Prime Reciprocals Diverge!

Gautam Manohar

20 May 2018

This document originally appeared as a blog post on my website. Find it at gautammanohar.com/prime-reciprocals.

1 Introduction

In this post, we will investigate the infinite series

$$\sum_p \frac{1}{p}, \tag{1}$$

taken over all primes p. In particular, we will show that this series diverges.

2 Calculus and the Prime Number Theorem

First, we'll show that this problem falls to the incredible might of the following inequalty, which bounds the value of the n-th prime:

$$n \log n - n \log \log n \le p_n \le n \log n + n \log \log n \quad (n > 5) \tag{2}$$

This identity comes in quite handy when, say, you're solving some Project Euler problems about primes, and you need to generate at least n primes. It's derived from the Prime Number Theorem.

Using the upper bound, we know that $\frac{1}{p} \ge \frac{1}{n \log n + n \log \log n}$ for $n > 5$. And so

$$\sum_p \frac{1}{p} \ge \sum_{n=6}^{\infty} \frac{1}{n \log n + n \log \log n}. \tag{3}$$

Because $\log x$ is a monotonically increasing function, $\log \log x \le \log x$ for all x. And so $n \log n + n \log n \ge n \log n + n \log \log n$. This means that

$$\sum_{n=6}^{\infty} \frac{1}{n \log n + n \log \log n} \ge \sum_{n=6}^{\infty} \frac{1}{2n \log n}. \tag{4}$$

Here's where the calculus comes in. The integral test says that if a function f is non-negative, continuous, and monotonically decreasing on a certain interval $[a, b+1]$, then the sum of $f(n)$ over integers $a \leq n \leq b$ is greater than the integral of f over that interval:

$$\int_a^{b+1} f(x) \, \mathrm{d}x < \sum_{\substack{n \in \mathbb{Z} \\ a \leq n \leq b}} f(n) \tag{5}$$

We know x is increasing and continuous, as is $\log x$. And so $x \log x$ is increasing and continuous, which means that $\frac{1}{x \log x}$ is decreasing. On $[6, \infty)$, it's also non-negative.

This means that

$$\sum_p \frac{1}{p} \geq \sum_{n=6}^{\infty} \frac{1}{2n \log n} \geq \frac{1}{2} \int_6^{\infty} \frac{1}{x \log x} \, \mathrm{d}x. \tag{6}$$

Substituting $u = \log x, \mathrm{d}u = \frac{1}{x} \, \mathrm{d}x$ gives

$$\int_{\log 6}^{\infty} \frac{\mathrm{d}u}{u}, \tag{7}$$

which we evaluate as $[\log u]_{\log 6}^{\infty} = \infty$. But wait! This means that the sum of the reciprocals of the primes is greater than infinity, which of course means that it cannot converge to any finite value.

3 Combinatorics

In my opinion, this combinatorially inspired method is way cooler.

3.1 Squares and Squarefree Products

First, every integer can be written uniquely as the product of a square and a squarefree integer (that is, a number which is not divisible by any square greater than 1—such numbers are the products of distinct primes). To see this, consider the prime factorization of a number

$$n = \prod_{p_i | n} p_i^{e_i}, \tag{8}$$

where e_i is the greatest positive integer such that $p_i^{e_i}$ divides n. Then for each odd e_i, factor out p_i. Then each exponent in the product is even, which means it is a square, and the term outside the product consists of primes with exponent 1.

3.2 Unleashing the Combo Punches

Second, each term in

$$\prod_p \left(1 + \frac{1}{p}\right) \tag{9}$$

is formed by multiplying either 1 or $\frac{1}{p}$ for each p, which means each term is the product of distinct primes. Furthermore, because the product is over all primes, every possible combination of primes is represented here exactly once. It is clear that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \tag{10}$$

contains the reciprocal of each square exactly once. Its value is the subject of the Basel problem, which Euler famously solved: $\frac{\pi^2}{6}$. All we really need is that the series converges to some finite value C. The product of (9) and (10) is

$$\left(\sum_{n=1}^{\infty} \frac{1}{n^2}\right) \prod_p \left(1 + \frac{1}{p}\right) = \sum_{n=1}^{\infty} \frac{1}{n}, \tag{11}$$

the harmonic series. This is because each term consists of a selection from the first sum (of reciprocals of squares) and the second sum (of reciprocals of squarefree numbers). Because all squares and squarefree numbers are here, the reciprocal of each number is represented. This fact is not integral to the divergence proof, but because the representation of a number as a product of a square and a squarefree number is unique, each reciprocal is represented exactly once, so the product is exactly equal to the harmonic series. The important fact here is that the harmonic series diverges.

3.2.1 The Divergence of the Harmonic Series

To see this, we can use the integral test again (with the function $\frac{1}{x}$), but we can also do it without calculus. Suppose the harmonic series converges to some value H. Then

$$H = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots \tag{12}$$

Because $n + 1 > n$ for all n, we have $\frac{1}{n+1} < \frac{1}{n}$ for all n. Replacing terms in a sum with smaller terms means the resulting sum is smaller. So

$$\begin{aligned}
H &> \frac{1}{2} + \frac{1}{2} + \frac{1}{4} + 14 + \frac{1}{6} + \frac{1}{6} + \cdots \\
&= 1 + \frac{1}{2} + \frac{1}{3} + \cdots \\
&= H.
\end{aligned} \tag{13}$$

We end up with $H > H$, which is certainly not true for any finite value. And so the harmonic series diverges.

3.3 Finishing Touches

Third, we'll use the identity

$$1 + x \leq e^x. \tag{14}$$

For three different proofs of this fact, see my post on comparing e^π and π^e. Applying (14) to (11) gives

$$\sum_{n=1}^{\infty} \leq \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right) \prod_p \left(1 + \frac{1}{p} \right) \leq C \prod_p e^{-p}. \tag{15}$$

Let's use some exponent rules to clean things up.

$$\sum_{n=1}^{\infty} \leq C e^{\sum_p \frac{1}{p}}$$

$$\log \sum_{n=1}^{\infty} \leq \log C + \sum_p \frac{1}{p}. \tag{16}$$

Becuase C is finite, $\log C$ is also finite, and so we can see that the desired sum is greater than the harmonic series, which establishes its divergence.

4 Bounding the Beast

By slightly modifying the above arguments—for example, change infinite sums over all primes or natural numbers to partial sums over all terms less than k—we obtain the following bound:

$$\log \sum_{n=1}^{k} \frac{1}{n} - \log \frac{\pi^2}{6} \leq \sum_{p \leq k} \frac{1}{p}. \tag{17}$$

We can make this even prettier by... using some calculus. Because $\frac{1}{x}$ is decreasing and continuous, we can use the integral test to bound the partial harmonic series

$$\log \sum_{n=1}^{k} \frac{1}{n} \geq \int_1^{k+1} \frac{1}{x} \, \mathrm{d}x = \log(k+1) - \log 1 = \log(k+1). \tag{18}$$

And so we obtain the very nice lower bound

$$\log \log(k+1) - \log \frac{\pi^2}{6} \leq \sum_{p \leq k} \frac{1}{p}. \tag{19}$$

5 Conclusion

So, calculus or combinatorics? I think both methods are cool in their own right—one uses continuous methods to analyze prime numbers, which are discrete entities, while the other is just mind-numbingly brilliant. Perhaps the calculus arguments are cleaner, but the combinatorics ones are much more inspiring and hide some deeper insights.