11. The Chebyshev Functions Theta and Psi

11.1. Definition (Prime number function). For real $x > 0$ we denote by $\pi(x)$ the number of all primes $p \leq x$. This can be also written as

$$\pi(x) = \sum_{p \leqslant x} 1.$$

$\pi(x)$ is a step function with jumps of height 1 at all primes. Of course $\pi(x) = 0$ for all $x < 2$. Some other values are

x	10	100	1000	10^4	10^5	10^6	10^7
$\pi(x)$	4	25	168	1229	9592	78498	664579

The prime number theorem, which we will prove in chapter 13, describes the asymptotic behavior of $\pi(x)$ for $x \to \infty$, namely

$$\pi(x) \sim \frac{x}{\log x},$$

meaning that the quotient $\pi(x)/\frac{x}{\log x}$ converges to 1 for $x \to \infty$. For the proof of the prime number theorem, some other functions, introduced by Chebyshev, are useful.

11.2. Definition (Chebyshev theta function). This function is defined for real $x > 0$ by

$$\vartheta(x) = \sum_{p \leqslant x} \log p.$$

(Of course this has nothing to do with the theta series and theta functions considered in the previous chapter.)
We will see that the prime number theorem is equivalent to the fact that the asymptotic behavior of the Chebyshev theta function is $\vartheta(x) \sim x$ for $x \to \infty$.

A first rough estimate is given by the following proposition.

11.3. Proposition. *For all $x > 0$ one has*

$$\vartheta(x) < x \log 4,$$

in particular $\vartheta(x) = O(x)$ for $x \to \infty$.

Proof. Of course it suffices to prove the assertion for $x = n \in \mathbb{N}_1$. The assertion is equivalent to

$$F(n) := \prod_{p \leqslant n} p \; < \; 4^n.$$

We will prove this by induction on n. It is obviously true for $n \le 3$.
For the induction step let $N \ge 4$ and assume that the assertion is true for all integers $n < N$.
First case: N even. Obviously $F(N) = F(N-1)$. For $F(N-1)$ we can use the induction hypothesis and obtain $F(N) = F(N-1) < 4^{N-1} < 4^N$.
Second case: N odd. We write N as $N = 2n+1$. Consider the binomial coefficient

$$\binom{2n+1}{n} = \frac{(2n+1)\cdot 2n \cdot (2n-1)\cdot \ldots \cdot (n+2)}{1\cdot 2\cdot 3\cdot \ldots \cdot n}.$$

Clearly, for every prime p with $n+2 \le p \le 2n+1$ one has

$$p \;\Big|\; \binom{2n+1}{n},$$

hence

$$\prod_{n+1<p\leqslant 2(n+1)} p \;\le\; \binom{2n+1}{n}.$$

Now $\binom{2n+1}{n} = \binom{2n+1}{n+1}$ are the two central terms in the binomial expansion of $(1+1)^{2n+1}$, therefore

$$\binom{2n+1}{n} < \tfrac{1}{2}(1+1)^{2n+1} = 4^n.$$

By induction hypothesis $\prod\limits_{p\leqslant n+1} p < 4^{n+1}$, hence

$$F(2n+1) = \prod_{p\leqslant 2n+1} p \;<\; 4^{n+1}\binom{2n+1}{n} < 4^{n+1}4^n = 4^{2n+1}, \qquad \text{q.e.d.}$$

11.4. Lemma (Abel summation II). *Let n_0 be an integer, $(a_n)_{n\geqslant n_0}$ a sequence of complex numbers and $A : [n_0, \infty[\, \to \mathbb{C}$ the function defined by*

$$A(x) := \sum_{n_0\leqslant n\leqslant x} a_n.$$

Further let $f : [n_0, \infty[\, \to \mathbb{C}$ be a continuously differentiable function. Then for all real $x \ge n_0$ the following formula holds

$$\sum_{n_0\leqslant k\leqslant x} a_k f(k) = A(x)f(x) - \int_{n_0}^{x} A(t)f'(t)dt.$$

Proof. We consider first the case when $x = n$ is an integer and prove the formula by induction on n. For $n = n_0$ both sides are equal to $a_{n_0} f(n_0)$.

Induction step $n \to n+1$. Denoting by $L(x)$ the left hand side and by $R(x)$ the right hand side of the asserted formula we have

$$L(n+1) - L(n) = a_{n+1} f(n+1)$$

and

$$\begin{aligned} R(n+1) - R(n) &= A(n+1)f(n+1) - A(n)f(n) - \int_n^{n+1} A(n)f'(t)dt \\ &= A(n+1)f(n+1) - A(n)f(n) - A(n)(f(n+1) - f(n)) \\ &= A(n+1)f(n+1) - A(n)f(n+1) \\ &= a_{n+1}f(n+1) = L(n+1) - L(n). \end{aligned}$$

This proves the induction step.

In the general case when x is not necessarily an integer, set $n := \lfloor x \rfloor$. Then

$$L(x) - L(n) = 0$$

and

$$\begin{aligned} R(x) - R(n) &= A(x)f(x) - A(n)f(n) - \int_n^x A(n)f'(t)dt \\ &= A(n)f(x) - A(n)f(n) - A(n)(f(x) - f(n)) = 0, \quad \text{q.e.d.} \end{aligned}$$

11.5. Theorem. *The following relations hold between the prime number function and the Chebyshev theta function:*

a) $$\pi(x) = \frac{\vartheta(x)}{\log x} + \int_2^x \frac{\vartheta(t)}{t \log^2 t}\, dt = \frac{\vartheta(x)}{\log x} + O\Big(\frac{x}{\log^2 x}\Big),$$

b) $$\vartheta(x) = \pi(x)\log x - \int_2^x \frac{\pi(t)}{t}\, dt = \pi(x)\log x + O\Big(\frac{x}{\log x}\Big).$$

Proof. a) Let $(a_n)_{n\geqslant 2}$ be the sequence defined by

$$a_n := \begin{cases} 1, & \text{if } n \text{ is prime,} \\ 0 & \text{otherwise,} \end{cases}$$

$b_n := a_n \log n$, and $f(x) = 1/\log x$. Then

$$\pi(x) = \sum_{2\leqslant n\leqslant x} a_n = \sum_{2\leqslant n\leqslant x} b_n f(n).$$

Since $\sum_{n\leqslant x} b_n = \vartheta(x)$ and $f'(x) = -1/(x\log^2 x)$, Abel summation (lemma 11.4) yields

$$\pi(x) = \vartheta(x)f(x) - \int_2^x \vartheta(t)f'(t)dt = \frac{\vartheta(x)}{\log x} + \int_2^x \frac{\vartheta(t)}{t\log^2 t}\, dt.$$

To estimate the integral, we use the result of theorem 11.3 that $|\vartheta(t)/t| \le \log 4$. Hence it remains to show that

$$\int_2^x \frac{dt}{\log^2 t} = O\Big(\frac{x}{\log^2 x}\Big).$$

This can be seen as follows (we may assume $x > 4$):

$$\begin{aligned}\int_2^x \frac{dt}{\log^2 t} &= \int_2^{\sqrt{x}} \frac{dt}{\log^2 t} + \int_{\sqrt{x}}^x \frac{dt}{\log^2 t} \\ &\le \frac{\sqrt{x}}{(\log 2)^2} + \frac{x}{(\log \sqrt{x})^2} = O(\sqrt{x}) + \frac{4x}{\log^2 x} = O\Big(\frac{x}{\log^2 x}\Big).\end{aligned}$$

b) With a_n as defined in a) we have $\vartheta(x) = \sum\limits_{2\leqslant n\leqslant x} a_n \log(n)$. Abel summation yields

$$\vartheta(x) = \pi(x)\log(x) - \int_2^x \frac{\pi(t)}{t}\, dt.$$

From $\vartheta(x) = O(x)$ and a) it follows that $\pi(x) = O(x/\log x)$, hence

$$\int_2^x \frac{\pi(t)}{t}\, dt = O\Big(\int_2^x \frac{dt}{\log t}\Big).$$

The last integral is estimated by the same trick as used in a)

$$\begin{aligned}\int_2^x \frac{dt}{\log t} &= \int_2^{\sqrt{x}} \frac{dt}{\log t} + \int_{\sqrt{x}}^x \frac{dt}{\log t} \\ &\le \frac{\sqrt{x}}{\log 2} + \frac{x}{\log \sqrt{x}} = O\Big(\frac{x}{\log x}\Big).\end{aligned}$$

11.6. Corollary. *The asymptotic relation*

$$\pi(x) \sim \frac{x}{\log x} \quad \text{for } x \to \infty \quad \text{(prime number theorem)}$$

is equivalent to the asymptotic relation

$$\vartheta(x) \sim x \quad \text{for } x \to \infty.$$

11.7. Definition (Mangoldt function). The arithmetical function $\Lambda : \mathbb{N}_1 \to \mathbb{Z}$ is defined by

$$\Lambda(n) := \begin{cases} \log p, & \text{if } n = p^k \text{ is a prime power } (k \ge 1), \\ 0 & \text{otherwise.}\end{cases}$$

11.8. Theorem. *The Dirichlet series associated to the Mangoldt arithmetical function satisfies for all $s \in \mathbb{C}$ with* $\mathrm{Re}(s) > 1$

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

Proof. By theorem 4.7 one has for $\mathrm{Re}(s) > 1$

$$\log \zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k} \sum_{p \in \mathbb{P}} \frac{1}{p^{ks}}.$$

This can be written as

$$\log \zeta(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

with

$$a_n := \begin{cases} 1/k, & \text{if } n = p^k \text{ is a prime power } (k \geq 1) \\ 0 & \text{otherwise} \end{cases}$$

Since

$$\frac{d}{ds}\frac{1}{n^s} = \frac{d}{ds} e^{-s \log n} = -\log n \; e^{-s \log n} = -\frac{\log n}{n^s}$$

and $a_n \log n = \Lambda(n)$, we get

$$\frac{\zeta'(s)}{\zeta(s)} = \frac{d}{ds} \log \zeta(s) = -\sum_{n=1}^{\infty} \frac{a_n \log n}{n^s} = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}, \quad \text{q.e.d.}$$

11.9. Definition (Chebyshev psi function). This function is defined by

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n).$$

11.10. Theorem. *The Chebyshev psi function and the Chebyshev theta function are related in the following way.*

a) $\psi(x) = \sum_{k \geqslant 1} \vartheta(x^{1/k}) = \vartheta(x) + \vartheta(x^{1/2}) + \vartheta(x^{1/3}) + \ldots = \vartheta(x) + O(x^{1/2} \log x),$

b) $\vartheta(x) = \sum_{k \geqslant 1} \mu(k) \psi(x^{1/k}) = \psi(x) - \psi(x^{1/2}) - \psi(x^{1/3}) - \psi(x^{1/5}) + \psi(x^{1/6}) - + \ldots .$

Proof. a) By the definition of the Mangoldt function one has

$$\psi(x) = \sum_{k \geqslant 1} \sum_{p^k \leqslant n} \log p = \sum_{k \geqslant 1} \sum_{p \leqslant x^{1/k}} \log p = \sum_{k \geqslant 1} \vartheta(x^{1/k}).$$

Since $\vartheta(t) = 0$ for $t < 2$, we have $\vartheta(x^{1/k}) = 0$ for $k > \log x / \log 2$, hence

$$\sum_{k \geqslant 2} \vartheta(x^{1/k}) \leq \Big\lfloor \frac{\log x}{\log 2} \Big\rfloor \vartheta(x^{1/2}) = O(x^{1/2} \log x).$$

b) This is just another form of the Möbius inversion theorem

$$\begin{aligned} \sum_{k \geqslant 1} \mu(k) \psi(x^{1/k}) &= \sum_{k \geqslant 1} \mu(k) \sum_{\ell \geqslant 1} \vartheta(x^{1/k\ell}) \\ &= \sum_{n \geqslant 1} \sum_{k \mid n} \mu(k) \vartheta(x^{1/n}) = \sum_{n \geqslant 1} \delta_{1,n}\, \vartheta(x^{1/n}) = \vartheta(x). \end{aligned}$$

11.11. Corollary. *The asymptotic relation*

$$\pi(x) \sim \frac{x}{\log x} \quad \text{for } x \to \infty \quad \text{(prime number theorem)}$$

is equivalent to the asymptotic relation

$$\psi(x) \sim x \quad \text{for } x \to \infty.$$

Proof. Since by the preceding theorem $\vartheta(x) \sim x$ is equivalent to $\psi(x) \sim x$, this follows from corollary 11.6.

Remark. We will indeed use this equivalence when we prove the prime number theorem in chapter 13.

11.12. Lemma. *The prime decomposition of* $n!$ *is*

$$n! = \prod_p p^{\alpha_p}, \qquad \text{where} \quad \alpha_p = \sum_{k \geqslant 1} \Big\lfloor \frac{n}{p^k} \Big\rfloor.$$

Proof. ...

11.13. Theorem (Bertrand's postulate). *For every positive integer* n *there is at least one prime* p *with* $n < p \leq 2n$.

Proof. ...

11.14. Theorem.

$$\sum_{p \leqslant x} \frac{\log p}{p} = \log x + O(1).$$

Proof. ...

11.15. Theorem. *There exists a real constant* B *such that*

$$\sum_{p \leqslant x} \frac{1}{p} = \log \log x + B + O\Big(\frac{1}{\log x}\Big).$$

Proof. ...