Lecture 4: The Sequence of Prime Numbers

William Stein

Math 124 Harvard University **Fall 2001**

This lecture is about the following three questions:

1. Are there infinitely many primes? (yes)
2. Are there infinitely many primes of the form $ax + b$? (yes, if $\gcd(a, b) = 1$)
3. How many primes are there? (asymptotically $x/\log(x)$ primes less than x)

1 There are infinitely many primes

Theorem 1.1 (Euclid). *There are infinitely many primes.*

Note that this is not obvious. There are completely reasonable rings where it is false, such as

$$R = \left\{ \frac{a}{b} : a, b \in \mathbb{Z} \text{ and } \gcd(b, 30) = 1 \right\}$$

There are exactly three primes in R, and that's it.

Proof of theorem. Suppose not. Let $p_1 = 2, p_2 = 3, \ldots, p_n$ be all of the primes. Let

$$N = 2 \times 3 \times 5 \times \cdots \times p_n + 1$$

Then $N \neq 1$ so, as proved in Lecture 2,

$$N = q_1 \times q_2 \times \cdots \times q_m$$

with each q_i prime and $m \geq 1$. If $q_1 \in \{2, 3, 5, \ldots, p_n\}$, then $N = q_1 a + 1$, so $q_1 \nmid N$, a contradiction. Thus our assumption that $\{2, 3, 5, \ldots, p_n\}$ are all of the primes is false, which proves that there must be infinitely many primes. $\square$

If we were to try a similar proof in R, we run into trouble. We would let $N = 2 \cdot 3 \cdot 5 + 1 = 31$, which is a unit, hence not a nontrivial product of primes.

Joke (Lenstra). "There are infinitely many composite numbers. *Proof:* Multiply together the first n primes and don't add 1."

According to

```
http://www.utm.edu/research/primes/largest.html
```

the largest known prime is

$$p = 2^{6972593} - 1,$$

which is a number having over two million[1] decimal digits. Euclid's theorem implies that there definitely *is* a bigger prime number. However, nobody has yet found it *and proved that they are right.* In fact, determining whether or not a number is prime is an extremely interesting problem. We will discuss this problem more later.

2 Primes of the form $ax + b$

Next we turn to primes of the form $ax + b$. We assume that $\gcd(a, b) = 1$, because otherwise there is no hope that $ax + b$ is prime *infinitely* often. For example, $3x + 6$ is only prime for one value of x.

Proposition 2.1. *There are infinitely many primes of the form* $4x - 1$*.*

Why might this be true? Let's list numbers of the form $4x-1$ and underline the ones that are prime:

$$\underline{3}, \underline{7}, \underline{11}, 15, \underline{19}, \underline{23}, 27, \underline{31}, 35, 39, \underline{43}, \underline{47}, \ldots$$

It certainly looks plausible that underlined numbers will continue to appear. The following PARI program can be used to further convince you:

```
f(n, s=0) = for(x=1, n, if(isprime(4*x-1), s++); s
```

Proof. The proof is similar to the proof of Euclid's Theorem, but, for variety, I will explain it in a slightly different way.

Suppose $p_1, p_2, \ldots, p_n$ are primes of the form $4x - 1$. Consider the number

$$N = 4p_1 \times p_2 \times \cdots \times p_n - 1.$$

Then $p_i \nmid N$ for any i. Moreover, not every prime $p \mid N$ is of the form $4x+1$; if they all were, then N would also be of the form $4x + 1$, which it is not. Thus there is a $p \mid N$ that is of the form $4x-1$. Since $p \neq p_i$ for any i, we have found another prime of the form $4x - 1$. We can repeat this process indefinitely, so the set of primes of the form $4x - 1$ is infinite. □

Example 2.2. Set $p_1 = 3$, $p_2 = 7$. Then

$$N = 4 \times 3 \times 7 - 1 = \underline{83}$$

is a prime of the form $4x - 1$. Next

$$N = 4 \times 3 \times 7 \times 83 - 1 = \underline{6971},$$

[1]It has exactly 2098960 decimal digits.

which is a again a prime of the form $4x - 1$. Again:

$$N = 4 \times 3 \times 7 \times 83 \times 6971 - 1 = 48601811 = 61 \times \underline{796751}.$$

This time 61 is a prime, but it is of the form $4x + 1 = 4 \times 15 + 1$. However, 796751 is prime and $(796751 - (-1))/4 = 199188$. We are unstoppable

$$N = 4 \times 3 \times 7 \times 83 \times 6971 \times 796751 - 1 = \underline{5591} \times 6926049421.$$

This time the small prime, 5591, is of the form $4x - 1$ and the large one is of the form $4x + 1$. Etc!

Theorem 2.3 (Dirichlet). *Let a and b be integers with $\gcd(a, b) = 1$. Then there are infinitely many primes of the form $ax + b$.*

The proof is out of the scope of this course. You will probably see a proof if you take Math 129 from Cornut next semester.

3 How many primes are there?

There are infinitely many primes.

Can we say something more precise?

Let's consider a similar question:

Question 3.1. How many even integers are there?

Answer: *Half* of all integers.

Question 3.2. How many integers are there of the form $4x - 1$?

Answer: *One fourth* of all integers.

Question 3.3. How many perfect squares are there?

Answer: Zero percent of all numbers, in the sense that the limit of the proportion of perfect squares to all numbers converges to 0. More precisely,

$$\lim_{x \to \infty} \#\{n : n \leq x \text{ and } n \text{ is a perfect square }\}/x = 0,$$

since the numerator is roughly $\sqrt{x}$ and $\sqrt{x}/x \to 0$.

A better question is:

Question 3.4. How many numbers $\leq x$ are perfect squares, as a function of x?

Answer: Asymptotically, the answer is $\sqrt{x}$.

So a good question is:

Question 3.5. How many numbers $\leq x$ are prime?

Let

$$\pi(x) = \#\{\text{ primes } p \leq x\}.$$

For example,

$$\pi(6) = \#\{2, 3, 5\} = 3.$$

We can compute a few more values of $\pi(x)$ using PARI:

```
? pi(x, c=0) = forprime(p=2,x,c++); c;
? for(n=1,7,print(n*100,"\t",pi(n*100)))
100 25
200 46
300 62
400 78
500 95
600 109
700 125
```

Now draw a graph on the blackboard. It will look like a straight line...

Gauss spent some of his free time counting primes. By the end of his life, he had computed $\pi(x)$ for x up to 3 million.

$$\pi(3000000) = 216816.$$

(I don't know if Gauss got the right answer.) Gauss conjectured the following:

Theorem 3.6 (Hadamard, Vallée Poussin, 1896). *$\pi(x)$ is asymptotic to $x/\log(x)$, in the sense that*

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

I will not prove this theorem in this class. The theorem implies that $x/(\log(x)-a)$ can be used to approximate $\pi(x)$, for any a. In fact, $a = 1$ is the best choice.

```
? pi(x, c=0) = forprime(p=2,x,c++); c;
? for(n=1,10,print(n*1000,"\t",pi(n*1000),"\t",n*1000/(log(n*1000)-1)))
1000 168 169.2690290604408165186256278
2000 303 302.9888734545463878029800994
3000 430 428.1819317975237043747385740
4000 550 548.3922097278253264133400985
5000 669 665.1418784486502172369455815
6000 783 779.2698885854778626863677374
7000 900 891.3035657223339974352567759
8000 1007 1001.602962794770080754784281
9000 1117 1110.428422963188172310675011
10000 1229 1217.976301461550279200775705
```

Remark 3.7.

3.1 Counting Primes Today

People all over the world are counting primes, probably even as we speak. See, e.g.,

`http://www.utm.edu/research/primes/howmany.shtml`

`http://numbers.computation.free.fr/Constants/Primes/Pix/pixproject.html`

A huge computation:

$$\pi(10^{22}) = 201467286689315906290$$

(I don't know for sure if this is right...)

3.2 The Riemann Hypothesis

The function

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\log(x)} dx.$$

is also a good approximation to $\pi(x)$.

The famous **Riemann Hypothesis** is equivalent to the assertion that

$$\pi(x) = \mathrm{Li}(x) + O(\sqrt{x}\log(x)).$$

(This is another \$1000000 prize problem.)

```
pi(10^22)      = 201467286689315906290
Li(10^22)      = 201467286691248261498.1505...    (using Maple)
Log(x)/(x-1)   = 201381995844659893517.7648...    (pari)
```