Chapter 7

The Prime number theorem for arithmetic progressions

7.1 The Prime number theorem

We denote by $\pi(x)$ the number of primes $\leqslant x$. We prove the Prime Number Theorem.

Theorem 7.1. *We have* $\pi(x) \sim \dfrac{x}{\log x}$ *as* $x \to \infty$.

Before giving the detailed proof, we outline our strategy. Define the functions

$$\theta(x) := \sum_{p \leqslant x} \log p, \quad \psi(x) := \sum_{k,p\colon p^k \leqslant x} \log p = \sum_{n \leqslant x} \Lambda(n),$$

where Λ is the von Mangoldt function, given by $\Lambda(n) = \log p$ if $n = p^k$ for some prime p and some $k \geqslant 1$, and $\Lambda(n) = 0$ otherwise.

• By Lemma 3.14 from Chapter 3 we have

$$\sum_{n=1}^{\infty} \Lambda(n) n^{-s} = -\frac{\zeta'(s)}{\zeta(s)} \quad \text{for } \operatorname{Re} s > 1.$$

By applying Theorem 6.3 (the Tauberian theorem for Dirichlet series) to the latter we obtain

$$\frac{\psi(x)}{x} = \frac{1}{x} \sum_{n \leqslant x} \Lambda(n) \to 1 \quad \text{as } x \to \infty.$$

- We prove that $\psi(x) - \theta(x)$ is small. This gives $\theta(x)/x \to 1$ as $x \to \infty$.
- Using partial summation, we deduce $\pi(x) \log x/x \to 1$ as $x \to \infty$.

We first verify the conditions of the Tauberian theorem.

Lemma 7.2. *there is an open set U containing $\{s \in \mathbb{C} : \operatorname{Re} s \geqslant 1\}$ such that $\sum_{n=1}^{\infty} \Lambda(n) n^{-s}$ can be extended to a function analytic on $U \setminus \{1\}$, with a simple pole with residue 1 at $s = 1$.*

Proof. Recall that $\sum_{n=1}^{\infty} \Lambda(n) n^{-s} = -\zeta'(s)/\zeta(s)$ for $s \in \mathbb{C}$ with $\operatorname{Re} s > 1$ (see Lemma 3.15). By Theorem 5.2, $\zeta(s)$ is analytic on $\{s \in \mathbb{C} : \operatorname{Re} s > 0\} \setminus \{1\}$, with a simple pole at $s = 1$. Further, by Corollary 5.4 and Theorem 5.5, $\zeta(s)$ is non-zero on $A := \{s \in \mathbb{C} : \operatorname{Re} s \geqslant 1\}$, and hence also non-zero on an open set U containing A. So by Lemma 0.29, $\zeta'(s)/\zeta(s)$ is analytic on $U \setminus \{1\}$, with a simple pole with residue -1 at $s = 1$. This proves Lemma 7.2. □

Lemma 7.3. *(i) $\theta(x) = O(x)$ as $x \to \infty$.*
(ii) $\psi(x) = \theta(x) + O(\sqrt{x})$ as $x \to \infty$.
(iii) $\psi(x) = O(x)$ as $x \to \infty$.

Proof. (i) By homework exercise 3a, we have $\prod_{p \leqslant x} p \leqslant 4^x$ for $x \geqslant 2$. This implies

$$\theta(x) = \sum_{p \leqslant x} \log p \leqslant x \log 4 = O(x) \quad \text{as } x \to \infty.$$

(ii) We have

$$\begin{aligned} \psi(x) &= \sum_{p,k:\, p^k \leqslant x} \log p = \sum_{p \leqslant x} \log p + \sum_{p^2 \leqslant x} \log p + \sum_{p^3 \leqslant x} \log p + \cdots \\ &= \theta(x) + \theta(x^{1/2}) + \theta(x^{1/3}) + \cdots \end{aligned}$$

Notice that $\theta(t) = 0$ if $t < 2$. So $\theta(x^{1/k}) = 0$ if $x^{1/k} < 2$, that is, if $k > \log x/\log 2$. Hence

$$\begin{aligned} \psi(x) - \theta(x) &= \sum_{k=2}^{[\log x/\log 2]} \theta(x^{1/k}) \leqslant \theta(\sqrt{x}) + \sum_{k=3}^{[\log x/\log 2]} \theta(\sqrt[3]{x}) \\ &\leqslant \theta(\sqrt{x}) + \Big(\frac{\log x}{\log 2} - 3\Big)\theta(\sqrt[3]{x}) = O\big(\sqrt{x} + \sqrt[3]{x} \cdot \log x\big) \\ &= O(\sqrt{x}) \quad \text{as } x \to \infty. \end{aligned}$$

(iii) Combine (i) and (ii). □

Proof of Theorem 7.1. Lemmas 7.2 and 7.3 (iii) imply that $L_\Lambda(s) = \sum_{n=1}^\infty \Lambda(n)n^{-s}$ satisfies all conditions of Theorem 6.3, with $\sigma = 1, \alpha = 1$. Hence

$$\frac{\psi(x)}{x} = \frac{1}{x}\sum_{n\leqslant x}\Lambda(n) \to 1 \quad \text{as } x \to \infty.$$

From Lemma 7.3 (ii) we infer

$$\frac{\theta(x)}{x} = \frac{\psi(x) + O(\sqrt{x})}{x} = \frac{\psi(x)}{x} + O(x^{-1/2}) \to 1 \ \text{ as } x \to \infty.$$

We now apply partial summation to obtain our result for $\pi(x)$. Thus,

$$\begin{aligned}\pi(x) &= \sum_{p\leqslant x} 1 = \sum_{p\leqslant x}\log p \cdot \frac{1}{\log p} = \theta(x)\frac{1}{\log x} - \int_2^x \theta(t)\cdot\Big(\frac{1}{\log t}\Big)' dt \\ &= \frac{\theta(x)}{\log x} + \int_2^x \frac{\theta(t)}{t\log^2 t} dt.\end{aligned}$$

By Lemma 7.3 (i) there is a constant $C > 0$ such that $\theta(t) \leqslant Ct$ for all $t \geqslant 2$. Together with homework exercise 1, this implies

$$\int_2^x \frac{\theta(t)}{t\log^2 t}\cdot dt \leqslant C\cdot\int_2^x \frac{dt}{\log^2 t} = O\Big(\frac{x}{\log^2 x}\Big) \ \text{ as } x \to \infty.$$

Hence

$$\begin{aligned}\frac{\pi(x)\log x}{x} &= \frac{\theta(x)}{x} + O\Big(\frac{\log x}{x}\cdot\frac{x}{\log^2 x}\Big) \\ &= \frac{\theta(x)}{x} + O\Big(\frac{1}{\log x}\Big) \ \to 1 \ \text{ as } x \to \infty.\end{aligned}$$

□

7.2 The Prime number theorem for arithmetic progressions

Let q, a be integers with $q \geqslant 2$, $\gcd(a, q) = 1$. Define

$$\pi(x; q, a) := \text{number of primes } p \leqslant x \text{ with } p \equiv a \pmod{q}.$$

Theorem 7.4. *We have* $\pi(x;q,a) \sim \frac{1}{\varphi(q)} \cdot \frac{x}{\log x}$ *as* $x \to \infty$.

The proof is very similar to that of the Prime number theorem. Define the quantities

$$\begin{aligned} \theta(x;q,a) &:= \sum_{p \leqslant x,\, p \equiv a \,(\mathrm{mod}\, q)} \log p, \\ \psi(x;q,a) &:= \sum_{p,k,\, p^k \leqslant x,\, p^k \equiv a \,(\mathrm{mod}\, q)} \log p = \sum_{n \leqslant x,\, n \equiv a \,(\mathrm{mod}\, q)} \Lambda(n). \end{aligned}$$

Let $f(n) := \Lambda(n)$ if $n \equiv a \,(\mathrm{mod}\, q)$, $f(n) = 0$ otherwise. Then

$$L_f(s) = \sum_{n=1,\, n \equiv a \,(\mathrm{mod}\, q)}^{\infty} \Lambda(n) n^{-s}.$$

Let $G(q)$ be the group of characters modulo q.

Lemma 7.5. *For* $s \in \mathbb{C}$ *with* $\mathrm{Re}\, s > 1$ *we have*

$$L_f(s) = -\frac{1}{\varphi(q)} \cdot \sum_{\chi \in G(q)} \overline{\chi(a)} \cdot \frac{L'(s,\chi)}{L(s,\chi)}.$$

Proof. By homework exercise 6a we have for $\chi \in G(q)$ and for $s \in \mathbb{C}$ with $\mathrm{Re}\, s > 1$, since $\chi \in G(q)$ is a strongly multiplicative arithmetic function and $L(s,\chi)$ converges absolutely,

$$\frac{L'(s,\chi)}{L(s,\chi)} = -\sum_{n=1}^{\infty} \chi(n)\Lambda(n) n^{-s}.$$

Using Theorem 4.11 (ii) (one of the orthogonality relations for characters) we obtain for $s \in \mathbb{C}$ with $\mathrm{Re}\, s > 1$,

$$\begin{aligned} \sum_{\chi \in G(q)} \overline{\chi(a)} \cdot \frac{L'(s,\chi)}{L(s,\chi)} &= -\sum_{\chi \in G(q)} \overline{\chi(a)} \sum_{n=1}^{\infty} \chi(n)\Lambda(n) n^{-s} \\ &= -\sum_{n=1}^{\infty} \left(\sum_{\chi \in G(q)} \overline{\chi(a)} \chi(n) \right) \Lambda(n) n^{-s} = -\varphi(q) \cdot \sum_{n=1,\, n \equiv a \,(\mathrm{mod}\, q)}^{\infty} \Lambda(n) n^{-s}. \end{aligned}$$

□

Lemma 7.6. *There is an open set U containing $\{s \in \mathbb{C} : \operatorname{Re} s \geqslant 1\}$ such that $L_f(s)$ can be continued to a function analytic on an open set containing $U \setminus \{1\}$, with a simple pole with residue $\varphi(q)^{-1}$ at $s = 1$.*

Proof. By Theorem 5.3 (iii), $L(s, \chi_0^{(q)})$ is analytic on $\{s \in \mathbb{C} : \operatorname{Re} s > 0\} \setminus \{1\}$, with a simple pole at $s = 1$. Further, by Corollary 5.4 and Theorem 5.5, $L(s, \chi_0^{(q)})$ is non-zero on $A := \{s \in \mathbb{C} : \operatorname{Re} s \geqslant 1\}$, and hence also non-zero on an open set U containing A. Therefore, $L'(s, \chi_0^{(q)})/L(s, \chi_0^{(q)})$ is analytic on $U \setminus \{1\}$. Further, by Lemma 0.29, it has a simple pole with residue -1 at $s = 1$.

Let χ be a character mod q with $\chi \neq \chi_0^{(q)}$. By Theorem 5.3 (ii), $L(s, \chi)$ is analytic on an open set containing A, and by Corollary 5.4 and Theorems 5.5, 5.7, it is non-zero on A, hence on an open set containing A for which we may take U by shrinking U if necessary.. Therefore, $L'(s, \chi)/L(s, \chi)$ is analytic on U.

Now by Lemma 7.5, $L_f(s)$ is analytic on $U \setminus \{1\}$, with a simple pole with residue $\chi_0^{(q)}(a)/\varphi(q) = \varphi(q)^{-1}$ at $s = 1$. $\square$

Lemma 7.7. *(i) $\theta(x; q, a) = O(x)$ as $x \to \infty$.*
(ii) $\psi(x; q, a) - \theta(x; q, a) = O(\sqrt{x})$ as $x \to \infty$.
(iii) $\psi(x; q, a) = O(x)$ as $x \to \infty$.

Proof. (i) We have $\theta(x; q, a) \leqslant \theta(x) = O(x)$ as $x \to \infty$.

(ii) We have

$$\begin{aligned}\psi(x; q, a) - \theta(x; q, a) &= \sum_{k,p,\, k \geqslant 2, p^k \leqslant x,\, p^k \equiv a \,(\mathrm{mod}\, q)} \log p \\ &\leqslant \sum_{k,p,\, k \geqslant 2, p^k \leqslant x} \log p = \psi(x) - \theta(x) = O(\sqrt{x}) \quad \text{as } x \to \infty.\end{aligned}$$

(iii) Obvious. $\square$

Proof of Theorem 7.4. Lemmas 7.7 (iii) and 7.6 imply that $f(n)$, $L_f(s)$ satisfy the conditions of Theorem 6.3, with $\sigma = 1$, $\alpha = \varphi(q)^{-1}$. Hence

$$\frac{\psi(x; q, a)}{x} \to \frac{1}{\varphi(q)} \quad \text{as } x \to \infty,$$

and then by Lemma 7.7 (ii),

$$\frac{\theta(x;q,a)}{x} = \frac{\psi(x;q,a)\ + O(\sqrt{x})}{x} \to \frac{1}{\varphi(q)} \quad \text{as } x \to \infty.$$

By partial summation we have

$$\pi(x;q,a) = \sum_{2\leqslant p\leqslant x,\, p\equiv a\,(\mathrm{mod}\,q)} \log p \cdot \frac{1}{\log p} = \frac{\theta(x;q,a)}{\log x} + \int_2^x \frac{\theta(t;q,a)}{t\log^2 t} \cdot dt.$$

Now

$$0 \leqslant \int_2^x \frac{\theta(t;q,a)}{t\log^2 t} \cdot dt \leqslant \int_2^x \frac{\theta(t)}{t\log^2 t} \cdot dt = O\Big(\frac{x}{\log^2 x}\Big) \quad \text{as } x \to \infty,$$

using the estimate from the proof of Theorem 7.1. So

$$\frac{\pi(x;q,a)\log x}{x} = \frac{\theta(x;q,a)}{x} + O\Big(\frac{1}{\log x}\Big) \to \frac{1}{\varphi(q)} \quad \text{as } x \to \infty.$$

This completes our proof. □

7.3 Related results

Riemann sketched a proof, and von Mangoldt gave the complete proof, of the following result, that relates the distribution of primes to the distribution of the zeros of the Riemann zeta function. Define

$$\psi_0(x) := \begin{cases} \psi(x) = \sum_{n\leqslant x} \Lambda(n) & \text{if } x \text{ is not a prime power,} \\ \psi(x) - \frac{1}{2}\Lambda(x) & \text{if } x \text{ is a prime power.} \end{cases}$$

Theorem 7.8. *We have for $x > 1$,*

$$\psi_0(x) = x - \lim_{T\to\infty} \sum_{|\mathrm{Im}\,\rho|<T} \frac{x^\rho}{\rho} - \log 2\pi - \tfrac{1}{2}\log(1-x^{-2}),$$

where the sum is over all zeros ρ of $\zeta(s)$ with $0 < \mathrm{Re}\,\rho < 1$ and $|\mathrm{Im}\,\rho| < T$.

For a proof of this theorem, see Chapter 17 and some previous preparatory chapters in 'Multiplicative Number Theory' of H. Davenport.

Knowing that an arithmetic progression contains infinitely many primes, one would like to know when the first prime in such a progression occurs, i.e., the smallest x such that $\pi(x;q,a) > 0$. The following estimate is due to Linnik (1944).

Theorem 7.9. *Denote by $P(q,a)$ the smallest prime number p with $p \equiv a \pmod q$. There are absolute constants c, L such that for every integer $q \geqslant 2$ and every integer a with $\gcd(a,q) = 1$ we have $P(q,a) \leqslant cq^L$.*

The exponent L is known as 'Linnik's constant.' Since the appearance of Linnik's paper, various people have tried to estimate it. The present record is $L = 5.18$, due to Xylouris (2011).

Using information about the distribution of the zeros of $\zeta(s)$, such as knowledge of a zero-free region, one can obtain a good estimate for $|\psi(x) - x|$ and from that, using partial summation techniques similar to those discussed above, for $|\pi(x) - \mathrm{Li}(x)|$. This leads to the sharp versions of the Prime Number Theorem with error term, mentioned in Chapter 1.

There are similar refinenents of the Prime number theorem for arithmetic progressions with an estimate for the error $|\pi(x;q,a) - \mathrm{Li}(x)/\varphi(q)|$. The simplest case is when we fix q and let $x \to \infty$, but for applications it is important to have also versions where q is allowed to move in some range when we let $x \to \infty$.

The following result was proved by Walfisz in 1936, with important preliminary work by Landau and Siegel.

Theorem 7.10. *For every $A > 0$ there is a constant $C(A) > 0$ such that for every real $x \geqslant 3$, every integer $q \geqslant 2$ and every integer a with $\gcd(q,a) = 1$, we have*

$$\left| \pi(x;q,a) - \frac{1}{\varphi(q)} \mathrm{Li}(x) \right| \leqslant C(A) \frac{x}{(\log x)^A}.$$

The constant $C(A)$ is *ineffective*, this means that by going through the proof of the theorem one cannot compute the constant, but only show that it exists.

Note that if $\varphi(q) > (\log x)^A$ then Theorem 7.10 is trivial because then $\pi(x;q,a)$ and $\frac{1}{\varphi(q)}\mathrm{Li}(x)$ are of order of magnitude lower than $x/\varphi(q) < x/(\log x)^A$.

As we mentioned in Chapter 1, there is an intricate connection between the zero-free region of $\zeta(s)$ and estimates for $|\pi(x) - \mathrm{Li}(x)|$, where $\mathrm{Li}(x) = \int_2^x dt/\log t$. Similarly, there is a connection between zero-free regions of L-functions and estimates for $|\pi(x;q,a) - \mathrm{Li}(x)/\varphi(q)|$. We recall the following, rather complicated, result of Landau (1921) on the zero-free region of L-functions.

Theorem 7.11. *There is an absolute constant $c > 0$ such that for every integer $q \geqslant 2$ the following holds. Among all characters χ modulo q, there is at most one such that $L(s, \chi)$ has a zero in the region*

$$R(q) := \left\{ s \in \mathbb{C} : \operatorname{Re} s > 1 - \frac{c}{\log(q(1 + |\operatorname{Im} s|))} \right\}.$$

If such a character χ exists, it has not more than one zero in $R(q)$ and moreover, $\chi \neq \chi_0^{(q)}$, χ is a real character and the zero is real.

Any character χ modulo q having a zero in $R(q)$ is called an *exceptional* character mod q, and the zero of $L(s, \chi)$ in $R(q)$ is called an exceptional zero. The *Generalized Riemann Hypothesis* (GRH) asserts that if χ is a Dirichlet character modulo an integer $q \geqslant 2$, then the zeros of $L(s, \chi)$ in the critical strip $0 < \operatorname{Re} s < 1$ lie in fact on the line $\operatorname{Re} s = \frac{1}{2}$. A consequence of GRH is that exceptional characters do not exist.

In order to obtain Theorem 7.10, one needs an estimate for the real part of a possible exceptional zero of an L-function. The following result was proved by Siegel (1935).

Theorem 7.12. *For every $\varepsilon > 0$ there is a number $c(\varepsilon) > 0$ such that for every integer $q \geqslant 2$ the following holds: if χ is an exceptional character modulo q and β an exceptional zero of $L(s, \chi)$, then* $\operatorname{Re} \beta < 1 - c(\varepsilon) q^{-\varepsilon}$.

Theorems 7.11 and 7.12 imply (after a lot of work) Theorem 7.10. Proofs of Theorems 7.10–7.12 may be found in H. Davenport, Multiplicative Number Theory, Graduate texts in mathematics 74, Springer Verlag, 2nd ed., 1980.

Under assumption of GRH, one can show, similar to von Koch's result mentioned in Chapter 1, that

$$\left| \pi(x; q, a) - \frac{1}{\varphi(q)} \operatorname{Li}(x) \right| \leqslant C x^{1/2} (\log x)^2, \tag{7.1}$$

where C is an absolute constant, i.e., not depending on anything (see Davenport's book, Chapter 20). This result is trivial if $q > x^{1/2}$ since then both $\pi(x; q, a)$ and $\frac{1}{\varphi(q)} \operatorname{Li}(x)$ are of order of magnitude lower than $x^{1/2} (\log x)^2$.

There has been considerable interest in sums

$$\sum_{q \leqslant Q} \max_{1 \leqslant a < q,\, \gcd(a,q)=1} \left| \pi(x; q, a) - \frac{1}{\varphi(q)} \operatorname{Li}(x) \right|,$$

and the aim is to obtain good estimates for such sums, for as large as possible Q. From (7.1), that is, assuming GRH, one can trivially deduce that this sum is at most

$$CQx^{1/2}(\log x)^{1/2}. \tag{7.2}$$

Surprisingly, such a result can also be proved without assuming GRH, provided that Q is not too large.

Theorem 7.13 (Bombieri, Vinogradov, 1965). *For every reals θ, A with $0 < \theta < \frac{1}{2}$, $A > 0$ and for every $x \geqslant 3$ we have*

$$\sum_{q \leqslant x^\theta} \max_{1 \leqslant a < q,\, \gcd(a,q)=1} \left| \pi(x; q, a) - \frac{1}{\varphi(q)} \mathrm{Li}(x) \right| \leqslant C(A,\theta) x / (\log x)^A,$$

where $C(A,\theta)$ depends only on θ and A.

Notice that for $Q = x^\theta$ with $\theta < \frac{1}{2}$ this gives an estimate similar to (7.2). A proof of a related result can be found in the book of Iwaniec and Kowalski, Chapter 17. The result of Walfisz mentioned above is an important ingredient in its proof. The Elliott-Halberstam conjecture asserts that Theorem 7.13 should hold for all θ with $0 < \theta < 1$.

Results such as the Bombieri-Vinogradov theorem are used in studies on gaps between consecutive primes. Let $p_1 < p_2 < p_3 < \cdots$ be the sequence of consecutive primes. The *twin prime conjecture* asserts that there are infinitely many n with $p_{n+1} - p_n = 2$. In 2014, Yitai Zhang made a breakthrough by showing that there are infinitely many n with $p_{n+1} - p_n \leqslant 7 \times 10^7$, that is, there are infinitely many pairs of consecutive primes that are not further than 7×10^7 apart. Recently, Maynard (2015) improved this to 600. The Bombieri-Vinogradov Theorem is an important ingredient in his proof. Several mathematicians have been working collectively to further improve Maynard's bound. Today's record is 246, see the polymath wiki,

`http://michaelnielsen.org/polymath1` $\to$ polymath8