Sum of Squares (Continued)

Week 13

Here – at last – is the proof we have been aiming for.

Theorem 18.3 *If p is an odd prime and $p \equiv 1 \pmod 4$, then p is representable as the sum of two squares.*

Proof. Since $p \equiv 1 \pmod 4$, -1 is a quadratic residue $\pmod p$. Thus we can find an integer a satisfying $a^2 \equiv -1 \pmod p$.

- **How?** By Lemma 18.4, $\left[\left(\frac{p-1}{2}\right)!\right]^2 \equiv -1 \pmod p$, so we may choose a to be the least residue of $\left(\frac{p-1}{2}\right)! \pmod p$.

Then $(a, p) = 1$ (for any common divisor of a and p must also divide -1) and by Thue's Lemma (Lemma 18.5), the congruence

$$ax \equiv y \pmod p$$

is satisfied by x_0, y_0 such that $0 < |x_0|, |y_0| < \sqrt{p}$, i.e., such that

$$0 < x_0^2 + y_0^2 < 2p. \tag{18.4}$$

But then

$$\begin{array}{lrcll} & (ax_0)^2 & \equiv & y_0^2 \pmod p & \text{because } ax_0 \equiv y_0 \pmod p \\ \text{i.e.} & a^2x_0^2 & \equiv & y_0^2 \pmod p & \\ \text{i.e.} & -x_0^2 & \equiv & y_0^2 \pmod p & \text{because } a^2 \equiv -1 \pmod p. \\ \text{i.e.} & x_0^2 + y_0^2 & \equiv & 0 \pmod p & \end{array}$$

Thus there exists an integer k such that

$$x_0^2 + y_0^2 = kp. \tag{18.5}$$

By (18.4), (18.5) is only possible if $k = 1$, which proves the theorem. ■

★ It can also be proved that except for the signs of the integers x_0 and y_0, and the order of the summands, this representation of the **prime** p as the sum of two squares is unique.

Example: We illustrate the above method by using it to write

$$13000 = 2^3 5^3 13 = 2 \cdot 10^2 \cdot 5 \cdot 13$$

as the sum of two squares.

- Note that
$$5 = 1^2 + 2^2 \quad \text{and} \quad 13 = x_0^2 + y_0^2 = 2^2 + 3^2.$$
- Now we apply Lemma 18.1:
$$5 \cdot 13 = (1^2 + 2^2)(2^2 + 3^2) = (1 \cdot 2 + 2 \cdot 3)^2 + (1 \cdot 3 - 2 \cdot 2)^2 = 8^2 + 1^2,$$
and again:
$$2 \cdot (5 \cdot 13) = (1^2 + 1^2)(8^2 + 1^2) = 9^2 + 7^2$$
and then Lemma 18.2:
$$10^2(2 \cdot (5 \cdot 13)) = 10^2(9^2 + 7^2) = 90^2 + 70^2,$$
which is what we want.

- Note: Instead of $5 \cdot 13 = (1^2+2^2)(2^2+3^2)$, we could have used $5 \cdot 13 = (1^2+2^2)(3^2+2^2) = 7^2 + 4^2$ to obtain $13000 = 110^2 + 30^2$.

18.3.4 Summary

Someone asks you to write the integer n as the sum of two squares, if this is possible. Proceed as follows.

1. Write n in prime-power decomposition as $n = q_1^{e_1} q_2^{e_2} ... q_k^{e_k}$.
2. Look at the q_i for which e_i is odd.
   (a) If $q_i \equiv 3 \pmod 4$, then n is not the sum of two squares.
   (b) If each such q_i is equal to 2 or congruent to 1 $\pmod 4$, continue.
3. Write n in the form
$$n = k^2 p_1 p_2 ... p_r \quad \text{or} \quad n = 2k^2 p_1 p_2 ... p_r,$$
for some k and some r, where the p_i are all distinct and $p_i \equiv 1 \pmod 4$ for each i.

4. For each p_i, write p_i as the sum of two squares.

5. Use Lemma 18.1 repeatedly (and the fact that $2 = 1^2+1^2$) to write $p_1p_2...p_r$ or $2p_1p_2...p_r$ as the sum of two squares.

6. Finish by using Lemma 18.2 if required, that is, if $k \neq 1$.

18.3.5 Sum of three squares

Every positive integer n can be written as the sum of three squares, except if n is of the form $n = 4^e(8k + 7)$ for some integers e and k.

18.3.6 Sum of four squares

Every positive integer n can be written as the sum of four squares. Assumed to be true by Diophantus, eventually proved (over a long period of time) mostly by Euler with Lagrange proving the last missing case.

18.3.7 Sum of cubes

Every positive integer can be written as the sum of nine cubes.

18.3.8 Sum of k^{th} powers

Conjecture: Every positive integer can be written as the sum of $2^k + \lfloor (3/2)^k \rfloor - 2$ k^{th} powers.

18.3.9 Goldbach Conjecture

($\$10^6$ Problem) Every even integer greater than two can be written as the sum of two primes.

Section 19

More About Primes

19.1 The Prime Number Theorem

Consider the question:

How many prime numbers are there?

Well, we know Euclid proved that there are infinitely many primes. So let us change the question a little:

How many prime numbers are there, relative to other natural numbers?

- If we consider the first ten positive integers, we notice that there are no fewer than four primes among them. There are eight primes among the first 20 positive integers, but only 10 primes among the first 30 positive integers, 25 primes less than 100, and 168 primes less than 1000.
- We see that the difference between two primes can be as small as 1 (well, between only one pair of primes, 2 and 3), or 2, between twin primes, but we don't know whether there are infinitely many twin primes.
- We also know that the difference between consecutive primes can be arbitrary: for any $n \in \mathbb{Z}^+$, the numbers $n!+2,\ n!+3,\ \cdots,\ n!+n$ are all composite; thus if p is the largest prime less than $n!+2$, and q is the smallest prime greater than $n!+n$, then p and q are consecutive primes such that $q-p \geq n$, that is, there are at least $n-1$ composite numbers between p and q.
- Thus it seems that as we consider increasingly large numbers, the primes become increasingly scarce.

We rephrase the second question:

x	$\pi(x)$	$x/\ln x$	$\frac{\pi(x)\ln x}{x}$
10	4	4.343	0.921 03
20	8	6.676	1.198 3
50	15	12.781	1.173 6
100	25	21.715	1.151 3
200	46	37.748	1.218 6
300	62	52.597	1.178 8
400	78	66.762	1.168 3
500	95	80.456	1.180 8
700	125	106.85	1.169 8
1000	168	144.76	1.160 5
5000	669	587.05	1.139 6
10000	1229	1085.7	1.132 0
100000	9592	8685.9	1.104 3

Table 19.1: The number of primes less than x

What is the density of the primes among the natural numbers?

Euler tried to discover "the secret of the primes" but it eluded him. In 1792, when only 15 years old, Gauss proposed that there are approximately $n/\ln n$ prime numbers less than n. Legendre independently conjectured the same ratio in 1798. Neither, however, could prove it.

For a real number $x > 1$, let $\pi(x)$ denote the number of primes less than x. The number $\pi(x)$ and the ratio $\frac{\pi(x)}{x/\ln x}$ for values of x up to $100,000$ are shown in Table 19.1.

Gauss and Legendre conjectured, therefore, what later became known as the

Theorem 19.1 **The Prime Number Theorem:** $\lim_{x\to\infty} \pi(x)/\dfrac{x}{\ln x} = 1.$

The first person to show that $\pi(x)$ has order of magnitude $x/\ln x$ was Tchebychef in 1852, who showed that

$$B < \frac{\pi(x)}{x/\ln x} < \frac{6B}{5},$$

where $B \approx 0.92129$ and $\frac{6B}{5} \approx 1.10555$. His argument was entirely elementary (meaning, it didn't use analysis, not that it was easy) and made use of properties of factorials. He also showed that $\lim_{x\to\infty} \pi(x)/\frac{x}{\ln x}$, if it existed, had to equal 1. Over the years there were several improvements of Tchebychef's result.

The first proof of the prime number theorem was given independently by Hadamard and de la Vallée Poussin in 1896. The proof was not elementary and made use of Hadamard's theory of integral functions applied to the Riemann zeta function. The general feeling was that an elementary proof was unlikely to exist.

In 1948 the mathematical world was stunned when Paul Erdös announced that he and Atle Selberg had found a truly elementary (but not easy) proof of the prime number theorem that used only the simplest properties of the logarithm function. Unfortunately, this announcement and subsequent events led to a bitter priority dispute between these two mathematicians, each accusing the other of using each other's results.

For his work on the elementary proof of the Prime Number Theorem, the zeros of the Riemann zeta function, and the development of the Selberg sieve method, Selberg received the Fields Medal (as prestigious as the Nobel Prize) in 1950. The Selberg sieve method, a cornerstone in elementary number theory, is the basis for Jing-Run Chen's spectacular proof that every positive even integer is the sum of a prime and a number having at most two prime factors.

Paul Erdös was a legendary eccentric and arguably the most prolific mathematician of the 20th century, in terms of both the number of problems he solved and the number of problems he convinced others to tackle. His work provided the foundations for graph and hypergraph theory and the probabilistic method, and had applications in combinatorics and number theory. For his work on the Prime Number Theorem he received the Frank Nelson Cole Prize, which recognizes a notable paper in number theory published during the preceding six years, in 1951.

19.2 Primes in Arithmetic Progression

We saw in Section 19.1 that there are arbitrarily long sequences of composite numbers. On the other hand, there are also

arbitrarily long sequences of prime numbers in arithmetic progression

(but not infinitely long such sequences). In number theory, the phrase "*primes in arithmetic progression*" refers to at least three prime numbers that are consecutive terms in an arithmetic progression, for example the primes $3, 7, 11$ (it does not matter that 5 is also prime). It is easy to see that any arithmetic progression of primes has finite length. Consider any prime number p, any integer $a > 1$ and the arithmetic progression

$$p,\ p+a,\ p+2a,\ \ldots,\ p+pa.$$

Then $p|(p+pa)$, hence $p+pa$ is not prime. On the other hand, the proof that there are arbitrarily long (finite) sequences of prime numbers in arithmetic progression is very difficult and has a long history. It is not even clear when the existence of such sequences was first conjectured, but Lagrange and Waring mentioned the problem in the 1770s. It was finally settled in 2004 by Ben Green, while he was a PIMS postdoctoral fellow at UBC in Vancouver, and Terence Tau.

Theorem 19.2 **The Green-Tau Theorem:** *The prime numbers contain infinitely many arithmetic progressions of length k for all k.*

The problem of finding long arithmetic progressions in the primes has also attracted the interest of computational mathematicians. As of September 2019, the longest and largest known sequence of primes in arithmetic progression is 27, namely the numbers

$$224584605939537911 + 81292139 \cdot 23 \cdot n, \quad \text{for} \quad n = 0, 1, ..., 26,$$

found in 2019. The first sequence of 26 primes in arithmetic progression,

$$43142746595714191 + 23681770 \cdot 223092870 \cdot n, \text{ for } n = 0, 1, ..., 25,$$

was found on April 12, 2010 by Benoãt Perichon.

Dirichlet proved a related result on arithmetic progressions of not necessarily prime numbers, but that contain prime numbers.

Theorem 19.3 **Dirichlet's Prime Number Theorem:** *If $(a, d) = 1$, then the arithmetic progression*

$$a, \ a + d, \ a + 2d, \ ...,$$

contains infinitely many primes.

Hence if $(a, d) = 1$, then there are infinitely many primes congruent to $a \pmod{d}$. This theorem extends Euclid's theorem that there are infinitely many prime numbers.

19.3 Twin Primes

For x a real number, let $\pi_2(x)$ denote the number of primes $p \leq x$ for which $p + 2$ is also prime. The density of twin primes can be expressed as follows.

Theorem 19.4 *For $x \geq 3$,*

$$\pi_2(x)/x = O\frac{(\ln \ln x)^2}{(\ln x)^2}.$$

It follows from Theorem 19.4 that the sum of the reciprocals of twin primes converges to a finite value, known as **Brun's constant**, as stated in the theorem that bears his name**. Brun's Theorem** was proved in 1919 (100 years ago) by the Norwegian mathematician Viggo Brun.

Theorem 19.5 **Brun's Theorem** *The sum, taken over all primes p such that $p + 2$ is also prime,*

$$\sum(\frac{1}{p} + \frac{1}{p+2}) = (\frac{1}{3} + \frac{1}{5}) + (\frac{1}{5} + \frac{1}{7}) + (\frac{1}{11} + \frac{1}{13}) + \cdots$$

either has finitely many terms, or it has infinitely many terms but converges. Its value is known as Brun's constant.

Because the sum of the reciprocals of the twin primes converges, it is not possible to conclude from this result whether there are finitely many or infinitely many twin primes. Brun's constant could be an irrational number only if there are infinitely many twin primes. It is still unknown whether there are infinitely many twin primes, but it is conjectured that this is the case.

Conjecture 19.6 **Weak Twin Prime Conjecture** *There are infinitely many twin primes.*

Conjecture 19.7 **Strong Twin Prime Conjecture** *For any real number x, the number of twin primes less than or equal to x is approximated by the expression*

$$2 \prod_{\substack{p \geq 3 \\ p \text{ prime}}} \frac{p(p-2)}{(p-1)^2} \int_2^x \frac{dx}{(\ln x)^2} \approx 1.320323632 \int_2^x \frac{dx}{(\ln x)^2}.$$

The American mathematician Yitang Zhang showed in 2013 that there are an infinite number of prime pairs differing by no more than 70 million. By refining Zhang's techniques this bound was improved to 246 in 2014. The infinite product

$$\prod_{\substack{p \geq 3 \\ p \text{ prime}}} \frac{p(p-2)}{(p-1)^2} = \prod_{\substack{p \geq 3 \\ p \text{ prime}}} \left(1 - \frac{1}{(p-1)^2}\right)$$

is essentially a constant, called the **twin primes constant,** and is approximately 0.6601618158...

.