1 The Primes Aren't That Spread Out

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

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

Infinite Series 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 **??** has an easy short proof.

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

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

This series does converge. Appendix **??** walks through Euler's historic and rather adventurous proof showing it converges to $\frac{\pi^2}{6}$.

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

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

Sum Of Prime Reciprocals

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

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

Because 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.

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

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

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

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.

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

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

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

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

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

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

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

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

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

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.