Chapter 8

The Prime Number Theorem

The Prime Number Theorem makes predictions about the growth of the prime counting function

$$\pi(x) = \{p \le x : \ p \text{ prime}\}.$$

The following diagrams from Zagier's article "The first 50 million primes" (Mathematical Intelligencer (1977), 1-19) show how $\pi(x)$ behaves for $x \le 100$:

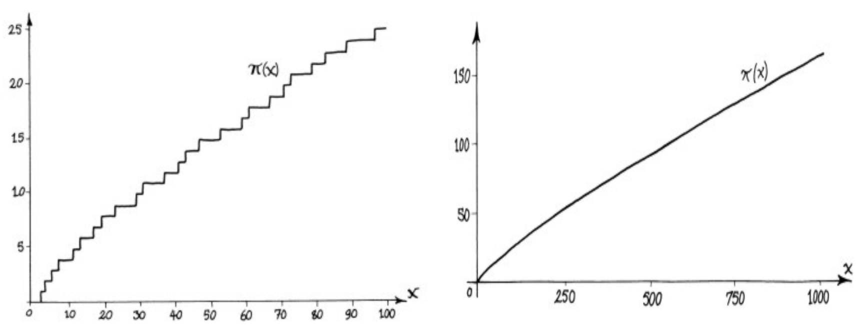

Figure 8.1: $\pi(x)$ for $x \le 100$ and $x \le 1000$

It seems clear that $\pi(x)$ behaves like a smooth function for large values of x. Is there an analytic function $f(x)$ such that $\pi(x) \sim f(x)$, i.e., $\lim \frac{\pi(x)}{f(x)} = 1$? Legendre and Gauss conjectures that $f(x) = \frac{x}{\log x}$ would do it, and this was proved a hundred years later independently by Hadamard and de la Vallee-Poussin using complex analysis (essentially, the truth of the prime number theorem rests upon the fact that the Riemann zeta function defined by $\zeta(s) = \sum_{n \ge 1} n^{-s}$ in the half plane Re $s > 1$ can be extended meromorphically to the whole complex plane and does not have any zeros on the line Re $s = 1$).

A first step in the direction of a proof of the prime number theorem was done by Chebyshev, who proved with his bare hands that there exist constants $c_1, c_2 > 0$ with

$$c_1 \frac{x}{\log x} < \pi(x) < c_2 \frac{x}{\log x}.$$

The closer the constants c_j are to 1, the more technical the proof becomes. Here we will show that $c_1 = \frac{\log 2}{2}$ and $c_2 = 6 \log 2$ do work.

Chebyshev's proof uses elementary properties of the binomial coefficients $\binom{2m}{m} = \frac{(2m)!}{m!m!}$. Here are some of them:

1. $\frac{2^{2n}}{2n} \le \binom{2n}{n} \le 2^{2n}$: this comes from $(1+1)^{2n} = \sum_{m=0}^{2n} \binom{2m}{m}$.
2. $\binom{2n}{n}$ is not divisible by any primes $p > 2n$.
3. $\binom{2n}{n}$ is divisible by all primes $n < p \le 2n$.

8.1 The Upper Bound

Properties (2) and (1) of the middle binomial coefficient imply that

$$n^{\pi(2n)-\pi(n)} \le \prod_{n<p\le 2n} p \le \binom{2n}{n} \le 2^{2n}.$$

Taking the log gives $\pi(2n) - \pi(n) \le 2 \log 2 \frac{n}{\log n}$. Using induction we now easily see that

$$\pi(2^k) \le 3 \cdot \frac{2^k}{k}.$$

In fact, this is checked directly for $k \le 5$; for $k > 5$ we find

$$\pi(2^{k+1}) \le \pi(2^k) + \frac{2^{k+1}}{k} \le \frac{3 \cdot 2^k}{k} + \frac{2 \cdot 2^k}{k} = \frac{5 \cdot 2^k}{k} \le \frac{3 \cdot 2^{k+1}}{k+1}.$$

Now we exploit the fact that $f(x) = \frac{x}{\log x}$ is monotonely increasing for $x \ge e$. Thus if $4 \le 2^k < x \le 2^{k+1}$, then

$$\pi(x) \le \pi(2^{k+1}) \le 6 \frac{2^k}{k+1} \le 6 \log 2 \frac{2^k}{\log 2^k} \le 6 \log 2 \frac{x}{\log x}.$$

Since $\pi(x) \le 6 \log 2 \frac{x}{\log x}$ for $x \le 4$, the proof is now complete.

8.2 The Lower Bound

Here we will have to work slightly harder. First we prove

Lemma 8.1. *Let $v_p(n)$ denote the exponent of the maximal power of p dividing n. Then*

$$v_p(n!) = \sum_{m \ge 1} \left\lfloor \frac{n}{p^m} \right\rfloor.$$

Proof. Among the numbers 1, 2, ..., n, exactly $\lfloor \frac{n}{p} \rfloor$ are multiples of p and thus contribute 1 to the exponent; moreover, exactly $\lfloor \frac{n}{p^2} \rfloor$ are multiples of p^2 and contribute another 1 to the exponent ... $\square$

Now put $N = \binom{2n}{n}$. By the lemma above we have

$$v_p(N) = \sum_{m \geq 1} \Big(\Big\lfloor \frac{2n}{p^m} \Big\rfloor - 2 \Big\lfloor \frac{n}{p^m} \Big\rfloor \Big).$$

Now we use

Lemma 8.2. *For all $x \in \mathbb{R}$ we have $\lfloor 2x \rfloor - 2\lfloor x \rfloor \in \{0, 1\}$.*

Proof. Write $x = \lfloor x \rfloor + \{x\}$. If the fractional part $\{x\} < \frac{1}{2}$, then $2x = 2\lfloor x \rfloor + \{2x\}$, hence $\lfloor 2x \rfloor - 2\lfloor x \rfloor = 0$. If $\{x\} \geq \frac{1}{2}$, then we get $\lfloor 2x \rfloor - 2\lfloor x \rfloor = 1$. $\square$

It is also clear that if $m > \frac{\log 2n}{\log p}$, then $\lfloor \frac{2n}{p^m} \rfloor - 2\lfloor \frac{n}{p^m} \rfloor = 0$. Thus we find $v_p(N) \leq \lfloor \frac{\log 2n}{\log p} \rfloor$, and now

$$\begin{aligned}
2n \log 2 - \log 2n &\leq \log \binom{2n}{n} && \text{because } \frac{2^{2n}}{2n} \leq \binom{2n}{n} \\
&\leq \sum_{p \leq 2n} \Big\lfloor \frac{\log 2n}{\log p} \Big\rfloor \log p && \text{because } N = \prod p^{v_p(N)} \\
&\leq \sum_{p \leq 2n} \log 2n && \text{because } \lceil x \rceil \leq x \\
&= \pi(2n) \log 2n
\end{aligned}$$

This yields the lower bound

$$\pi(2n) \geq \log 2 \frac{2n}{\log 2n} - 1.$$

We claim that this implies

$$\pi(x) \geq \frac{\log 2}{2} \frac{x}{\log x}$$

for all $x \geq 2$. This inequality can be checked directly for $x \leq 16$, hence it is sufficient to prove it for $x > 16$. Pick an integer n with $16 \leq 2n < x \leq 2n + 2$. Then

$$\frac{2n}{\log 2n} - \frac{n+1}{\log 2n} = \frac{n-1}{\log 2n} \geq \frac{7}{4 \log 2} \geq \frac{1}{\log 2},$$

hence

$$\pi(x) \geq \pi(2n) \geq \log 2 \frac{2n}{\log 2n} - 1 \geq \frac{(n+1)\log 2}{\log(2n+2)} \geq \frac{\log 2}{2} \frac{x}{\log x}.$$

8.3 Bertrand's Postulate

By improving the constants c_1 and c_2 above it is clear that Bertrand's conjecture that $\pi(2n) - \pi(n) > 0$ for all integers n must follow: there is at least one prime between n and $2n$ for all $n \in \mathbb{N}$. In fact a simple calculation shows that Bertrand's conjecture follows as soon as we can find bounds that satisfy $2c_1 > c_2$.

Here is a direct proof. Put $\Theta(x) = \prod_{p \le x} p$.

Proposition 8.3. *We have $\Theta(x) \le 4^x$.*

Proof. It is clearly sufficient to prove this for integers $x \ge 4$. Observe that $\binom{2m+1}{m} = \binom{2m+1}{m+1} < 2^{2m} = 4^m$. This gives

$$\prod_{m+2 \le p \le 2m+1} p \mid \binom{2m+1}{m} < 4^m.$$

Now we prove the claim by induction; assume it is true for all $n < k$. If k is even, then $\Theta(k) = \Theta(k-1) < 4^{k-1} < 4^k$ by induction assumption and the fact that k is not prime. Assume therefore that $k = 2m+1$. Then

$$\Theta(k) = \Theta(m+1) \prod_{m+2 \le p \le 2m+1} p < 4^{m+1} \cdot 4^m = 4^k.$$

□

We have already seen that $N = \binom{2n}{n}$ is divisible by all primes p with $n < p \le 2n$ (if there are any). Now we claim that the primes p with $\frac{2}{3}n < p \le n$ do not divide N. In fact we have $2n < 3p \le p^2$, hence $2 \le \frac{2n}{p} < 3$, hence $\lceil \frac{2n}{p} \rceil = 2$ and $\lceil \frac{n}{p} \rceil = 1$, and this implies that $v_p(N) = 2 - 2 = 0$.

Now we prove Bertrand's postulate by contradiction. Assume there is an integer n such that the interval $(n, 2n]$ does not contain any prime. By the discussion above this implies that $N = \binom{2n}{n}$ is not divisible by any prime $p > \frac{2}{3}n$. Now consider primes $p \mid N$ with $v_p(N) > 1$. They satisfy $p^2 \le p^{v_p} \le 2n$, hence we must have $p \le \sqrt{2n}$ for such primes. The number of such primes is clearly bounded by $\sqrt{2n}$. Now we find

$$\frac{2^{2n}}{2n} \le \binom{2n}{n} \le \prod_{v_p > 1} p^{v_p} \cdot \prod_{v_p = 1} p \le (2n)^{\sqrt{2n}} \cdot \Theta\Big(\frac{2n}{3}\Big) \le (2n)^{\sqrt{2n}} 2^{4n/3}.$$

Taking the log we get

$$2n \log 2 \le 3(1 + \sqrt{2n}\,) \log 2n.$$

Since $\frac{x}{\log x}$ is monotonically increasing for $x > 3$, this inequality must be false for all sufficiently large values of n. In fact, it is false for $n \ge 512$. For $n < 512$, Bertrand's postulate is proved by looking at the sequence of primes 7, 13, 23, 43, 83, 163, 317, 631.