Almost Ready for Prime Time: Elementary Results Towards the PNT

Carlo Francisco E. Adajar

Ross Mathematics Program

Summer 2024

Prime numbers

Definition

A prime number is an element of $\mathbb{Z}^+$ that has exactly two distinct divisors in $\mathbb{Z}^+$, namely 1 and itself.

Example

The first few prime numbers are $2, 3, 5, 7, 11, 13, \ldots$ (OEIS A000040)

How many are there?

You might notice that if you list a few more, they seem to be getting harder and harder to find. Think about why this should be the case.

Despite that, there are indeed infinitely many of them!

Theorem

There are infinitely many prime numbers.

The prime-counting function $\pi(x)$

For $x > 0$, let us denote by $\pi(x)$ the number of primes in the interval $[1, x]$.

x	$\pi(x)$	x	$\pi(x)$
10	4	10^6	78498
10^2	25	10^7	664579
10^3	168	10^8	5761455
10^4	1229	10^9	50847534
10^5	9592	10^{10}	455052511

Table: A list of some values of $\pi(x)$.

The prime-counting function $\pi(x)$

Here's a plot of $\pi(x)$ for $x \leq 10^8$:

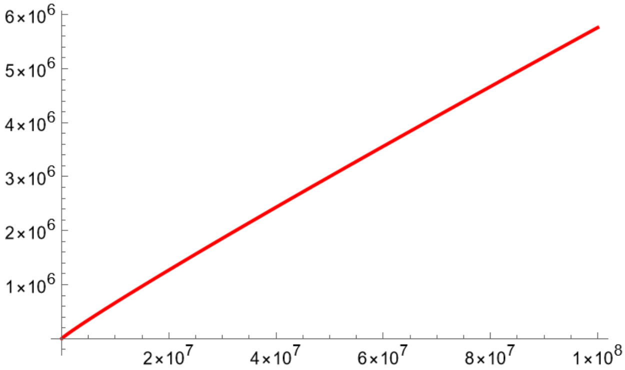


Figure: A plot of $\pi(x)$.

Legendre's Approximation of $\pi(x)$

In 1797 Legendre conjectured, based on tables by Vega, Chernac, and Burkhardt, that

$$\pi(x) \approx \frac{x}{A\log(x) + B}$$

for some constants A, B. Here and elsewhere, log will be used to denote the natural (base e) logarithm.

Legendre's Approximation of $\pi(x)$

In 1808 he chose the values $A = 1, B = -1.08366$. This looks pretty good for our range:

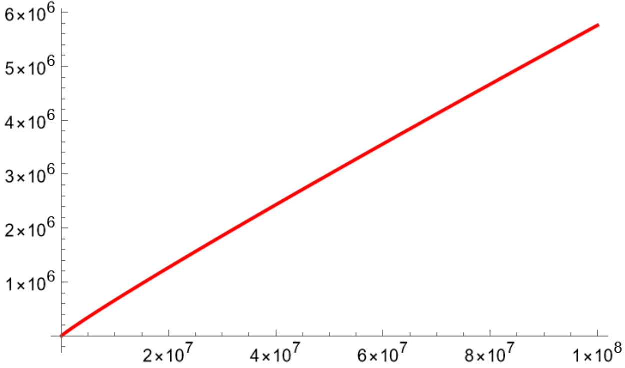


Figure: A plot of Legendre's approximation (in green) compared to $\pi(x)$ (in red).

Legendre's Approximation of $\pi(x)$

In 1808 he chose the values $A = 1, B \approx -1.08366$. This looks pretty good for our range:

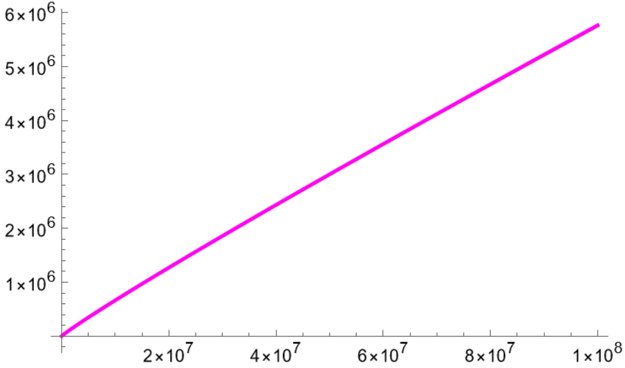


Figure: A plot of Legendre's approximation (in magenta) compared to $\pi(x)$ (in red).

Legendre's Approximation of $\pi(x)$

If (understandably so) you were not convinced by these poor graphs, perhaps a table of calculations will be more to your liking.

x	$\pi(x)$	$x/(\log(x) - 1.08366)$	Rel. Error
10^3	168	171	$2.2 \cdot 10^{-2}$
10^4	1229	1231	$1.23 \cdot 10^{-3}$
10^5	9592	9588	$-3.75 \cdot 10^{-4}$
10^6	78498	78543	$5.76 \cdot 10^{-4}$
10^7	664579	665140	$8.44 \cdot 10^{-4}$
10^8	5761455	5768004	$1.14 \cdot 10^{-3}$
10^{10}	455052511	455743004	$1.52 \cdot 10^{-3}$

Table: Table showcasing the accuracy of Legendre's approximation.

Gauss' Approximation of $\pi(x)$

Gauss, in an 1849 letter to the astronomer Encke, had given the approximation

$$\pi(x) \approx \int_0^x \frac{\mathrm{d}t}{\log t} =: \mathrm{li}(x).$$

Gauss also suggested that Legendre's constant 1.0833 was not quite right for larger x (more on this later).

Gauss' Approximation of $\pi(x)$

Gauss' approximation turns out to be pretty good, too.

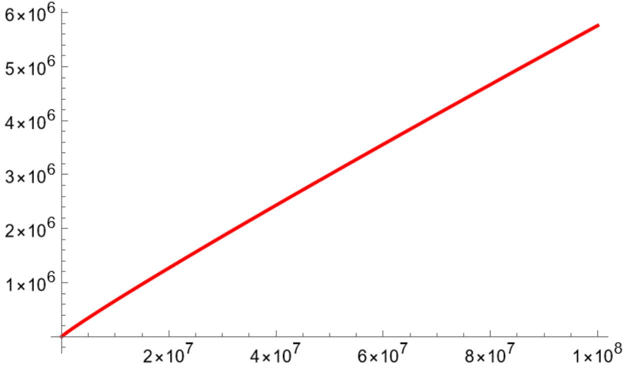


Figure: A plot of Gauss' approximation (in orange) compared to $\pi(x)$ (in red).

Gauss' Approximation of $\pi(x)$

Gauss' approximation turns out to be pretty good, too.

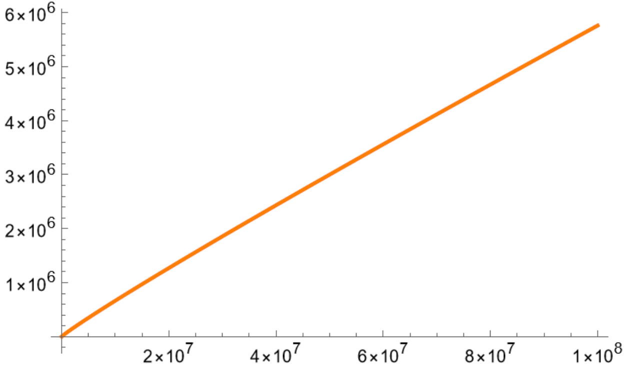


Figure: A plot of Gauss' approximation (in orange) compared to $\pi(x)$ (in red).

Gauss's Approximation of $\pi(x)$

Here's a table similar to earlier, for $\text{li}(x)$:

x	$\pi(x)$	$\text{li}(x)$	Rel. Error
10^3	168	178	$5.72 \cdot 10^{-2}$
10^4	1229	1246	$1.39 \cdot 10^{-2}$
10^5	9592	9630	$3.94 \cdot 10^{-3}$
10^6	78498	78628	$1.65 \cdot 10^{-3}$
10^7	664579	664918	$5.11 \cdot 10^{-4}$
10^8	5761455	5762209	$1.31 \cdot 10^{-4}$
10^{10}	455052511	455055615	$6.82 \cdot 10^{-6}$

Table: Table showcasing the accuracy of Gauss' approximation.

The Prime Number Theorem (PNT)

The preceding tables and graphs suggest that the relative errors of Legendre's and Gauss's approximations to $\pi(x)$ go to 0 as x goes to ∞.

Equivalently,

$$\pi(x) \sim \frac{x}{\log(x)} \sim \operatorname{li}(x)$$

where we write $f(x) \sim g(x)$ to mean

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1.$$

The Prime Number Theorem (PNT)

This is the celebrated Prime Number Theorem, first proved independently by Hadamard and de la Vallée Poussin in 1896.

Theorem (Prime Number Theorem)

For $x \geq 0$, let $\pi(x)$ denote the number of primes less than or equal to x. Then

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

Some equivalent forms and consequences of the PNT

Denote the nth prime number by p_n. Then the PNT is equivalent to the statement that $p_n \sim n \log n$.

The PNT also implies the following corollary:

Corollary (Primes in short intervals)

Let $\epsilon > 0$. Then there exists x_0 depending on ϵ such that for all $x \geq x_0$, there exists a prime in the interval $[x, (1+\epsilon)x]$.

The case $\epsilon = 1, x_0 = 1$ is known as Bertrand's postulate, after Bertrand, who conjectured it in 1845.

Theorem (Bertrand's postulate)

For all $x \geq 1$, there exists a prime in the interval $[x, 2x]$.

Some equivalent forms and consequences of the PNT

A more explicit version of the prime number theorem, with an asymptotic bound on the error term, is known.

Theorem (PNT with error term)

We have, for some absolute constant $c > 0$,

$$\pi(x) = \mathrm{li}(x) + O\left(xe^{-c\sqrt{\log(x)}}\right)$$

as $x \to \infty$.

We write $f(x) = O(g(x))$ to mean that for sufficiently large x, $|f(x)| \leq Mg(x)$ for some absolute constant M.

Some equivalent forms and consequences of the PNT

A consequence of the version of the PNT with error term is as follows:

Corollary (Legendre's constant)

We have

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

This tells us that the "correct value" for B in Legendre's estimate should be -1.

The PNT is hard

Hadamard's and de la Vallée Poussin's proofs of the Prime Number Theorem were involved and used difficult results from complex analysis.

They used ideas from Riemann's seminal 1859 paper, which drew connections between the zeros of the Riemann zeta function ζ and the distribution of the prime numbers.

The Riemann Hypothesis

Conjecture (Riemann Hypothesis)

Suppose $\zeta(\sigma + it) = 0$ with $\sigma \in \mathbf{R}$, $t \in \mathbf{R}^{\times}$. Then $\sigma = \frac{1}{2}$.

Corollary

If the Riemann Hypothesis is true, then

$$\pi(x) = \mathrm{li}(x) + O(\sqrt{x} \log^2(x)).$$

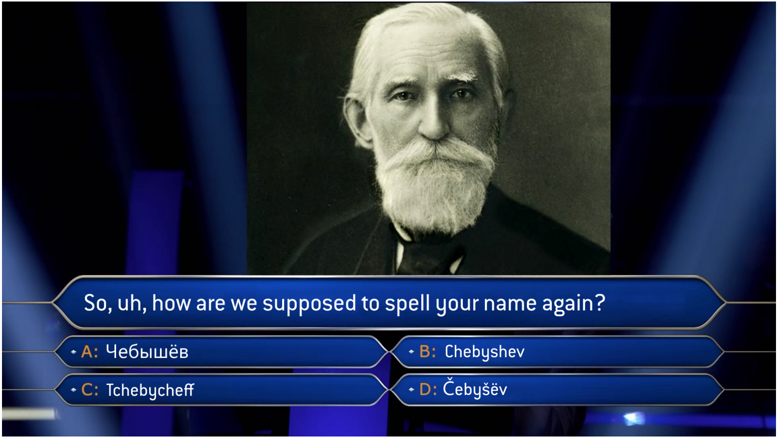
The PNT is hard

In 1948, Erdős and Selberg independently came up with proofs of the prime number theorem that did not require complex analysis, but were still quite technical.

Since then there have been other "elementary" proofs of the PNT, but none of them are "easy."

The PNT is hard

Enter Chebyshev



Enter Chebyshev

Chebyshev never got to the Prime Number Theorem, but he did get several results adjacent to it, often using much more rudimentary methods.

In particular, while Chebyshev did not quite get the asymptotic statement, he showed that $\pi(x)$ was always bounded between two multiples of $x/\log x$. In other words, $x/\log(x)$ is the "right order of growth" for $\pi(x)$.

Enter Chebyshev

Theorem (Weak Prime Number Theorem)

- *There exist constants* $0 < c_1 < c_2$, $x_0 > 0$ *such that*

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

for all $x \geq x_0$.

- *If*

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log x}$$

exists, then it must be equal to 1.

Chebyshev got constants $c_1 \approx 0.9212$, $c_2 \approx 1.11056$, and his proof was a bit technical; however, with a bit less effort we can establish the theorem albeit with worse constants.

Other results of Chebyshev

Chebyshev also gave a proof of Bertrand's postulate, which can be deduced from his theorem.

In addition, Chebyshev proved that if Legendre had the "right" form $x/(\log x + B)$ for the estimate of $\pi(x)$ (i.e. the limit in the corollary exists), then we in fact would have $B = -1$. His proof was quite difficult and used complex-analytic tools. Pintz gave a simpler proof in 1980.

Basic properties of $x/\log x$

To prove Chebyshev's (first) theorem, we start by establishing some elementary properties of the function $x/\log x$.

Lemma

Let $f : (1, \infty) \to \mathbb{R}$ be given by $f(x) = x/\log x$. Then we have

1. $f(x)$ *is increasing for* $x > e$;
2. $f(x-2) > \dfrac{1}{2} f(x)$ *for* $x \geq 4$;
3. $f\left(\dfrac{x+2}{2}\right) < \dfrac{15}{16} f(x)$ *for* $x \geq 8$.

Basic properties of $x/\log x$

Proof.

1. The proof of (1) is a simple derivative calculation.
2. To prove (2), we use $x - 2 \geq x/2$ for $x \geq 4$.
3. To prove (3), we use that $x/2 \leq x^{2/3}$ and $x + 2 \leq 5x/4$ for $x \geq 8$.

$\square$

A proof of (a version of) Chebyshev's theorem

We are now ready to prove a version of Chebyshev's theorem.

Theorem (Weak Chebyshev's Theorem)

For $x \geq 8$, we have

$$\frac{\log(2)}{4} < \frac{\pi(x)}{x/\log x} < 30\log(2).$$

The constants are worse than in Chebyshev's original proof, but this proof is a bit less technical.

We prove the lower bound; the upper bound can be done similarly.

A proof of (a version of) Chebyshev's theorem

Proof.

We begin by considering the binomial coefficients

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} = \frac{(2n)(2n-1)\dots(n+1)}{n(n-1)\dots 1}.$$

Note that any prime $p \in (n, 2n]$ must divide the numerator of $\binom{2n}{n}$, but not the denominator. Letting

$$P_n := \prod_{n<p\leq 2n} p$$

be the product of the primes $p \in (n, 2n]$, we then have that $P_n \mid \binom{2n}{n}$.

A proof of (a version of) Chebyshev's theorem

Proof. (cont'd.)

Now, we have that each prime $p \in (n, 2n]$ is greater than n, and moreover, there are exactly $\pi(2n) - \pi(n)$ of them. Thus

$$n^{\pi(2n)-\pi(n)} < P_n \leq \binom{2n}{n}.$$

Then, we consider for each prime $p \leq 2n$ the unique integer r_p such that

$$p^{r_p} \leq 2n < p^{r_p+1}$$

(in other words, $r_p := \lfloor \log(2n)/\log(p) \rfloor$).

A proof of (a version of) Chebyshev's theorem

Proof. (cont'd.)

For each prime p, we know that the exponent $v_p := v_p(\binom{2n}{n})$ of p in the prime factorization of $\binom{2n}{n}$ is given by

$$\begin{aligned} v_p &= v_p((2n)!) - 2v_p(n!) \\ &= \sum_{j=1}^{r_p} \left(\left\lfloor \frac{2n}{p^j} \right\rfloor - 2 \left\lfloor \frac{n}{p^j} \right\rfloor \right). \end{aligned}$$

Since $0 \le \lfloor 2\alpha \rfloor - 2\lfloor \alpha \rfloor \le 1$ for all $\alpha \in \mathbf{R}$ (check!) we then have $v_p \le r_p$, and so

$$Q_n := \prod_{p \le 2n} p^{r_p}$$

is divisible by $\binom{2n}{n}$.

A proof of (a version of) Chebyshev's theorem

Proof. (cont'd.)

Now, note that for each p, we have $p^{r_p} \leq 2n$, and so $Q_n \leq (2n)^{\pi(2n)}$. Putting it together, we have

$$\binom{2n}{n} \leq Q_n \leq (2n)^{\pi(2n)}.$$

We now estimate the size of $\binom{2n}{n}$. Note that by the binomial theorem,

$$2^{2n} = (1+1)^{2n} > \binom{2n}{n}$$

for all $n \geq 1$.

A proof of (a version of) Chebyshev's theorem

Proof. (cont'd.)

We also have

$$\binom{2n}{n} = \prod_{k=0}^{n-1} \frac{2n-k}{n-k} \geq 2^n.$$

Thus, we get

$$2^n \leq (2n)^{\pi(2n)}$$

for all $n \geq 1$. Taking logarithms yields

$$n \log(2) \leq \pi(2n) \log(2n).$$

A proof of (a version of) Chebyshev's theorem

Proof. (cont'd.)

Writing once again $f(x) = x/\log(x)$, we get that for $x \geq 5$,

$$\begin{aligned}
\pi(x) &\geq \pi\left(2\left\lfloor \frac{x}{2} \right\rfloor\right) \\
&\geq \frac{\left\lfloor \frac{x}{2} \right\rfloor \log(2)}{\log\left(2\left\lfloor \frac{x}{2} \right\rfloor\right)} \\
&= \frac{\log(2)}{2} f\left(2\left\lfloor \frac{x}{2} \right\rfloor\right) \\
&> \frac{\log(2)}{2} f(x-2) \\
&\geq \frac{\log(2)}{4} f(x).
\end{aligned}$$

This gives us the claimed lower bound.

A proof of (a version of) Chebyshev's theorem

Proof. (cont'd.)

To get the upper bound, we start with the inequality

$$n^{\pi(2n)-\pi(n)} < 2^{2n}$$

and take logarithms and rearrange to get.

$$\pi(2n) < (2\log 2)f(n) + \pi(n).$$

We then use the Well-Ordering Principle to show that for all $n > 1$, we have

$$\pi(2n) < 32\log(2)f(n).$$

We finish off by using the lemma. $\square$

Some lemmata

A similar analysis of the central binomial coefficients gets us the proof of Bertrand's postulate.

Lemma

For all $n \geq 1$, we have

$$\binom{2n}{n} \geq \frac{4^n}{2n}.$$

Proof.

Write

$$(1+1)^{2n} = 2 + \sum_{k=1}^{2n-1} \binom{2n}{k} \leq 2n\binom{2n}{n}.$$

$\square$

Some lemmata

We also observe that $\binom{2n}{n}$ has no odd prime factors between $2n/3$ and n.

Lemma

Let n be a positive integer. Then for any prime $p > 2$ with $p \in (2n/3, n]$, we have $p \nmid \binom{2n}{n}$.

Proof.

Exercise! $\square$

Some lemmata

For this final lemma, we estimate the *primorials*

$$n\# = \prod_{p \leq n} p.$$

(The product is taken over prime p.)

Lemma

For all $n \geq 1$, $n\# < 4^n$.

Proof.

First, verify it for $n \leq 2$. Then, it suffices to verify it for n prime. (Why?) Since we only care about the case $n > 2$, we can write $n = 2m + 1$ odd.

Some lemmata

Proof. (cont'd.)

Consider the set $S \subseteq \mathbf{Z}^+$ of n for which $n\# \geq 4^n$. Suppose S is non-empty. By the Well-Ordering Principle, S contains a minimal element n_0, and in particular, we can write $n_0 = 2m_0 + 1$ for some $m_0 \in \mathbf{Z}^+$.

Now, we note that every prime in $(m_0 + 1, 2m_0 + 1]$ divides $\binom{2m_0+1}{m_0}$ and so $(2m_0 + 1)\#/(m_0 + 1)\sharp$ divides $\binom{2m_0+1}{m_0}$.

Some lemmata

Proof. (cont'd.)

We then observe that

$$\binom{2m_0+1}{m_0} < \sum_{j=0}^{m_0} \binom{2m_0+1}{j}$$
$$= \frac{1}{2} \cdot 2^{2m_0+1} = 4^{m_0}.$$

Thus $(2m_0+1)\#/(m_0+1)\sharp < 4^{m_0}$. Thus, we must have $(m_0+1)\# > 4^{m_0+1}$, and so in fact $m_0+1 \in S$. But $m_0+1 < 2m_0+1 = n_0$, contradicting the minimality of n_0 in S.

Thus S is empty, and the inequality holds for all $n \geq 1$. $\square$

A proof of Bertrand's postulate

We are now ready to prove Bertrand's postulate.

Proof.

Suppose for the sake of contradiction that there exists $n \in \mathbf{Z}^+$ for which there are no primes in $[n, 2n]$. We can easily check that $n > 4$. Consider the prime factors of $\binom{2n}{n}$. Clearly none of them exceed $2n$, none of them are in $[n, 2n]$ (because we assumed that there are no primes in this interval), and in fact by our lemma, none of them are in $(2n/3, n]$. That is, all primes dividing $\binom{2n}{n}$ are at most $2n/3$.

A proof of Bertrand's postulate

Proof. (cont'd.)

Since $n > 4$, we have $\sqrt{2n} < 2n/3$, and so we can write $\binom{2n}{n} = P_1 P_2$ where

$$P_1 = \prod_{p \leq \sqrt{2n}} p^{v_p}$$

$$P_2 = \prod_{\sqrt{2n} < p \leq 2n/3} p^{v_p}$$

Recall that from our proof of Chebyshev's theorem, $p^{v_p} \leq 2n$ for all prime powers $p^{v_p} \mid \binom{2n}{n}$. Thus $P_1 \leq (2n)^{\sqrt{2n}}$.

A proof of Bertrand's postulate

Proof. (cont'd.)

For the same reason, for P_2, we have that $v_p = 1$ for all $p \mid P_2$ (otherwise $p^{v_p} > 2n$). Thus $P_2 < (2n/3)\# < 4^{2n/3}$.
Finally, from our first lemma, we have that

$$\frac{4^n}{2n} \leq \binom{2n}{n} = P_1 P_2 \leq (2n)^{\sqrt{2n}} 4^{2n/3}.$$

Taking logarithms and simplifying yields

$$(\sqrt{2n}+1)\log(2n) \geq \frac{2\log(2)}{3} n.$$

A proof of Bertrand's postulate

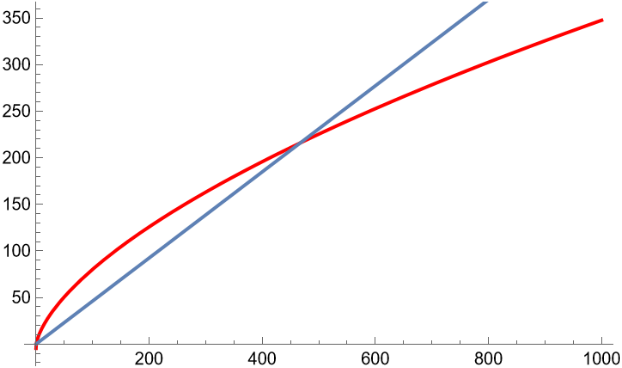


Figure: A plot of the two sides of the inequality $(\sqrt{2n}+1)\log(2n) \geq (2\log(2)/3)n$. LHS in red, RHS in blue.

A proof of Bertrand's postulate

Proof. (cont'd.)

Verify that the left-hand side is concave down on $(0, \infty)$ and the inequality is thus true on some closed bounded interval; specifically, it is true for all $n \leq 467$, but false for $n > 468$. Thus $n \leq 467$.

However, $n \leq 467$ cannot be a counterexample, because of the sequence of primes

$$2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 631$$

with each less than twice the one preceding it (OEIS A006992). Thus there are in fact no counterexamples, and Bertrand's postulate is true for all $n \geq 1$. $\square$

References I

G. E. Andrews, *Number theory*, W. B. Saunders Co., Philadelphia, Pa.-London-Toronto, Ont., 1971. MR 309838

Math Circle at UW, *Proof of Bertrand's Postulate*.

J. Bertrand, *Mémoire sur le nombre de valeurs que peut prendre une fonction quand on y permute les lettres qu'elle renferme*, J. l'École Roy. Polytech. **18** (1845), 123–140 (fr).

C. de la Vallée Poussin, *Sur la fonction $\zeta(s)$ de Riemann et le nombre des nombres premiers inférieurs à une limite donnée*, Mémoires de l'Académie royale de Belgique **59** (1899), no. 1, 1–74 (fr).

References II

P. Erdős, *Beweis eines satzes von Tschebyschef*, Acta Litt. Sci. Szeged **5** (1932), 194–198 (de).

______, *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem*, Proc. Nat. Acad. Sci. U.S.A. **35** (1949), 374–384. MR 29411

D. Goldfeld, *The elementary proof of the prime number theorem: an historical perspective*, Number theory (New York, 2003), Springer, New York, 2004, pp. 179–192. MR 2044518

L. J. Goldstein, *A history of the prime number theorem*, Amer. Math. Monthly **80** (1973), 599–615. MR 313171

References III

J. Hadamard, *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques*, Bulletin de la Société Mathématique de France **24** (1896), 199–220 (fr).

W. Narkiewicz, *The development of prime number theory*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000, From Euclid to Hardy and Littlewood. MR 1756780

OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, 2024, Published electronically at `http://oeis.org`.

A. Selberg, *An elementary proof of the prime-number theorem*, Ann. of Math. (2) **50** (1949), 305–313. MR 29410

References IV

P. L. Tchebichef, *Mémoire sur les nombres premiers.*, Journal de Mathématiques Pures et Appliquées (1852), 366–390 (fr).