A Tale of Twin Primes

Mukund Madhav Mishra[1], Mir Samreen[2]

AFFILIATIONS:

[1]Mukud Madhav Mishra, Assistant Professor, department of Mathematics, Hansraj College, University of Delhi, Delhi, IND, email : **mukund@hrc.du.ac.in**)

[2]Mir Samreen, Undergrad student, department of Mathematics, Hansraj College, University of Delhi, Delhi, IND, email: **mirsamreen12@gmail.com**)

Abstract

The twin prime conjecture says that there are infinitely many pairs of primes (p, q) such that $q - p = 2$. This conjecture has been an open problem since it was first proposed and in year 2014, Y. Zhang made a significant progress towards the proof of the conjecture by proving existence of infinitely many prime pairs (p, q) that satisfy $q - p < C$ for certain constant C. We give a summarized account of the history of, and the contemporary progress towards the twin prime conjecture.

Keywords: Prime Numbers, Twin Prime Conjecture, Hardy-Littlewood theorem, Brun's conjecture, Sieve theory.

Introduction:

A positive integer is called a prime if it has precisely two factors, the number itself and 1. This definition of a prime number intentionally excludes 1 (1 has only factor) and makes 2 the smallest prime. 2 is the only even prime and all other primes are odd numbers. Prime numbers are interesting and worth studying because these numbers work as building blocks for integers in the sense that every positive integer splits in certain unique way as a product of prime numbers. [1]

Euclid of Alexandria first proved that there are infinitely many prime numbers, which, in other words means that no matter how large positive integer N we choose, there shall be a prime number p that is larger than N (see Niven et. al., 1991). But Euclid's proof is not constructive in the sense that it just gives arguments towards existence of arbitrary large primes, but does not help finding them. In fact, finding large primes is itself an interesting problem and needs a full-fledged theory in itself.

When we are looking at a prime p, in view of Euclid's theorem, one may look for a prime larger than p, or the prime next to p. If $q > p$ are both primes, we call them consecutive primes if every integer between p and q is composite. Distribution of prime numbers inside the set of positive integers is peculiar. We will see that what this 'peculiar' stands for, but asymptotically, the distribution of primes is known to certain extent. More precisely, the celebrated Prime Number Theorem says that for large positive integer N, the number of prime numbers less than or equal to N is nearly $\frac{N}{\log N}$ (for the proof, see Apostol, 1976). The peculiar part is the distance between primes. Though the number of primes up to a given natural number can be approximately told, finding the distances between consecutive primes is still a partially known open problem. In the light of the Prime Number Theorem, it is apparent that the average gap between primes goes larger as N grows. But the average gap between primes does not ensure that beyond certain point, smaller gaps between consecutive primes would not exist. But the following result, that has a simple proof, is a necessary condition for the Prime Number Theorem to be true (Neale, 2017).

Theorem 1.1 Given any positive integer, there exist consecutive primes p and q such that $q - p \geq N$.

Proof: Let p be the largest prime less than or equal to $N! + 1$. Now, the numbers $N! + 2, N! + 3, \ldots, N! + N$ are divisible by $2, 3, \ldots, N$ respectively. Therefore these $N - 1$ consecutive numbers are all composite. If q is the prime next to p, q can not be equal to any of these composite numbers and hence $q \geq p + N$.

Theorem 1.1 itself implies that there are infinitely many such pairs of primes (p, q) for each choice of N. But it has been observed that for very large values of N, consecutive primes exhibit smaller gaps as well. It has been a very difficult problem to determine whether small gaps between consecutive primes keep appearing or not.

Since no even number larger than 2 is prime, except for the pair (2,3), any two consecutive primes must have a composite number between them. In other words, if p_1 and p_2 are consecutive primes and $p_2 > p_1 > 2$ then $p_2 - p_1 \geq 2$.The pairs (3,5), (5,7), (11,13), (17, 19) and (27, 29) are some examples where the gap between consecutive primes is 2. Such prime pairs are called **twin primes**. *The twin prime conjecture* says that given any positive integer N, there exist consecutive primes $q > p > N$ such that $q - p = 2$ (Hardy and Wright, 2008).

1. HISTORICAL REMARKS

Prime numbers are often studied by classifying them into categories based on the remainders they leave after dividing by other numbers.

That all primes except 2 are odd is one such property. Similarly, it is easy to see that a prime number other than 2 and 3 must leave remainder 1 or 5 when divided by 6, i.e.if $p \geq 5$ is a prime then p must be of the form $6n \pm 1$ for some n. If, for some n, both $6n-1$ and $6n+1$ are primes, we have a pair of twin primes in hand. We prove an easy result that, to some extent suggests that there should be infinitely many twin primes.

Theorem 2.1There are infinitely many primes of types $6n-1$and $6n+1$.

Proof: We first prove that there are infinitely many primes of the type $6n-1$. Note that if $N_1, N_2, \dots, N_k$ are numbers of the form $6n+1$ or $6n+3$ then the product $N_1\ N_2 \dots N_k$ is also of the form $6n+1$ or $6n+3$. Now, on the contrary, we assume that there are finitely many primes $p_1, p_2, \dots, p_r$ of type $6n-1$. Consider $Q = 6\ p_1 p_2 \dots p_r - 1$. By our assumption, Qcannot be a prime. Also, none of 2, 3 or $p_i's$ can divide Q. Thus Qmust have a prime divisor, say$q$, that is not equal to 2 or 3, and is not of the form $6n-1$. So q must be a prime of the form $6n+1$. Q, therefore, is a product of primes of the form $6n+1$, and therefore Q itself should be of the form $6n+1$ or $6n+3$, which is a contradiction.

Thus there are infinitely many primes of type $6n+1$ can be concluded from the Dirichlet's theorem. The theorem is in itself very strong and guarantees existence of infinitely many primes in an arithmetic progression.

Theorem 2.2(Dirichlet). Let$a$ and d be relatively prime positive integers. Then the arithmetic progression $a+nd$ contains infinitely many primes.

For the proof of Dirichlet theorem, refer to (J. P. Serre, 1973).

Though there are evidences that twin primes were known to early Greeks since the time of Euclid, the first instance where twin prime conjecture is stated and discussed dates to 1849 when A. de Polignac published his article "Six Arithmetical Propositions Deduced from the Sieve of Eratosthenes (originally in French)". The conjecture of Polignac is rather a more general form of the twin prime conjecture.

Conjecture 2.1(de Polignac, 1849). Every even number is the difference of two consecutive prime numbers in infinitely many ways.

If the even number in Polignac's conjecture is taken to be 2, we have the twin prime conjecture.Over the years there have been several attempts to prove or disprove the assertion of twin prime conjecture however it still remains unsolved till date. Even though the problem isn't solved, lot of insights has been established in pursuing its solutions. All these insights presently provide us with several properties and information on twin primes, if not the complete answer. For more on history of the Twin Prime Conjecture, see (Nazardonyavi (2012)).

2. EARLY DEVELOPMENTS

Brun (reference), Hardy and Littlewood (reference),were among the first contributors who proposed and proved significant results about the distributional bounds for twin primes. Results and conjectures pertaining to these bounds had some implications on infinitude of twin primes.

3.1 Sieve Theory. The sieve of Eratosthenes dates back to as early as second century A. D., and yet is a powerful method of filtering out all the prime numbers up to a given positive integer. Method of sieving, which literally means filtering using a mesh, has been of wide utility in determining primes and their distribution.

Thanks to the evolution of set theory, modern mathematics has been able to give a concrete meaning to the term 'sieving'. Sieve methods are used in estimating the size of sifted (removed) sets of integers. For example, consider set $\mathcal{A}$ being the sequence of integers and $\mathcal{P}$as sequence of prime numbers and a number $t \geq 2$. If we sift out such elements from set $\mathcal{A}$ that are divisible by the primes in $\mathcal{P}$, then only if the unsifted elements of $\mathcal{A}$ are large, they can have prime divisors from $\mathcal{P}$. Also, each unsifted element of $\mathcal{A}$ may have just few such prime divisors provided t is not too small when compared to max $|a|$,where $a \in \mathcal{A}$. The objective of doing so is to estimate the number S($\mathcal{A}$; $\mathcal{P}, t$) of unsifted elements of $\mathcal{A}$ where,S($\mathcal{A}$; $\mathcal{P}, t$) is number of elements of $\mathcal{A}$ that are co-prime to all primes in $\mathcal{P}$ that are not greater than t.

The information we gather from sieves is theoretical in nature. If we look at the sieving procedure described above, possibly a way to calculate the number of non-eliminated elements of set $\mathcal{A}$ might be concocted. It means if we are able to find an upper bound to these numbers, it would imply there cannot be too many primes, we may still not knowwhetherprimes are finite or infinite in number, but we will certainly be ensured that the primes become rare and scarce as we move on. On the other hand, a lower bound to this number would imply that there exist infinitely many prime numbers.

3.2 Brun's Theorem: The first one to devise an effective sieve method was a Norwegian mathematician Viggo Brun(Brun, 1919).Brun's theorem was a very important step towards the progress on twin prime conjecture. His theorem claimed that the sum of the reciprocals of twin primes converges.One requires Brun's methods of sieving to arrive at this remarkable conclusion.

Theorem 3.1. Let $A = \{(p_i,\ q_i) : p_i\ , q_i \text{are primes}, q_i - p_i = 2\}$be the set of alltwin prime pairs. Then

$$\sum_{(p_i,q_i)\in A} \left(\frac{1}{p_i} + \frac{1}{q_i}\right) = \left(\frac{1}{3} + \frac{1}{5}\right) + \left(\frac{1}{5} + \frac{1}{7}\right) + \left(\frac{1}{11} + \frac{1}{13}\right) \dots < \infty$$

The proof ofTheorem 3.1 relies onTheorem 3.2 belowand the Brun's method of sieving which is beyond the scope of this exposition. We refer the reader to (LeVeque, 1996) for the complete proof.

The real number to which the above series converges is called the Brun's constant.Not much is known about this number except for the fact that the value is pretty close to 2 and 1.9021605824 is an approximate. If twin prime conjecture is proven, irrationality of the Brun's constant would be a corollary.

Theorem 3.2. There exists a positive constant C so that $\pi_2(x)$, the number of twin primes not exceeding x, satisfies.

$$\pi_2(x) < C.x.(\frac{\log\log x}{\log x})^2, \quad x > 3.$$

This means that the twin primes occur less frequently as compared to prime numbers by nearly a logarithmic factor. Soon after establishing this bound for twin primes, Brun announced a stronger bound where he showed that,

$$\pi_2(x) = O\left(\frac{x}{log^2 x}\right), where\ O\ is\ Big\ 'O'$$

Theorem 3.1 follows from establishment of the bound in Theorem 3.2. Convergence of Brun's series does not provide us with information on existence of twin primes finitely or infinitely, if the series would have diverged, it would have implied that there are infinitely many twin primes. The convergence implies that the twin primes are scarce, the question that are they infinite or not becomes harder now.

3.3 Hardy-Littlewood Conjecture.In 1923, G.H. Hardy and J. E. Littlewood proposed a conjecture concerning the number of primes in intervals. The statement that an integer near x has a probability of $\frac{1}{\log x}$ of being a prime number is equivalent to the celebrated Prime Number Theorem. If we wish to compute probability that $p\ and\ p+2$ are both prime where p is near x, we simply multiply the probability of each being individually prime to get the probability equal to $\frac{1}{log^2 x}$ then, summing up over all primes up to the x leads the way to Hardy-Littlewood Conjecture (Hardy and Littlewood, 1923).

Conjecture 3.2. For any $k > 0$ there are infinitely many prime pairs $p, p+2k$, and the number $\pi_{2k}(x)$ of such pairs less than x is

$$\pi_2(x) \sim 2\Pi_2 \int_2^x \frac{dt}{log\,(t)^2}$$

Here Π_2 refers to twin prime constant which is calculated as:

$$\Pi_2 = \prod_{p>3} \frac{p(p-2)}{(p-1)^2} \approx 0.66106181548686957 39 \ldots$$

Conjecture 2.1 focuses on consecutive primes which differ by an even number whereas Conjecture 3.2 concerns any pair of prime that differ by an even number. Formulation of Conjecture 3.1 is done using Brun's sieve techniques. What Hardy-Littlewood did was that they considered a set of primes with properties that correlated with the properties of the size of intervals between consecutive numbers. Then Brun's sieve helped in generalizing those properties which were based on primes themselves. Hardy-Littlewood Conjecture is at times substituted as the strong twin prime conjecture since it is understood that its proof will require or involve the integral going to infinity and thus making π_2 going to infinity and hence the affirmation to twin prime conjecture. As a matter of fact, computing $\pi_2(x)$ to the higher values of x has got various attempts following the work of Hardy and Littlewood. Significant calculations have been established by Thomas Nicely (1996) and Pascal Sebah (2002). Current record holder value is that of Sebah's up to 10^{16} which gives

$$\pi_2(x) = 10{,}304{,}195{,}697{,}298$$

Also, an estimate using Hardy Littlewood leads us to expect it around

$$\pi_2(x) \approx 10{,}304{,}192{,}554{,}496.$$

The precision of the strong twin prime conjecture can be easily visualized, which means that to not believe in the conjecture to be true becomes impossible.

3. THE BREAKTHROUGH: Y. ZHANG

The breakthrough towards the proof of the twin prime conjecture was a result by Y. Zhang published in 2014. Before we present Zhang's result, it is a must to look at other major contributions of the 21st century. One of them is popularly known as GPY after the names of three authors.

4.1 GPY Result. GPY result is another remarkable approach to prove the Twin Prime Conjecture worked out very recently by Daniel Goldstone, CemYildirin and JanozPintz. The original proof was given by Goldstone and Yildirin in 2003 but their proof was flawed which they corrected with Pintz in 2005 (Soundarajan (2006), Goldston et. al. (2009)). Before their work, it was only known that there are infinitely many gaps which are about quarter the size of an average gap. However, GPY theorem revealed some exciting results such as:

- There are infinitely many prime pairs for which the gap is as small as we want in comparison to the average gap between consecutive primes.
- It can lead towards famous twin prime conjecture that says that gap 2 occurs infinitely many times, the smallest possible gap between primes.
- It gives connection between the distribution of primes in arithmetic progression and small gaps between primes, for example based on certain difficult conjectures on such distribution they are able to prove existence of infinite prime pairs which differ by at most 16.

From the celebrated Prime Number Theorem we establish that as we look at primes around size x where x is very large, the average gap between consecutive primes is about $\log x$. It means the average gap between p and p_{next} is about $\log p$. We can now introduce GPY theorem.

Theorem 4.1.We have,

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

This theorem caught a lot of attraction, mainly the part concerning the behavior of $p_{n+1} - p_n$. Both the lower and the upper bounds were studied and found out in order to find largest and smallest possible gaps between consecutive primes. In order to know the lead towards Conjecture 2.1 we focus more on the small gaps. Erdos was the first one to show that,

$$\lim_{n\to\infty} \inf \frac{p_{n+1} - p_n}{\log p_n} < 1$$

This was further reduced to the limit being less than or equal to 0.25. For the twin prime conjecture to be true, there must exist infinitely many n for which $p_{n+1} - p_n = 2$ where limit infimum is clearly 0. The GPY method came surprisingly close to prove the Conjecture 2.1 but eventually met failure.

4.2 The Theorem of Yitang Zhang.Y. Zhang, a professor of mathematics at the University of California, USA works in the broad area of analytic number theory. He was the first person to come up with a finite bound on small gaps between consecutive prime numbers that are infinitely many. The inspiration for the proof came from the GPY theorem. It was evident from the theorem that a finite bound to the small gaps does exist, so, many researchers had tried to find it but met failure and GPY sieve was the tool which many mathematicians realized, had a potential to establish the result only if the level of distribution of the primes could be shown $\geq 1/2$. But the level of distribution from the GPY sieve is known to be at least $1/2$. Zhang worked alone in building some connection between GPY result and the conjecture of bounded prime gaps, but it has been said that even he failed for about 3 years in pursuit of an accurate proof . Onlywhenhe recognized that the GPY sieve wasn't going to work for him and he felt a requisition of a modified version of it, in which the sieves filter out numbers with no large factors rather than every number, which gave him convenience in allowing his sets of arguments to work, he was finally able to present his theorem.

Theorem 4.2 (Y. Zhang, (2014)). There exists an even integer $h \geq 2$ with the property that there exist infinitely many pairs of prime numbers of the form $(p, p + h)$. In fact, there exits such an h with $h \leq 7 \times 10^7$.

Equivalently, if $p_n, n \geq 1$, denotes the n^{th} prime number, we have

$$\lim_{n\to\infty} \inf(p_{n+1} - p_n) < \infty$$

or more precisely,

$$\lim_{n\to\infty} \inf(p_{n+1} - p_n) < 7 \times 10^7$$

In other words, there are infinitely many distinct prime pairs (p, q) such that

$$|p - q| < 7 \times 10^7.$$

This was a major step towards the celebrated Twin Prime Conjecture. The choice of 70 million made the proof simple but soon it was realized that this huge number can be drastically reduced as well. But as per Daniel Goldstone, to push is as far as to the Twin Prime Conjecture is quite unlikely. With strongest possible assumptions made on the level of distribution, the best possible result which can be drawn from GPY method would be that there are infinitely many prime pairs that differ by 16 or less, but to have it reduced to 2 might not be achieved.

Even though Zhang's proof couldn't prove the Twin Prime Conjecture, it is still a remarkable breakthrough in the history of number theory. Also, very huge step forward in someday solving the question of twin primes as well.

4. AFTERMATHS OF ZHANG'S CONTRIBUTION

The great advances witnessed in attempts to prove the Twin Prime Conjecture were shortly after the Zhang's proof came in April 2013. Australian mathematician Terence Tao initiated an internet-based project called Polymath8 project in June 2013. This project is an open online collaboration which aims to find more precise and better estimate for the bound and this attracted lot of people who gave their sets of contribution. They worked on the same small prime gap given by

$$\mathrm{H}_1 := \lim_{n\to\infty} \inf(p_{n+1} - p_n)$$

where p_n denotes the n^{th} prime. In more general setting,

$$H_m \coloneqq \lim_{n\to\infty} \inf(p_{n+m} - p_n)$$

In the paper of Goldston, Pintz and Yildirin, under a strong hypothesis of the Elliot-Halberstam Conjecture which is a conjecture about distribution of primes in arithmetic progression, the bound $H_1 \leq 16$, was obtained. But it remained difficult to achieve until Zhang proved the bound $H_1 \leq 70{,}000{,}000$. The Polymath project improved this to 4680 by July 2013. A spectacular progress was achieved by James Maynard in November 2013 (Maynard (2015)). He gave out his independent proof, inspired by Zhang's Proof and pushed down the gap to 600 which corresponds to 12 in Elliot-Halberstam Conjecture, furthermore bounds on H_m for higher m were first time calculated i.e.

$$H_m \ll m^3 e^{4m} \; for\ all\ m \geq 1.$$

To arrive at this, he introduced new multidimensional Selberg sieve. The Polymath then combined their techniques with Maynard's to arrive at a further lower bound. Recently, on 14 April, 2014 the bound has been reduced down to $H_1 \leq 246$ by Maynard and Tao, which corresponds to $H_1 \leq 6$ in Elliot-Halberstam Conjecture. Progresses made by Polymath8 project have been published in (Polymath, 20). Mathematicians are working now on pushing the gap down to 2. It will be astonishing if that happens and one of the biggest problems in the history of number theory will be thus solved.

References

1. Apostol, Tom M, Introduction to analytic number theory. Springer(1976).
2. Brun, Viggo, La série1/5+1/7+1/11+1/13+1/17+1/19+1/29+1/31+1/41+1/43+1/59+1/61+..., où les dénominateurs sont nombres premiers jumeaux est convergente ou finie. Bulletin des Sciences Mathématiques. 1919;43:100–104, 124–128.
3. Goldston, D. A. , Pintz, J. and Yıldırım, C. Y., Primes in tuples. I. Annals of Mathematics. 2009;170(2): 819–862.
4. Hardy, G. H., Littlewood, J. E., Some Problems of 'PartitioNumerorum.' III. On the Expression of a Number as a Sum of Primes. Acta Mathematica, 1923;Vol 44: 1–70.
5. Hardy, G.H., Wright, E. M., An introduction to the theory of numbers, 6th Ed. Oxford University Press. London (2008).
6. LeVeque, William Judson, Fundamentals of Number Theory. Dover Publishing.New York (1996).
7. Maynard, James, Small gaps between primes. Annals of Mathematics, 2015;181:1-31.
8. Nazardonyavi, S, Some history about Twin Prime Conjecture. arXiv:1205.0774, 2012.
9. Neale, Vicky, Closing the gap: the quest to understand prime numbers. The Oxford University Press (2017).
10. Niven, Ivan, Zuckerman, Herbert S. and Montgomery, Hugh L., An introduction to the theory of numbers. John Wiley & Sons Inc (1991).
11. De Polignac, A., Six propositions arithmologiquesdéduitesdu cribled'Ératosthène. Nouvelles annales de mathématiques 1re série, 1849;8:423-429.
12. Polymath, D.H.J., New equidistribution estimates of Zhang type. Algebra & Number Theory.2014;8(9);2067–2199.
13. Serre, Jean-Pierre, A course in arithmetic. Springer-Verlag. New York (1973).
14. Soundarajan, K., Small gaps between prime numbers: the work of Goldstone-Pintz-Yildirim. Bulletin of American Mathematical Society, 2006;44(1):1-18.
15. Zhang, Y., Bounded gaps between primes. Annals of Mathematics, 2014;179:1121-1174.