Introduction to Analytic Number Theory

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

1. Some Elementary Theorems on the Distribution of Prime Numbers
   - Chebyshev's functions $\psi(x)$ and $\vartheta(x)$
   - Relations connecting $\vartheta(x)$ and $\pi(x)$
   - Some Equivalent Forms of the Prime Number Theorem
   - Inequalities for $\pi(n)$ and p_n
   - Shapiro's Tauberian Theorem
   - Applications of Shapiro's Theorem
   - An Asymptotic Formula for the Partial Sums $\sum_{p\leq x}(1/p)$
   - The Partial Sums of the Möbius Function
   - Selberg's Asymptotic Formula

Subsection 1

Chebyshev's functions $\psi(x)$ and $\vartheta(x)$

Chebyshev's ψ-Function

Definition

For $x > 0$ we define **Chebyshev's ψ-function** by the formula

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

- Recall $\Lambda(n) = 0$, unless n is a prime power.

  So we get

$$\psi(x) = \sum_{n \le x} \Lambda(n) = \sum_{\substack{m=1 \\ p^m \le x}}^{\infty} \Lambda(p^m) = \sum_{m=1}^{\infty} \sum_{p \le x^{1/m}} \log p.$$

Chebyshev's ψ-Function

- We got
$$\psi(x) = \sum_{m=1}^{\infty} \sum_{p \leq x^{1/m}} \log p.$$
The sum on m is actually a finite sum.
Suppose $m > \log_2 x$. Then $m > \frac{\log x}{\log 2}$.
So $\frac{1}{m} \log x < \log 2$. Hence, $x^{1/m} < 2$.
It follows that, in this case, the sum on p is empty.
So
$$\psi(x) = \sum_{m \leq \log_2 x} \sum_{p \leq x^{1/m}} \log p.$$

Chebyshev's ϑ-Function

Definition

If $x > 0$ we define **Chebyshev's ϑ-function** by the equation

$$\vartheta(x) = \sum_{p \leq x} \log p,$$

where p runs over all primes $\leq x$.

- The last formula for $\psi(x)$ can now be restated as follows:

$$\psi(x) = \sum_{m \leq \log_2 x} \sum_{p \leq x^{1/m}} \log p = \sum_{m \leq \log_2 x} \vartheta(x^{1/m}).$$

Relation Between $\frac{\psi(x)}{x}$ and $\frac{\vartheta(x)}{x}$

Theorem

For $x > 0$, we have

$$0 \leq \frac{\psi(x)}{x} - \frac{\vartheta(x)}{x} \leq \frac{(\log x)^2}{2\sqrt{x}\log 2}.$$

Note: This inequality implies that

$$\lim_{x\to\infty}\left(\frac{\psi(x)}{x} - \frac{\vartheta(x)}{x}\right) = 0.$$

In other words, if one of $\frac{\psi(x)}{x}$ or $\frac{\vartheta(x)}{x}$ tends to a limit then so does the other, and the two limits are equal.

Proof of the Theorem

- We saw
$$\psi(x) = \sum_{m \le \log_2 x} \vartheta(x^{1/m}).$$
Therefore,
$$0 \le \psi(x) - \vartheta(x) = \sum_{2 \le m \le \log_2 x} \vartheta(x^{1/m}).$$
From the definition of $\vartheta(x)$ we have the trivial inequality
$$\vartheta(x) = \sum_{p \le x} \log p \le \sum_{p \le x} \log x \le x \log x.$$

Proof of the Theorem

- We have
  - $0 \leq \psi(x) - \vartheta(x) = \sum_{2 \leq m \leq \log_2 x} \vartheta(x^{1/m})$;
  - $\vartheta(x) \leq x \log x$.

  Combining, we get

$$\begin{array}{rcl} 0 & \leq & \psi(x) - \vartheta(x) \\ & \leq & \sum_{2 \leq m \leq \log_2 x} x^{1/m} \log(x^{1/m}) \\ & \leq & (\log_2 x)\sqrt{x} \log \sqrt{x} \\ & = & \frac{\log x}{\log 2} \cdot \frac{\sqrt{x}}{2} \log x \\ & = & \frac{\sqrt{x}(\log x)^2}{2 \log 2}. \end{array}$$

  Now divide by x to obtain the conclusion.

Subsection 2

Relations connecting $\vartheta(x)$ and $\pi(x)$

Abel's Identity

- Both functions $\pi(x)$ and $\vartheta(x)$ are step functions with jumps at the primes:
  - $\pi(x)$ has a jump 1 at each prime p;
  - $\vartheta(x)$ has a jump of $\log p$ at each prime p.
- Sums involving step functions of this type can be expressed as integrals by means of the following theorem.

Theorem (Abel's Identity)

For any arithmetical function $a(n)$, let $A(x) = \sum_{n \leq x} a(n)$, where $A(x) = 0$ if $x < 1$. Assume f has a continuous derivative on the interval $[y, x]$, where $0 < y < x$. Then we have

$$\sum_{y < n \leq x} a(n) f(n) = A(x) f(x) - A(y) f(y) - \int_y^x A(t) f'(t) dt.$$

Proof of Abel's Identity

- Let $k = [x]$ and $m = [y]$, so that $A(x) = A(k)$ and $A(y) = A(m)$. Then

$$\begin{aligned}
\sum_{y<n\leq x} a(n)f(n) &= \sum_{n=m+1}^{k} a(n)f(n) \\
&= \sum_{n=m+1}^{k}\{A(n) - A(n-1)\}f(n) \\
&= \sum_{n=m+1}^{k} A(n)f(n) - \sum_{n=m}^{k-1} A(n)f(n+1) \\
&= \sum_{n=m+1}^{k-1} A(n)\{f(n) - f(n+1)\} + A(k)f(k) - A(m)f(m+1) \\
&= -\sum_{n=m+1}^{k-1} A(n)\int_{n}^{n+1} f'(t)dt + A(k)f(k) - A(m)f(m+1) \\
&= -\sum_{n=m+1}^{k-1} \int_{n}^{n+1} A(t)f'(t)dt + A(k)f(k) - A(m)f(m+1) \\
&= -\int_{m+1}^{k} A(t)f'(t)dt + A(x)f(x) - \int_{k}^{x} A(t)f'(t)dt \\
&\qquad - A(y)f(y) - \int_{y}^{m+1} A(t)f'(t)dt \\
&= A(x)f(x) - A(y)f(y) - \int_{y}^{x} A(t)f'(t)dt.
\end{aligned}$$

Alternate Proof of Abel's Identity

- A shorter proof of Abel's Identity is available to readers familiar with Riemann-Stieltjes integration.

  $A(x)$ is a step function with jump $a(n)$ at each integer n.

  So the sum in $\sum_{y<n\leq x} a(n)f(n)$ can be expressed as a Riemann-Stieltjes integral

  $$\sum_{y<n\leq x} a(n)f(n) = \int_y^x f(t)dA(t).$$

  Integration by parts gives us

  $$\begin{aligned} \textstyle\sum_{y<n\leq x} a(n)f(n) &= f(x)A(x) - f(y)A(y) - \textstyle\int_y^x A(t)df(t) \\ &= f(x)A(x) - f(y)A(y) - \textstyle\int_y^x A(t)f'(t)dt. \end{aligned}$$

A Special Case

- We have $A(t) = 0$ if $t < 1$.
- So, when $y < 1$, Abel's Identity takes the form

$$\sum_{y<n\leq x} a(n)f(n) = A(x)f(x) - \int_1^x A(t)f'(t)dt.$$

Euler's Summation Formula

- Suppose $a(n) = 1$, for all $n \geq 1$.
- Then $A(x) = [x]$.
- Thus, we get

$$\sum_{y<n\leq x} f(n) = f(x)[x] - f(y)[y] - \int_y^x [t]f'(t)dt.$$

- Integration by parts gives

$$\int_y^x tf'(t)dt = xf(x) - yf(y) - \int_y^x f(t)dt.$$

Euler's Summation Formula (Cont'd)

- Now we obtain Euler's summation formula.

$$\begin{aligned}\sum_{y<n\le x} f(n) &= f(x)[x]-f(y)[y]-\int_y^x [t]f'(t)dt \\ &= f(x)[x]-f(y)[y]+\int_y^x (t-[t])f'(t)dt \\ &\quad -\int_y^x tf'(t)dt \\ &= f(x)[x]-f(y)[y]+\int_y^x (t-[t])f'(t)dt \\ &\quad -(xf(x)-yf(y)-\int_y^x f(t)dt) \\ &= \int_y^x f(t)dt+\int_y^x (t-[t])f'(t)dt \\ &\quad +f(x)([x]-x)-f(y)([y]-y).\end{aligned}$$

An Integral Formula for $\vartheta(x)$

Theorem

For $x \geq 2$,

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

- Let $a(n)$ denote the characteristic function of the primes,

$$a(n) = \begin{cases} 1, & \text{if } n \text{ is a prime} \\ 0, & \text{otherwise} \end{cases}.$$

Then we have

$$\begin{aligned} \pi(x) &= \sum_{p \leq x} 1 = \sum_{1 < n \leq x} a(n); \\ \vartheta(x) &= \sum_{p \leq x} \log p = \sum_{1 < n \leq x} a(n) \log n. \end{aligned}$$

An Integral Formula for $\vartheta(x)$ (Cont'd)

- We obtained
  - $\vartheta(x) = \sum_{1<n\leq x} a(n)\log n$;
  - $\pi(x) = \sum_{1<n\leq x} a(n)$.

  Taking $f(x) = \log x$ in Abel's Identity with $y = 1$, we obtain

$$\sum_{1<n\leq x} a(n)\log n = \pi(x)\log x - \int_1^x \frac{\pi(t)}{t}dt.$$

  So we have

$$\begin{array}{rcl}\vartheta(x) & = & \sum_{1<n\leq x} a(n)\log n \\ & = & \pi(x)\log x - \int_1^x \frac{\pi(t)}{t}dt \\ & \overset{\pi(t)=0,\ t<2}{=} & \pi(x)\log x - \int_2^x \frac{\pi(t)}{t}dt.\end{array}$$

An Integral Formula for $\pi(x)$

Theorem

For $x \geq 2$,

$$\pi(x) = \frac{\vartheta(x)}{\log x} - \int_2^x \frac{\vartheta(t)}{t \log^2 t} dt.$$

- Let

$$a(n) = \begin{cases} 1, & \text{if } n \text{ is a prime} \\ 0, & \text{otherwise} \end{cases}.$$

  Define

$$b(n) = a(n) \log n.$$

  Write

$$\begin{aligned} \pi(x) &= \sum_{3/2<n\leq x} a(n) = \sum_{3/2<n\leq x} b(n) \tfrac{1}{\log n}; \\ \vartheta(x) &= \sum_{n\leq x} b(n). \end{aligned}$$

An Integral Formula for $\pi(x)$ (Cont'd)

- We have:
  - $\pi(x) = \sum_{3/2 < n \leq x} b(n) \frac{1}{\log n}$;
  - $\vartheta(x) = \sum_{n \leq x} b(n)$.

  Take $f(x) = \frac{1}{\log x}$ in Abel's Identity with $y = \frac{3}{2}$.

$$\begin{aligned} \pi(x) &= \frac{\vartheta(x)}{\log x} - \frac{\vartheta(3/2)}{\log 3/2} + \int_{3/2}^{x} \frac{\vartheta(t)}{t \log^2 t} dt \\ &\overset{\vartheta(t) = 0,\ t < 2}{=} \frac{\vartheta(x)}{\log x} - \int_{2}^{x} \frac{\vartheta(t)}{t \log^2 t} dt. \end{aligned}$$

Subsection 3

Some Equivalent Forms of the Prime Number Theorem

Equivalent Forms of the Prime Number Theorem

Theorem

The following relations are logically equivalent:

$$\lim_{x\to\infty}\frac{\pi(x)\log x}{x}=1,\quad \lim_{x\to\infty}\frac{\vartheta(x)}{x}=1,\quad \lim_{x\to\infty}\frac{\psi(x)}{x}=1.$$

- By the preceding theorem,

$$\frac{\vartheta(x)}{x}=\frac{\pi(x)\log x}{x}-\frac{1}{x}\int_2^x\frac{\pi(t)}{t}dt,\quad \frac{\pi(x)\log x}{x}=\frac{\vartheta(x)}{x}+\frac{\log x}{x}\int_2^x\frac{\vartheta(t)dt}{t\log^2 t}.$$

  To see that the first limit implies the second it suffices to see that it implies

$$\lim_{x\to\infty}\frac{1}{x}\int_2^x\frac{\pi(t)}{t}dt=0.$$

Equivalent Forms of the Prime Number Theorem (Cont'd)

- Suppose $\lim_{x\to\infty}\frac{\pi(x)\log x}{x}=1$.

  This implies
  $$\frac{\pi(t)}{t}=O\left(\frac{1}{\log t}\right),\quad \text{for } t\geq 2.$$
  So we get
  $$\frac{1}{x}\int_2^x\frac{\pi(t)}{t}dt=O\left(\frac{1}{x}\int_2^x\frac{dt}{\log t}\right).$$
  Now
  $$\begin{array}{rcl}\int_2^x\frac{dt}{\log t} & = & \int_2^{\sqrt{x}}\frac{dt}{\log t}+\int_{\sqrt{x}}^x\frac{dt}{\log t}\\ & \leq & \frac{\sqrt{x}}{\log 2}+\frac{x-\sqrt{x}}{\log\sqrt{x}}.\end{array}$$
  So $\frac{1}{x}\int_2^x\frac{dt}{\log t}\to 0$, as $x\to\infty$.

Equivalent Forms of the Prime Number Theorem (Cont'd)

- Suppose $\lim_{x\to\infty}\frac{\vartheta(x)}{x}=1$.
  We must show that this implies $\lim_{x\to\infty}\frac{\log x}{x}\int_2^x\frac{\vartheta(t)dt}{t\log^2 t}=0$.
  It does imply $\vartheta(t)=O(t)$.
  So
  $$\frac{\log x}{x}\int_2^x\frac{\vartheta(t)dt}{t\log^2 t}=O\left(\frac{\log x}{x}\int_2^x\frac{dt}{\log^2 t}\right).$$
  Now
  $$\int_2^x\frac{dt}{\log^2 t}=\int_2^{\sqrt{x}}\frac{dt}{\log^2 t}+\int_{\sqrt{x}}^x\frac{dt}{\log^2 t}\leq\frac{\sqrt{x}}{\log^2 2}+\frac{x-\sqrt{x}}{\log^2\sqrt{x}}.$$
  Hence,
  $$\frac{\log x}{x}\int_2^x\frac{dt}{\log^2 t}\to 0,\ \text{as } x\to\infty.$$
  This proves that the first and second limits are equivalent.
  By a previous theorem, the second and third limits are equivalent.

Prime Number Theorem and Asymptotic Value of p_n

Theorem

Let p_n denote the n-th prime. Then the following asymptotic relations are logically equivalent:

$$\lim_{x\to\infty}\frac{\pi(x)\log x}{x}=1,\quad \lim_{x\to\infty}\frac{\pi(x)\log\pi(x)}{x}=1,\quad \lim_{n\to\infty}\frac{p_n}{n\log n}=1.$$

- We show $1\to 2\to 3\to 2\to 1$.

Proof (1→2)

- Assume $\lim_{x\to\infty} \frac{\pi(x)\log x}{x} = 1$.
  Taking logarithms we obtain
  $$\lim_{x\to\infty} [\log \pi(x) + \log\log x - \log x] = 0.$$
  Equivalently,
  $$\lim_{x\to\infty} \left[\log x \left(\frac{\log \pi(x)}{\log x} + \frac{\log\log x}{\log x} - 1\right)\right] = 0.$$
  But $\log x \to \infty$ as $x \to \infty$.
  It follows that
  $$\lim_{x\to\infty} \left(\frac{\log \pi(x)}{\log x} + \frac{\log\log x}{\log x} - 1\right) = 0.$$
  This yields $\lim_{x\to\infty} \frac{\log \pi(x)}{\log x} = 1$.
  This gives $\lim_{x\to\infty} \frac{\pi(x)\log \pi(x)}{x} = 1$.

Proof (2→3→2)

2→3 Suppose $\lim_{x\to\infty} \frac{\pi(x)\log\pi(x)}{x} = 1$.
Assume $x = p_n$. Then $\pi(x) = n$. So $\pi(x)\log\pi(x) = n\log n$.
So 2 implies $\lim_{n\to\infty} \frac{n\log n}{p_n} = 1$.

3→2 Suppose $\lim_{n\to\infty} \frac{n\log n}{p_n} = 1$.
Given x, define n by the inequalities

$$p_n \leq x < p_{n+1}.$$

Then $n = \pi(x)$. Dividing by $n\log n$, we get

$$\frac{p_n}{n\log n} \leq \frac{x}{n\log n} < \frac{p_{n+1}}{n\log n} = \frac{p_{n+1}}{(n+1)\log(n+1)} \cdot \frac{(n+1)\log(n+1)}{n\log n}.$$

Now let $n \to \infty$ and use the hypothesis to get

$$\lim_{n\to\infty} \frac{x}{n\log n} = 1.$$

Equivalently, $\lim_{x\to\infty} \frac{x}{\pi(x)\log\pi(x)} = 1$.

Proof (2→1)

2→1 Suppose $\lim_{x\to\infty}\frac{\pi(x)\log\pi(x)}{x}=1$.
Taking logarithms, we obtain

$$\lim_{x\to\infty}(\log\pi(x)+\log\log\pi(x)-\log(x))=0.$$

Equivalently,

$$\lim_{x\to\infty}\left[\log\pi(x)\left(1+\frac{\log\log\pi(x)}{\log\pi(x)}-\frac{\log x}{\log\pi(x)}\right)\right]=0.$$

But $\log\pi(x)\to\infty$. It follows that

$$\lim_{x\to\infty}\left(1+\frac{\log\log\pi(x)}{\log\pi(x)}-\frac{\log x}{\log\pi(x)}\right)=0.$$

Equivalently,

$$\lim_{x\to\infty}\frac{\log x}{\log\pi(x)}=1.$$

This gives $\lim_{x\to\infty}\frac{\pi(x)\log x}{x}=1$.

Subsection 4

Inequalities for $\pi(n)$ and p_n

An Estimate for $\binom{2n}{n}$

Lemma

For every integer $n \geq 0$, we have

$$2^n \leq \binom{2n}{n} < 4^n.$$

- The left inequality is by induction.
  - For $n = 0$, it holds that $1 \leq \frac{0!}{0!0!}$.
  - Suppose the inequality holds for some $k \geq 0$.
  - We have

$$\begin{aligned} 2^{k+1} = 2 \cdot 2^k &\leq 2\frac{(2k)!}{k!k!} = 2\frac{(2k)!(k+1)(k+1)}{k!k!(k+1)(k+1)} \\ &\leq \frac{(2k)!(2k+1)(2k+2)}{(k+1)!(k+1)!} = \frac{(2(k+1))!}{(k+1)!(k+1)!}. \end{aligned}$$

The right follows from $4^n = (1+1)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} > \binom{2n}{n}$.

Inequalities for $\pi(n)$

Theorem

For every integer $n \geq 2$, we have

$$\frac{1}{6}\frac{n}{\log n} < \pi(n) < 6\frac{n}{\log n}.$$

- Taking logarithms in $2^n \leq \binom{2n}{n} < 4^n$, we get

$$n \log 2 \leq \log (2n)! - 2 \log n! < 4 \log 4.$$

A previous theorem implies that

$$\log n! = \sum_{p \leq n} \alpha(p) \log p,$$

where the sum is extended over primes and $\alpha(p)$ is given by $\alpha(p) = \sum_{m=1}^{[\frac{\log n}{\log p}]} [\frac{n}{p^m}]$.

Inequalities for $\pi(n)$ (Cont'd)

- Hence, we get

$$\log(2n)! - 2\log n! = \sum_{p \le 2n} \sum_{m=1}^{\left[\frac{\log 2n}{\log p}\right]} \left\{ \left[\frac{2n}{p^m}\right] - 2\left[\frac{n}{p^m}\right] \right\} \log p.$$

Now note that $[2x] - 2[x]$ is either 0 or 1.
So the inequality $n \log 2 \le \log(2n)! - 2\log n!$ implies

$$n\log 2 \le \sum_{p \le 2n} \left(\sum_{m=1}^{\left[\frac{\log 2n}{\log p}\right]} 1 \right) \log p \le \sum_{p \le 2n} \log 2n = \pi(2n)\log 2n.$$

Taking into account $\log 2 > \frac{1}{2}$, this gives

$$\pi(2n) \ge \frac{n \log 2}{\log 2n} = \frac{2n}{\log 2n}\frac{\log 2}{2} > \frac{1}{4}\frac{2n}{\log 2n}.$$

Inequalities for $\pi(n)$ (Cont'd)

- For odd integers, we have

$$\begin{array}{rcl}\pi(2n+1) & \geq & \pi(2n) \\ & > & \frac{1}{4}\frac{2n}{\log 2n} \\ & > & \frac{1}{4}\frac{2n}{2n+1}\frac{2n+1}{\log(2n+1)} \\ & \geq & \frac{1}{6}\frac{2n+1}{\log(2n+1)},\end{array}$$

since $\frac{2n}{2n+1} \geq \frac{2}{3}$.

This, together with $\pi(2n) > \frac{1}{4}\frac{2n}{\log 2n}$, gives

$$\pi(n) > \frac{1}{6}\frac{n}{\log n}, \quad \text{for all } n \geq 2.$$

Inequalities for $\pi(n)$ (Cont'd)

- Consider again

$$\log(2n)! - 2\log n! = \sum_{p\le 2n} \sum_{m=1}^{\left[\frac{\log 2n}{\log p}\right]} \left\{ \left[\frac{2n}{p^m}\right] - 2\left[\frac{n}{p^m}\right] \right\} \log p.$$

Extract the term corresponding to $m = 1$.

The remaining terms are nonnegative, whence

$$\log(2n)! - 2\log n! \ge \sum_{p\le 2n} \left\{ \left[\frac{2n}{p}\right] - 2\left[\frac{n}{p}\right] \right\} \log p.$$

For those primes p, with $n < p \le 2n$, we have $\left[\frac{2n}{p}\right] - 2\left[\frac{n}{p}\right] = 1$.

So

$$\log(2n)! - 2\log n! \ge \sum_{n<p\le 2n} \log p = \vartheta(2n) - \vartheta(n).$$

Inequalities for $\pi(n)$ (Cont'd)

- It follows that

$$\log{(2n)!} - 2\log n! < n\log 4 \quad \text{implies} \quad \vartheta(2n) - \vartheta(n) < n\log 4.$$

If $n = 2^r$,

$$\vartheta(2^{r+1}) - \vartheta(2^r) < 2^r\log 4 = 2^{r+1}\log 2.$$

Summing on $r = 0, 1, \ldots, k$, we find

$$\vartheta(2^{k+1}) < 2^{k+2}\log 2.$$

Now we choose k so that $2^k \leq n < 2^{k+1}$.

We obtain

$$\vartheta(n) \leq \vartheta(2^{k+1}) < 2^{k+2}\log 2 \leq 4n\log 2.$$

Inequalities for $\pi(n)$ (Conclusion)

- If $0 < \alpha < 1$, we have
$$(\pi(n) - \pi(n^\alpha)) \log n^\alpha < \sum_{n^\alpha < p \leq n} \log p \leq \vartheta(n) < 4n \log 2.$$
Hence
$$\begin{array}{rcl} \pi(n) & < & \frac{4n \log 2}{\alpha \log n} + \pi(n^\alpha) \\ & < & \frac{4n \log 2}{\alpha \log n} + n^\alpha \\ & = & \frac{n}{\log n}(\frac{4 \log 2}{\alpha} + \frac{\log n}{n^{1-\alpha}}). \end{array}$$
If $c > 0$ and $x \geq 1$, $f(x) = x^{-c} \log x$ attains its maximum at $x = e^{1/c}$. So, for $n \geq 1$, $n^{-c} \log n \leq (e^{1/c})^{-c} \log (e^{1/c}) \leq \frac{1}{ce}$. Taking $\alpha = \frac{2}{3}$ in the last inequality for $\pi(n)$, we find
$$\pi(n) < \frac{n}{\log n} \left(6 \log 2 + \frac{3}{e} \right) < 6 \frac{n}{\log n}.$$

Inequalities for p_n

Theorem

For $n \geq 1$, the n-th prime p_n satisfies the inequalities

$$\frac{1}{6} n \log n < p_n < 12 \left(n \log n + n \log \frac{12}{e} \right).$$

- Consider $\frac{1}{6} \frac{n}{\log n} < \pi(n) < 6 \frac{n}{\log n}$.
  Take $k = p_n$. Then $k \geq 2$ and $n = \pi(k)$.
  Thus, we have
  $$n = \pi(k) < 6 \frac{k}{\log k} = 6 \frac{p_n}{\log p_n}.$$
  It follows that
  $$p_n > \frac{1}{6} n \log p_n > \frac{1}{6} n \log n.$$

Inequalities for p_n (Cont'd)

- For the upper bound, by the preceding theorem,

$$n = \pi(k) > \frac{1}{6}\frac{k}{\log k} = \frac{1}{6}\frac{p_n}{\log p_n}.$$

So $p_n < 6n \log p_n$.
But, for $x \geq 1$, $\log x \leq \frac{2}{e}\sqrt{x}$. So we have $\log p_n \leq \frac{2}{e}\sqrt{p_n}$.
It follows that $p_n < 6n\frac{2}{e}\sqrt{p_n}$. I.e., $\sqrt{p_n} < \frac{12}{e}n$.
Therefore,

$$\frac{1}{2}\log p_n < \log n + \log\frac{12}{e}.$$

Used in $p_n < 6n \log p_n$, this gives

$$p_n < 6n\left(2\log n + 2\log\frac{12}{e}\right).$$

Note: The upper bound shows once more that the series $\sum_{n=1}^{\infty}\frac{1}{p_n}$ diverges, by comparison with $\sum_{n=2}^{\infty}\frac{1}{n\log n}$.

Subsection 5

Shapiro's Tauberian Theorem

Tauberian Theorems

- We have seen that the prime number theorem is equivalent to the asymptotic formula
$$\frac{1}{x}\sum_{n\leq x}\Lambda(n)\sim 1, \text{ as } x\to\infty.$$
- We have also derived a related asymptotic formula,
$$\sum_{n\leq x}\Lambda(n)\left[\frac{x}{n}\right]=x\log x-x+O(\log x).$$
- Both sums above are weighted averages of the function $\Lambda(n)$.
  - Each term $\Lambda(n)$ is multiplied by a weight factor $\frac{1}{x}$ in the first and by $[\frac{x}{n}]$ in the second.
- Theorems relating different weighted averages of the same function are called **Tauberian theorems**.

Shapiro's Tauberian Theorem

- Shapiro's Tauberian Theorem relates, for nonnegative $a(n)$:
  - Sums of the form $\sum_{n\leq x} a(n)$;
  - Sums of the form $\sum_{n\leq x} a(n)[\frac{x}{n}]$

Theorem

Let $\{a(n)\}$ be a nonnegative sequence such that

$$\sum_{n\leq x} a(n)\left[\frac{x}{n}\right] = x\log x + O(x), \text{ for all } x \geq 1.$$

(a) For $x \geq 1$, we have

$$\sum_{n\leq x} \frac{a(n)}{n} = \log x + O(1).$$

(I.e., dropping the square brackets leads to a correct result.)

Shapiro's Tauberian Theorem (Cont'd)

Theorem (Cont'd)

(b) There is a constant $B > 0$, such that

$$\sum_{n \leq x} a(n) \leq Bx, \text{ for all } x \geq 1.$$

(c) There is a constant $A > 0$ and an $x_0 > 0$, such that

$$\sum_{n \leq x} a(n) \geq Ax, \text{ for all } x \geq x_0.$$

Shapiro's Tauberian Theorem (Part (b))

- Let $S(x) = \sum_{n \leq x} a(n)$ and $T(x) = \sum_{n \leq x} a(n)[\frac{x}{n}]$.

  Claim: $S(x) - S(\frac{x}{2}) \leq T(x) - 2T(\frac{x}{2})$.

  Write

$$\begin{array}{rcl} T(x) - 2T(\frac{x}{2}) & = & \sum_{n \leq x}[\frac{x}{n}]a(n) - 2\sum_{n \leq x/2}[\frac{x}{2n}]a(n) \\ & = & \sum_{n \leq x/2}([\frac{x}{n}] - 2[\frac{x}{2n}])a(n) + \sum_{x/2 < n \leq x}[\frac{x}{n}]a(n). \end{array}$$

  But $[2y] - 2[y]$ is either 0 or 1.

  So the first sum is nonnegative, giving

$$\begin{array}{rcl} T(x) - 2T(\frac{x}{2}) & \geq & \sum_{x/2 < n \leq x}[\frac{x}{n}]a(n) \\ & = & \sum_{x/2 < n \leq x} a(n) \\ & = & S(x) - S(\frac{x}{2}). \end{array}$$

Shapiro's Tauberian Theorem (Part (b) Cont'd)

- By the hypothesis,

$$\begin{aligned} T(x) - 2T(\tfrac{x}{2}) &= x\log x + O(x) - 2(\tfrac{x}{2}\log\tfrac{x}{2} + O(x)) \\ &= O(x). \end{aligned}$$

  Hence, by the claim,

$$S(x) - S\left(\frac{x}{2}\right) = O(x).$$

  This means that there is some constant $K > 0$, such that

$$S(x) - S\left(\frac{x}{2}\right) \leq Kx, \text{ for all } x \geq 1.$$

Shapiro's Tauberian Theorem (Part (b) Cont'd)

- We showed that, if $S(x) = \sum_{n \leq x} a(n)$, there is some constant $K > 0$, such that
$$S(x) - S\left(\frac{x}{2}\right) \leq Kx, \text{ for all } x \geq 1.$$
Replace x successively by $\frac{x}{2}, \frac{x}{4}, \ldots$ to get
$$\begin{array}{rcl} S(\frac{x}{2}) - S(\frac{x}{4}) & \leq & K\frac{x}{2}, \\ S(\frac{x}{4}) - S(\frac{x}{8}) & \leq & K\frac{x}{4}, \\ & \vdots & \end{array}$$
Note that, if $2^n > x$, then $S(\frac{x}{2^n}) = 0$.
Adding these inequalities we get
$$S(x) \leq Kx\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = 2Kx.$$
This proves (b) with $B = 2K$.

Shapiro's Tauberian Theorem (Part (a))

- Write $[\frac{x}{n}] = \frac{x}{n} + O(1)$.

  Then

$$\begin{array}{rcl} T(x) & = & \sum_{n\leq x}[\frac{x}{n}]a(n) \\ & = & \sum_{n\leq x}(\frac{x}{n}+O(1))a(n) \\ & = & x\sum_{n\leq x}\frac{a(n)}{n}+O(\sum_{n\leq x}a(n)) \\ & \overset{\text{Part (b)}}{=} & x\sum_{n\leq x}\frac{a(n)}{n}+O(x). \end{array}$$

  Hence

$$\sum_{n\leq x}\frac{a(n)}{n} = \frac{1}{x}T(x)+O(1) \overset{\text{hyp.}}{=} \log x + O(1).$$

Shapiro's Tauberian Theorem (Part (c))

- Let $A(x) = \sum_{n \leq x} \frac{a(n)}{n}$.
  Then, by Part (a),
  $$A(x) = \log x + R(x),$$
  where $R(x)$ is the error term, with $R(x) = O(1)$.
  Thus, for some $M > 0$, we have $|R(x)| \leq M$.
  Choose α to satisfy $0 < \alpha < 1$ (we shall specify α more exactly in a moment) and consider the difference
  $$A(x) - A(\alpha x) = \sum_{\alpha x < n \leq x} \frac{a(n)}{n} = \sum_{n \leq x} \frac{a(n)}{n} - \sum_{n \leq \alpha x} \frac{a(n)}{n}.$$

Shapiro's Tauberian Theorem (Part (c) Cont'd)

- If $x \geq 1$ and $\alpha x \geq 1$, we can apply the asymptotic formula for $A(x)$ to write

$$\begin{aligned} A(x) - A(\alpha x) &= \log x + R(x) - (\log \alpha x + R(\alpha x)) \\ &= -\log \alpha + R(x) - R(\alpha x) \\ &\geq -\log \alpha - |R(x)| - |R(\alpha x)| \\ &\geq -\log \alpha - 2M. \end{aligned}$$

Shapiro's Tauberian Theorem (Part (c) Cont'd)

- We got
$$A(x) - A(\alpha x) \geq -\log \alpha - 2M.$$
Now choose α so that $-\log \alpha - 2M = 1$. I.e., $\alpha = e^{-2M-1}$.
Note that $0 < \alpha < 1$. For this α, if $x \geq \frac{1}{\alpha}$, $A(x) - A(\alpha x) \geq 1$.
But
$$A(x) - A(\alpha x) = \sum_{\alpha x < n \leq x} \frac{a(n)}{n} \leq \frac{1}{\alpha x} \sum_{n \leq x} a(n) = \frac{S(x)}{\alpha x}.$$
Hence, if $x \geq \frac{1}{\alpha}$, $\frac{S(x)}{\alpha x} \geq 1$.
Therefore, if $x \geq \frac{1}{\alpha}$, $S(x) \geq \alpha x$.
This proves Part (c), with $A = \alpha$ and $x_0 = \frac{1}{\alpha}$.

Subsection 6

Applications of Shapiro's Theorem

Application to $\Lambda(n)$

- Recall the asymptotic formula
$$\sum_{n\leq x}\Lambda(n)\left[\frac{x}{n}\right]=x\log x-x+O(\log x).$$
- It implies that
$$\sum_{n\leq x}\Lambda(n)\left[\frac{x}{n}\right]=x\log x+O(x).$$
- Since $\Lambda(n)\geq 0$, we can apply Shapiro's Theorem with $a(n)=\Lambda(n)$.

Application to $\Lambda(n)$ (Cont'd)

Theorem

For all $x \geq 1$, we have

$$\sum_{n \leq x} \frac{\Lambda(n)}{n} = \log x + O(1).$$

Also, there exist positive constants c_1 and c_2, such that

$$\psi(x) \leq c_1 x, \text{ for all } x \geq 1,$$

and

$$\psi(x) \geq c_2 x, \text{ for all sufficiently large } x.$$

- Follows directly by Shapiro's Theorem.

Application to $\Lambda_1(n)$

- We also showed that

$$\sum_{p \le x} \left[\frac{x}{p}\right] \log p = x \log x + O(x).$$

This can be written as

$$\sum_{n \le x} \Lambda_1(n) \left[\frac{x}{n}\right] = x \log x + O(x),$$

where

$$\Lambda_1(n) = \begin{cases} \log p & \text{if } n \text{ is a prime } p, \\ 0, & \text{otherwise.} \end{cases}$$

- Since $\Lambda_1(n) \ge 0$, the hypothesis of Shapiro's Theorem is satisfied with $a(n) = \Lambda_1(n)$.

Application to $\Lambda_1(n)$ (Cont'd)

Theorem

For all $x \geq 1$, we have

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

Also, there exist positive constants c_1 and c_2, such that

$$\vartheta(x) \leq c_1 x, \quad \text{for all } x \geq 1,$$

and

$$\vartheta(x) \geq c_2 x, \quad \text{for all sufficiently large } x.$$

- This follows diectly from Shapiro's Theorem, taking into account that, by definition, $\vartheta(x) = \sum_{n \leq x} \Lambda_1(n)$.

Partial Sum Formulas for $\psi(x)$ and $\vartheta(x)$

- We also proved that

$$\sum_{n\leq x} f(n)\left[\frac{x}{n}\right]=\sum_{n\leq x} F\left(\frac{x}{n}\right),$$

for any arithmetical $f(n)$ with $F(x)=\sum_{n\leq x} f(n)$.

- Consider again the asymptotic formulas

$$\begin{array}{rcl}\sum_{n\leq x}\Lambda(n)[\frac{x}{n}] & = & x\log x-x+O(\log x),\\ \sum_{n\leq x}\Lambda_1(n)[\frac{x}{n}] & = & x\log x+O(x).\end{array}$$

Take, also, into account that

$$\psi(x)=\sum_{n\leq x}\Lambda(n)\quad\text{and}\quad\vartheta(x)=\sum_{n\leq x}\Lambda_1(n).$$

Partial Sum Formulas for $\psi(x)$ and $\vartheta(x)$ (Cont'd)

Theorem

For all $x \geq 1$, we have

$$\begin{aligned} \sum_{n \leq x} \psi\left(\frac{x}{n}\right) &= x \log x - x + O(\log x), \\ \sum_{n \leq x} \vartheta(x) &= x \log x + O(x). \end{aligned}$$

- By the preceding identities, we get

$$\begin{aligned} \sum_{n \leq x} \psi\left(\frac{x}{n}\right) &= \sum_{n \leq x} \Lambda(n) \left[\frac{x}{n}\right] \\ &= x \log x - x + O(\log x); \\ \sum_{n \leq x} \vartheta(x) &= \sum_{n \leq x} \Lambda_1(n) \left[\frac{x}{n}\right] \\ &= x \log x + O(x). \end{aligned}$$

Subsection 7

An Asymptotic Formula for the Partial Sums $\sum_{p\leq x}(1/p)$

Asymptotic Formula for the Partial Sums $\sum_{p\leq x}(1/p)$

Theorem

There is a constant A such that

$$\sum_{p\leq x}\frac{1}{p}=\log\log x+A+O\left(\frac{1}{\log x}\right),\ \text{for all } x\geq 2.$$

- Let $A(x)=\sum_{p\leq x}\frac{\log p}{p}$. Let $a(n)=\begin{cases}1, & \text{if } n \text{ is prime}\\ 0, & \text{otherwise}\end{cases}.$

  Then we have

$$\sum_{p\leq x}\frac{1}{p}=\sum_{n\leq x}\frac{a(n)}{n}\quad\text{and}\quad A(x)=\sum_{n\leq x}\frac{a(n)}{n}\log n.$$

Asymptotic Formula (Cont'd)

- Abel's identity

$$\sum_{y<n\leq x} a(n)f(n) = A(x)f(x) - A(y)f(y) - \int_y^x A(t)f'(t)dt$$

yields

$$\begin{aligned}\sum_{y<n\leq x} \frac{a(n)}{n}\log n\frac{1}{\log n} &= A(x)\frac{1}{\log x} + \int_2^x A(t)\frac{1}{t\log^2 t}dt\\ \sum_{p\leq x}\frac{1}{p} &= \frac{A(x)}{\log x} + \int_2^x A(t)\frac{1}{t\log^2 t}dt.\end{aligned}$$

From a previous theorem, we have

$$A(x) = \log x + R(x), \quad R(x) = O(1).$$

Thus,

$$\begin{aligned}\sum_{p\leq x}\frac{1}{p} &= \frac{\log x + O(1)}{\log x} + \int_2^x \frac{\log t + R(t)}{t\log^2 t}dt\\ &= 1 + O\left(\frac{1}{\log x}\right) + \int_2^x \frac{dt}{t\log t} + \int_2^x \frac{R(t)}{t\log^2 t}dt.\end{aligned}$$

Asymptotic Formula (Cont'd)

We found

$$\sum_{p\leq x}\frac{1}{p}=1+O\left(\frac{1}{\log x}\right)+\int_2^x\frac{dt}{t\log t}+\int_2^x\frac{R(t)}{t\log^2 t}dt.$$

Now note the following:

- $\int_2^x\frac{dt}{t\log t}=\log\log x-\log\log 2$;
- $\int_2^x\frac{R(t)}{t\log^2 t}dt=\int_2^\infty\frac{R(t)}{t\log^2 t}dt-\int_x^\infty\frac{R(t)}{t\log^2 t}dt$, the existence of the improper integral being assured by the condition $R(t)=O(1)$;
- $\int_x^\infty\frac{R(t)}{t\log^2 t}dt=O(\int_x^\infty\frac{dt}{t\log^2 t})=O(\frac{1}{\log x})$.

Hence,

$$\sum_{p\leq x}\frac{1}{p}=\log\log x+1-\log\log 2+\int_2^\infty\frac{R(t)}{t\log^2 t}dt+O\left(\frac{1}{\log x}\right).$$

This proves the theorem with $A=1-\log\log 2+\int_2^\infty\frac{R(t)}{t\log^2 t}dt$.

Subsection 8

The Partial Sums of the Möbius Function

Partial Sums of the Möbius Function

Definition

If $x \geq 1$, we define

$$M(x) = \sum_{n \leq x} \mu(n).$$

- The exact order of magnitude of $M(x)$ is not known.
  - Mertens' Conjecture stated that $|M(x)| < \sqrt{x}$ if $x > 1$.
  - However, the conjecture has been disproven.
  - A new conjecture asserts that $M(x)$ grows as $\sqrt{x}(\log\log\log x)^{5/4}$.
- We prove that the weaker statement

$$\lim_{x \to \infty} \frac{M(x)}{x} = 0$$

is equivalent to the prime number theorem.

$M(x)$ and $H(x)$

Definition

If $x \geq 1$, we define

$$H(x) = \sum_{n \leq x} \mu(n) \log n.$$

Theorem

We have

$$\lim_{x \to \infty} \left(\frac{M(x)}{x} - \frac{H(x)}{x \log x} \right) = 0.$$

- Taking $f(t) = \log t$ in Abel's Identity, we obtain

$$H(x) = \sum_{n \leq x} \mu(n) \log n = M(x) \log x - \int_1^x \frac{M(t)}{t} dt.$$

$M(x)$ and $H(x)$

- Taking $x > 1$, and dividing by $x \log x$, we get

$$\frac{M(x)}{x} - \frac{H(x)}{x \log x} = \frac{1}{x \log x} \int_1^x \frac{M(t)}{t} dt.$$

Therefore, to prove the theorem we must show that

$$\lim_{x \to \infty} \frac{1}{x \log x} \int_1^x \frac{M(t)}{t} dt = 0.$$

Trivially, $M(x) = O(x)$.
So

$$\int_1^x \frac{M(t)}{t} dt = O\left(\int_1^x dt\right) = O(x).$$

This implies

$$\lim_{x \to \infty} \frac{1}{x \log x} \int_1^x \frac{M(t)}{t} dt = 0.$$

Prime Number Theorem Implies the Limit

Theorem

The prime number theorem implies

$$\lim_{x\to\infty} \frac{M(x)}{x} = 0.$$

- We use the Prime Number Theorem in the form

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

We prove that $\frac{H(x)}{x\log x} \to 0$ as $x \to \infty$.

Prime Number Theorem Implies the Limit (Claim)

- Claim: $-H(x) = -\sum_{n \leq x} \mu(n) \log n = \sum_{n \leq x} \mu(n) \psi(\frac{x}{n})$.
  We proved the following formula for $\Lambda(n)$.
  $$\Lambda(n) = -\sum_{d|n} \mu(d) \log d.$$
  Applying Möbius inversion, we get
  $$-\mu(n) \log n = \sum_{d|n} \mu(d) \Lambda\left(\frac{n}{d}\right).$$
  Summing over all $n \leq x$ and using the formula for the partial sums of a Dirichlet product, with $f = t$, $g = \Lambda$, we get
  $$\sum_{n \leq x} -\mu(n) \log n = \sum_{n \leq x} (\mu * \Lambda)(n) = \sum_{n \leq x} \mu(n) \psi\left(\frac{x}{n}\right).$$

Prime Number Theorem Implies the Limit (Cont'd)

- We know that $\psi(x) \sim x$.

  Thus, given $\varepsilon > 0$ is given, there is a constant $A > 0$, such that

$$\left|\frac{\psi(x)}{x} - 1\right| < \varepsilon, \text{ whenever } x \geq A.$$

In other words,

$$|\psi(x) - x| < \varepsilon x, \quad \text{whenever } x \geq A.$$

Choose $x > A$, let $y = [\frac{x}{A}]$ and split the sum on the right of the Claim into two parts,

$$\sum_{n \leq y} \mu(n)\psi\left(\frac{x}{n}\right) + \sum_{y < n \leq x} \mu(n)\psi\left(\frac{x}{n}\right).$$

Prime Number Theorem Implies the Limit (First Sum)

- We look at $\sum_{n\leq y}\mu(n)\psi(\frac{x}{n})$, where $y=[\frac{x}{A}]$.
  We have $n\leq y$. So $n\leq\frac{x}{A}$. Hence, $\frac{x}{n}\geq A$.
  Therefore, for $n\leq y$, we have $|\psi(\frac{x}{n})-\frac{x}{n}|<\varepsilon\frac{x}{n}$.
  Thus,

$$\begin{array}{rcl}\sum_{n\leq y}\mu(n)\psi(\frac{x}{n}) & = & \sum_{n\leq y}\mu(n)(\frac{x}{n}+\psi(\frac{x}{n})-\frac{x}{n}) \\ & = & x\sum_{n\leq y}\frac{\mu(n)}{n}+\sum_{n\leq y}\mu(n)(\psi(\frac{x}{n})-\frac{x}{n}).\end{array}$$

  It follows that

$$\begin{array}{rcl}|\sum_{n\leq y}\mu(n)\psi(\frac{x}{n})| & \leq & x|\sum_{n\leq y}\frac{\mu(n)}{n}|+\sum_{n\leq y}|\psi(\frac{x}{n})-\frac{x}{n}| \\ & < & x+\varepsilon\sum_{n\leq y}\frac{x}{n} \\ & < & x+\varepsilon x(1+\log y) \\ & < & x+\varepsilon x+\varepsilon x\log x.\end{array}$$

Prime Number Theorem Implies the Limit (Second Sum)

- We look at $\sum_{y<n\leq x}\mu(n)\psi(\frac{x}{n})$, where $y=[\frac{x}{A}]$.
  In this sum, $y<n\leq x$ So $n\geq y+1$.
  As $y\leq\frac{x}{A}<y+1$, we get
  $$\frac{x}{n}\leq\frac{x}{y+1}<A.$$
  But $\frac{x}{n}<A$ implies
  $$\psi\left(\frac{x}{n}\right)\leq\psi(A).$$
  Therefore,
  $$\sum_{y<n\leq x}\mu(n)\psi\left(\frac{x}{n}\right)\leq x\psi(A).$$

Prime Number Theorem Implies the Limit (Sum)

- Now, regarding the full sum, we have, for $\varepsilon < 1$,

$$\begin{array}{rcl}\sum_{n \leq x} \mu(n)\psi(\frac{x}{n}) & < & (1+\varepsilon)x + \varepsilon x \log x + x\psi(A) \\ & < & (2+\psi(A))x + \varepsilon x \log x.\end{array}$$

So, given any $0 < \varepsilon < 1$, we have, for $x > A$,

$$|H(x)| < (2+\psi(A))x + \varepsilon x \log x.$$

Equivalently,

$$\frac{|H(x)|}{x \log x} < \frac{2+\psi(A)}{\log x} + \varepsilon.$$

Choose $B > A$ so that $x > B$ implies $\frac{2+\psi(A)}{\log x} < \varepsilon$.
Then, for $x > B$, we have

$$\frac{|H(x)|}{x \log x} < 2\varepsilon.$$

This shows that $\frac{H(x)}{x \log x} \to 0$ as $x \to \infty$.

Little Oh Notation

Definition

The notation

$$f(x) = o(g(x)) \quad \text{as } x \to \infty$$

(read: $f(x)$ is **little oh of** $g(x)$) means that

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

An equation of the form

$$f(x) = h(x) + o(g(x)) \quad \text{as } x \to \infty$$

means that

$$f(x) - h(x) = o(g(x)) \quad \text{as } x \to \infty.$$

Some Observations

- The equation
$$\lim_{x\to\infty} \frac{M(x)}{x} = 0$$
states that $M(x) = o(x)$ as $x \to \infty$.
- The Prime Number Theorem, expressed in the form
$$\psi(x) \sim x,$$
can also be written as $\psi(x) = x + o(x)$ as $x \to \infty$.
- More generally, an asymptotic relation
$$f(x) \sim g(x) \text{ as } x \to \infty$$
is equivalent to
$$f(x) = g(x) + o(g(x)) \text{ as } x \to \infty.$$
- Finally note that $f(x) = O(1)$ implies $f(x) = o(x)$ as $x \to \infty$.

$M(x)$ and $\psi(x)$

Theorem

The relation $M(x) = o(x)$ as $x \to \infty$ implies $\psi(x) \sim x$ as $x \to \infty$.

Claim: We have

$$\psi(x) = x - \sum_{\substack{q,d \\ qd \leq x}} \mu(d) f(q) + O(1),$$

where:

- f is given by $f(n) = \sigma_0(n) - \log n - 2C$, with C Euler's constant;
- $\sigma_0(n) = d(n)$ is the number of divisors of n.

$M(x)$ and $\psi(x)$ (Claim Cont'd)

- We start with the identities

$$[x] = \sum_{n \leq x} 1, \quad \psi(x) = \sum_{n \leq x} \Lambda(n), \quad 1 = \sum_{n \leq x} \left[\frac{1}{n}\right].$$

Now we take into account

$$\sigma_0(n) = \sum_{d \mid n} 1, \quad \log n = \sum_{d \mid n} \Lambda(n), \quad 1 = \sum_{d \mid n} \left[\frac{1}{d}\right].$$

Applying Möbius inversion, we express each summand in the first sums as a Dirichlet product involving the Möbius function

$$1 = \sum_{d \mid n} \mu(d)\sigma_0\left(\frac{n}{d}\right), \quad \Lambda(n) = \sum_{d \mid n} \mu(d) \log \frac{n}{d}, \quad \left[\frac{1}{n}\right] = \sum_{d \mid n} \mu(d).$$

$M(x)$ and $\psi(x)$ (Claim Cont'd)

- We have

$$1=\sum_{d\mid n}\mu(d)\sigma_0\left(\frac{n}{d}\right),\quad \Lambda(n)=\sum_{d\mid n}\mu(d)\log\frac{n}{d},\quad \left[\frac{1}{n}\right]=\sum_{d\mid n}\mu(d).$$

Then

$$\begin{aligned}[x]-\psi(x)-2C &= \textstyle\sum_{n\leq x}\{1-\Lambda(n)-2C[\frac{1}{n}]\}\\ &= \textstyle\sum_{n\leq x}\sum_{d\mid n}\mu(d)\{\sigma_0(\frac{n}{d})-\log\frac{n}{d}-2C\}\\ &= \textstyle\sum_{\substack{q,d\\ qd\leq x}}\mu(d)\{\sigma_0(q)-\log q-2C\}\\ &= \textstyle\sum_{\substack{q,d\\ qd\leq x}}\mu(d)f(q).\end{aligned}$$

For the proof of the theorem we must show

$$\sum_{\substack{q,d\\ qd\leq x}}\mu(d)f(q)=o(x),\quad \text{as } x\to\infty.$$

$M(x)$ and $\psi(x)$ (Another Claim)

Claim: We have

$$F(x) = \sum_{n \leq x} f(n) = O(\sqrt{x}).$$

We use Dirichlet's formula

$$\sum_{n \leq x} \sigma_0(n) = x \log x + (2C - 1)x + O(\sqrt{x})$$

together with the relation

$$\sum_{n \leq x} \log n = \log [x]! = x \log x - x + O(\log x).$$

$M(x)$ and $\psi(x)$ (Another Claim Cont'd)

- These give us

$$\begin{array}{rcl} F(x) & = & \sum_{n\leq x}(\sigma_0(n)-\log n-2C) \\ & = & \sum_{n\leq x}\sigma_0(n)-\sum_{n\leq x}\log n-2C\sum_{n\leq x}1 \\ & = & x\log x+(2C-1)x+O(\sqrt{x}) \\ & & -(x\log x-x+O(\log x))-2Cx+O(1) \\ & = & O(\sqrt{x})+O(\log x)+O(1)=O(\sqrt{x}). \end{array}$$

Therefore, there is a constant $B>0$, such that

$$|F(x)|\leq B\sqrt{x},\quad \text{for all } x\geq 1.$$

$M(x)$ and $\psi(x)$ (Final Claim)

Claim: We have

$$\sum_{\substack{q,d \\ qd\le x}} \mu(d)f(q) = o(x), \quad \text{as } x \to \infty.$$

We use the generalized formula for the partial sums of Dirichlet products, where a and b are any positive numbers, with $ab = x$, and $F(x) = \sum_{n\le x} f(n)$.

$$\sum_{\substack{q,d \\ dq\le x}} f(d)g(q) = \sum_{n\le a} f(n)G\left(\frac{x}{n}\right) + \sum_{n\le b} g(n)F\left(\frac{x}{n}\right) - F(a)g(b).$$

It gives

$$\sum_{\substack{q,d \\ qd\le x}} \mu(d)f(q) = \sum_{n\le b} \mu(n)F\left(\frac{x}{n}\right) + \sum_{n\le a} f(n)M\left(\frac{x}{n}\right) - F(a)M(b).$$

$M(x)$ and $\psi(x)$ (Final Claim Cont'd)

- By the last Claim, in the first sum on the right, we get that, for some constant $A > B > 0$,

$$\left|\sum_{n\leq b}\mu(n)F\left(\frac{x}{n}\right)\right| \leq B\sum_{n\leq b}\sqrt{\frac{x}{n}} \leq A\sqrt{xb} = \frac{Ax}{\sqrt{a}}.$$

Let $\varepsilon > 0$ be arbitrary and choose $a > 1$, such that $\frac{A}{\sqrt{a}} < \varepsilon$.

Then, for all $x \geq 1$,

$$\left|\sum_{n\leq b}\mu(n)F\left(\frac{x}{n}\right)\right| < \varepsilon x.$$

Note that a depends on ε and not on x.

$M(x)$ and $\psi(x)$ (Final Claim Cont'd)

- Now $M(x) = O(x)$ as $x \to \infty$.

  So, for the same ε, there exists $c > 0$ (depending only on ε), such that, for $x > c$,

  $$\frac{|M(x)|}{x} < \frac{\varepsilon}{K},$$

  where K is any positive number (to be specified).

  The second sum satisfies

  $$\left|\sum_{n \leq a} f(n) M\left(\frac{x}{n}\right)\right| \leq \sum_{n \leq a} |f(n)| \frac{\varepsilon}{K} \frac{x}{n} = \frac{\varepsilon x}{K} \sum_{n \leq a} \frac{|f(n)|}{n},$$

  provided $\frac{x}{n} > c$, for all $n \leq a$.

  Therefore, the inequality holds if $x > ac$.

$M(x)$ and $\psi(x)$ (Final Claim Cont'd)

- Now take $K = \sum_{n \leq a} \frac{|f(n)|}{n}$.
  Then the inequality implies, for $x > ac$,
  $$\left| \sum_{n \leq a} f(n) M\left(\frac{x}{n}\right) \right| < \varepsilon x.$$
  The last term on the right of the sum, $F(a)M(b)$, is dominated by
  $$|F(a)M(b)| \leq A\sqrt{a}|M(b)| < A\sqrt{a}b < \varepsilon\sqrt{a}\sqrt{a}b = \varepsilon ab = \varepsilon x.$$
  Now we get, for $x > ac$,
  $$\left| \sum_{\substack{q,d \\ qd \leq x}} \mu(d) f(q) \right| < 3\varepsilon x,$$
  where a and c depend only on ε.

The Limit Implies the Prime Number Theorem

Theorem

Let $A(x) = \sum_{n \leq x} \frac{\mu(n)}{n}$.
$A(x) = o(1)$, as $x \to \infty$, implies the prime number theorem.
In other words, the prime number theorem is a consequence of the statement that the series $\sum_{n=1}^{\infty} \frac{\mu(n)}{n}$ converges and has sum 0.

- Using Abel's identity, we get

$$M(x) = \sum_{n \leq x} \mu(n) = \sum_{n \leq x} \frac{\mu(n)}{n} n = xA(x) - \int_1^x A(t)dt.$$

So

$$\frac{M(x)}{x} = A(x) - \frac{1}{x} \int_1^x A(t)dt.$$

To complete the proof, we show $\lim_{x \to \infty} \frac{1}{x} \int_1^x A(t)dt = 0$.

The Limit Implies the Prime Number Theorem (Cont'd)

- We show $\lim_{x\to\infty} \frac{1}{x}\int_1^x A(t)dt = 0$.
  If $\varepsilon > 0$ is given, there exists a c (depending only on ε) such that
  $$|A(x)| < \varepsilon, \quad \text{for } x \geq c.$$
  Since $|A(x)| \leq 1$, for all $x \geq 1$, we have
  $$\left|\frac{1}{x}\int_1^x A(t)dt\right| \leq \left|\frac{1}{x}\int_1^c A(t)dt\right| + \left|\frac{1}{x}\int_c^x A(t)dt\right| \leq \frac{c-1}{x} + \frac{\varepsilon(x-c)}{x}.$$
  Letting $x \to \infty$, we find
  $$\limsup_{x\to\infty} \left|\frac{1}{x}\int_1^x A(t)dt\right| \leq \varepsilon.$$
  Since ε is arbitrary we get $\lim_{x\to\infty} \frac{1}{x}\int_1^x A(t)dt = 0$.

Subsection 9

Selberg's Asymptotic Formula

An Inversion Formula

Theorem

Let F be a real- or complex-valued function defined on $(0,\infty)$, and let

$$G(x) = \log x \sum_{n\leq x} F\left(\frac{x}{n}\right).$$

Then

$$F(x)\log x + \sum_{n\leq x} F\left(\frac{x}{n}\right)\Lambda(n) = \sum_{d\leq x} \mu(d) G\left(\frac{x}{d}\right).$$

An Inversion Formula (Cont'd)

- First we write $F(x)\log x$ as a sum,

$$F(x)\log x = \sum_{n\leq x}\left[\frac{1}{n}\right] F\left(\frac{x}{n}\right)\log\frac{x}{n} = \sum_{n\leq x} F\left(\frac{x}{n}\right)\log\frac{x}{n}\sum_{d|n}\mu(d).$$

Recall that

$$\Lambda(n) = \sum_{d|n}\mu(d)\log\frac{n}{d}.$$

Now we get

$$\sum_{n\leq x} F\left(\frac{x}{n}\right)\Lambda(n) = \sum_{n\leq x} F\left(\frac{x}{n}\right)\sum_{d|n}\mu(d)\log\frac{n}{d}.$$

An Inversion Formula (Cont'd)

- We got

$$\begin{aligned} F(x)\log x &= \sum_{n\le x} F(\tfrac{x}{n})\log\tfrac{x}{n}\sum_{d|n}\mu(d) \\ \sum_{n\le x} F(\tfrac{x}{n})\Lambda(n) &= \sum_{n\le x} F(\tfrac{x}{n})\sum_{d|n}\mu(d)\log\tfrac{n}{d}. \end{aligned}$$

Adding the two equations we find

$$\begin{aligned} &F(x)\log x + \sum_{n\le x} F(\tfrac{x}{n})\Lambda(n) \\ &= \sum_{n\le x} F(\tfrac{x}{n})\sum_{d|n}\mu(d)\{\log\tfrac{x}{n} + \log\tfrac{n}{d}\} \\ &= \sum_{n\le x}\sum_{d|n} F(\tfrac{x}{n})\mu(d)\log\tfrac{x}{d}. \end{aligned}$$

In the last sum we write $n = qd$ to obtain

$$\begin{aligned} \sum_{n\le x}\sum_{d|n} F(\tfrac{x}{n})\mu(d)\log\tfrac{x}{d} &= \sum_{d\le x}\mu(d)\log\tfrac{x}{d}\sum_{q\le x/d} F(\tfrac{x}{qd}) \\ &= \sum_{d\le x}\mu(d)G(\tfrac{x}{d}). \end{aligned}$$

Selberg's Asymptotic Formula

Theorem (Selberg's Asymptotic Formula)

For $x > 0$ we have

$$\psi(x)\log x + \sum_{n\leq x}\Lambda(n)\psi\left(\frac{x}{n}\right) = 2x\log x + O(x).$$

- We shall first apply the previous theorem to the function $F_1(x) = \psi(x)$, taking into account that, by a previous theorem,

$$\sum_{n\leq x}\psi\left(\frac{x}{n}\right) = x\log x - x + O(\log x).$$

$$\begin{aligned} G_1(x) &= \log x\sum_{n\leq x}\psi(\tfrac{x}{n}) \\ &= x\log^2 x - x\log x + O(\log^2 x). \end{aligned}$$

Selberg's Asymptotic Formula (Cont'd)

- We will then apply the previous theorem to the function $F_2(x) = x - C - 1$, where C is Euler's constant.

$$\begin{aligned} G_2(x) &= \log x \sum_{n \le x} F_2\left(\frac{x}{n}\right) \\ &= \log x \sum_{n \le x}\left(\frac{x}{n} - C - 1\right) \\ &= x \log x \sum_{n \le x} \frac{1}{n} - (C+1) \log x \sum_{n \le x} 1 \\ &= x \log x\left(\log x + C + O\left(\frac{1}{x}\right)\right) - (C+1) \log x (x + O(1)) \\ &= x \log^2 x - x \log x + O(\log x). \end{aligned}$$

Comparing the formulas for $G_1(x)$ and $G_2(x)$, we see that

$$G_1(x) - G_2(x) = O(\log^2 x).$$

We shall only use the weaker estimate $G_1(x) - G_2(x) = O(\sqrt{x})$.

Selberg's Asymptotic Formula (Cont'd)

- Apply the preceding theorem to each of F_1 and F_2 and subtract the two relations so obtained.

  The difference of the two right members is

$$\begin{aligned}\sum_{d\leq x}\mu(d)\{G_1(\tfrac{x}{d})-G_2(\tfrac{x}{d})\} &= O(\sum_{d\leq x}\sqrt{\tfrac{x}{d}}) \\ &= O(\sqrt{x}\sum_{d\leq x}\tfrac{1}{\sqrt{d}}) \\ &= O(x).\end{aligned}$$

  Therefore the difference of the two left members is also $O(x)$.

$$\{\psi(x)-(x-C-1)\}\log x+\sum_{n\leq x}\left\{\psi\left(\frac{x}{n}\right)-\left(\frac{x}{n}-C-1\right)\right\}\Lambda(n)=O(x).$$

Selberg's Asymptotic Formula (Cont'd)

- We got

$$\{\psi(x)-(x-C-1)\}\log x+\sum_{n\leq x}\left\{\psi\left(\frac{x}{n}\right)-\left(\frac{x}{n}-C-1\right)\right\}\Lambda(n)=O(x).$$

Rearranging terms, we find that

$$\psi(x)\log x+\textstyle\sum_{n\leq x}\psi(\frac{x}{n})\Lambda(n)$$
$$=(x-C-1)\log x+\textstyle\sum_{n\leq x}(\frac{x}{n}-C-1)\Lambda(n)+O(x).$$

By a previous theorem,

$$\sum_{n\leq x}\frac{\Lambda(n)}{n}=\log x+O(1).$$

Now we conclude that

$$\psi(x)\log x+\sum_{n\leq x}\psi\left(\frac{x}{n}\right)\Lambda(n)=2x\log x+O(x).$$