THE THEOREMS OF CHEBYSHEV AND MERTENS

Introduction. We will now present two of the important precursors to the prime number theorem, namely some results by Chebyshev (1848) and Mertens (1874). Here we will be somewhat liberal in our usage of the term "Chebyshev's theorem", by which we will mean the asymptotic estimate $\psi(x) \asymp x$; using Abel summation in the usual way, we see that this is equivalent to the assertion that $\pi(x) \asymp x/\log x$.

Our starting point is the "weak version" of Stirling's formula:

$$\Sigma(x) := \sum_{n\le x} \log n = x\log x - x + O(\log x). \tag{1}$$

Recalling that

$$\log n = \sum_{d|n} \Lambda(d),$$

we see that $\Sigma(x)$ can be written as

$$\Sigma(x) = \sum_{qd\le x} \Lambda(d). \tag{2}$$

This sum can be expressed in two different ways, depending on the order of summation. Remarkably, we are thus taken respectively to Chebyshev's theorem and the Mertens theorems.

Chebyshev's theorem. Summing first with respect to d in (2), we get

$$\Sigma(x) = \sum_{q\le x} \psi(x/q).$$

Hence, taking into account (1), we get

$$\Sigma(x) = \sum_{q\le x} \psi(x/q) = x\log x - x + O(\log x).$$

We are interested in estimating $\psi(x)$ from this expression. To this end, we observe that if we instead consider the function $\Sigma(x) - 2\Sigma(x/2)$, we get

$$\psi(x) - \psi(x/2) + \psi(x/3) - \psi(x/4) + \cdots = x\log 2 + O(\log x).$$

Since we have alternating signs and decreasing magnitude of the terms on the left-hand side, this yields immediately

$$\psi(x) \ge x\log 2 + O(\log x). \tag{3}$$

For the same reason, we also get

$$\psi(x) - \psi(x/2) \le x\log 2 + O(\log x).$$

Summing this inequality for $\psi(x) - \psi(x/2)$, $\psi(x/2) - \psi(x/4)$, $\psi(x/4) - \psi(x/8)$, etc., we find that

$$\psi(x) \le \sum_{0\le k\le (\log x)/\log 2} \left(2^{-k} x\log 2 + O\left(\log\left(2^{-k}x\right)\right)\right) = x\log 4 + O\left((\log x)^2\right). \tag{4}$$

From (3) and (4), we see that we have now proved "Cheyshev's theorem" in the form

$$x\log 2 + O(\log x) \le \psi(x) \le x\log 4 + O\left((\log x)^2\right).$$

Mertens's theorems. Summing next with respect to q in (2), we get

$$\Sigma(x) = \sum_{d\le x} \Lambda(d) \left[\frac{x}{d}\right] = \sum_{d\le x} \Lambda(d) \left(\frac{x}{d} + O(1)\right) = x \sum_{d\le x} \frac{\Lambda(d)}{d} + O\big(\psi(x)\big).$$

Plugging in the bound $\psi(x) \asymp x$ from the preceding subsection and using again (1), we get from this

$$\sum_{d\le x} \frac{\Lambda(d)}{d} = \log x + O(1). \tag{5}$$

We now use an argument we have seen before:

$$\sum_{d\le x} \frac{\Lambda(d)}{d} = \sum_{p\le x} \frac{\log p}{p} + \sum_{p^2\le x} \frac{\log p}{p^2} + \sum_{p^3\le x} \frac{\log p}{p^3} + \cdots.$$

Here it suffices to estimate the higher order terms quite crudely. For each prime p, we have

$$\sum_{k=2}^{\infty} \frac{1}{p^k} \le 2p^{-2},$$

and hence

$$\sum_{d\le x} \frac{\Lambda(d)}{d} = \sum_{p\le x} \frac{\log p}{p} + O\left(\sum_{p} \frac{\log p}{p^2}\right) = \sum_{p\le x} \frac{\log p}{p} + O(1).$$

Combining this with (5), we then have

$$\sum_{p\le x} \frac{\log p}{p} = \log x + O(1), \tag{6}$$

which is known as the first theorem of Mertens.

By Abel summation, we get from (6),

$$\begin{aligned}\sum_{p\le x} \frac{1}{p} &= 1 + O(1/\log x) + \int_2^x \big(\log y + O(1)\big) \frac{dy}{y(\log y)^2} \\ &= \log\log x + 1 - \log\log 2 + \int_2^\infty O(1) \frac{dy}{y(\log y)^2} + O(1/\log x) \\ &= \log\log x + M + O(1/\log x),\end{aligned}$$

where $M = 0.261...$ is a constant that is named after Mertens (sometimes called the Meissel–Mertens constant).

The third theorem of Mertens,

$$\prod_{p\le x} (1 - 1/p) = \frac{e^{-\gamma}}{\log x} \big(1 + O(1/\log x)\big),$$

can be deduced from (6) as well, where γ is Euler's constant.