Lecture 7

January 15, 2018

1 Facts about primes

Since this is an "Effective Number Theory" school, we should list some effective results.

Proposition 1. *(i) The inequality*

$$\frac{x}{\log x - 0.5} < \pi(x) \qquad \textit{holds for all } x \geq 67.$$

(ii) The inequality

$$\pi(x) < \frac{x}{\log x - 1.5} \qquad \text{holds for all } x > e^{3/2}.$$

You already know that the Prime Number Theorem implies that $p_n = n \log n + o(n \log n)$, where p_n is the nth prime. More specifically, the following result holds.

Proposition 2. *Let p_n be the nth prime. Then the inequality*

$$n(\log n - \log\log n - 1.5) < p_n \qquad \textit{holds for all } n \geq 2,$$

and

$$p_n < n(\log n - \log\log n - 0.5) \qquad \textit{holds for all } n \geq 20.$$

Problem 1. *Find the largest n such that the interval $[n, 2n)$ contains less than 100 primes.*

Problem 2. *The Fibonacci sequence $\{F_n\}_{n\geq 1}$ is given by $F_1 = F_2 = 1$ and $F_{n+2} = F_{n+1} + F_n$ for all $n \geq 1$. Show that the interval $[F_n, F_{n+1})$ contains a prime for all $n \geq 2$.*

Let $2 = p_1 < p_2 < \ldots < p_n < \ldots$ be the sequence of all prime numbers. A postulate due to Bertrand and proved by Chebyshev assert that the interval $[x, 2x)$ contains a prime for all $x > 1$. In particular, $p_{n+1} \leq 2p_n$ or $p_{n+1} - p_n \leq p_n$. By probabilistic reasoning, Cramer was lead in 1936 to conjecture that

$$\limsup_{n\to\infty} \frac{p_{n+1} - p_n}{(\log p_n)^2} \leq 1.$$

The best known upper bound on $p_{n+1} - p_n$ is $p_{n+1} - p_n < p_n^{0.525}$ for all n sufficiently large and it was proved by Baker, Harman and Pintz in 2001.

What about small gaps between consecutive primes? It is conjectured that $p_{n+1} - p_n = 2$ holds for infinitely many n. The primes p and $p + 2$ are said to form a *twin pair*. If we assume that the primes are *randomly* distributed and an integer x is prime with an "expectation" of about $1/\log x$, then we would suspect that p and $p + 2$ be simultaneously primes with the expectation $1/(\log p)^2$. Thus, we expect about $x/(\log x)^2$ primes $p \leq x$ with $p + 2$ also prime. A more careful heuristic suggests that there are about $Cx/(\log x)^2$ such primes p, where $C > 0$ is a certain constant ($C \neq 1$). In 1920, Viggo Brun proved an upper bound for the number of twin primes $p \leq x$ of the above shape. In the same paper he also proved that there are infinitely many primes p with $p + 2$ a P_9. The record belongs to Chen who in 1973 showed that there are more than $0.6x/(\log x)^2$ primes $p < x$ such that $p + 2$ is a P_2, that it is it is prime or a product of two primes, if x is sufficiently large.

Until recently it wasn't even known that

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{\log p_n} = 0,$$

but this was recently proved by Goldston, Pintz and Yildirim in 2005 who in fact proved the stronger result that

$$\liminf_{n\to\infty} \frac{p_{n+1} - p_n}{(\log p_n)^{1/2}(\log\log p_n)^2} < \infty.$$

In 2013, Yitang Zhang introduced new ideas into the proof of Goldston, Pinz and Yildirim and showed that

$$p_{n+1} - p_n < 70,000,000$$

for infinitely many n. Since then Maynard reduced the number 7×10^7 to 600 and Terrence Tao and a Polymath Project reduced it further to 246.

Maynard has also shown that for each $m \geq 1$, there exists a constant $c(m)$ such that

$$p_{n+m} - p_n < c(m)$$

holds for infinitely many n. The function $c(m)$ can be taken to be Cm^3e^{4m} for some absolute constant C which Maynard did not compute.

In 1955, Ricci proved that the set of cluster points of $\{(p_{n+1}-p_n)/\log p_n : n \in \mathbb{Z}\}$ has positive measure. The only points of accumulation of the above set which are known are 0 and 1.

In the 1930's, Erdős proved that for infinitely many positive integers n,

$$p_{n+1} - p_n > c_1 \log p_n \frac{\log\log p_n}{(\log\log\log p_n)^2}$$

holds with some positive constant c_1. A few years later Rankin added a factor of $\log\log\log\log p_n$. Last year, in 2017, Ford, Green, Konyagin, Maynard and Tao removed a factor of $\log\log\log p_n$ from the denominator!

In conclusion, we have much work to do. We don't know too many primes. In fact,

$$\sum_{p \text{ known prime}} \frac{1}{p} < 4.$$

Primes of the form $2^p - 1$ are called *Mersenne primes*. In the preface of his 1644 book *Cogitata Physica-Matematica* the French priest Marin Mersenne claimed that the primes $p \leq 257$ such that $2^p - 1$ is prime are precisely $\{2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257\}$. The list was wrong. The correct list was not obtained until 1947. Nevertheless, the numbers of the form $2^p - 1$ are called Mersenne numbers and denoted M_p. In 1876, Edouard Lucas [**?**] proved that M_{127} is prime. Note that $M_{127} = 2^{127} - 1 > (2^{10})^{12} = (1024)^{12} > 10^{36}$ has at least 36 digits. Lucas discovered a property of the Mersenne numbers which makes them easy to test for primality. The largest primes nowadays are Mersenne primes which are tested for primality using Lucas's primality test. There are only 50 Mersenne primes known. The largest Mersenne prime is $2^{77232917} - 1$ and was found by John Pace two weeks ago (on December 26, 2017). In fact, if we assume that the probability that $2^p - 1$ is prime is $1/\log(2^p - 1)$ then the number of Mersenne primes with $p \leq x$ should be say

$$\sum_{p \leq x} \frac{1}{\log(2^p - 1)} \sim \frac{1}{\log 2} \sum_{p \leq x} \frac{1}{p}.$$

The above series tends to infinity but not very quickly. In fact, this heuristic is perhaps totally bogus. We recommend the book by Crandall and Pomerance on "Prime Numbers; a computational perspective" for a more "believable" heuristic. At any rate, while we expect infinitely many Mersenne primes, we do not expect them to be too numerous. A similar heuristic can be applied to conjecture that there should be only finitely many Fermat primes.

Let $b_1 < \cdots < b_k$ be positive integers. For each prime p, let $\nu(p)$ be the number of residue classes modulo p occupied by the integers $b_1, \ldots, b_k$. That is

$$\nu(p) = \#\{b_i \pmod p \ : \ i = 1, \ldots, k\}.$$

The following conjecture of Dickson is known as the *Prime k-tuplets Conjecture.*

Conjecture 1. *If $\nu(p) < p$ holds for all prime numbers p, then there exist infinitely many positive integers n such that*

$$n + b_1, \ n + b_2, \ \ldots, \ n + b_k \tag{1}$$

are all primes.

For $k = 2, \ b_1 = 0, \ b_2 = 2$ we get the twin primes conjecture. Note that the condition $\nu(p) < p$ is necessary if we want that all numbers from list (**??**) to be primes for infinitely many n. (Why?) Note also that it suffices to check this condition only when $p < k$. Almost 30 years later, Hardy and Littlewood proposed a quantitative version of Conjecture **??**.

Hardy and Littlewood also conjectured the following inequality.

Conjecture 2.

$$\pi(x + y) \le \pi(x) + \pi(y) \qquad \textit{for all } x > 1, \ y > 1. \tag{2}$$

Some evidence for the above inequality is provided by the Prime Number Theorem: If we were to replace the function $\pi(x)$ by the function $x/\log x$, inequality (**??**) would read

$$\frac{x + y}{\log(x + y)} < \frac{x}{\log x} + \frac{y}{\log y} \qquad \text{for all } x > 1, \ y > 1,$$

which is evident. However, in 1972, Hensley and Richards proved that the two conjectures **??** and **??** above are incompatible. Their proof is the following:

Call an increasing sequence of positive integers $b_1 < \cdots < b_k$ *admissible* if $\nu(p) < p$ for all primes p. Let

$$\rho^*(y) = \max\{k \in \mathbb{Z}_{\geq 0} \ : \ x < b_1 < \cdots < b_k \leq x + y \text{ admissible}\}.$$

Thus, $\rho^*(y)$ is the maximal length ($\leq y$) of an admissible sequence of positive integers squeezed in an interval of length y. Then Hensley and Richards proved that

$$\lim_{y \to \infty} (\rho^*(y) - \pi(y)) = \infty.$$

Let us assume that y is such that $\rho^*(y) > \pi(y)$ and let $x < b_1 < \cdots < b_k \leq x + y$ be admissible, where $k = \rho^*(y)$. If Conjecture **??** is true, then for some n we have that $n + b_1 < \cdots < n + b_k$ are all primes. Since they belong to $(n + x, n + x + y]$, we get that

$$\rho^*(y) = k \leq \pi(x + n + y) - \pi(x + n),$$

or

$$\pi(x + n + y) \geq \pi(x + n) + \rho^*(y) > \pi(x + n) + \pi(y),$$

contradicting inequality (**??**).

It is believed that Conjecture **??** holds. This conjecture was extended by Schinzel in a joint paper with Sierpiński. This conjecture is known as *Schinzel's Hypothesis H.*

Conjecture 3. *Let $f_1(X), \ldots, f_k(X) \in \mathbb{Z}[X]$ be non-constant polynomials of positive leading terms. Assume that*

(i) $f_i(X)$ is irreducible for all $i = 1, \ldots, k$.

(ii) there does not exist a prime p such that $p \mid f_1(n) f_2(n) \cdots f_k(n)$ for all $n \geq 0$.

Then there exist infinitely many positive integers n such that

$$f_1(n), \ f_2(n), \ \ldots, \ f_k(n)$$

are all prime numbers.

Bateman and Horn proposed a quantitative version of this conjecture. Namely, for each prime p let

$$\omega(p) = \#\{0 \leq n \leq p - 1 \ : \ f_1(n) f_2(n) \cdots f_k(n) \equiv 0 \pmod{p}\}.$$

Put

$$\pi_{f_1,\ldots,f_k}(x) = \#\{n \le x \ : \ f_1(n),\ \ldots,\ f_k(n) \text{ are all primes}\}.$$

Then $\pi_{f_1,\ldots,f_k}(x)$ should asymptotically equal

$$C(f_1,\ldots,f_k)\,\frac{1}{d_1\cdots d_k}\,\frac{x}{(\log x)^k},$$

where $d_i = \deg(f_i)$ and the constant $C(f_1,\ldots,f_k)$ equals

$$C(f_1,\ldots,f_k) = \prod_{p\ge 2}\frac{1-\omega(p)/p}{(1-1/p)^k}. \tag{3}$$

It is not even clear from the above formula (**??**) that $C(f_1,\ldots,f_k)$ represents a product convergent to a limit > 0. I leave it as an exercise to you to prove that this is indeed so when $f_i(X) = a_iX+b_i$ ($a_i > 0$, $\gcd(a_i,b_i) = 1$) is linear for all $i = 1,\ldots,k$, and, in particular, to discover the Hardy and Littlewood effective form of the Prime k-tuplets Conjecture **??**.

2 Homework

Problem 3. *Let p_n be the nth prime. Show that*

$$\frac{1}{4}n\log n < p_n.$$

Problem 4. *Show that*

$$p_n < 12\left(n\log n + \log\left(\frac{12}{e}\right)\right).$$

Problem 5. *Show that there do not exist polynomials $P(x)$ and $Q(x)$ such that $\pi(x) = P(x)/Q(x)$.*

Problem 6. *Show that for every $n > 1$ there exist n consecutive composite numbers.*

Problem 7. *Let $S_n = \sum_{i=1}^{n} p_i$. Prove that the interval $[S_n, S_{n+1}]$ contains a perfect square.*

Problem 8. *Show using the Prime k-tuples Conjecture that for every positive integer K there exists a positive integer A such that $n^2 - n + A$ is prime for all $n = 0,\ 1,\ \ldots, K$ (Hint: Use $b_k = k^2 - k$).*