Calculus, Group Theory, and the Infinitude of Primes

Theorem (Euclid, ca. 300 B.C.E.)**.** There are infinitely many positive prime numbers.

First proof of the infinitude of primes. Let $p_1, p_2, \ldots, p_k$ be the first k prime numbers. Consider the new integer $n = p_1 p_2 \ldots p_k + 1$, which is obviously greater than any of $p_1, p_2, \ldots, p_k$. It is either prime or it is composite. If it is prime, then we have found a prime larger than any of $p_1, p_2, \ldots, p_k$. If it is composite, then it is divisible by a prime p. But p cannot be any of $p_1, p_2, \ldots, p_k$, because the remainder of n upon division by any of $p_1, \ldots, p_k$ is 1, not 0. In this case, the prime divisor p of n must be a prime different from any of $p_1, p_2, \ldots, p_k$. In either case, we have found a prime that is different from all of $p_1, \ldots, p_k$. Thus, the collection of primes must be infinite. □

Remark. The first value of k for which the integer n above is not prime is $k = 6$. One can check that $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 30031 = 59 \cdot 509$.

Our second proof will use some ideas from Calculus, namely, limits and infinite series.

Second proof of the infinitude of primes. Given a positive real number x, define $\pi(x)$ to be the number of primes that are less than or equal to x. We want to show that $\lim_{x\to\infty} \pi(x) = \infty$. Recall that the natural logarithm function $\ln(x)$ is defined by the equation $\ln(x) = \int_1^x \frac{1}{t}\,dt$. We are going to use a left endpoint rectangular approximation for the integral $\int_1^x \frac{1}{t}\,dt$ to come up with a lower bound estimate for the size of $\pi(x) + 1$.

If $n \le x \le n + 1$, then

$$\ln(x) \le 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n-1} + \frac{1}{n}.$$

The prime divisors of $1, 2, 3, \ldots, n$ are all less than or equal to x. Thus, we can say $\ln(x) \le \sum \frac{1}{m}$, where we now sum over all $m \in \mathbb{N}$ such that the prime divisors of m are $\le x$. We want to find a new way to express this (much larger) new sum.

Example. Suppose $5 \le x \le 6$. Then we are looking at the estimate $\ln(x) \le 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5}$, and $k = \pi(x) = 3$. The new sum we consider is

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10} + \frac{1}{12} + \frac{1}{15} + \frac{1}{16} + \frac{1}{18} + \frac{1}{20} + \frac{1}{24} + \frac{1}{25} + \frac{1}{27} + \frac{1}{30} + \cdots,$$

which is equal to

$$\left(1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots\right)\left(1 + \frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \frac{1}{3^4} + \cdots\right)\left(1 + \frac{1}{5} + \frac{1}{5^2} + \frac{1}{5^3} + \frac{1}{5^4} + \cdots\right).$$

Back to the proof. Set $k = \pi(x)$. So there are k distinct prime numbers $\le x$. (In the above example, $k = 3$.) Write them as $p_1, p_2, \ldots, p_k$. Then

$$\ln(x) \le \sum \frac{1}{m} = \prod_{i=1}^{k} \left(\sum_{r=0}^{\infty} \frac{1}{p_i^r} \right).$$

Applying the summation formula for geometric series,

$$\ln(x) \le \prod_{i=1}^{k} \left(\frac{1}{1 - \frac{1}{p_i}} \right) = \prod_{i=1}^{k} \left(\frac{p_i}{p_i - 1} \right).$$

Now, it's easy to see that $p_i \geq i+1$. Then $p_i - 1 \geq i$, and

$$\frac{p_i}{p_i - 1} = 1 + \frac{1}{p_i - 1} \leq 1 + \frac{1}{i} = \frac{i+1}{i}.$$

Then

$$\ln(x) \leq \prod_{i=1}^{k} \frac{i+1}{i} = \frac{2}{1} \cdot \frac{3}{2} \cdot \frac{4}{3} \cdots \frac{k+1}{k} = k+1.$$

But $k = \pi(x)$. So $\ln(x) \leq \pi(x) + 1$. We know that $\lim_{x\to\infty} \ln(x) = \infty$. So then also

$$\lim_{x\to\infty} \pi(x) = \infty. \qquad \square$$

Perhaps the appearance of the logarithm function in the above proof seems like a coincidence or a convenience, but in fact the logarithm function is intimately connected to prime numbers. Indeed, the famous Prime Number Theorem states that $\pi(x)$ is asymptotic to $x/\ln(x)$, that is

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln(x)} = 1,$$

so in some sense the function $x/\ln(x)$ provides an approximation for the number of prime numbers less than or equal to x. Here are a few values of $\pi(x)$ and $x/\ln(x)$ for "small" values of x.

x	$\pi(x)$	$x/\ln(x)$	$\pi(x)/(x/\ln(x))$
10^3	168	145	1.159...
10^6	78498	72382	1.084...
10^9	50847534	48254942	1.053...

Now we look at a group theoretic proof of the infinitude of prime numbers. Roughly speaking, a **group** is a collection of symmetries. More accurately, every (closed) collection of symmetries of a geometric object forms a group.

Example. Look at the rigid motion symmetries of the square.

1. Describe some of the rigid motion symmetries of the square.
2. How many distinct rigid motion symmetries are there?
3. What are the orders of the various symmetries? Are they divisors of the answer to #2?

Here are some of the important properties of the rigid motion symmetries of the square:

1. There is a symmetry that does nothing (the identity e).
2. If you perform one symmetry, and then another, you get another symmetry.
3. For every symmetry, there is another that undoes it.

The above observations are a very rough approximation for the axioms defining a **group**. We can think of composing symmetries as a kind of multiplication. If r stands for the symmetry that rotates the square by 90°, then $r \cdot r \cdot r \cdot r = r^4 = e$. If s represents a reflection across a diagonal, then $s \cdot s = s^2 = e$. What is $s \cdot r \cdot s$?

Notice that we can deform the shape of the square (say by adding small spikes in the center of each side) without affecting the group of symmetries. Thus, there are many different ways in which this group of symmetries can be realized. Group theorists study the structure of collection of symmetries, without caring about the particular way in which those symmetries are realized. The group theorist only cares about the abstract properties of the group.

Example (The integers modulo p)**.** Let p be a prime. Given an integer n, let $[n]$ represent the remainder when n is divided by p. (For those of you who have taken Math 3200, $[n]$ really stands for the equivalence class of n under the relation of congruence modulo p.) Set $\mathbb{Z}_p^* = \{[1], [2], \ldots, [p-1]\}$, the set of nonzero remainders for division by p. Define a multiplication on $\mathbb{Z}_p^*$ by $[a] \cdot [b] = [ab]$. Then this multiplication makes $\mathbb{Z}_p^*$ a group. In particular, $([a] \cdot [b]) \cdot [c] = [a] \cdot ([b] \cdot [c]) = [abc]$.

Example (The case $p = 5$)**.** Work out the multiplication table for the case $p = 5$.

*	[1]	[2]	[3]	[4]
[1]	[1]	[2]	[3]	[4]
[2]	[2]	[4]	[1]	[3]
[3]	[3]	[1]	[4]	[1]
[4]	[4]	[3]	[2]	[1]

The reason why the product of two remainders is another remainder is that p is a prime. If a and b are integers and if p divides ab, then p must divide one of a or b. In our example, the element $[2]$ is like the reflection r of the square, in that $[2] \cdot [2] \cdot [2] \cdot [2] = [2]^4 = [16] = [1]$.

Theorem (Lagrange's Theorem)**.** Let G be a finite group. Then the order of each element in G divides $|G|$.

Third proof of the infinitude of primes. Suppose to the contrary that there are only finitely many prime numbers. Let p be the largest prime, and consider the *Mersenne number* $2^p - 1$. Let q be a prime factor of $2^p - 1$. Then $[2^p] = [1]$ in $\mathbb{Z}_q^*$. But $[2^p] = [2]^p$, so $[2]^p = [1]$, i.e., $[2]$ has order p in $\mathbb{Z}_q^\times$. Then p divides the order of the group $\mathbb{Z}_q^*$, which is $q - 1$. So $p \mid (q-1)$, and hence $p < q$, a contradiction, because p was assumed to be the largest prime. □