The sum of the reciprocals of the primes

In 1737 Leonard Euler published the following result:

> *Summa serier reciprocae numerorum primorum*
>
> $$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\textit{etc.}$$
>
> *est infinite magna, infinities tanem minor quam summa serier harmonicae*
>
> $$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\textit{etc.}$$
>
> *Atque illius summa est huius summae quasi logarithmus.*

What this means is that the series of the reciprocals of the prime numbers diverges with the partial sums growing at the same rate as the logarithm of the partial sums of the harmonic series. In this note, I want to investigate what this means.

You've already seen in your calculus class that the harmonic series diverges. This means that the partial sums grow without bound. Let's show that the partial sums, in fact, grow as the logarithm function. I'll show the following.

THEOREM 1. For every positive integer N,

$$\ln N \le \sum_{k=1}^{N}\frac{1}{k} \le \ln N + 1.$$

Proof. Consider first Figure 1. The rectangles form a left-hand Riemann sum approximation to $\int_1^{N+1} 1/x\ dx$. All the rectangles have base of length 1. The first rectangle has height 1; the second, height $\frac{1}{2}$, and so on. The total area inside all the rectangles is the partial sum $\sum_{k=1}^{N} 1/k$. By comparing this area with the area under the graph of $1/x$, you can readily see that

$$\ln(N+1) = \int_1^{N+1}\frac{1}{x}\,dx \le \sum_{k=1}^{N}\frac{1}{k}.$$

Since the logarithm function is increasing, we have $\ln N \le \ln(N+1)$ which gives the left-hand inequality.

Consider now Figure 2. Here the rectangles form a right-hand Riemann sum approximation to $\int_1^{N} 1/x\ dx$. The total area inside the rectangles is the partial sum $\sum_{k=2}^{N} 1/k$. By comparing this area with the area under the graph of $1/x$, you can readily see that

$$\sum_{k=2}^{N}\frac{1}{k} \le \int_1^{N}\frac{1}{x}\,dx = \ln N.$$

from which it follows that

$$\sum_{k=1}^{N}\frac{1}{k} = 1 + \sum_{k=2}^{N}\frac{1}{k} \le 1 + \ln N,$$

which is the right–hand inequality. $\square$

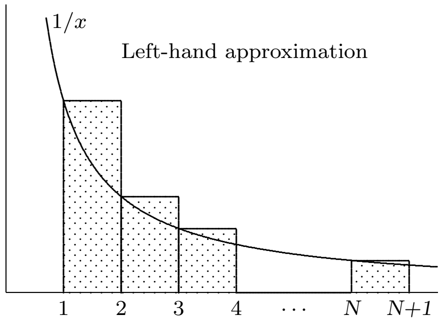

Figure 1.

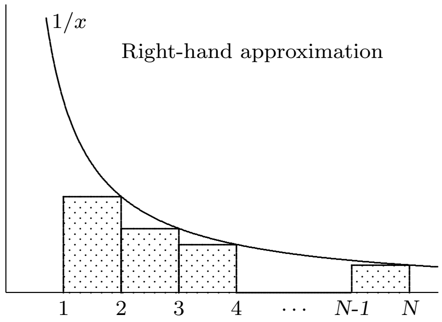

Figure 2.

We now move on to the series of reciprocals of the primes.

You certainly remember that a prime number is an integer which is divisible only by itself and by 1. For technical reasons, we exclude the number 1 from being a prime. The only other fact you'll need to recall is that any integer can be factored into primes numbers.

The first thing to see is that the series of reciprocals of primes is indeed an *infinite* series; that is, that there are infinitely many primes. This fact was known to the ancient Greeks 3000 years ago. The following proof is by Euclid.

THEOREM 2. *There exist infinitely many primes.*

Proof. The proof is by contradiction. Suppose there are only finitely many primes and list them in increasing order starting with $p_1 = 2$, $p_2 = 3$, up to the largest one p_N. Now look at the number $M = (p_1 p_2 \cdots p_N) + 1$. Since M is one more than a multiple of p_1, it's not a multiple of p_1. Similarly, it's not a multiple of p_2 or of any other prime! But M can be factored into primes. This is impossible! You are forced to conclude that the initial supposition, that there are only finitely many primes, is wrong, proving the result. □

Next, I'll show that the series of reciprocals of the primes diverges and give a a lower bound for the rate at which the partial sums are growing. More exactly, I'll show the following.

THEOREM 3. *For any positive integer* $N \geq 2$,

$$\sum_{p \leq N} \frac{1}{p} \geq \ln\ln N - \frac{1}{2}.$$

Proof. I'll start with a couple of observations. First note that if p is any prime, then the convergence of the geometric series gives

$$1 + \frac{1}{p} + \frac{1}{p^2} + \cdots = \frac{1}{1 - (1/p)}.$$

Next take any collection of primes $p_1, p_2, \ldots, p_k$ and look at

$$\prod_{i=1}^{k} \left(1 + \frac{1}{p_i} + \frac{1}{p_i^2} + \cdots\right).$$

Note that this product is finite by the first observation. If you were to multiply this out, what would the terms look like? Each one would have a factor from within the first set of parentheses, one from within the second set of parentheses, and so on. Therefore, it would look like

$$\frac{1}{p_1^{e_1}} \frac{1}{p_2^{e_2}} \cdots \frac{1}{p_k^{e_k}},$$

where the e_i's are various powers. Now, if N is a positive integer whose prime factors are among $p_1, p_2, \ldots, p_k$, then $1/N$ appears somewhere among the terms in the product. Therefore, for $N \geq 2$,

$$\prod_{p \leq N} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right) \geq \sum_{k=1}^{N} \frac{1}{k} \geq \ln N,$$

so

$$\prod_{p \leq N} \frac{1}{1 - (1/p)} \geq \ln N.$$

Taking the log of both sides gives

$$\sum_{p \leq N} -\ln\left(1 - \frac{1}{p}\right) \geq \ln\ln N.$$

Using the Taylor expansion of $\ln(1 - x)$ we get [*]

$$\begin{aligned} -\ln\left(1 - \frac{1}{p}\right) = \sum_{k=1}^{\infty} \frac{1}{kp^k} &\leq \frac{1}{p} + \frac{1}{2} \sum_{k=2}^{\infty} \frac{1}{p^k} \\ &= \frac{1}{p} + \frac{1}{2p^2} \sum_{k=0}^{\infty} \frac{1}{p^k} \\ &= \frac{1}{p} + \frac{1}{2p^2} \frac{1}{1 - (1/p)} \\ &= \frac{1}{p} + \frac{1}{2} \frac{1}{p(p-1)} \end{aligned}$$

---

[*] As a reminder, $\ln(1 - x) = -\sum_{k=1}^{\infty} x^n/n$, with the expansion valid for $-1 < x < 1$.

These inequalities together give us

$$\begin{aligned}\ln\ln N &\leq \sum_{p\leq N} -\ln\left(1-\frac{1}{p}\right)\\ &= \sum_{p\leq N}\frac{1}{p}+\frac{1}{2}\sum_{p\leq N}\frac{1}{p(p-1)}\end{aligned}$$

Hence

$$\ln\ln N \leq \sum_{p\leq N}\frac{1}{p}+\frac{1}{2}\sum_{k=1}^{\infty}\frac{1}{k(k-1)}.$$

By partial fractions $\frac{1}{k(k+1)}=\frac{1}{k}-\frac{1}{k+1}$ so

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{1}{k(k-1)} &= \sum_{k=1}^{\infty}\frac{1}{k}-\frac{1}{k+1}\\ &= (1-\frac{1}{2})+(\frac{1}{2}-\frac{1}{3})+(\frac{1}{3}-\frac{1}{4})+\cdots\\ &= 1.\end{aligned}$$

Therefore

$$\ln\ln N \leq \frac{1}{2}+\sum_{p\leq N}\frac{1}{p}$$

from which the desired result follows. □

Using similar techniques, the following result can also be proven.

THEOREM 4. *For any positive integer* $N\geq 2$

$$\sum_{p\leq N}\frac{1}{p}\leq 3\ln\ln N+6.$$