1 Introduction

1.1 background information

The twin prime conjecture is one of the most popular open problems, stating that there are infinitely many twin primes i.e there are infinitely many prime pairs with $|p-q|=2$. A lot of progress has been made on this in this century, and we shall review the main milestones.

First, a lot of the arguments for things close to the twin prime conjecture are probabilistic. The "probability" of a number being prime is important to know for this. This is found from the famous prime number theorem[1], which states that

$$\lim_{x\to\infty}\frac{\pi(x)}{x/\log(x)}=1$$

Where $\pi(x)$ is the prime counting function, i.e counts primes less than x. This tells us in some sense that there is a $\frac{1}{\log(x)}$ chance of x being prime. Ofcourse a number being prime is deterministic so this is not literal, but it does allow us to say for instance how many primes we expect between $[x,2x]$ and etc.

With this we can have a very rough estimate of having a twin prime pair. If we estimate the probabilities of being prime to be independent among different numbers, then the probability that both x and $x+2$ are primes is $\frac{1}{\log(x)\log(x+2)}$, and so we expect $\pi_2(x)\sim\frac{x}{\log(x)^2}$. We define π_2 to be the number of twin primes less than x. Now this is obviously a very handwavy estimate, but this does give us a reason to believe the twin prime conjecture.

In reality the probabilities are not independent and that has to be taken into account, a modifying constant makes this true.

1.2 GPY and the Salberg seive

One of the first major breakthroughs in this century was due to Goldston, Pintz and Yıldırım in 2003[5]. They proved that there are infinitely many prime pairs with $|p-q| < \epsilon \log(p)$. The idea is as follows, lets say we have k numbers $h_1 \dots h_k$ in increasing order, and we wanna see when all of $n+h_1 \dots n+h_k$ are primes. The Hardy-Littlewood conjecture says this happens infinitely often, twin primes being a specail case of it. Specifically we want to count how many times these are primes in the interval $[x,2x]$.So we try to find a function $a(n)$ so that

$$\sum_{\substack{x \leq n \leq 2x \\ n+h_j = \text{prime}}} a(n) > \frac{1}{k} \sum_{x \leq n \leq 2x} a(n)$$

And then summing over all j would give us

$$\sum_{x \leq n \leq 2x} \#\{j : n+h_j = \text{prime}\} a(n) > \sum_{x \leq n \leq 2x} a(n)$$

so that atleast one of the coefficients is more than 1, that is there is a prime pair in $[x,2x]$ with $|p-q| \leq h_k - h_1$. So the goal is to construct such a function, and the method is the Salberg sleive.

Essentially we try to write $a(n) = \left(\sum_{d|(n+h_1)\dots(n+h_k)} \lambda_d \right)^2$ with $\lambda_1 = 1$ and $\lambda_d = 0$ if $d > R$ for some R yet to be decided. Now GPY could get the ratio to be very close to the desired $\frac{1}{k}$ but never equal or bigger than it. They tried a specific class of λ_d and the contraints did not allow them go bigger than $R \leq x^{1/4}$, which came from counting residue classes. Now this was due to the Bombieri-Vinogradov theorem. However there are conjectures stronger than this, namely the Elliott-Halberstam conjecture, that would allow us to take $R \leq x^{1/2-\epsilon}$ and then we could get infinitely many primes of distance atmost 16. Recently in 2015 it has been found that assuming Elliot-Halberstam conjecture, we could take the distance down to 12[2].

We were really close to proving a finite bound on the gap of primes, and so it turns out this is enough to bound with $\epsilon \log(p)$. We wont go into the details of the computation, but this was the strongest bound for the time and would be for 10 more years.

1.3 Zhang's finite bound for prime gap

The next breakthrough came in 2013, when a relatively unknown Zhang published a paper with the proof that there are infinitely many prime pairs $|p-q| < 7\times 10^7$[3]. The way Zhang proved this is by almost the exact same method GPY did their old result, including the same $a(n)$. Zhang essentially managed to extend the Bombieri-Vinogradov so that $R = x^{1/4+\epsilon}$ could be allowed, and this was enough to just get over the $\frac{1}{k}$ ratio, but he had to use $k > 3500000$ and ultimately this led to the given bound.

This led to the famous polymath project initiated by Tao to improve this bound. This polymath project used a different $a(n)$ with the improvement of Zhang's and slowly wittled this bound down to a remarkable 246.

1.4 applications

Now the twin prime conjecture itself is not the most applicable theorem. However the attempt to solve this theorem has allowed the math community to be able to say more about the distribution in primes, and specifically about small gaps in prime numbers. Progress on conjectures like Hardy-Littlewood has been made to understand the distribution of gaps of primes better. Progress on conjectures like Elliot-Halberstam will be made to better understand distributions of primes in arithmetic sequences. These all helps draw a slightly better picture of how primes are distributed, and this can have,albeit probably very small, an effect on cryptography. If we can figure out where primes are more concentrated etc we can find better prime factoring algorithms and hence be able to do more with cryptography. The twin prime conjecture itself wont do this, but one of the theorems along the way might.

2 Elliot-Halberstam

Both Elliott-Halberstam and Bombieri-Vinogradov are bounds on error terms of $\left|\pi(x;q,a) - \frac{\pi(x)}{\phi(q)}\right| \leq E(x;q)$. Here $\pi(x;q,a)$ is counting the number of primes less than x thats a mod q. The idea is that the primes are roughly equidistributed to the $\phi(q)$ possible residue classes. Now Elliott-Halberstam states that for each $\theta < 1$ and $A > 0$ there is a constant C such that

$$\sum_{1\leq q\leq x^\theta} E(x;q) \leq \frac{Cx}{\log(x)^A}$$

Bombieri-Vinogradov proved this but only for $\theta < \frac{1}{2}$ and as we have seen it alone is not enough to make finite bounds for prime gaps.

We will try to find some bounds for the sum in hopes of finding some new values of θ for which this works.

Trying to find some bounds

First we try a weak bound.

Theorem. *There is a C such that*

$$\sum_{1 \leq q \leq x^\theta} E(x;q) \leq \frac{Cx^{1+\theta}}{\log(x)}$$

Proof. Notice that $E(x;q)$ can be atmost $\left(1 - \frac{1}{\phi(p)}\right)\pi(x) \leq \pi(x)$, corrosponding to each prime going to a single residue class. This means the sum is bounded above by

$$\sum_{1 \leq q \leq x^\theta} E(x;q) \leq x^\theta \pi(x) \leq \frac{Cx^{1+\theta}}{\log(x)}$$

$\square$

This is obviously a very weak bound, lets try making it better.

Theorem. *There is a C such that*

$$\sum_{1 \leq q \leq x^\theta} E(x;q) \leq \frac{Cx^{1+\theta}}{\log(x)} - C\theta x$$

Proof. Notice that $E(x;q)$ can be atmost $\left(1-\frac{1}{\phi(p)}\right)\pi(x)$, corrosponding to each prime going to a single residue class. This time we do a stronger bound, now clearly $\phi(n) \leq n$ so that $\left(1-\frac{1}{\phi(p)}\right)\pi(x) \leq \pi(x) - \frac{\pi(x)}{n}$. Now this means

$$\sum_{1\leq q\leq x^\theta} E(x;q) \leq (x^\theta - H_{x^\theta})\pi(x) \leq \frac{Cx^{1+\theta}}{\log(x)} - C\theta x$$

Here H_n denote the harmonic numbers and we used that $H_n \geq \log(n)$. □

We can use this line of reasoning to make one final stronger bound:

Theorem. *There is a C such that*

$$\sum_{1\leq q\leq x^\theta} E(x;q) \leq \frac{Cx^{1+\theta}}{\log(x)} - C\frac{\zeta(2)\zeta(3)}{\zeta(6)}\theta x$$

where ζ is the riemann zeta function.

Proof. Notice that $E(x;q)$ can be atmost $\left(1-\frac{1}{\phi(p)}\right)\pi(x)$, corrosponding to each prime going to a single residue class. This time we do a stronger bound, $\left(1-\frac{1}{\phi(p)}\right)\pi(x)$.Then we use[4] :

$$\sum_{q\leq x}\frac{1}{\phi(x)} \geq \frac{\zeta(2)\zeta(3)}{\zeta(6)}\log(x)$$

Now this means

$$\sum_{1\leq q\leq x^\theta} E(x;q) \leq (x^\theta - \sum_{q\leq x^\theta}\frac{1}{\phi(x)})\pi(x) \leq \frac{Cx^{1+\theta}}{\log(x)} - C\frac{\zeta(2)\zeta(3)}{\zeta(6)}\theta x$$

$\square$

3 conclusion and thoughts

We made some bounds however they were nowhere near as strong as even the already proven Bombieri-Vinogradov. We used almost a worst case analysis, when all the primes belonged in the same residue class. This does not really give us information on how evenly the primes are distributed in the residue classes hence. To understand this distribution better and make a dent in the theorem, undoubtedly more advanced techniques would have to be used than those of elementary number theory. This lack of results really just tell us how non-trivially the error term is going to behave, and that is to be expected.

The Elliot-Halberstam is probably the next step on the way to twin prime conjecture, it would improve from the current bound 20 fold. Zhang used a weaker form of Elliot-Halberstam to prove his result, so it stands to reason the next big change is due to this theorem. And then there is the general Elliot-Halberstam theorem, which would imply the bound is 6. This would very quickly start to zero-in on the required bound of 2. Most papers suggest that this would be the end of how much one could derive from these theorems, and that for the actual twin prime conjectures a different approach would have to be found. However proving these would also add a lot to our information about prime distributions even if not as strong as the twin primes. [4]

References

[1] Tom M. Apostol. *Introduction to analytic number theory.* Springer, 2011.

[2] James Maynard. Small gaps between primes. *Annals of Mathematics*, page 383–413, 2015.

[3] James Maynard. The twin prime conjecture. *Japanese Journal of Mathematics*, 14(2):175–206, 2019.

[4] R. Sitaramachandrarao. On an error term of landau-ii. *Rocky Mountain Journal of Mathematics*, 15(2), 1985.

[5] K. Soundararajan. Small gaps between prime numbers: The work of goldston-pintz-yildirim. *Bulletin of the American Mathematical Society*, 44(1):1–18, 2006.