The Riemann zeta function and the prime number theorem

Ben Kelly, David Bass, Katherine Hennessy

University of Virginia

30 November 2022

The Riemann zeta function

For complex numbers s with $Re(s) > 1$ the **Riemann zeta function** is defined by

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

The Euler product expansion of the Riemann zeta function

The Riemann zeta function Euler's identity

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1-p^{-s}}.$$

Recall that the geometric series $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$, for $r \in \mathbb{C}$ when $|r| < 1$.

$$\begin{aligned}\prod_{p \text{ prime}} \frac{1}{1-p^{-s}} &= \prod_{p \text{ prime}} \sum_{n=0}^{\infty} \left(p^{-s}\right)^n \\ &= \prod_{p \text{ prime}} \sum_{n=0}^{\infty} \left(p^n\right)^{-s}\end{aligned}$$

The Euler product expansion of the Riemann zeta function

$$\prod_{p \text{ prime}} \sum_{n=0}^{\infty} (p^n)^{-s} = (1 + 2^{-s} + 4^{-s} + \cdots)(1 + 3^{-s} + 9^{-s} + \cdots)\cdots$$
$$= 1 + 2^{-s} + 3^{-s} + 4^{-s} + 6^{-s} + 9^{-s} + 12^{-s} + \cdots .$$

By the fundamental theorem of arithmetic we conclude that

$$\prod_{p \text{ prime}} \sum_{n=0}^{\infty} (p^n)^{-s} = \sum_{n=1}^{\infty} n^{-s} = \zeta(s).$$

$\square$

The reciprocal of the Riemann zeta function

The reciprocal of the Riemann zeta function can be defined as

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

The Euler product expansion of the Riemann zeta function shows that

$$\begin{aligned}
\frac{1}{\zeta(s)} &= \frac{1}{\prod_{p \text{ prime}} \frac{1}{1-p^{-s}}} = \prod_{p \text{ prime}} (1 - p^{-s}) \\
&= \left(1 - 2^{-s}\right)\left(1 - 3^{-s}\right)\left(1 - 5^{-s}\right)\cdots \\
&= 1 - 2^{-s} - 3^{-s} - 5^{-s} + 6^{-s} + 10^{-s} + 15^{-s} - 30^{-s} + \cdots
\end{aligned}$$

The reciprocal of the Riemann zeta function

Again by the fundamental theorem of arithmetic it is clear that

$$\prod_{p \text{ prime}} (1 - p^{-s}) = k_1 \frac{1}{1} + k_2 \frac{1}{2^s} + k_3 \frac{1}{3^s} + k_4 \frac{1}{4^s} + \cdots = \sum_{n=1}^{\infty} \frac{k_n}{n^s}.$$

In this sum, k_n is determined by the prime factorization of n. Specifically, if n is square-free and $\omega(n)$ is odd, then $k_n = -1$, if n is square-free and $\omega(n)$ is even, then $k_n = 1$, and if n is not square-free then $k_n = 0$, so $k_n = \mu(n)$. Then

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{k_n}{n^s} = \sum_{1 \leq n \text{ square-free}} (-1)^{\omega(n)} n^{-s} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.$$

□

Another expression involving the Riemann zeta function

We can now show that

$$\frac{\zeta(s)}{\zeta(2s)} = \sum_{1 \le n \text{ square-free}} \frac{1}{n^s} = \sum_{n \ge 1} \frac{|\mu(n)|}{n^s}.$$

with some basic arithmetic we can show

$$\begin{aligned}
\frac{\zeta(s)}{\zeta(2s)} &= \left(\prod_{p \text{ prime}} \frac{1}{1-p^{-s}} \right) \Big/ \left(\prod_{p \text{ prime}} \frac{1}{1-p^{-2s}} \right) \\
&= \prod_{p \text{ prime}} \frac{1-p^{-2s}}{1-p^{-s}} \\
&= \prod_{p \text{ prime}} \frac{(1+p^{-s})(1-p^{-s})}{1-p^{-s}} \\
&= \prod_{p \text{ prime}} (1+p^{-s}) \\
&= 1 + 2^{-s} + 3^{-s} + 5^{-s} + 6^{-s} + 7^{-s} + 10^{-s} + 11^{-s} + \cdots .
\end{aligned}$$

Another expression involving the Riemann zeta function

$$\frac{\zeta(s)}{\zeta(2s)} = 1 + k_2 2^{-s} + k_3 3^{-s} + k_4 + 4^{-s} + k_5 5^{-s} + k_6 6^{-s} + \cdots = \sum_{n=1}^{\infty} \frac{k_n}{n^s}$$

As in the previous proof, for n^{-s} to appear in the sum, n must be square-free, and since there are no negatives in the product terms $(1 + (p_n)^{-s})$, k_n can only be positive. Thus, if n is not square-free, then $k_n = 0$, and if n is square-free, then $k_n = 1$. Then $k_n = |\mu(n)|$, so

$$\frac{\zeta(s)}{\zeta(2s)} = \prod_{p \text{ prime}} \left(1 + p^{-s}\right) = \sum_{1 \leq n \text{ square-free}} n^{-s} = \sum_{n=1}^{\infty} \frac{|\mu(n)|}{n^s}.$$

$\square$

The proportion of kth power-free integers

Recall that the proportion of positive integers that are 2^k-free is $(1-2^{-k})$. Moreover, the proportion of positive integers that are the free of the kth powers of relatively prime positive integers $n_1, \ldots, n_m$ is

$$\prod_{j=1}^{m}\left(1-(n_m)^{-k}\right).$$

It follows that the proportion of positive integers that are kth power-free, i.e. are not divisible by any primes to the power of k, is

$$\prod_{p \text{ prime}}\left(1-p^{-k}\right)=\frac{1}{\zeta(k)}.$$

For example, the proportion of square-free integers is

$$\frac{1}{\zeta(2)}=\frac{1}{\pi^2/6}=\frac{6}{\pi^2}.$$

Approximating the proportion of square-free integers

We can use a truncated version of the Euler product expansion of the Riemann zeta function to compute the proportions of square-free integers up to a specific integer.

Let $F(X)$ equal the Euler product expansion of the Riemann zeta function truncated to the Xth term, i.e.

$$F(X) := \prod_{p \text{ prime}}^{X} \left(1 - p^{-2}\right),$$

so $F(X)$ gives the proportion of square-free integers from 1 to X.

Approximating the proportion of square-free integers

In the following table, we calculate the value of $F(x)$ and their percent errors relative to the proportion of all positive integers that are square-free, $\frac{6}{\pi^2} \approx 0.607927102$.

X	$F(X)$	Percent error
100	0.60903373	0.18%
200	0.60838189	0.075%
300	0.608225602	0.049%
400	0.608144905	0.036%
500	0.608093228	0.027%
600	0.608065267	0.023%
700	0.608041737	0.019%
800	0.608026443	0.016%
900	0.608013722	0.014%
1000	0.608004307	0.013%

We see that, as X increases, $F(X)$ quickly becomes an excellent approximation of the proportion of square-free integers.

Approximating the proportion of square-free integers

We can conclude that the density of square-free integers from 1 to 1000 is approximately equal to the density of square-free integers.

Additionally, approximating $\frac{1}{\zeta(s)}$ with the truncated product $F(X)$ is fairly accurate.

The von Mangoldt function

We define the **von Mangoldt function** $\Lambda : \mathbb{Z}^+ \to \mathbb{R}$ as

$$\Lambda(n) = \begin{cases} 0 & \text{if } n = 1, \\ 0 & \text{if } n > 1 \text{ is not a prime power}, \\ \log p & \text{if } n = p^a. \end{cases}$$

The von Mangoldt function

Its first twelve values are
$\Lambda(1) = 0,$
$\Lambda(2) = \log 2,$
$\Lambda(3) = \log 3,$
$\Lambda(4) = \log 2,$
$\Lambda(5) = \log 5,$
$\Lambda(6) = 0,$
$\Lambda(7) = \log 7,$
$\Lambda(8) = \log 2,$
$\Lambda(9) = \log 3,$
$\Lambda(10) = 0,$
$\Lambda(11) = \log 11,$
$\Lambda(12) = 0.$

The natural logarithm in terms of the von Mangoldt function

We can show that

$$\log n = \sum_{1 \leq d | n} \Lambda(d).$$

Proof: Choose $n \in \mathbb{Z}^+$. In the case that $n = 1$, $\log n = 0 = \Lambda(1) = \sum_{1 \leq d | 1} \Lambda(d)$. Hence, we will suppose that $n > 1$. The fundamental theorem of arithmetic gives that n has a unique prime factorization of the form $n = \prod_{i=1}^{m} p_i^{a_i}$, for $m \geq 1$, $p_1, \ldots, p_m$ distinct primes, and $a_1, \ldots, a_m \in \mathbb{Z}^+$.

The natural logarithm in terms of the von Mangoldt function

Let X be the set of prime powers greater than 1 that divide n, i.e. $X = \{p_i^a : 1 \leq i \leq m, 1 \leq a \leq a_i\}$. Λ evaluates to 0 on divisors of n not in X, so

$$\begin{aligned}
\sum_{1 \leq d | n} \Lambda(d) &= \sum_{d \in X} \Lambda(d) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{a_i} \Lambda(p_i^j) \\
&= \sum_{i=1}^{m} a_i \log p_i \\
&= \sum_{i=1}^{m} \log p_i^{a_i}.
\end{aligned}$$

The natural logarithm in terms of the von Mangoldt function

$$\sum_{1 \leq d | n} \Lambda(d) = \sum_{i=1}^{m} \log p_i^{a_i}$$
$$= \log \Pi_{i=1}^{m} p_i^{a_i}$$
$$= \log n.$$

n was chosen arbitrarily, so we are done. □

Applying the Möbius inversion formula

We can apply the Möbius inversion formula to this result to show that, for $n \in \mathbb{Z}^+$,

$$\Lambda(n) = -\sum_{1 \leq d | n} \mu(d) \log d.$$

Proof: Recall that the Möbius inversion formula gives that, for arithmetic functions f and g such that, $\forall n \in \mathbb{Z}^+$

$$g(n) = \sum_{1 \leq d | n} f(d),$$

it holds that

$$f(n) = \sum_{1 \leq d | n} \mu(d) g(n/d).$$

Applying the Möbius inversion formula

Thus,

$$\begin{aligned}\Lambda(n) &= \sum_{1\le d|n} \mu(d)\log\left(\frac{n}{d}\right)\\ &= \sum_{1\le d|n} \mu(d)(\log n - \log d)\\ &= \sum_{1\le d|n} \mu(d)\log n - \sum_{1\le d|n} \mu(d)\log d\\ &= \log n \sum_{1\le d|n} \mu(d) - \sum_{1\le d|n} \mu(d)\log d\\ &= -\sum_{1\le d|n} \mu(d)\log d.\end{aligned}$$

Note that first term in the penultimate line vanishes because, in the case that $n = 1$, $\log n = 0$, and in the case that $n > 1$, $\sum_{1\le d|n} \mu(d) = 0$. $\square$

Relation between $\Lambda(n)$ and $\zeta(s)$

The proof of the prime number theorem relies on the identity

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

Proof: Recall the Euler product expansion:

$$\zeta(s) = \prod_{p \text{ prime}} \frac{1}{1-p^{-s}} = \prod_{p \text{ prime}} \left(1-p^{-s}\right)^{-1}.$$

Relation between $\Lambda(n)$ and $\zeta(s)$

Applying the Mercator series

$$\log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k,$$

we get that

$$\begin{aligned}
\log \zeta(s) &= \log \left(\prod_{p \text{ prime}} (1 - p^{-s})^{-1} \right) \\
&= \sum_{p \text{ prime}} \log \left((1 - p^{-s})^{-1} \right) \\
&= - \sum_{p \text{ prime}} \log(1 - p^{-s}) \\
&= - \sum_{p \text{ prime}} \left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (-p^{-s})^k \right)
\end{aligned}$$

Relation between $\Lambda(n)$ and $\zeta(s)$

$$
\begin{aligned}
\log \zeta(s) &= -\sum_{p \text{ prime}} \left(-\sum_{k=1}^{\infty} \frac{(p^{-s})^k}{k} \right) \\
&= \sum_{p \text{ prime}} \sum_{k=1}^{\infty} \frac{p^{-ks}}{k} \\
&= \sum_{p \text{ prime}} \sum_{k=1}^{\infty} \frac{1}{k} \cdot \frac{1}{p^{ks}} \\
&= \sum_{p^k \text{ prime power}} \frac{1}{k} \cdot \frac{1}{(p^k)^s} \\
&= \sum_{p^k \text{ prime power}} \frac{\log p}{k \log p} \cdot \frac{1}{(p^k)^s}
\end{aligned}
$$

Relation between $\Lambda(n)$ and $\zeta(s)$

$$\log \zeta(s) = \sum_{p^k \text{ prime power}} \frac{\log p}{\log p^k} \cdot \frac{1}{(p^k)^s}$$
$$= \sum_{n=1}^{\infty} \frac{\Lambda(n)}{\log n} \cdot \frac{1}{n^s}.$$

The last equality holds because $\Lambda(n)$ equals $\log p$ if n is an integer power of prime p and equals 0 if n is not a prime power.

Relation between $\Lambda(n)$ and $\zeta(s)$

We can now take the logarithmic derivative of $\zeta(s)$ to get

$$\begin{aligned}
\frac{\zeta'}{\zeta}(s) &= \frac{d}{ds}\log(\zeta(s)) \\
&= \frac{d}{ds}\left(\sum_{n=1}^{\infty}\frac{\Lambda(n)}{\log n}\cdot\frac{1}{n^s}\right) \\
&= \sum_{n=1}^{\infty}\frac{d}{ds}\left(\frac{\Lambda(n)}{\log n}\cdot\frac{1}{n^s}\right) \\
&= \sum_{n=1}^{\infty}\frac{\Lambda(n)}{\log n}\cdot\frac{d}{ds}\left(\frac{1}{n^s}\right)
\end{aligned}$$

Relation between $\Lambda(n)$ and $\zeta(s)$

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

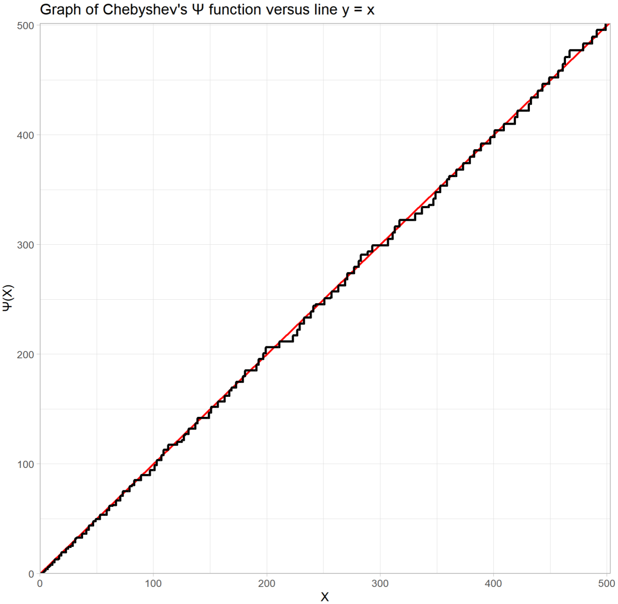
Multiplying both sides of the equation by -1, we are done. $\square$

The Chebyshev function

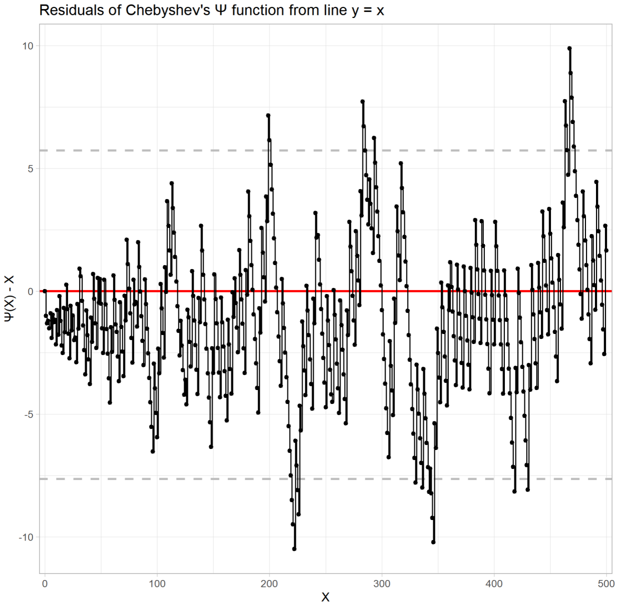
For $X \in \mathbb{R}$, the **Chebyshev function** is defined by

$$\Psi(x) := \sum_{p^k \leq x} \log p.$$

Plotting the Chebyshev function



Plotting the Chebyshev function



Alternative formula for $\Psi(x)$

We can give an exact formula for $\Psi(x)$ in terms of the zeros of the Riemann zeta function:

$$\Psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2}\log(1 - x^{-2}),$$

where ρ iterates over the non-trivial zeros of the Riemann zeta function.

Prime number theorem

Recall the Prime Number Theorem:

$$\text{As } x \to \infty, \text{ we have that } \pi(x) \sim \frac{x}{\log x},$$

where $\pi(x) := |\{p \leq x : p \text{ is prime}\}|$ is the **prime-counting function**.

We can show that the Riemann hypothesis applied to our alternative formula for $\Psi(x)$ implies the prime number theorem.

Prime number theorem

Step 1: We want to show that for large x, $\Psi(x) \sim x$.

As $x \to \infty$, we see that the last two terms of our formula for $\Psi(x)$ are negligible:

$$\lim_{x \to \infty} \frac{1}{2} \log(1 - x^{-2}) = \frac{1}{2} \log 1 = 0$$

and

$$-\frac{\zeta'}{\zeta}(0) = \log 2\pi,$$

which is negligible as $x \to \infty$. Thus, as $x \to \infty$,

$$\Psi(x) \sim x - \sum_{\rho} \frac{x^{\rho}}{\rho}.$$

Prime number theorem

It remains to show that as $x \to \infty$, $\sum_\rho \frac{x^\rho}{\rho} \to 0$.

Assuming the Riemann hypothesis, we have that, for any non-trivial zeros of the Riemann zeta function, ρ, $Re(\rho) = \frac{1}{2}$. Let ρ be a non-trivial zero of the Riemann zeta function and let $\gamma_\rho := Im(\rho)$. Then

$$\begin{aligned}\left|\sum_\rho \frac{x^\rho}{\rho}\right| &= \left|\sum_\rho \frac{x^{\frac{1}{2}+i\gamma_\rho}}{\rho}\right| \\ &= x^{\frac{1}{2}}\left|\sum_\rho \frac{x^{i\gamma_\rho}}{\rho}\right| \\ &\leq \sqrt{x}\sum_\rho \frac{1}{|\rho|},\end{aligned}$$

where the inequality holds via the triangle inequality because $|x^{i\gamma_\rho}| = 1$.

Prime number theorem

Define $N(T) := |\{\rho : \zeta(\rho) = 0,\ 0 < \gamma_\rho < T, \text{ and } T \in \mathbb{R}\}|$, i.e. $N(T)$ denotes the number of non-trivial zeros of the Riemann zeta function whose imaginary part is positive and less than some real number T.

Then $N(T) \sim \left(\frac{T}{2\pi}\right) \log\left(\frac{T}{2\pi}\right)$, so $\sum_\rho \frac{1}{|\rho|}$, which is a sum over all non-trivial zeros of the Riemann zeta function, has approximately $\lim_{T\to\infty} \left(\frac{T}{2\pi}\right) \log\left(\frac{T}{2\pi}\right)$ many terms.

Prime Number Theorem

Recall that the Tth harmonic number is less than or equal to $\log T$, i.e.

$$\sum_{n=1}^{T} \frac{1}{n} \leq \log T.$$

Let (ρ_n) be the sequence of the nontrivial zero of the Riemann zeta function with positive imaginary part in order of increasing imaginary part. Let (b_n) be the sequence defined by $b_n = |\rho_n|$, $\forall n \in \mathbb{Z}^+$. Since the real parts of all elements of (ρ_n) are equal and the imaginary part of each element of (ρ_n) is greater than that of its predecessor, (b_n) is an increasing sequence.

Prime number theorem

For sufficiently large T,

$$\begin{aligned}\sum_{n=1}^{N(T)} \frac{1}{|\rho_n|} &= \sum_{n=1}^{N(T)} \frac{1}{b_n} \\ &\leq \log\left(\left(\frac{T}{2\pi}\right)\log\left(\frac{T}{2\pi}\right)\right) \\ &= \log\left(\frac{T}{2\pi}\right) + \log^2\left(\frac{T}{2\pi}\right) \\ &= \log^2 T + \log T - \log^2(2\pi) - \log(2\pi) \\ &\sim O\left(\log^2(T)\right).\end{aligned}$$

Prime number theorem

It follows that, for sufficiently large T,

$$\begin{aligned}\left|\sum_{\rho}\frac{x^{\rho}}{\rho}\right| &\leq \sqrt{x}\sum_{\rho}\frac{1}{|\rho|}\\ &\leq \sqrt{x}\cdot O(\log^2 T)\\ &\leq O(\sqrt{x}\log^2 T),\end{aligned}$$

which is what we wanted to show. We now have:

$$\text{As } x\to\infty,\ \Psi(x) = x - O(\sqrt{x}\log^2 T)\sim x.$$

Prime number theorem

Step 2: We want to use the von Mangoldt function, $\Lambda(n)$, to approximate $\pi(x)$.

Consider

$$\frac{\Lambda(n)}{\log n} = \begin{cases} 0 & \text{if } n > 1 \text{ is not a prime power,} \\ 1 & \text{if } n \text{ is prime,} \\ \frac{\log p}{\log n} & \text{if } n = p^a, \end{cases}$$

for $n > 1$. As $n \to \infty$, $\frac{\log p}{\log n} \to 0$ because, on average, $n \gg p$.

Prime number theorem

So, we get that

$$\pi(x) \sim \sum_{n=2}^{x} \frac{\Lambda(n)}{\log n}$$

because the sum adds 0 if n is not a prime power, adds $\frac{\log n}{\log n} = 1$ if n is prime, and adds $\frac{\log p}{\log n}$ if n is an integer power of prime p.

Prime number theorem

Abel's summation formula: for a sequence (a_n) of complex numbers, $A(t) := \sum_{n=0}^{t} a_n$, $x, y \in \mathbb{R}$ such that $x < y$, and continuously differentiable function $\phi \in C^1([x, y])$, we have the following:

$$\sum_{n>x}^{y} a_n \phi(n) = A(y)\phi(y) - A(x)\phi(x) - \int_x^y A(u)\phi'(u)du.$$

Prime number theorem

Note that

$$\Psi(x) = \sum_{n=1}^{x} \Lambda(n),$$

by our original definition.

Applying Abel's summation formula with $(\Lambda(n))$ as the sequence of complex numbers and $\phi(n) = \frac{1}{\log n}$ as the function in $C^1([2, x])$, we get:

$$\begin{aligned}
\sum_{n=2}^{x} \frac{\Lambda(n)}{\log n} &= \frac{\Psi(x)}{\log x} - \frac{\Psi(2)}{\log 2} + \int_2^x \frac{\Psi(u)}{u \log^2 u} du \\
&= \frac{\Psi(x)}{\log x} - \frac{\log 2}{\log 2} + \int_2^x \frac{\Psi(u)}{u \log^2(u)} du \\
&\sim \frac{\Psi(x)}{\log x} + \int_2^x \frac{\Psi(u)}{u \log^2 u} du.
\end{aligned}$$

Prime number theorem

$$
\begin{aligned}
\sum_{n=2}^{x} \frac{\Lambda(n)}{\log n} &\sim \frac{x + O(\sqrt{x}\log^2 x)}{\log x} + \int_2^x \frac{u + O(\sqrt{u}\log^2 u)}{u \log^2 u} du \\
&= \frac{x}{\log x} + \frac{O(\sqrt{x}\log^2 x)}{\log x} + \int_2^x \left(\frac{u + O(\sqrt{u}\log^2 u)}{u} \cdot \frac{1}{\log^2 u} \right) du \\
&= \frac{x}{\log x} + O\left(\frac{\sqrt{x}\log^2 x}{\log x} \right) + O\left(\int_2^x \frac{1}{\log^2 u} du \right) \\
&= \frac{x}{\log x} + O\left(\sqrt{x}\log x \right) + O\left(\int_2^x \frac{1}{\log^2 u} du \right) \\
&\sim \frac{x}{\log x}
\end{aligned}
$$

Prime number theorem

We conclude that as $x \to \infty$

$$\pi(x) \sim \sum_{n=2}^{x} \frac{\Lambda(n)}{\log n} \sim \frac{x}{\log x},$$

which is precisely what we set out to prove. $\square$