Chapter 12

Bounded gaps between primes

12.1 Introduction

One often hears that the twin primes conjecture is an old conjecture, possibly dating back thousands of years to the time of the ancient Greeks. In reality, the first-recorded instance of the twin primes conjecture comes from a textbook written by Alphonse de Polignac, who was the son of the prime minister to Charles X of France.

Conjecture 12.1.1 (de Polignac, 1849)**.** *For all even integers h, there are infinitely many prime pairs $(p, p+h)$.*

In the particular case where $h = 2$, this is the famous twin primes conjecture. However, it actually says a lot more: it posits that every even integer can be a gap between pairs of primes infinitely often! Contrast this with the situation for odd gaps between primes: each odd gap can occur at most once, since 2 is the only even prime.

Recall that in Section 8.6.3, we gave a heuristic argument for the count of twin prime pairs. We saw that the heuristic strongly suggests that:

$$\#\{p \leqslant x : p, p+2 \text{ prime}\} \approx C_2 \frac{x}{(\log x)^2},$$

where

$$C_2 = 2 \prod_{\substack{q \text{ prime} \\ q \geqslant 3}} \frac{(1-2/q)}{1-1/q)^2} \approx 1.3203236...$$

Our heuristic relied on the assumption that the primes p are uniformly distributed $\bmod q$. If the primes *were* uniformly distributed, we would expect

$$\pi(x; q, a) \approx \frac{\pi(x)}{\varphi(q)},$$

since we would expect there to be a $1/\varphi(q)$ chance of landing in a given residue class $a \pmod q$ for all $q \leqslant x$ with $\gcd(a, q) = 1$.

In Chapter 11, we encountered the Bombieri-Vinogradov Theorem. We will use the version that is stated in terms of $\pi(x; q, a)$:

Theorem 12.1.2 (Bombieri-Vinogradov, 1965/6)**.** *For every constant $A > 0$, there exists a constant $B = B(A)$ such that*

$$\sum_{q \leqslant Q} \max_{\substack{a \bmod q \\ (a,q)=1}} \left| \pi(x; q, a) - \frac{\mathrm{Li}(x)}{\varphi(q)} \right| \ll_A \frac{x}{(\log x)^A},$$

where $Q = \frac{x^{1/2}}{(\log x)^B}$.

Observe that we can take Q almost up to $x^{1/2}$, but not quite. The fact that we do not know if the primes are well-distributed $\bmod q$ for values of q beyond $x^{1/2}$ is often referred to as the "square root barrier." There is a conjecture of Elliott and Halberstam which says that the same principle *should* hold for larger values of Q. In particular, it says:

Conjecture 12.1.3 (Elliott-Halberstam, 1968)**.** *The Bombieri-Vinogradov Theorem still holds with $Q = x^\theta$, for every $\theta < 1$.*

The parameter θ is called the *level of distribution* of the set of primes. Note that we cannot take Q as large as $Q = \frac{x}{(\log x)^B}$, since it was shown by Friedlander and Granville [3] in 1989 that the result is false for such Q.

12.2 Admissible k-tuples

As a generalization of de Polignac's conjecture, it is natural to wonder whether there are infinitely many k-tuples of primes:

$$(p + h_1, p + h_2, ..., p + h_k).$$

Some k-tuples clearly fail. For example:

Example 1. $p, p + 2, p + 4$ cannot be simultaneously prime infinitely often, since one of these numbers must be divisible by 3.

The problem that we encountered in the Example 1 is that it "covered" all of the residue classes modulo 3. Thus, one of the three elements in the tuple was always bound to be divisible by three, regardless of the choice of p. Such numbers clearly cannot be prime. In order to exclude the tuples with so-called "congruence class obstructions" that prevent them from being prime infinitely often, we define the concept of admissibility.

Definition 12.2.1. A k-tuple $(h_1, ..., h_k)$ of ordered nonnegative integers is *admissible* if it does not cover all residue classes $\bmod\, p$ for any prime p.

Example 2. The tuple $(0, 2, 6, 8, 12)$ is an admissible 5-tuple. Observe that the following remainders are not covered:

$$\begin{aligned} 1 &\pmod 2 \\ 1 &\pmod 3 \\ 4 &\pmod 5 \end{aligned}$$

(For moduli larger than 5, there will always be at least one residue class that is not covered, since there are more residue classes than elements in the tuple.)

Prior to the groundbreaking work of Zhang, Maynard, and Tao in 2013, there was only a conditional proof that there exists a finite number that occurs infinitely often as a gap between pairs of primes:

Theorem 12.2.2 (Goldston, Pintz, Yıldırım, 2009)**.** *If $(h_1, ..., h_k)$ is admissible and the Elliot-Halberstam conjecture holds with $Q = x^{1/2+\eta}$ (for some $\eta > 0$ depending on k) then there exist infinitely many integers n such that at least 2 of $n + h_1, ..., n + h_k$ are prime.*

12.3 The GPY Argument

In what follows, we will outline the argument given by Goldston, Pintz, and Yıldırım (GPY) in [5], highlighting the places where sieve methods are used. This chapter largely follows the proof outlined in Section 4 of [6].

Let $\mathcal{H} = (h_1, ..., h_k)$ be an admissible k-tuple. Let $x > h_k$. Our goal is to choose a nonnegative weight ν (where $\nu(n) \geqslant 0$ for all n) such that

$$\sum_{x<n\leqslant 2x} \nu(n) \left(\sum_{i=1}^{k} \theta(n+h_i) - \log 3x \right) > 0, \tag{12.3.1}$$

where

$$\theta(m) = \begin{cases} \log m & \text{if } m \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases}$$

By introducing the sieve weight $\nu(n)$, we are massively biasing $(n+h_1), ..., (n+h_k)$ towards having few prime divisors. If (12.3.1) holds then there is at least one positive term, n_0, in the sum. Then $\nu(n_0) > 0$ since $\nu(n) \geqslant 0$ for all n. Hence,

$$\sum_{i=1}^{k} \theta(n_0 + h_i) > \log 3x.$$

But each $n + h_i \leqslant 2x + h_k < 2x + x$, so each $\theta(n + h_i) < \log 3x$. Recall that $\theta(n + h_i) = \log(n + h_i)$ if $n + h_i$ is prime. By the last inequality, at least two of the $\theta(n + h_i)$'s are nonzero, hence at least two of the $n + h_1, ..., n + h_k$ are prime.

The main difficulty in this approach is to choose a weight ν that satisfies the conditions above for which (12.3.1) is easy to compute.

Key Idea: Take

$$\nu(n) := \left(\sum_{\substack{d|\mathcal{P}(n) \\ d\leqslant z}} \lambda_d \right)^2,$$

where $\lambda_d := \mu(d)G\left(\frac{\log d}{\log z}\right)$, $G\left(\frac{\log d}{\log z}\right)$ is a measurable bounded function supported only on $[0, 1]$, and $\mathcal{P}(n) := (n + h_1)(n + h_2) \cdots (n + h_k)$.

The square of the sum of λ_d's should look familiar. That's because we are using a version of Selberg's sieve. Observe that, because $\mu(d) = 0$ when d is divisible by a square, then λ_d is only supported on squarefree positive integers $\leqslant z$. Using our Key Idea in (12.3.1) and expanding, we obtain:

$$\sum_{\substack{d_1,d_2\leqslant z\\ d:=\mathrm{lcm}[d_1,d_2]}} \lambda_{d_1}\lambda_{d_2}\left(\sum_{i=1}^{k}\sum_{\substack{x<n\leqslant 2x\\ d\mid\mathcal{P}(n)}}\theta(n+h_i) - \log 3x \sum_{\substack{x<n\leqslant 2x\\ d\mid\mathcal{P}(n)}} 1\right). \tag{12.3.2}$$

Let $\Omega(d) := \{m \pmod{d} : d \mid P(m)\}$ and $\Omega_i(d) := \{m \in \Omega(d) : \gcd(d, m+h_i) = 1\}$. Then the expression in the parentheses in (12.3.2) can be re-written as:

$$\sum_{i=1}^{k}\sum_{m\in\Omega_i(d)}\sum_{\substack{x<n\leqslant 2x\\ n\equiv m \bmod d}}\theta(n+h_i) - \log 3x \sum_{m\in\Omega(d)}\sum_{\substack{x<n\leqslant 2x\\ n\equiv m \bmod d}} 1, \tag{12.3.3}$$

since $P(n) \equiv P(m) \pmod{d}$ if and only if $n \equiv m \pmod{d}$. The inner sum in the subtracted term is straightforward to compute:

$$\sum_{\substack{x<n\leqslant 2x\\ n\equiv m \bmod d}} 1 = \frac{x}{d} + O(1).$$

We will take $z \leqslant x^{1/4-o(1)}$ so that the error terms are negligible. In order to compute

$$\sum_{\substack{x<n\leqslant 2x\\ n\equiv m \pmod{d}}} \theta(n+h_i),$$

we apply the Bombieri-Vinogradov Theorem, which allows us to conclude that, on average,

$$\theta(x; d, m+h_i) \sim \frac{x}{\varphi(d)}$$

for $d < x^{1/2-o(1)}$. So, on average, we have

$$\theta(2x; d, m+h_i) - \theta(x; d, m+h_i) \sim \frac{x}{\varphi(d)}$$

for $d < x^{1/2-o(1)}$. Recall that we assumed that $z \leqslant x^{1/4-o(1)}$. Goldston, Pintz, and Yıldırım showed that if one can take z just beyond $x^{1/4}$ then it would already be enough to prove that there are bounded gaps between primes. In other words, if we could take the level of distribution of the primes up to $1/2+\eta$ for some $0 < \eta < 1/2$, we would have our desired result. Let's examine $\Omega(d)$ and $\Omega_i(d)$:

- We can construct $\Omega(d)$ and $\Omega_i(d)$ from $\Omega(p)$ and $\Omega_i(p)$ via the Chinese Remainder Theorem. Let $\omega(d) := |\Omega(d)|$. Then ω is multiplicative.
- Each $|\Omega_i(p)| = \omega(p) - 1 := \omega^*(p)$. Thus, we can extend ω^* to a multiplicative function on d: $|\Omega_i(d)| = \omega^*(d)$.

In this proof sketch, we will ignore the accumulated error terms and focus on understanding the size of the main terms. By doing so, we see that the equation (12.3.3) can be rewritten as

$$\begin{aligned} &= k\omega^*(d)\frac{x}{\varphi(d)} - (\log 3x)\omega(d)\frac{x}{d} \\ &= x\left(k\frac{\omega^*(d)}{\varphi(d)} - (\log 3x)\frac{\omega(d)}{d}\right). \end{aligned}$$

Notice that this difference is usually negative, which is why we cannot take the λ_d's to all be positive. Therefore, equation (12.3.2) is equivalent to:

$$x\Big(k \sum_{\substack{d_1,d_2\leqslant z \\ d:=\mathrm{lcm}[d_1,d_2]}} \lambda_{d_1}\lambda_{d_2}\frac{\omega^*(d)}{\varphi(d)} - (\log 3x) \sum_{\substack{d_1,d_2\leqslant z \\ d:=\mathrm{lcm}[d_1,d_2]}} \lambda_{d_1}\lambda_{d_2}\frac{\omega(d)}{d}\Big). \tag{12.3.4}$$

In order to evaluate these two sums, we can either apply Perron's Formula or Selberg's combinatorial approach. Details can be found on p. 18 - 23 in [6].

Recall that our goal was to show that the equation above is > 0. We can show this provided that there exists $0 < \eta < 1/2$ (depending on k) such that

$$1 + 2\eta > \Big(1 + \frac{1}{2\ell+1}\Big)\Big(1 + \frac{2\ell+1}{k}\Big) \tag{12.3.5}$$

when $k + \ell = m$. For $k = (2\ell+1)^2$, any ℓ that is large enough (depending on $\eta > 0$) will do.

In conclusion, if the primes have level of distribution $1/2 + \eta$ (i.e., if the Elliot-Halberstam Conjecture holds) and if $\ell \in \mathbb{Z}$ satisfies equation (12.3.5), then for every admissible k-tuple $h_1, ..., h_k$ with $k = (2\ell+1)^2$ there are infinitely many positive integers n such that $n + h_1$,...,$n + h_k$ contains at least two primes.

Recall that sieve methods do not distinguish between primes and almost primes (the so-called *parity problem*). Sieving lets us work in a sample space consisting of primes and almost-primes. But the primes have positive density within this sample space, so showing that equation (12.3.4) is positive still means that we have caught some tuples with primes in them.

12.4 Zhang's Approach

In May 2013, Yitang Zhang [19] stunned the mathematical community by posting a paper on the arXiv that claimed to prove that there are bounded gaps between primes infinitely often. One of the remarkable features of his proof was that he mainly uses the ideas that were already proposed by Goldston, Pintz, and Yildirim. However, he devised a clever way to get around the square root barrier.

Definition 12.4.1. Let $P^+(n)$ denote the largest prime factor of an integer n. An integer n is y-smooth (or y-friable) if $P^+(n) \leqslant y$.

In the GPY setup, Zhang takes

$$\lambda_d := \mu(d)\frac{1}{m!}\Big(\frac{\log(z/d)}{\log z}\Big)^m,$$

where $d \in \mathscr{D}$ is the subset of squarefree integers in $\{1, .., z\}$ that are y-smooth, and $\lambda_0 = 0$ otherwise. He takes $z < x^{1/3}$. With these extra conditions, he was able to obtain a result like the Bombieri-Vinogradov Theorem that goes just beyond $x^{1/2}$:

Theorem 12.4.2 (Zhang, 2013)**.** *There exist constants $\eta, \delta > 0$ such that for any given $a \in \mathbb{Z}$,*

$$\sum_{\substack{q\leqslant Q \\ (q,a)=1 \\ q \ y-\text{smooth} \\ q \ \text{squarefree}}} \left|\theta(x;q,a) - \frac{x}{\varphi(q)}\right| \ll_A \frac{x}{(\log x)^A},$$

where $Q = x^{1/2+\eta}$ and $y = x^{\delta}$.

If we take $\eta/2 = \delta = \frac{1}{1168}$ in Theorem 12.4.2 then we obtain:

Corollary 12.4.3 (Zhang, 2013)**.** *There exist infinitely many pairs of primes that differ by at most 70 million.*

After Zhang's paper appeared, Terence Tao organized a group of mathematicians to work on "optimizing" the arguments in order to obtain a bound smaller than 70 million. This group operated under the name Polymath8 (there were other Polymath groups formed with the goal of bringing together large groups of people from around the globe to chip away at difficult proofs). By crowdsourcing the technical details, Polymath8 was able to get the bound down to 4680.

12.5 Maynard and Tao's Approach

In the Goldston, Pintz, and Yıldırım approach, which was the basis for Zhang's proof, they studied divisors d such that $d \mid (n+h_1)\cdots(n+h_k)$, with $d \leqslant z$. In November 2013, Maynard [9] (and Tao, independently) instead studied k-tuples of divisors $d_1, ..., d_k$ such that

$$d_1 \mid (n+h_1), ..., d_k \mid (n+h_k),$$

with $d_1 d_2 \cdots d_k \leqslant z$. Rather than taking sieve weights λ_d as in GPY and Zhang's papers, Maynard uses multidimensional sieve weights. He takes

$$\lambda_{d_1,\ldots,d_k} \approx \Big(\prod_{i=1}^{k} \mu(d_i)\Big) f(d_1, ..., d_k)$$

for a suitable smooth function f. The idea is then to look at:

$$\sum_{a \in \Omega(m)} \sum_{\substack{x<n\leqslant 2x \\ n\equiv a \pmod m}} \Big(\sum_{j=0}^{k} \theta(n+h_j) - c\log 3x\Big) \times \Big(\sum_{\substack{d_i \mid n+h_i \\ 1\leqslant i\leqslant k}} \lambda_{d_1,\ldots,d_k}\Big)^2.$$

The goal is to show that this sum is positive. This is where the approaches of Maynard and Tao differ. Tao uses Fourier analysis, while Maynard creates a higher-dimensional version of Selberg's sieve (which involves choosing optimal sieve weights using the method of Lagrange multipliers). One important feature of Maynard's approach is that he does not need to go beyond the square root barrier. He just uses the Bombieri-Vinogradov Theorem with $z = x^{1/4-o(1)}$. Although their approaches to showing that the sum is positive differ, both Maynard and Tao arrived at the same conclusion:

Theorem 12.5.1 (Maynard-Tao, 2013)**.** *Let $m \geqslant 2$. Then for any admissible k-tuple $h_1, ..., h_k$ with k "large enough" relative to m, there are infinitely many n such that at least m of $n+h_1$,...,$n+h_k$ are prime.*

In Maynard's original paper (which was partly responsible for his 2022 Fields Medal), he was able to shrink the gap between primes down to 600, without relying on the optimizations performed by Polymath8. Afterwards, Polymath8 used a combination of methods from both Maynard and Zhang in order to prove the current state-of-the-art:

Theorem 12.5.2 (D. H. J. Polymath, 2014)**.** *There exist infinitely many pairs of primes that are at most* 246 *apart.*

By the Pigeonhole Principle, there is at least one even integer between 2 and 246 that occurs infinitely often as a gap between primes. Note that this gap can be reduced to 6 by assuming a very strong form of the Elliot-Halberstam conjecture. Currently, there are not even any conditional proofs that would allow one to conclude that prime gaps of size 2 or 4 occur infinitely often. Sadly, twin primes remain out of reach.

12.6 Bounded Gaps Recipe

In this section, we present a recipe for producing infinitely many primes with bounded gaps between them, following the Maynard-Tao approach. Our goal is to find values of $n \in [N, 2N]$ for which $n + h_1, ..., n + h_k$ contains several primes. (For somewhat technical reasons, we want the endpoints on our interval to have the same order of magnitude; this is why we are looking in the interval $[N, 2N]$ instead of $[1, N]$.)

Our first objective is to create a sample space of plausible values of n for which $n + h_1, ..., n + h_k$ contains several primes. Let $W := \prod_{p \leqslant \log_3 N} p$, where $\log_3 N := \log\log\log N$. Since $h_1, ..., h_k$ are admissible, then we can choose $\nu \in \mathbb{Z}$ such that $\gcd(\nu + h_i, W) = 1$ for all $1 \leqslant i \leqslant k$. Next, we pre-sieve the interval so that just those n satisfying $n \equiv \nu \pmod W$ remain. This step has come to be known as the "W-trick." After performing the W-trick, our sample space becomes:

$$\Omega := \{N < n \leqslant 2N : n \equiv \nu \pmod W\}.$$

Now, let $\omega(n)$ be nonnegative weights, and let $\chi_{\mathscr{P}}$ denote the characteristic function of the set of primes $\mathscr{P}$. Moreover, define the following two sums:

$$S_1 := \sum_{\substack{N < n \leqslant 2N \\ n \equiv \nu \pmod W}} \omega(n), \tag{12.6.1}$$

$$S_2 := \sum_{\substack{N < n \leqslant 2N \\ n \equiv \nu \pmod W}} \Big(\sum \chi_{\mathscr{P}}(n + h_i)\Big)\omega(n). \tag{12.6.2}$$

(The subscript on the inner sum depends a bit on the problem at hand. We leave it out for simplicity.)

Then $\frac{S_2}{S_1}$ is a weighted average of the number of primes among $n+h_1$,...,$n+h_k$.

Key Idea: If $\frac{S_2}{S_1} > (m-1)$ for some $m \in \mathbb{Z}^+$ then at least m of the $n+h_1$,...,$n+h_k$ are prime, for some $n \in \Omega$.

Therefore, we need to select weights $\omega(n)$ such that:

1. S_1 and S_2 can be estimated without too much difficulty using tools that we have at hand.

2. S_2/S_1 is large.

Unlike Zhang's approach, which was useful precisely for overcoming the square root barrier so that he could push through the GPY machinery, Maynard and Tao's approach has proven to be much more flexible. The results from Maynard's paper have been generalized in a number of ways. For example, they have been extended to the number field and function field settings by Castillo, Hall, Lemke Oliver, Pollack, and Thompson [2]. In a somewhat different direction, Thorner [18] showed that there are bounded gaps between primes in Chebotarev sets[1], from which a number of interesting results can be derived. We will say that a set of primes $q_1, q_2, ...$ has the *bounded gaps property* if $\liminf_{n\to\infty} q_{n+m} q_n < \infty$ for every m. Some examples of sets with the bounded gaps property include:

- The set of primes $p \equiv 1 \pmod 3$ for which 2 is a cube $\pmod p$.
- Fix $n \in \mathbb{Z}^+$. The set of primes expressible in the form $x^2 + ny^2$.
- Let τ be the Ramanujan tau function[2]. The set of primes p for which

$$\tau(p) \equiv 0 \pmod d$$

for any positive integer d.

[1]These are primes p such that Frob_p lies in a specified conjugacy-invariant subset of $\mathrm{Gal}(K/\mathbb{Q})$ for Galois extensions $K/\mathbb{Q}$. These sets of primes have positive relative density, according to the Chebotarev Density Theorem.

[2]Let $\mathbb{H} = \{z \in \mathbb{C} \;:\; \mathfrak{Re}(z) > 0\}$ be the complex upper half plane. Note that the function $\Delta\colon \mathbb{H} \to \mathbb{C}$ defined by $\Delta(z) := \mathrm{e}(z)\prod_{n=1}^{\infty}(1-\mathrm{e}(nz))^{24}$ is periodic with period 1, and so it has a Fourier series expansion: $\Delta(z) = \sum_{n=1}^{\infty}\tau(n)e(nz)$. The Fourier coefficients $\tau(n)$ define the *Ramanujan tau function*.

- The set of primes p for which $\#E(\mathbb{F}_p) \equiv p+1 \pmod{d}$ for any positive integer d, where $\#E(\mathbb{F}_p)$ is the number of $\mathbb{F}_p$-rational points on an elliptic curve over the finite field $\mathbb{F}_p$.

All of these sets are Chebotarev sets, so the fact that they possess the bounded gaps property follows from Thorner's paper.

The Maynard-Tao approach has also been used for a number of problems on "runs of consecutive primes" (sequences of consecutive primes that possess a given property). Some examples of these types of theorems that have been proven using the Maynard-Tao machinery include:

- For each $k \in \mathbb{Z}^+$, let $d_k = p_{k+1} - p_k$. In 1948, Erdős and Turán conjectured that the sequence $\{d_k\}$ contains arbitrarily long runs of consecutive values in the sequence that are increasing, as well as arbitrarily long runs that are decreasing. This was proven by Banks, Freiburg, and Turnage-Butterbaugh [1] in 2013.
- Let $S_{10}(p_n)$ be the base-10 sum of digits of the prime p_n. In the 1960's, Erdős and Sierpiński were interested in determining whether there are arbitrarily long runs of primes p_n for which $s_{10}(p_n)$ is increasing/decreasing/constant. In 2015, Pollack and Thompson [16] proved that the answer is "yes" and that it can be generalized to any base g.

In all of these problems, the proof follows the main ideas from the bounded gaps recipe. One of the key differences lies in the pre-sieving that occurs when the "W-Trick" is applied. Choosing good sieve weights is one of the more-technical aspects of Maynard's work. Fortunately, in most cases, it has been possible to push through these arguments using the same sieve weights that Maynard found!