Chapter 5

Prime Numbers

5.1 The fundamental theorem of arithmetic

Definition: An integer $p > 1$ is said to be *prime* if its only positive divisors are 1 and p itself. All other integers greater than 1 are called composite.

A composite number n can be written as a product $n = ab$ of two strictly smaller numbers $1 < a, b < n$. Note that, by convention, 1 is neither prime nor composite. Here are all primes below 100:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97.$$

Given a prime p and another integer a, either a is a multiple of p or $\gcd(p, a) = 1$. Indeed, $\gcd(p, a)$ divides p, so it must be either 1 or p, and since $\gcd(p, a)$ also divides a then either $\gcd(p, a) = 1$ or a is a multiple of p. This can be used to prove a very important property of primes:

Theorem 5.1.1. *Let p be a prime.*

(a) Given two integers a and b, if $p|ab$ then either $p|a$ or $p|b$.

(b) Given k integers $a_1, a_2, \ldots, a_k$, if $p| \prod_{i=1}^{k} a_i$ then $p|a_i$ for some $1 \leq i \leq k$.

Proof.

(a) If $p|a$ we are done. Otherwise $\gcd(p, a) = 1$ and by Bezout's identity there exist linear coefficients u and v for which $1 = ua + vp$. Multiplying both sides by b we get $b = uab + vpb$. Since p divides ab, p divides the whole sum $uab + vpb$. Therefore $p|b$.

(b) The proof proceeds by induction. The case $k = 1$ is trivial and $k = 2$ is handled in part (a). So we assume that the claim holds for some $k > 1$ and prove that it also holds for $k + 1$. Given that $p| \prod_{i=1}^{k+1} a_i$, we put $b = \prod_{i=1}^{k} a_i$. Since $p|ba_{k+1}$, part (a) implies that either $p|a_{k+1}$ or $p|b$. In both cases the claim holds, in the latter case by the induction hypothesis. This completes the proof by induction.

□

Theorem 5.1.1 can be used to derive a fundamental theorem of number theory. It is so fundamental it has “fundamental” in its name.

Theorem 5.1.2 (Fundamental Theorem of Arithmetic)**.** *Every positive integer can be represented in a unique way as a product of primes,*

$$n = p_1 p_2 \cdots p_k \qquad\qquad (p_1 \leq p_2 \leq \ldots \leq p_k).$$

Proof. We first prove existence and then uniqueness. Actually, we already proved existence in one of the previous lectures as an illustration of strong induction, but give the prove here again for completeness. So, to prove that every integer can be represented as a product of primes we use strong induction. The base case $n = 1$ holds because the *empty product*, as we previously discussed, is defined to equal 1. The induction hypothesis assumes that for some $n > 1$, all positive integers $k < n$ can be represented as a product of primes. If n is prime, then it is trivially a product of primes. Otherwise it can be written as $n = ab$, for $1 < a, b < n$. By the induction hypothesis, both a and b are products of primes, so their product n is also a product of primes. This proves existence.

The proof that the above representation is unique proceeds by contradiction. Assume then that there exists some positive integer that can be represented as a product of primes in (at least) two ways. By the well-ordering principle, there is a smallest such integer n. It holds that $n = p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l$, where $p_1 \leq p_2 \leq \ldots \leq p_k$, $q_1 \leq q_2 \leq \ldots \leq q_l$, and $p_i \neq q_i$ for some i. By Theorem 5.1.1(b), since $p_i | q_1 q_2 \cdots q_l$, there must exist some q_j for which $p_i | q_j$. Since q_j is prime and $p_i > 1$, this can only occur when $p_i = q_j$. Thus we can eliminate p_i and q_j from the equation $p_1 p_2 \cdots p_k = q_1 q_2 \cdots q_l$ and get two distinct representations of the positive integer number n/p_i as a product of primes. This contradicts the assumption that n is the smallest positive integer with this property, and concludes the proof of uniqueness. □

5.2 The infinity of primes

Here is another fundamental result with a proof from Euclid’s *Elements*:

Theorem 5.2.1. *There are infinitely many primes.*

Proof. Assume for the sake of contradiction that there is only a finite set of primes, $p_1, p_2, \ldots, p_n$. Consider the number

$$p = p_1 p_2 \ldots p_n + 1.$$

By Theorem 5.1.2, p has a prime divisor, which has to be p_i, for some $1 \leq i \leq n$. Since p_i divides both p and $p_1 p_2 \ldots p_n$, it also divides $p - p_1 p_2 \ldots p_n = 1$. However, this is impossible since $p_i > 1$. This contradiction proves the theorem. □

Let’s get some more mileage out of Euclid’s proof. The results below show that not only do the primes never stop, but the number of primes $p \leq x$ is at least a certain natural function of x, namely at least $\log \log x$. (Here the base of the logarithm is 2.)

Theorem 5.2.2. *The n-th prime p_n satisfies $p_n \leq 2^{2^{n-1}}$ for all $n \geq 1$.*

Proof. We proceed using strong induction. For the base case, the first prime is $2 = 2^{2^0}$. Assume that the claim holds for all primes p_1 through p_k. Consider $p = p_1 p_2 \dots p_k + 1$. As in the above proof, p has a prime factor that is not one of the first k primes. This prime factor is thus at least as large as p_{k+1}, which implies

$$\begin{aligned} p_{k+1} \leq p = p_1 p_2 \dots p_k + 1 &\leq 2^{2^0} 2^{2^1} \cdots 2^{2^{k-1}} + 1 \\ &= 2^{1+2+4+\ldots+2^{k-1}} + 1 \\ &= 2^{2^k-1} + 1 \\ &= \frac{1}{2} 2^{2^k} + 1 \\ &\leq 2^{2^k}. \end{aligned}$$

This is precisely the induction step we needed, and concludes the proof by strong induction. □

Denote by $\pi(x)$ the number of primes $p \leq x$.

Corollary 5.2.3. *For $x \geq 2$, $\pi(x) \geq \lfloor \log \log x \rfloor + 1$.*

Proof. Plugging $n = \lfloor \log \log x \rfloor + 1$ into Theorem 5.2.2 implies that the n-th prime is at most x. Thus there are at least n primes below x. □

For general education, you should know that this is by far not the best possible estimate. A celebrated achievement in number theory is the Prime Number Theorem due to Hadamard and de la Vallée Poussin, which states that $x/\ln x$ (here we use the natural logarithm) is the "right" bound, in the sense that

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