UNSOLVED PROBLEMS IN NUMBER THEORY

Myung-Hwan Kim

Department of Mathematics
Seoul National University, Seoul 151-742, Korea

Two Main Themes in Number Theory :

1. Study on Prime Numbers

2. Study on Diophantine Equations (A Diophantine equation is an equation with integer coefficients involving arbitrary number of variables such that its integral solutions are sought.)

Almost all branches of modern number theory stem from these two themes.

1. Prime Numbers

It is well known that every positive integer can be uniquely factored into a product of primes. So the prime numbers are the smallest units generating positive integers.

1-1. How many primes are there?

• Theorem. There are infinitely many primes.

(Euclid's proof) Suppose there are only finitely many primes, say $p_1, p_2, \cdots, p_n$, then one can deduce an easy contradiction from the number $p_1 p_2 \cdots p_n + 1$.

(Euler's Proof) $\sum_{p:\text{prime}} \frac{1}{p} = \infty$.

• Primes in Arithmetic Progression (Dirichlet, 1837) Let a, d be relatively prime positive integers. Then there are infinitely many primes in the arithmetic progression : $a, a+d, a+2d, a+3d, \cdots$.

Roughly $\frac{1}{\phi(d)}$ of primes are in the a.p.

For $d \geq 2$, let $p(a,d)$ be the smallest prime in the a.p. above and define $p(d) = \max\{p(a,d) \,|\, 1 \leq a \leq d, \gcd(a,d) = 1\}$. Linnik (1944) proved that $p(d) < d^L$ for some constant L, called Linnik's constant, independent of d. Kanold (1963) conjectured $L = 2$. Chen (1979) proved $L \leq 17$, which is the best result so far.

• Prime Number Theorem (Poussin, Hadamard, 1896 - independently) Let $\pi(x)$ be the number of primes $p \leq x$. Then, $\pi(x) \sim \dfrac{x}{\log x}$.

Note that PNT implies that $p_n \sim n \log n$, where p_n is the n-th prime counted from $p_1 = 2$.

1-2. Twin Prime Conjecture

• Twin Prime Conjecture : There are infinitely many twin primes. Two primes p, q are called twin primes if their difference is 2.

• The followings are elementary results deduced from Wilson's Theorem ($(p-1)! \equiv -1 \pmod{p}$ if p is a prime) :

(Clements, 1949) $n, n+2$ are twin primes $\Longleftrightarrow 4((n-1)!+1)+n \equiv 0 \pmod{n(n+2)}$.

(Sergusov, 1971) $n, n+2$ are twin primes $\Longleftrightarrow \phi(m)\sigma(m) = (m-3)(m+1)$, where $m = n(n+2)$.

• Let $\pi_2(x)$ be the number of primes p such that $p \leq x$ and $p+2$ is also a prime. Then it is known :

$$\pi_2(x) \leq C_1 C_2 \frac{x}{(\log x)^2} \left(1 + O\big(\frac{\log\log x}{\log x}\big)\right)$$

where $C_2 = \prod_{p>2}(1-(p-1)^{-2}) = 0.66016...$ is the twin-prime constant. Another constant C_1 is conjectured to be 2 by Hardy and Littlewood, but the best result so far is $C_1 = 7 + \epsilon$ obtained by Bombieri, Friedlander, and Iwaniec (1986).

• Brun in 1919 proved an interesting and important result as follows :

$$B = \sum_{p,p+2 \,:\, \text{twin primes}} \big(\frac{1}{p} + \frac{1}{p+2}\big) < \infty.$$

B is now called the Brun's constant. ($B = 1.90216054\cdots$.)

• Chen (1966, published in 1978) : There are infinitely many primes p such that $p+2 \in P_2$, where P_2 is the set of positive integers which are primes or products of two primes.

• Polignac's conjecture : For any given positive even integer $2k$, there are infinitely many primes p such that $p+2k$ is the next prime.

The twin prime conjecture is a special case of Polignac's conjecture, which is also open.

In fact, even the following simple question has not been answered yet : For any given positive integer $2k$, do there exist two primes p, q (not necessarily consecutive) whose difference is $2k$?

• Other interesting results on twin primes are :

(Brent, 1976) $\pi_2(10^{11}) = 224,376,048$.

(Dubner, 1985) $107570463 \times 10^{2250} \pm 1$ are twin primes.

1-3. Goldbach Conjecture

• Goldbach Conjecture (1742) : Every integer > 5 is a sum of three primes. $\Longleftrightarrow$ Every even integer ≥ 4 is a sum of two primes.

• Early contributions to Goldbach conjectures are :

(Schinzel, 1959) For any $k, m \geq 2$, there exist infinitely many primes p, q for which $2k \equiv p+q \pmod{m}$.

(Hardy-Littlewood, 1923 with RH, Vinogradov, 1937 without RH) Every odd integer $n \gg 0$ is a sum of three primes.

• Chen (1966, published in 1978) made a breakthrough : $2k = p+m$ for a prime p and $m \in P_2$ if $2k \gg 0$.

• Montgomery-Vaughan (1975) : Let $G(x)$ be the number of positive integers $2n \leq x$ such that $2n$ is not a sum of two primes. Then $G(x) < x^{1-\alpha}$ for some α $(0 < \alpha < 1)$ if $x \gg 0$. Note that this implies $\lim_{x\to\infty} \dfrac{G(x)}{x} = 0$.

Chen and Pan (1980) proved : $1 - \alpha = 1/100$ is possible.

Stein and Stein (1965) : Goldbach conjecture is true up to 10^8.

1-4. Riemann Hypothesis

• Consider the Riemann zeta function $\zeta(s) = \sum_n n^{-s},\ s \in \mathbf{C},\ Re(s) > 1$.

It is well known that this function has a meromorphic continuation to the whole complex plane $\mathbf{C}$ with a unique simple pole at $s = 1$. We still write this extended function by $\zeta(s)$.

This function has zeroes at $-2, -4, -6, \cdots$, which are called trivial zeros of the function. The famous Riemann Hypothesis is that all the nontrivial zeros of $\zeta(s)$ are on the line $\frac{1}{2} + it$, which is called the critical line. It is well known that there are no nontrivial zeros outside region $S = \{\sigma + it \in \mathbf{C} \mid 0 \leq \sigma \leq 1\}$, which is called the critical strip.

Zeroes of $\zeta(s)$ are closely related to the distribution of primes. But this is just one of thousands of applications that RH could provide if proved affirmative. If you study mathematics, you can see this RH pops up everywhere.

• Hardy (1914) : There are infinitely many zeroes of $\zeta(s)$ on the critical line.

Selberg (1942) improved this : $N(T) > C\frac{T}{2\pi}\log\left(\frac{T}{2\pi}\right)$ for some constant C $(0 < C < 1)$ if $T \gg 0$. Here N(T) is the number of zeros of $\zeta(s)$ on the critical line $\frac{1}{2} + it$ such that $0 < t < T$.

• So far, no nontrivial zeroes of $\zeta(s)$ has been found outside the critical line.

(Levinson, 1974) At least one-third of nontrivial zeros of $\zeta(s)$ are on the critical line. This has been improved to two-fifths by Conrey (1989).

(Van de Lune, Riele, Winter, 1986) Computed the first 1,500,000,001 nontrivial zeroes of $\zeta(s)$ and showed that they are all simple zeroes.

• Let $n = 1, 2, 3, \cdots$. It is well known that $\zeta(2n) = \frac{(-1)^{n+1}(2\pi)^{2n}B_{2n}}{2(2n)!} \notin \mathbf{Q}$, where B_{2n} is the $2n$-th Bernouille number.

Ramanujan (1914) gave an ingenious formula for $\zeta(2n+1)$: If $\alpha, \beta > 0, \alpha\beta = \pi^2$, then

$$\alpha^{-n}\Big(\frac{1}{2}\zeta(2n+1) + \sum_{k=1}^{\infty} \frac{k^{-2n-1}}{e^{2\alpha k}-1}\Big) = \beta^{-n}\Big(\frac{1}{2}\zeta(2n+1) + \sum_{k=1}^{\infty} \frac{k^{-2n-1}}{e^{2\beta k}-1}\Big) - 2^{2n}\sum_{k=0}^{n+1}(-1)^k \frac{B_{2k}B_{2n+2-2k}}{(2k)!(2n+2-2k)!}\alpha^{n+1-k}\beta^k.$$

He never proved it, rather Berndt gave a complete proof (1977).

Apéry (1981) proved that $\zeta(3)$ is irrational. For other positive odd integers this is not known.

1-5. Other Interesting Open Problems

• Mersenne number : $M_n = 2^n - 1,\ n = 1, 2, 3, \cdots$. If M_n is a prime, then it is called a Mersenne prime, and in this case n should be a prime too. Concerning Mersenne numbers, the followings are open questions :

(1) Are there infinitely many Mersenne primes?

(2) Are there infinitely many composite Mersenne numbers?

(3) Is every Mersenne number square-free?

(4) Are $C_1 = M_2 = 3, C_2 = M_{C_1} = M_3 = 7, C_3 = M_{C_2} = M_7 = 127, C_4 = M_{C_3} = M_{127}, \cdots, C_{n+1} = M_{C_N}, \cdots$ primes? (Catalan's conjecture)

Note that $M_{13} = 8191$ is a prime, but M_{8191} is a multiple of 338193759479.

As is well known, Mersenne primes are in one to one correspondence with even perfect numbers : $M_p = 2^p - 1 \longleftrightarrow 2^{p-1}(2^p - 1)$. A positive integer is called a perfect number if its positive proper divisors sum up to itself. $6, 28, 496, 8128, \cdots$ are even perfect numbers. No odd perfect number has been found yet. So here is another open question.

(5) Are there odd perfect numbers?

I'd rather stop introducing open questions on this subject - there are too many. Also there are bunches of results on these questions some are beautiful and some are ugly while some introduce another conjecture $\cdots$

There are 31 known Mersenne primes now and M_{216091} is the largest one. So, there are 31 known even perfect numbers and the largest one is $2^{216090}(2^{216091}-1)$.

• Fermat number :$F_n = 2^{2^n}+1,\ n = 0, 1, 2, \cdots$. If F_n is a prime, it is called a Fermat prime. One can easily see that if $2^m + 1$ is a prime, then m should be of the form 2^n. Fermat thought that F_n are primes for all n. $F_0 = 3, F_1 = 5, F_2 = 17, F_3 = 257, F_4 = 65537$ are primes and these five are all he could compute. (F_5 is a 10 digit number.)

However, $F_5 = 641 \times 6700417$ (Euler, 1732), $F_6 = 274177 \times 67280421310721$ (Landry, 1880), $\cdots$, and no other Fermat primes are found after the first five.

Questions (1), (2), (3) for Mersenne numbers can be asked for Fermat numbers and all of them are unanswered yet.

Gauss' famous result concerning ruler and compass construction is closely related to Fermat primes. His theorem reads : For $n \geq 3$, a regular n-gon is ruler and compass constructible $\iff$ $n = 2^k p_1 p_2 \cdots p_t$, where $k, t \geq 0$ and p_i are distinct Fermat primes.

• Here are some more open questions on special types of primes:

Are there infinitely many primes whose digits are all 1's? $Rn = \dfrac{10^n - 1}{9}$ are called repunits. If Rn is a prime, then n is a prime. Only known prime repunits are $R2, R19, R23, R317, R1031$, and the last one was found by Dubner (1986).

Are there infinitely many Wilson primes $W(p) = \dfrac{(p-1)! + 1}{p}$? Only known Wilson primes are $W(5), W(13), W(563)$, and the last one was found by Goldberg (1953).

Are there infinitely many primes of the form $m^2 + 1$? Or, more generally, of the form $am^2 + bm + c$?

Are there finitely many primes of the form $m^m + 1$? (Sierpinski) If so, then one can conclude that there are infinitely many composite Fermat numbers.

• Artin's conjecture concerning primitive roots modulo odd primes, Dickson's conjecture concerning linear polynomials (generalized by Sierpinski and Schinzel concerning irreducible polynomials), prime producing functions, $\cdots$

Concerning prime producing function, the following strange polynomial of degree 25 in 26 variables was found by Jones, Sato, Wada, and Wiens (1976) :

$$\begin{aligned}
P = (k+2)\{&1-[wz+h+j-q]^2-[(gk+2g+k+1)(h+j)+h-z]^2\\
&-[2n+p+q+z-e]^2-[16(k+1)^3(k+2)(n-1)^2+1-f^2]^2\\
&-[e^3(e+2)(a+1)^2+1-o^2]^2-[(a^2-1)y^2+1-x^2]^2\\
&-[16r^2y^4(a^2-1)+1-u^2]^2-[((a+u^2(u^2-a))^2-1)(n+4dy)^2+1-(x-cu)^2]^2\\
&-[n+l+v-y]^2-[(a^2-1)l^2+1-m^2]^2-[ai+k+1-l-i]^2\\
&-[p+l(a-n-1)+b(2an+2a-n^2-2n-2)-m]^2\\
&-[q+y(a-p-1)+s(2ap+2a-p^2-2p-2)-x]^2\\
&-[z+pl(a-p)+t(2ap-p^2-1)-pm]^2\}.
\end{aligned}$$

The positive range of this polynomial is equal to the set of all primes!!! ($R = 2 + \frac{1}{2}\{(P-2)+|P-2|\}$.)

• Effective Primality Tests and Factorizations. These are very closely related to Cryptology and Coding Theory.

2. Diophantine Equations

From the definition of Diophantine equations, you can easily make a conjecture that there will be infinitely many conjectures on these equations... I will introduce only a few of them including Fermat's Last Theorem.

2-1. Sums of Squares

• Is $x^2+y^2=n$ solvable? How many solutions are there? How about $x^2+y^2+z^2=n$ or $x^2+y^2+z^2+w^2=n$? These questions are answered by Euler (1754 ; Fermat claimed he had proved it in 1640), Gauss(1801 ; Legendre 1798?), and Lagrange (1772). Lagrange theorem reads : Every positive integers is a sum of four integer squares.

• Pell's Equations : $x^2-dy^2=1$, where d is a square free integer. Lagrange (1766) proved that there are infinitely many solutions. (Fermat also claimed a proof..., and Pell had little to do with Pell's equation...) Pell's equations are very closely related to the

study of quadratic number fields and the famous Gauss Conjecture on class numbers of these fields. It also has a close relation with continued fractions.

2-2. Waring's Problem

• The minimum number of squares to express all positive integers is four (Lagrange). How about the minimum number of cubes to do the same? $\cdots$ More generally, what is the minimum number of k-th powers necessary to express all positive integers? This is the Waring's Problem.

• Hilbert (1909) proved that such a minimum number exists for any $k \geq 1$, and denoted the number by $g(k)$. Waring (1782) himself conjectured that $g(3) \leq 9,\ g(4) \leq 19$. Note that $g(1) = 1,\ g(2) = 4$.

• Euler (1862, son of the Euler) proved $g(3) \geq 9,\ g(4) \geq 19,\ g(5) \geq 37,\ g(6) \geq 73, \cdots$

Wieferich (1909) : $g(3) \leq 9$ to settle $g(3) = 9$; Balasubramanian (1986) : $g(4) \leq 19$ to settle $g(4) = 19$; Chen (1964) : $g(5) \leq 37$ to settle $g(5) = 37$; Pillai (1940) : $g(6) \leq 73$ to settle $g(6) = 73$; $\cdots$

According to the works by Dickson (1936), Rubugunday (1942), Niven (1944), one can determine $g(k)$ for $k \geq 7$: possibly except for finitely many $k \geq 7,\ g(k) = [(3/2)^k] - 2 + 2^k$.

• Define $G(k)$ be the minimum number of k-th powers necessary to express all positive integers $\gg 0$. Then $G(k) \leq g(k)$.

$G(1) = 1,\ G(2) = 4,\ G(4) = 16$ are proved. But for other k's, $G(k)$ are unknown.

Most interesting question is of course about $G(3)$. It is known that $4 \leq G(3) \leq 7$. The lower bound is obtained by Maillet (1895) and the upper bound by Linnik (1943).

$6 \leq G(5) \leq 21,\ 9 \leq G(6) \leq 31, \cdots$

2-3. Fermat's Last Theorem

• (Fermat's Last Theorem, 1630?) $x^n + y^n = z^n$ has no solutions in nonzero integers if $n \geq 3$.

One can easily reduce FLT to : $x^p + y^p = z^p$ has no primitive solutions in nonzero integers if p is a prime ≥ 3. A solution $x = a, y = b, z = c$ is called primitive if $\gcd(a, b, c) = 1$. (Fermat himself gave a proof of this when $n = 4$.)

FLT is sometimes divided into two cases : (Case I) is about the solutions a, b, c with $p \nmid abc$ and (Case II) is about the solutions a, b, c with $p \mid abc$. Case II is much harder.

• Pierre de Fermat (1601-1655) was a lawyer and a councilor of Toulous, France. Although he was an amateur mathematician, he became one of the greatest mathematicians of all time.

Here are some euphuistic descriptions of Fermat : pioneer in analytic geometry and differential calculus with Decartes ; founder of probability with Pascal ; and above all the Father of modern number theory.

Fermat carried a Latin translation of 'Arithmetika' written (around A.D.250) by Diophantus of Alexandria. He solved numerous unsolved problems in the book or claimed so and also proposed new problems. He jotted down his proofs and problems on the margins of the book or on letters sent to famous mathematicians of his time. After his death, his son Samuel collected Fermat's works and published two books : 'Diophanti' (1670) and 'Varia Opera Mathematica' (1679).

By early eighteenth century, the problems that Fermat proved (rather, claimed so) without proofs or newly proposed were all resolved except F**L**T. In 1630(?) he wrote in a margin of his copy of Arithmetika :

"$\cdots$ to divide a cube into two cubes, a fourth power, or in general any power whatever into two powers of the same exponent above the second is impossible, and I have assuredly found an admirable proof of this, but the margin is too narrow to contain it."

• Here's a brief history of challenging FLT after Fermat:

1770 : Euler proved FLT for $p = 3$.

1816 : The French Academy announced a prize for solutions to FLT.

1820 : Sophie Germain proved Case I of FLT for p, where $2p+1$ is also a prime (such a prime p is called a Sophie Germain prime and it is not known yet whether there are infinitely many Sophie Germain primes).

1825 : Dirichlet, Legendre independently proved FLT for $p = 5$.

1839 : Lamé proved FLT for $p = 7$.

1847 : Lamé and Cauchy announced a false proof for FLT.

1850 : Kummer proved FLT for all regular primes. He approached FLT by observing $x^p = z^p - y^p = (z-y)(z-\zeta_p y)(z-\zeta_p^2 y)\cdots(z-\zeta_p^{p-1}y)$, where $\zeta_p = e^{2\pi i/p}$. So, the coefficients are not integers anymore, rather they are algebraic integers in $\mathbf{Q}(\zeta_p)$.

The ordinary integers has a unique factorization property (into primes). But algebraic integers of $\mathbf{Q}(\zeta_p)$ may not. As a sort of measurement how badly the unique factorization property is broken, one can define the class number $h_p \in \mathbf{N}$ of $\mathbf{Q}(\zeta_p)$.

Now a prime is called a regular prime if $p \nmid h_p$, and irregular prime otherwise. It is known that there are infinitely many irregular primes. But it is not known how many regular primes exist. Experiments with small numbers show there are more regular primes than irregular primes by ratio about $6:4$, but you never know what will happen if the numbers blow up.

Another criterion for the regularity of a prime is : p is regular $\iff$ p does not divide $B_2, B_4, \cdots, B_{p-3}$. Only irregular primes ≤ 100 are $37, 59, 67$.

Kummer devoted his whole life to FLT. Although he could give only a partial solution, that was enough to bring him an honorable medal from the French Academy. But the most important thing about Kummer's work is that it paved the way for algebraic number theory.

1908 : The Wolfskehl prize for a solution to FLT was announced.

1909 : Wieferich proved Case I of FLT for p, where p does not divide the Fermat quotient $(2^{p-1}-1)/p$. (It is not known how many such prime exist. A prime p which divide the Fermat quotient is called a Wieferich prime. It is not known either how many Wieferich primes exist.)

1920 : Vandiver proved FLT for certain irregular primes. In particular $p = 37, 59, 67$ cases were cleared to establish FLT for all $p \leq 100$.

1983 : Faltings proved Mordell Conjecture : Any algebraic curve of genus ≥ 2 defined over $\mathbf{Q}$ (or over a number field K) can have only finitely many primitive solutions.

Whatever that means, Fermat curve $x^p + y^p = z^p$ has genus $g = (p-1)(p-2)/2 \geq 6$ if $p \geq 5$. So Faltings' result implies that the number of primitive solutions is finite. By

this work, Faltings surprised the mathematical world and received a Fields medal in 1986.

1988 : Miyaoka announced and withdrew a false proof of FLT.

Recently, FLT was checked to be true for primes ($\neq 2$) less than $4,000,003$ by using computer, and thereby the probability that FLT be false is $< 10^{-24,000,000}$. (It is known that if FLT is true for all p, $3 \leq p \leq N$, then the probability that FLT be false is $< N^{-(N-3)}$.)

• From 1985, a totally new approach to FLT was attempted by Frey. Assuming p is a prime > 3 and (u, v, w) is a primitive solution of $x^p + y^p = z^p$ such that $u, v, w \neq 0$, Frey constructed an elliptic curve $y^2 = x(x - u^p)(x + v^p)$. (One may assume without loss of generality that $u \equiv 3 \pmod 4$ and v is even.) An elliptic curve is the solution set of an equation of the form $y^2 = ax^3 + bx^2 + cx + d$. Now Frey's idea was that his elliptic curve, which is now called a Frey curve, possesses too good properties to exist.

He looked at Shimura-Taniyama-Weil Conjecture (STW Conjecture), which has been one of the most famous problems in Elliptic Curve Theory since its birth in 1955 : Every elliptic curve is modular. More precisely, one can associate a modular form of weight 2 and level N to each elliptic curve, where N is a positive integer called the conductor of the elliptic curve.

He thought that one cannot associate such a modular form to his elliptic curve. (I will skip the detailed explanation on modular forms $\cdots$. Briefly, a modular form is a function on the upper half of the complex plane whose Fourier coefficients contain a lot of arithmetic information. And it satisfies a very special transformation property under linear fractional transformation $z \mapsto \dfrac{az+b}{cz+d}$ for any $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, namely $f\big(\dfrac{az+b}{cz+d}\big) = f(z)(cz+d)^k$, where k is a positive integer called the weight.

And there was Serre Conjecture : one can reduce the level N above to 2 if N is even. (In fact, the conjecture handles more general cases but it's been restricted to fit in our purpose. The conductor of Frey curve is $N = \prod_{q\,:\,\text{prime}\,|\,uvw} q$, which is even because v is even.)

But that a modular form of weight 2 and level 2 cannot exist is a well known fact.

So, combining these, Frey (1985) aided by Serre settled that STW Conjecture and Serre Conjecture together will imply FLT.

• Ribet (1986) proved Serre Conjecture. So now, STW Conjecture implies FLT. STW Conjecture was at first Taniyama Conjecture. But later Shimura and Weil added a partial answer to the original conjecture and reformulated it in a modern language.

• STW Conjecture was regarded as one of the most difficult and deep Conjecture in Elliptic Curve Theory. After Frey, Serre, and Ribet proved that it implies FLT, mathematicians thought that they cannot see its proof in their life time except one British mathematician in Princeton, Andrew Wiles.

• At the age of 10, he met FLT (his father introduced it to him) and was fascinated by FLT ever since. He said he wasted lots of time on FLT and exhausted before entering Oxford. After Ph.D. at Cambridge, he went to Harvard and settled at Princeton since early 1980's. He studied elliptic curves and STW conjecture. And in 1986, FLT was brought in front of him again by Frey, Serre, and Ribet. He immediately jumped into the problem and seven years later he come up with a proof!

He did not prove STW Conjecture in full. He proved a part of it. But that part was enough to settle FLT. From the start, he aimed exactly that. More precisely, he proved that there corresponds a modular form of weight 2 and level N to any semistable elliptic curve of conductor N. Now, one can easily check that Frey curves are semistable.