Notes for Lecture 3

These are notes about the proof that there are infinitely many primes, which is not covered in the reading material.

1 There are infinitely many primes

An integer $n \geq 2$ is *composite* if we can write n as a product of two positive integers smaller than n, that is, if there are a, b such that $n = a \cdot b$, where $0 < a < n$ and $0 < b < n$. An integer $n \geq 2$ is *prime* otherwise.

For example, 2 is prime, 3 is prime, $4 = 2 \cdot 2$ is composite, 5 is prime, $6 = 2 \cdot 3$ is composite, 7 is prime, $8 = 2 \cdot 4$ is composite, $9 = 3 \cdot 3$ is composite, and so on.[1]

We will prove that there are infinitely many prime numbers.

We will use the following two lemmas, whose proofs are discussed in the next section. Completely rigorous proofs of Lemma 1 and Lemma 2 require induction, which is the topic of the next lecture.

Lemma 1 (Prime Factorization) *Every integer $n \geq 2$ can be written as a product of primes, meaning that either n is prime itself, or it can be written as a product of two or more prime numbers, not all necessarily distinct.*

For example we can write $60 = 2 \cdot 2 \cdot 3 \cdot 5$, where 2, 3 and 5 are prime.

Lemma 2 (How Division Works) *For every nonnegative integer $n \geq 0$ and positive integer d there exists an integer $q \geq 0$ (the quotient) and an integer r (the remainder) such that $0 \leq r \leq d - 1$ and*

$$n = qd + r$$

[1] Note that, if a number is composite, there may be more than one way of writing it as a product of two smaller positive integers. For example $60 = 2 \cdot 30 = 3 \cdot 20 = 4 \cdot 15 = 5 \cdot 12 = 6 \cdot 10$.

From this point on we proceed completely rigorously. First, we can prove that the value of r in Lemma 2 is unique.

Lemma 3 *For every nonnegative integers $n \geq 0$ and $d \geq 1$, there is a unique value of r such that $0 \leq r \leq d-1$ and that we can write $n = qd + r$ for some $q \geq 0$.*

PROOF: We proceed by contradiction, so we assume that there are at least two ways of writing n as a multiple of d plus a non-negative remainder smaller than d, that is for some integers $q, q' \geq 0$ and $0 \leq r < r' \leq d-1$ we have

$$n = qd + r$$

$$n = q'd + r'$$

Let us subtract the two equations; we get

$$(q - q') \cdot d = r' - r$$

Note that $r \geq 0$, so $r' - r \leq r' \leq d - 1$ and that $r' > r$, so $r' - r > 0$, which means

$$0 < (q - q') \cdot d \leq d - 1$$

which is not possible, because a strictly positive multiple of d cannot be smaller than d. □

We are ready to prove our main result

Theorem 4 *There are infinitely many primes.*

PROOF: We proceed by contradiction, and so we assume that there are *finitely many* primes. Whenever we have assumptions stating the existence of objects with certain properties it is always good to give them names, so let us call k the number of primes, and let us call the set of all primes $\{p_1, \ldots, p_k\}$.

Now we want to argue that there is an integer that cannot be written as a product of the primes $p_1, \ldots, p_k$. Here comes the idea of the proof: define

$$N = 1 + p_1 \cdot p_2 \cdot \cdots \cdot p_k$$

Then N cannot be prime itself, because it is bigger than all the other primes, which means that it has to be divisible by one of the other primes. But consider the division of N by p_1: we have

$$N = p_1 \cdot (p_2 \cdots p_k) + 1$$

which means that we get a remainder of 1 when dividing N by p_1 and by Lemma 3 this is the only possible remainder, while if N were divisible by p_1 the remainder would be zero. So N is not divisible by p_1.

By the same reasoning, N is not divisible by p_2, nor by p_3, ..., nor by p_k, and we have reached a contradiction. □

After succeeding in devising a proof by contradiction, it is always good to stop and think if the same reasoning can give a more direct argument. We proved that if $p_1, \ldots, p_k$ are the first k primes, then the number $1 + p_1 \cdots p_k$ is not divisible by any of them. Does it mean that if $p_1, \ldots, p_k$ are the first k primes then the number $1+p_1 \cdots p_k$ is also prime? Our argument does not show it: it could be that $1+p_1 \cdots p_k$ is composite and its prime factors are all bigger than p_k.

But is it *true* that if $p_1, \ldots, p_k$ are the first k primes then $1 + p_1 \cdots p_k$ is also prime? A few test cases check out: $3 = 1+2$ is prime, and so is $7 = 1+2 \cdot 3$, $31 = 1+2 \cdot 3 \cdot 5$, $211 = 1 + 2 \cdot 3 \cdot 5 \cdot 7$, and $2311 = 1 + 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11$.

Unfortunately, $30031 = 1 + 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13$ is composite, because $30031 = 59 \cdot 509$.

This is too bad, because otherwise we would have a solution to the problem of finding a simple non-randomized way to generate large primes (just get the first few primes, say with the sieve of Eratosthenes, multiply them together, and add 1). Specifically, the following is an open problem whose solution would be a big breakthrough: devise a deterministic (that is, not randomized) algorithm that, on input n, runs in time polynomial in n and outputs a prime larger than 2^n.

2 Proving facts about division

Here is a proof of Lemma 1:

PROOF: Suppose by contradiction that there are integers n that are not products of primes (meaning that they are not primes and that cannot be written as a product of two or more primes.) Let $n_{\min}$ be the *smallest* such integer. Since $n_{\min}$ is not prime it must be composite, so let us write $n_{\min} = a \cdot b$ where a, b are positive integers smaller than $n_{\min}$. Since a and b are both smaller than $n_{\min}$, they can be written as products of primes. Multiplying the expression of a as a product of primes and the expression of b as a product of primes we get an expression of $n_{\min}$ as a product of primes, which is a contradiction.□

Isn't this a completely rigorous proof? Why did we say that the proof needs to be done by induction? The proof assumes that if we have a set C of positive integers

(the set of counterexamples to Lemma 1) and the set is nonempty, then it has a minimum element. Isn't this fact obvious? Well, a nonempty set of integers does not have to have a minimum (take all the integers), and a nonempty set of positive reals also does not have to have to minimum (take all the real numbers which are strictly greater than 3), so how can we be sure that there isn't some strange set of positive integers that does not have a minimum? In fact, what can we safely assume about positive integers and, for that matter, was it the *precise definition* of the set of positive integers? Mathematical induction answers all these questions.

Here is a proof of Lemma 2:

PROOF: Consider the sequence $n - i \cdot d$ for all integer values of $i \geq 0$, that is

$$n, n-d, n-2d, n-3d, \ldots$$

The first element of the sequence is $n \geq 0$, but for every $i > n$ we have $n - id < n - nd \leq 0$. So the sequence starts nonnegative, and from some point on is always negative. Let q be the index of the last nonnegative value, then we have

$$n - qd \geq 0$$

$$n - (q+1) \cdot d < 0$$

Now call $r = n - qd$. By the definition of r we have

$$n = qd + r$$

From the fact that $n - qd \geq 0$ we have

$$r \geq 0$$

And from the fact that $n - (q+1)d < 0$ we have $r < d$ and, since r and d are integers

$$r \leq d - 1$$

□

Again, what is not rigorous about the above proof? We looked at the set of all i such that $n - id \geq 0$ and we realized this is a nonempty set of integers all whose elements are at most n, from which we deduced that the set has a maximum element. This seems obvious but raises the same kind of "are we really sure it's true" questions that arise when you assume that every set of positive integers has a minimum.