftrightarrow \prod (1 - a_n) \text{ converges, for } 0 < a_n < 1}$$

So $\prod_2^\infty (1 - 1/n)$ diverges, because $\sum 1/n$ diverges.

Divergence To Zero

The logarithmic view of infinite products is useful because it turns a product into a sum, but it has an interesting side effect.

If the partial products tend to zero, then the logarithm diverges towards $-\infty$. This is why we say a **product diverges to zero**.

Convergence Criterion 3

The previous convergence criteria are for real values of a_n. It would be useful to have a criterion for complex a_n. To do that we need an intermediate result about $|a_n|$.

For complex a_n, we have $|a_n| > 0$ for all $a_n \neq 0$. This gives is an intermediate result.

$$\sum |a_n| \text{ converges } \Leftrightarrow \prod (1 + |a_n|) \text{ converges}$$

We are interested in products $\prod(1 + a_n)$ with complex a_n, not just $\prod(1 + |a_n|)$. Let's start with two partial products with complex a_n.

$$p_N = \prod^N (1 + a_n)$$

$$q_N = \prod^N (1 + |a_n|)$$

We need to assert $a_n \neq -1$ to ensure no zero-valued factors $(1 + a_n)$.

For $N > M \geq 1$, we can compare $|p_N - p_M|$ with $|q_N - q_M|$ with a little algebra.

$$\begin{aligned}
|p_N - p_M| &= |p_M| \cdot \left| \frac{p_N}{p_M} - 1 \right| \\
&= |p_M| \cdot \left| \prod_{M+1}^{N} (1 + a_n) - 1 \right| \\
&\leq |q_M| \cdot \left| \prod_{M+1}^{N} (1 + |a_n|) - 1 \right| \\
&= |q_M| \cdot \left| \frac{q_N}{q_M} - 1 \right| \\
|p_N - p_M| &\leq |q_N - q_M|
\end{aligned}$$

If $|q_N - q_M| < \epsilon$, where ϵ is as small as we want, then $|p_N - p_M| < \epsilon$ too. This the Cauchy criterion for convergence, and it tells us that if q_N converges, so does p_N.

So $\sum |a_n|$ converges means $\prod(1+|a_n|)$ converges, which we can now say means $\prod(1 + a_n)$ also converges. We finally have our third convergence criterion.

$$\boxed{\sum |a_n| \text{ converges} \implies \prod(1 + a_n) \text{ converges, for } a_n \neq -1}$$

Notice this criterion is one way. We can't say the sum converges if the product converges.

Why Convergence Is Non-Zero

Let's see why convergence according to these criteria mean the products converge to a **non-zero** value. We've already seen how $a_n \to 0$, which means that $|a_n| < 1/2$ except for a finite number of terms.

We use the useful inequality $1 + x \leq e^x$ again.

$$1 \leq \prod(1 + |a_n|) < e^{\sum |a_n|}$$

This tells us that if the sum $\sum |a_n|$ converges, then the product $\prod(1 + |a_n|)$ converges and is non-zero.

We can use another inequality $1 - x \geq e^{-2x}$ for $0 \leq x \leq 1/2$, and that $e^y > 0$ for all real y.

$$0 < e^{-2\sum |a_n|} \leq \prod(1 - |a_n|) \leq 1$$

This tells us that if the sum $\sum |a_n|$ converges, then the product $\prod(1 - |a_n|)$ converges and is non-zero.

Now we use another inequality $1 - |a_n| \leq |1 \pm a_n| \leq 1 + |a_n|$ to relate $\prod(1 - |a_n|)$ to $\prod(1 - a_n)$.

$$\prod(1 - |a_n|) \leq \left|\prod(1 \pm a_n)\right| \leq \prod(1 + |a_n|)$$

Assuming $\sum |a_n|$ converges, we can finally say $\prod(1 \pm a_n)$ is non-zero, because its absolute value is between two known non-zero values.

Riemman Zeta Function

The Riemann Zeta function can be written as an infinite product over primes. Here $s = \sigma + it$, and $\sigma > 1$.

$$\zeta(s) = \sum \frac{1}{n^s} = \prod \left(1 - \frac{1}{p^s}\right)^{-1}$$

It is natural to ask if $\zeta(s)$ has any zeros in the domain $\sigma > 1$.

None of the factors $(1 - 1/p^s)^{-1}$ is zero. That would require p^s to be zero, and that isn't possible.

$$|p^s| = \left|e^{s\ln(p)}\right| = e^{\sigma\ln(p)} > 0$$

We also need to check the infinite product doesn't diverge to zero. For the moment let's consider $1/\zeta(s) = \prod(1 - 1/p^s)$. Using the third convergence criterion, we check whether $\sum |1/p^s|$ converges.

$$\sum \left|\frac{1}{p^s}\right| = \sum \frac{1}{p^\sigma} \leq \sum \frac{1}{n^\sigma}$$

The reason $\sum 1/p^\sigma \leq \sum 1/n^\sigma$ is because there are fewer primes p than integers n. Because $\sum 1/n^\sigma$ converges for $\sigma > 1$, so does $\sum |1/p^s|$. This means $1/\zeta(s)$ converges to a non-zero value, and therefore so does $\zeta(s)$.

We can now say the Riemann Zeta function has no zeros in the domain $\sigma > 1$.

J. Abel's Partial Summation Formula

It is often easier to understand how a discrete function behaves if it can be expressed as a continuous function. **Abel's partial summation formula** allows us to write a discrete sum as continuous integral.

A Useful Object

The following is a sum over an arithmetic function $a(n)$ weighted by a smooth function $f(x)$.

$$\sum_{x_1<n\leq x_2} a(n)f(n)$$

This is a useful general object to find an integral form for, because it gives us flexibility in choosing the arithmetic and smooth functions.

To be more precise about the functions, $a(n)$ takes only positive integers $n \geq 1$, and $f(x)$ has a continuous derivative over the domain $[x_1, x_2]$. Both $a(n)$ and $f(x)$ can be complex.

Deriving An Integral Form

If you look ahead, the derivation of an integral form for the sum might look overwhelming, but it is just lots of simple algebra.

Because $a(n)$ is only defined over integers $n \geq 1$, we can clarify the sum by setting $m_1 = \lfloor x_1 \rfloor$ and $m_2 = \lfloor x_2 \rfloor$. Remember that $\lfloor x \rfloor$ is the largest integer up to, and including, x.

$$\sum_{x_1 < n \leq x_2} a(n)f(n) = \sum_{n=m_1+1}^{m_2} a(n)f(n)$$

Let's define $A(x) = \sum_{n \leq x} a(n)$. By definition $a(n) = A(n) - A(n-1)$ so we can replace $a(n)$.

$$\begin{aligned}\sum_{m_1+1}^{m_2} a(n)f(n) &= \sum_{m_1+1}^{m_2} \left[A(n) - A(n-1)\right] f(n) \\ &= \sum_{m_1+1}^{m_2} A(n)f(n) - \sum_{m_1}^{m_2-1} A(n)f(n+1)\end{aligned}$$

The two sums have different limits for n, but both cover $[m_1+1, m_2-1]$.

$$\begin{aligned}\sum_{m_1+1}^{m_2} a(n)f(n) = \sum_{m_1+1}^{m_2-1} & A(n)\left[f(n) - f(n+1)\right] \\ & + A(m_2)f(m_2) - A(m_1)f(m_1+1)\end{aligned}$$

Noticing that $\int_n^{n+1} f'(t)dt = f(n+1) - f(n)$ allows us to introduce the integral.

$$\sum_{m_1+1}^{m_2} a(n)f(n) = -\sum_{m_1+1}^{m_2-1} A(n)\int_n^{n+1} f'(t)dt$$
$$+ A(m_2)f(m_2) - A(m_1)f(m_1+1)$$

Now, because $A(t) = A(n)$ over the interval $[n, n+1)$, we can move $A(n)$ it inside the integral as $A(t)$.

$$\sum_{m_1+1}^{m_2} a(n)f(n) = -\sum_{m_1+1}^{m_2-1}\int_n^{n+1} A(t)f'(t)dt$$
$$+ A(m_2)f(m_2) - A(m_1)f(m_1+1)$$

That sum of integrals over consecutive intervals can be simplified to a single integral.

$$\sum_{x_1<n\leq x_2} a(n)f(n) = -\int_{m_1+1}^{m_2} A(t)f'(t)dt$$
$$+ A(m_2)f(m_2) - A(m_1)f(m_1+1)$$

We now need to adjust the integration limits back to x_1 and x_2, not forgetting the intervals to m_1 and m_2. Writing out the integrals that split the interval $[x_1, x_2]$ is helpful.

$$\int_{x_1}^{x_2} X dt = \int_{m_1+1}^{m_2} X dt + \int_{x_1}^{m_1+1} X dt + \int_{m_2}^{x_2} X dt$$

Rearranging, and noticing that $A(t) = A(x_1)$ over the interval $[x_1, m_1 + 1)$, and $A(t) = A(x_2)$ over the interval $[m_2, x_2]$, gives us the following.

$$\begin{aligned}
-\int_{m_1+1}^{m_2} A(t)f'(t)dt &= \int_{x_1}^{m_1+1} A(t)f'(t)dt + \int_{m_2}^{x_2} A(t)f'(t)dt \\
&\quad - \int_{x_1}^{x_2} A(t)f'(t)dt \\
&= A(x_1)\left[f(m_1+1) - f(x_1)\right] + A(x_2)\left[f(x_2) - f(m_2)\right] \\
&\quad - \int_{x_1}^{x_2} A(t)f'(t)dt
\end{aligned}$$

We plug this integral back into our object, then use $A(m_1) = A(x_1)$ and $A(m_2) = A(x_2)$.

$$\begin{aligned}
\sum_{x_1 < n \leq x_2} a(n)f(n) &= A(x_2)f(m_2) - A(x_1)f(m_1+1) \\
&\quad + A(x_1)\left[f(m_1+1) - f(x_1)\right] + A(x_2)\left[f(x_2) - f(m_2)\right] \\
&\quad - \int_{x_1}^{x_2} A(t)f'(t)dt
\end{aligned}$$

Many of these terms cancel out, leaving us with the much neater **Abel Identity**, as it is also called.

$$\boxed{\sum_{x_1 < n \leq x_2} a(n)f(n) = A(x_2)f(x_2) - A(x_1)f(x_1) - \int_{x_1}^{x_2} A(t)f'(t)dt}$$

In many cases n starts at 1, and so the formula reduces further.

$$\boxed{\sum_{1\leq n\leq x_2} a(n)f(n) = A(x_2)f(x_2) - \int_1^{x_2} A(t)f'(t)dt}$$

The lower limit of the integral is 1 because $A(t) = 0$ in the range $[0, 1)$.

Example: Growth of $\sum 1/n$

Abel's identity is particularly useful for clarifying the asymptotic behaviour of discrete functions. Let's illustrate its use to show the how the harmonic series $\sum_1^N 1/n$ grows.

We can choose $a(n) = 1$ and $f(x) = 1/x$, which means $A(x) = \lfloor x \rfloor$ and $f'(x) = -1/x^2$.

$$\begin{aligned}\sum_{n\leq N} \frac{1}{n} &= A(N)f(N) - \int_1^N A(t)f't) \\ &= \frac{N - \{N\}}{N} + \int_1^N \frac{t - \{t\}}{t^2} dt\end{aligned}$$

We've used $\lfloor x \rfloor = x - \{x\}$, where $\{x\}$ is the fractional part of x. This also lets us split the integral into manageable parts.

$$\sum_{0<n\leq N} \frac{1}{n} = 1 + \mathcal{O}\left(\frac{1}{N}\right) + \int_1^N \frac{1}{t} dt - \int_1^N \frac{\{t\}}{t^2} dt$$

Because $\{t\}$ is only ever in the range $[0, 1)$, the last integral is always less than $\int_1^N 1/t^2 dt$, that is, $\mathcal{O}\left([-1/t]_1^N\right)$.

$$
\begin{aligned}
\sum_{0<n\leq N} \frac{1}{n} &= 1 + \mathcal{O}\left(\frac{1}{N}\right) + \ln(N) + \mathcal{O}\left(1 - \frac{1}{N}\right) \\
&= \ln(N) + \mathcal{O}\left(1\right)
\end{aligned}
$$

This tells us the harmonic series grows like $\ln(N)$, a powerful result derived easily using the Abel identity. It also tells us the difference is bounded by 1. In fact, the difference tends to the Euler–Mascheroni constant $\gamma \approx 0.5772$ which pops up in several areas of number theory and analysis.

K. $\zeta(s)$ Has One Pole In $\sigma > 0$

The Riemann Zeta function represented by the series $\zeta(s) = \sum 1/n^s$ converges for $\sigma > 1$, and therefore has no poles in that domain.

We developed a new series for $\zeta(s)$ based on the eta function $\eta(s)$.

$$\zeta(s) = \frac{1}{1 - 2^{1-s}} \eta(s)$$

Because $\eta(s)$ converges for $\sigma > 0$ (Appendix G), any divergence must come from the factor $(1 - 2^{1-s})^{-1}$. The denominator $(1 - 2^{1-s})$ is zero at $s = 1 + 2\pi ia/\ln(2)$ for integer a, so the factor diverges at all these points.

Visualising $\zeta(s)$ suggested it had only one pole at $s = 1 + 0i$. If true, this would require $\eta(s)$ to have zeros at $s = 1 + 2\pi ia/\ln(2)$ for integers $a \neq 0$ to cancel out the other poles from $(1 - 2^{1-s})^{-1}$.

To prove this directly isn't easy, but there is a nice indirect path.

Yet Another Series for $\zeta(s)$

We start with a specially constructed Dirichlet series.

$$X(s) = \frac{1}{1^s} + \frac{1}{2^s} - \frac{2}{3^s} + \frac{1}{4^s} + \frac{1}{5^s} - \frac{2}{6^s} + \ldots$$

The pattern can be exploited to find yet another series for $\zeta(s)$.

$$\zeta(s) - X(s) = \frac{3}{3^s} + \frac{3}{6^s} + \frac{3}{9^s} + \ldots$$

$$= \frac{3}{3^s}\zeta(s)$$

$$\zeta(s) = \frac{1}{1 - 3^{1-s}} X(s)$$

Comparing Potential Poles

Since $X(s)$ converges for $\sigma > 0$ (Appendix G), any divergence must come from the factor $(1 - 3^{1-s})^{-1}$. The denominator $(1 - 3^{1-s})$ is zero when $s = 1 + 2\pi i b / \ln(3)$ for integer b.

We can equate the two expressions for where the poles of $\zeta(s)$ could be.

$$1 + \frac{2\pi i a}{\ln(2)} = 1 + \frac{2\pi i b}{\ln(3)}$$

$$\frac{a}{b} = \frac{\ln(2)}{\ln(3)}$$

There are no non-zero integers a and b which satisfy this because $\ln(2)/\ln(3)$ is irrational.

This leaves us with $s = 1 + 0i$ as the only pole for $\zeta(s)$ in the domain $\sigma > 0$.

L. Analytic Continuation

Let's first develop an intuitive understanding of analytic continuation before we discuss it more precisely.

Analytic Functions

Let's remind ourselves of some definitions from complex analysis.

An **analytic function** is one that can be represented by a convergent Taylor series at every point in its domain.

Because Taylor series can be differentiated infinitely, analytic functions are not just **continuous**, they are **smooth**. Smoothness is a stronger property because continuous just means no gaps.

If a complex function is analytic, then it is also infinitely differentiable. The reverse is also true, if a complex function is infinitely differentiable, then it is analytic.

A differentiable complex function is called **holomorphic**. The terms analytic and holomorphic are used interchangeably for complex functions because of this equivalence, but they do have distinct historical origins.

The equivalence of smoothness (infinitely differentiable) and analyticity (Taylor series) is a special property of complex functions. It is not always true for real functions.

Example 1 SHow zeta series is analytic.

Example 2 sss $\overline{\zeta(\overline{s})}$ ss

ss