The Divergence of $\sum \frac{1}{p}$

S. Ziskind

1 Overview

Two famous facts about infinite series are that $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges, but $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges. Both follow from the integral test, using $\int_1^{\infty} \frac{1}{x}$ and $\int_1^{\infty} \frac{1}{x^2}$ respectively.

The convergence can be viewed as a statement about the rarity of perfect squares, namely that if the reciprocals of all but the squares are removed, then so few remain that the previously divergent series now converges.

The same question can be asked about the rarity of the prime numbers: if the reciprocals of the composites are removed, are the remaining inverse primes convergent? Euler showed the answer to be: NO. As before, this can be viewed as a statement about the rarity of primes. Unlike the perfect squares, the primes occur frequently enough that removing the composites from $\sum \frac{1}{n}$ does not result in convergence. There are still plenty of primes.

The proof of this theorem proceeds in three steps. First, we establish a remarkable identity known as the Euler Product Formula, which equates an infinite sum to an infinite product taken over just the prime numbers. Second, we will justify taking the limit of these two things, both depending on a parameter s, as $s \to 1^+$. Finally, the resulting identity will be combined with a simple calculus fact to reach the desired conclusion.

Each step will be considered separately. After the proof we will consider how fast the sum is growing.

2 The Euler Product Formula

The Euler Product Formula begins with a function known as the Zeta Function, defined for $s > 1$ as this infinite sum:

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

This function plays a key role in many deep theorems of number theory. It was actually invented by Euler in the form shown here, but a century later Riemann extended the domain of definition to accept all complex numbers, hence it is now called the Riemann Zeta Function. We will use it here in Euler's original form.

The product formula states that:

$$\prod_{p, prime} \frac{1}{1 - p^{-s}} = \zeta(s) \tag{2}$$

This rather mystifying equality is best digested in two steps. We will first show that the formula is "reasonable", explaining how it arises, but without a proof. Then we will present Euler's original proof.

First, consider what happens when the following "numerical polynomials" are multiplied.

$$\begin{aligned}(1 + 2 + 2^2 + 2^3) \times (1 + 3 + 3^2 + 3^3)&\\ &= (1 + 2 + 4 + 8\\ &+ 3 + 6 + 12 + 24\\ &+ 9 + 18 + 36 + 72\\ &+ 27 + 54 + 108 + 216)\end{aligned}$$

Each of the powers of 2 is multiplied by each of the powers of 3. We would like to replace these sums of powers, currently comprising just 4 terms each, with all the powers of 2 and all the powers of 3, but that would lead to nonsense, because everything would add to $+\infty$.

To preclude nonsense, use the negative powers:

$$(1 + 2^{-1} + 2^{-2} + 2^{-3}) \times (1 + 3^{-1} + 3^{-2} + 3^{-3})$$

The sums of the powers of 2 and of 3, with this change, can be extended to get all the powers, and neither sum will diverge. The resulting product will be the sum of the inverses of all the integers that can be written as a power of 2 times a power of 3.

If we next expand the left side product to include all primes, not just 2 and 3, the right side will be the sum of the inverses of all the integers. Thus

$$\prod_{p, prime} (1 + p^{-1} + p^{-2} + p^{-3} + p^{-4} + \cdots) = \sum_{n=1}^{\infty} n^{-1} \tag{3}$$

Now we have a new problem. The sum of the inverses of the integers, the right hand side of this last equation, diverges (integral test). But we can make everything converge if the exponents on the integers in the sum change from -1 to -s, where $s > 1$. This convergence can be seen from the integral test.

This leads to the formula, with no divergence issues,

$$\prod_{p, prime} (1 + p^{-s} + p^{-2s} + p^{-3s} + p^{-4s} + \cdots) = \sum_{n=1}^{\infty} n^{-s} \tag{4}$$

Noting that each term in the left hand product is a convergent infinite geometric series, and applying the formula for summing such a series,

$$(1 + r + r^2 + r^3 + \cdots) = \frac{1}{1-r}$$

the original product in Euler's formula results.

Of course the preceding discussion hardly qualifies as a proof, so we'll now present one. Begin by writing the series for both $\zeta(s)$ and $2^{-s}\zeta(s)$.

$$\begin{aligned} \zeta(s) &= 1 + 2^{-s} + 3^{-s} + 4^{-s} + 5^{-s} + 6^{-s} + \cdots \\ 2^{-s}\zeta(s) &= 2^{-s} + 4^{-s} + 6^{-s} + 8^{-s} + 10^{-s} + \cdots \end{aligned}$$

Subtracting, we find that

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

This operation has removed from $\zeta(s)$ all of the "even" terms. If we next repeat the prior multiplication and subtraction using 3^{-s}, we will remove any terms that are multiples of 3, and be left with

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

Continuing with prime after prime, the RHS will equal 1 plus the tail end of a convergent series that starts only after the largest of the primes in the product. Because the limit of the tail of a convergent series equals 0, the limit of the RHS will be 1. Explicitly,

$$\zeta(s) \prod_{p,prime} (1-p^{-s}) = 1$$

Moving the product to the other side of the equality sign, we immediately have the product formula (2).

3 Taking limits in the product formula

Now we want to take the product formula (2) and see what happens when $s \to 1^+$. In particular we want to prove that

$$\prod_{p,prime} \left(1 - \frac{1}{p}\right) = 0 \tag{5}$$

To begin the proof, remembering the idea behind the integral test from calculus, note that $\zeta(s)$ is larger than

$$\sum_{n=1}^{\infty} \left(\int_n^{n+1} x^{-s} dx \right) = \int_1^{\infty} x^{-s} dx = \frac{1}{s-1} \to \infty$$

Consequently, the product on the LHS of equation (2) also tends to ∞ as $s \to 1^+$:

$$\lim_{s\to 1^+} \left(\prod_{p,prime} \frac{1}{1-p^{-s}} \right) = \infty \tag{6}$$

Now we want to take the limit from outside the product and apply it inside, term by term. In the most general of situations an interchange of product and limit cannot be done, but in our special case we can justify doing it because each term in the product increases monotonically as $s \to 1^+$.

Here is the justification.

Let us select any large number, M. We need to show that the LHS of equation (2) will exceed M when we make the value of s close enough to 1. Start by selecting a value of s for which $\frac{1}{s-1} - 1 > M$.

Because $\zeta(s)$ exceeds $\frac{1}{s-1}$, the LHS of (2) will also do so. Consequently there is some value N for which

$$\prod_{p<N,prime} \frac{1}{1-p^{-s}} \geq \frac{1}{s-1} - 1 > M$$

In this last finite product, however, each term will only become larger when s is replaced by 1. (This is the use of monotonicity.) That means that

$$\prod_{p<N,prime} \frac{1}{1-p^{-1}} > M$$

The infinite product of (2), however, is the limit of this last finite product as $N \to \infty$, and the lower bound of M was selected first, independently of N. This means that the infinite product in (2) exceeds any preselected value, as desired. Equation (5) immediately follows.

4 An interesting fact from calculus

In order to complete the proof of Euler's theorem about the divergence of the prime reciprocals we will establish a simple theorem about infinite products of the form $\prod_1^\infty (1 - x_n)$, where $\{x_n\}$ is a sequence of small positive numbers converging to zero. Because each factor in this infinite product is between 0 and 1, the partial products will get progressively smaller. The only mystery is whether the product will converge all the way to zero, or will converge to something positive. Here are two special examples that illustrate the issue.

Example 1:

$$\prod_{n=2}^{\infty}(1-\frac{1}{n}) = \frac{1}{2}\frac{2}{3}\frac{3}{4}\cdots \to 0$$

Example 2:

$$\prod_{n=2}^{\infty}(1-\frac{1}{n^2}) = \left(\frac{1\times 3}{2\times 2}\right)\left(\frac{2\times 4}{3\times 3}\right)\left(\frac{3\times 5}{4\times 4}\right)\cdots \to \frac{1}{2}$$

We wil show that the first example product converges to 0 because $1/n$ is a divergent series, whereas the second product converges to something positive because $1/n^2$ is convergent.

Lemma: If $0 \leq x \leq \frac{1}{2}$, then $-2x \leq log(1-x) \leq -x$.

Proof. Letting $f(x) = log(1-x)$, we find that $f'(x) = \frac{-1}{1-x} = \frac{1}{x-1}$ and that $f''(x) = \frac{-1}{(x-1)^2}$.

Thus f' is a decreasing function of x on $[0, \frac{1}{2}]$. (Actually on all of $(-\infty, 1]$). Thus $f'(\frac{1}{2}) \leq f'(x) \leq f'(0)$ on $[0, \frac{1}{2}]$. That is, $-2 \leq f'(x) \leq -1$.

But $\frac{f(x)-f(0)}{x-0} = \frac{log(1-x)}{x} = f'(a)$ for some a between 0 and x, by the Mean Value Theorem. So $-2 \leq \frac{log(1-x)}{x} \leq -1$. Because x is positive, we can multiply without changing the direction of the inequalities, and conclude that $-2x \leq log(1-x) \leq -x$, as desired. □

Theorem: Let x_k be a sequence of numbers, all between 0 and $\frac{1}{2}$. Then $\prod_1^\infty(1-x_k) = 0$ if and only if $\sum_1^\infty x_k = \infty$.

Proof. Let P_n denote $\prod_1^N (1 - x_k)$. Then the preceding lemma says that

$$-2 \sum_{k=1}^{N} x_k \leq log(P_N) \leq - \sum_{k=1}^{N} x_k$$

But now $P_N \to 0$ if and only if $log(P_N) \to -\infty$, and this happens if and only if $\sum_{k=1}^{N} x_k \to \infty$. $\square$

Now that we have this last theorem at our disposal, we combine it with equation (5). The divergence of the prime reciprocals is established.

5 Rate of Divergence

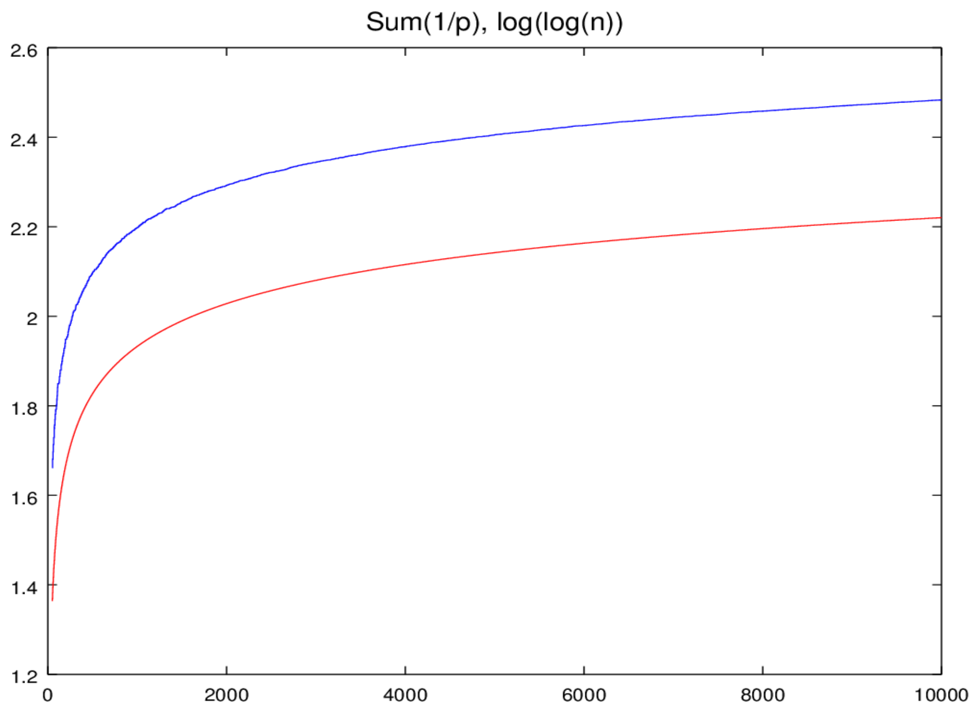
Now that we know that the sum of the prime reciprocals diverges it is natural to ask how fast the partial sums grow. In the simple case of the harmonic series, for example, we know that $\sum_1^n \frac{1}{k}$ looks roughly like $\int_1^n \frac{1}{t} dt = log(n)$. This is a slowly diverging function, so we can expect the restriction to primes to grow even slower. The following "suggestive" argument, which uses the Prime Number Theorem, gives some sense of what happens.

As in the case of the harmonic series we will approximate the sum by an integral, but we will include in the integral a factor that weights each value by the probability that it is prime. The Prime Number Theorem says that the probability that a large number, x, is prime is roughly $\frac{1}{log(x)}$. So the integral that, hopefully, approximates the sum, is

$$\int_2^n \frac{1}{x} \frac{1}{log(x)} dx$$

We know, however, that the integrand is the derivative of $log(log(x))$, so the integral is basically $log(log(n))$. The numerical experiment graphed below shows values for the partial sums of the prime reciprocals and $log(log(n))$, and it strongly indicates that the growth looks like what we expect, except for a small offset. The actual value, beyond the scope of this memo, is:

$$\sum_{p \leq n} \frac{1}{p} \sim c + log(log(n)) + o(1)$$



Judging from the graph, the constant c has a value of roughly 0.263. Note that $log(log(n))$ grows exceedingly slowly. When $n = 10^{1000}$ it has only grown to 7.74.