E-320: Teaching Math with a Historical Perspective **Oliver Knill, 2010**

Lecture 4: Number Theory

Number theory studies the structure of integers and equations with integer solutions. Gauss called it the "Queen of Mathematics". In this lecture, we look at a few theorems and open problems.

An integer larger than 1 which is divisible only by 1 and itself is called a **prime number**. The number $2^{43112609} - 1$ is the largest known prime number. It has 12978189 digits. **Euclid** was the first to prove that there are infinitely many primes: [Proof. Assume there are only finitely many primes $p_1 < p_2 < \dots < p_n$. Then $n = p_1 p_2 \cdots p_n + 1$ can not be divisible by $p_1, \dots, p_n$. Therefore, it is either a prime or divisible by a prime larger than p_n.] Primes become more sparse as larger as they get. One of the most important results is the **prime number theorem** which states that the n'th prime number has approximately the size $n \log(n)$. For example the $n = 10^{12}$'th prime is $p(n) = 29996224275833$ and $n \log(n) = 27631021115928.545...$ and $p(n)/(n \log(n)) = 1.0856...$ Many questions about prime numbers are unsettled: Here are four open problems on prime numbers. The third uses the notation $(\Delta a)_n = |a_{n+1} - a_n|$ to get the difference sequence of a given sequence: $\Delta^2(1, 4, 9, 16, 25...) = \Delta(3, 5, 7, 9, 11, ...) = (2, 2, 2, 2, ...)$.

Twin prime conjecture	there are infinitely many primes p such that $p + 2$ is prime.
Goldbach conjecture	every even integer $n > 2$ is a sum of two primes.
Gilbreaths conjecture	If p_n enumerates the primes, then $(\Delta^k p)_1 = 1$ for all $k > 0$.
Andrica conjecture:	The prime gap estimate $\sqrt{p_{n+1}} - \sqrt{p_n} < 1$ holds for all n.

If the sum of the proper divisors of a number is equal the number itself, it is called a **perfect number**. For example, 6 is a perfect number; its proper divisors $1, 2, 3$ sum up to 6. All currently known perfect numbers are even. The question whether odd perfect numbers exist, is not settled. It is one of the oldest known problems in mathematics. Perfect numbers were known already by Pythagoras and his followers. Calendar coincidences like the fact that we have 6 work days and the moon needs "perfect" 28 days to circle the earth could have helped to add some "mystical properties" to perfect number. **Euclid of Alexandria** (300-275 BC) was the first to realize that if $2^p - 1$ is prime then $k = 2^{p-1}(2^p - 1)$ is a perfect number: [Proof: let $\sigma(n)$ be the sum of **all** factors of n, including n. Now $\sigma(2^n - 1)2^{n-1}) = \sigma(2^n - 1)\sigma(2^{n-1}) = 2^n(2^n - 1) = 2 \cdot 2^n(2^n - 1)$ shows $\sigma(k) = 2k$ verifying k is perfect.] Around 100 AD, **Nicomachus of Gerasa** (60-120) classified in his work "Introduction to Arithmetic" numbers on the concept of perfect numbers and lists four perfect numbers. Only much later it was established that Euclid got all perfect numbers: Euler showed that all even perfect numbers are of the form $(2^n - 1)2^{n-1}$, where $2^n - 1$ is prime. The corresponding$2^n - 1$ is called a **Mersenne prime**. [Proof: Assume $N = 2^k m$ is perfect where m is odd and $k > 0$. Then $2^{k+1} m = 2N = \sigma(N) = (2^{k+1} - 1)\sigma(m)$. This gives $\sigma(m) = 2^{k+1} m/(2^{k+1} - 1) = m(1 + 1/(2^{k+1} - 1)) = m + m/(2^{k+1} - 1)$. Because $\sigma(m)$ and m are integers, also $m/(2^{k+1} - 1)$ is an integer. It must also be a factor of m. The only way that $\sigma(m)$ can be the sum of only two of its factors is that m is prime and so $2^{k+1} - 1 = m$.] The first 39 **known Mersenne primes** are of the form $2^n - 1$ with n = 2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213, 19937, 21701, 23209, 44497, 86243, 110503, 132049, 216091, 756839, 859433, 1257787, 1398269, 2976221, 3021377, 6972593, 13466917. There are 8 more known from which one does not know the rank of the corresponding Mersenne prime: n = 20996011, 24036583, 25964951, 30402457, 32582657, 37156667, 42643801, 43112609. The last 3 were found only in the 2 years (2008 and 2009). It is unknown whether there are infinitely many.

A polynomial equations for which all coefficients and variables are integers is called a **Diophantine equation**. The first Diophantine equation which was studied is $x^2 + y^2 = z^2$. A solution (x, y, z) of this equation in positive integers is called a **Pythagorean triple**. For example, $(3, 4, 5)$ is a Pythagorean triple. Since 1600 BC, it is known that all solutions to this equation are of the form $(x, y, z) = (2st, s^2 - t^2, s^2 + t^2)$ or $(x, y, z) = (s^2 - t^2, 2st, s^2 + t^2)$, where s, t are different integers. [Proof. Either x or y has to be even because if both are odd, then the sum $x^2 + y^2$ is even but not divisible by 4 but the right hand side is either odd or divisible by 4. Move the even one, say x^2 to the left and write $x^2 = z^2 - y^2 = (z - y)(z + y)$, then the right hand side contains a factor 4 and is of the form $4s^2t^2$. Therefore $2s^2 = z - y, 2t^2 = z + y$. Solving for z, y gives $z = s^2 + t^2, y = s^2 - t^2$, $x = 2st$.]
Analyzing Diophantine equations can be very difficult. Only recently, one has established that the **Fermat equation** $x^n + y^n = z^n$ has no solutions with $xyz \neq 0$ if $n > 2$. Here are some **open problems** for Diophantine equations. Are there nontrivial solutions to the following Diophantine equations?

$x^6 + y^6 + z^6 + u^6 + v^6 = w^6$	$x, y, z, u, v, w > 0$
$x^5 + y^5 + z^5 = w^5$	$x, y, z, w > 0$
$x^k + y^k = n!z^k$	$k \geq 2, n > 1$
$x^a + y^b = z^c, a, b, c > 2$	$\gcd(a, b, c) = 1$

The last equation is called the **Super Fermat** equation. A banker **Andrew Beals** once sponsored a prize of $100'000$ dollars for a proof or counter example to the statement: "If $x^p + y^q = z^r$ with $p, q, r > 2$, then $\gcd(x, y, z) > 1$."
Given a prime like 7 and a number n we can add or subtract multiples of 7 from n to get a number in $\{0, 1, 2, 3, 4, 5, 6\}$. We write for example $19 = 12 \bmod 7$ because 12 and 19 both leave the rest 5 when dividing by 7. Or $5 * 6 = 2 \bmod 7$ because 30 leaves the rest 2 when dividing by 7. The most important theorem in elementary number theory is **Fermat's little theorem** which tells that if a is an integer and p is prime then $a^p - a$ is divisible by p. For example $2^7 - 2 = 126$ is divisible by 7. [Proof: use induction. For $a = 0$ it is clear. The binomial expansion shows that $(a + 1)^p - a^p - 1$ is divisible by p. This means $(a + 1)^p - (a + 1) = (a^p - a) + mp$ for some m. By induction, $a^p - a$ is divisible by p and so $(a + 1)^p - (a + 1)$.] An other beautiful theorem is **Wilson's theorem** which allows to characterize primes: It tells that $(n - 1)! + 1$ is divisible by n if and only if n is a prime number. For example, for $n = 5$, we verify that $4! + 1 = 25$ is divisible by 5. [Proof: assume n is prime. There are then exactly two numbers $1, -1$ for which $x^2 - 1$ is divisible by n. The other numbers in $1, \dots, n - 1$ can be paired as (a, b) with $ab = 1$. Rearranging the product shows $(n - 1)! = -1$ modulo n. Conversely, if n is not prime, then $n = km$ with $k, m < n$ and $(n - 1)! = ...km$ is divisible by $n = km$.]

The solution to systems of linear equations like $x = 3 \pmod 5, x = 2 \pmod 7$ is given by the **Chinese remainder theorem**. To solve it, continue adding 5 to 3 until we reach a number which leaves rest 2 to 7: on the list 3, 8, 13, 18, 23, 28, 33, 38, the number 23 is the solution. Since 5 and 7 have no common divisor, the system of linear equations has a solution.
For a given n, how do we solve $x^2 - yn = 1$ for the unknowns y, x? A solution produces a square root x of 1 modulo n. For prime n, only $x = 1, x = -1$ are the solutions. For composite $n = pq$, more solutions $x = r \cdot s$ where $r^2 = -1 \bmod p$ and $s^2 = -1 \bmod q$ appear. Finding x is equivalent to factor n, because the greatest common divisor of $x^2 - 1$ and n is a factor of n. **Factoring is difficult** if the numbers are large. It assures that **encryption algorithms** work and keep bank accounts safe. Number theory, once the least applied discipline of mathematics has become one of the most useful one, in our daily life.