Erdős's proof of Bertrand's postulate

David Galvin*

May 1, 2015

Abstract

In 1845 Bertrand postulated that there is always a prime between n and $2n$, and he verified this for $n < 3 \times 10^6$. Tchebychev gave an analytic proof of the postulate in 1850. In 1932, in his first paper, Erdős gave a beautiful elementary proof using nothing more than a few easily verified facts about the middle binomial coefficient. We describe Erdős's proof and make a few additional comments, including a discussion of how the two main lemmas used in the proof very quickly give an approximate prime number theorem. We also describe a result of Greenfield and Greenfield that links Bertrand's postulate to the statement that $\{1, \dots, 2n\}$ can always be decomposed into n pairs such that the sum of each pair is a prime.

1 Introduction

Write $\pi(x)$ for the number of primes less than or equal to x. The Prime Number Theorem (PNT), first proved by Hadamard [5] and de la Vallée-Poussin [8] in 1896, is the statement that

$$\pi(x) \sim \frac{x}{\ln x} \quad \text{as } x \to \infty \tag{1}$$

(that is, $\lim_{n\to\infty}(\pi(x)\ln x)/x = 1$; here ln indicates natural logarithm).

A consequence of the PNT is that for all $\epsilon > 0$ there exists $n(\epsilon) > 0$ such that for all $n \geq n(\epsilon)$, there is always a prime in the interval $(n, (1+\epsilon)n]$, that is, there is a prime p satisfying

$$n < p \leq (1+\epsilon)n. \tag{2}$$

Indeed, by (1) we have

$$\pi((1+\epsilon)n) - \pi(n) \sim \frac{(1+\epsilon)n}{\ln(1+\epsilon)n} - \frac{n}{\ln n} \to \infty \quad \text{as } n \to \infty.$$

Using a more refined version of the PNT with an error estimate, we may prove the following theorem.

Theorem 1.1 *For all $n \geq 1$ there is a prime p such that $n < p \leq 2n$.*

*Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556; dgalvin1@nd.edu.

This is Bertrand's postulate, conjectured in the 1845, verified by Bertrand for all $N < 3\times 10^6$, and first proved by Tchebychev in 1850. (See [6, p. 25] for a discussion of the original references).

In his first paper Erdős [3] gave a beautiful elementary proof of Bertrand's postulate which uses nothing more than some easily verified facts about the middle binomial coefficient $\binom{2n}{n}$. We describe this proof in Section 2 and present some comments, conjectures and consequences in Section 3. One consequence is the following lovely theorem of Greenfield and Greenfield [4].

Theorem 1.2 *For $n \geq 1$ the set $\{1, \ldots, 2n\}$ can be partitioned into pairs*

$$\{a_1, b_1\}, \ldots, \{a_n, b_n\}$$

such that for each $1 \leq i \leq n$, $a_i + b_i$ is a prime.

This bears an amusing (completely superficial) resemblance to the famous Goldbach conjecture. Goldbach asserts that every even number at least 4 can be expressed as the sum of two primes; Greenfield and Greenfield say that the set of positive integers up to every even number can be partitioned into pairs that add to primes.

Another consequence of Erdős' proof is an elementary proof of an approximate version of (1).

Theorem 1.3 *There are constants $c, C > 0$ such that for all $x \geq 2$*

$$c\frac{x}{\ln x} \leq \pi(x) \leq C\frac{x}{\ln x}.$$

In particular, for all $\varepsilon > 0$ there is $n(\varepsilon) > 0$ such that for $n \geq n(\varepsilon)$

$$\frac{(1-\varepsilon)x}{\log_2 x} \leq \pi(x) \leq \frac{(4+\varepsilon)x}{\log_2 x}. \tag{3}$$

2 Erdős's proof

Our presentation draws on [1, Chapter 2]. We consider the middle binomial coefficient $\binom{2n}{n} = (2n)!/(n!)^2$. An easy lower bound is

$$\binom{2n}{n} \geq \frac{4^n}{2n+1}. \tag{4}$$

Indeed, $\binom{2n}{n}$ is the largest term in the $(2n+1)$-term sum $\sum_{i=0}^{2n}\binom{2n}{i} = (1+1)^{2n} = 4^n$. Erdős's proof proceeds by showing that if there is no prime p with $n < p \leq 2n$ then we can put an upper bound on $\binom{2n}{n}$ that is *smaller* than $4^n/(2n+1)$ unless n is small. This verifies Bertrand's postulate for all sufficiently large n, and we deal with small n by hand.

For a prime p and an integer n we define $o_p(n)$ to be the largest exponent of p that divides n. Note that $o_p(ab) = o_p(a) + o_p(b)$ and $o_p(a/b) = o_p(a) - o_p(b)$. Note also that

$$o_p(n!) = \sum_{i\geq 1}\left[\frac{n}{p^i}\right];$$

indeed, $[n/p]$ is the number of multiples of p that do not exceed n, each of which contributes 1 to $o_p(n!)$, $[n/p^2]$ is the number of multiples of p^2 that do not exceed n, each of which contributes an additional 1 to $o_p(n!)$, an so on.

The heart of the whole proof is the following simple observation.

$$\text{If } n \geq 3 \text{ and } \tfrac{2}{3}n < p \leq n \text{ then } o_p\left(\tbinom{2n}{n}\right) = 0 \text{ (i.e., } p \not| \tbinom{2n}{n}\text{)}. \tag{5}$$

Indeed, for such a p

$$o_p\left(\binom{2n}{n}\right) = o_p((2n)!) - 2o_p(n!) = 2 - 2.1 = 0.$$

So if $n \geq 3$ is such that there is no prime p with $n < p \leq 2n$, then all of the prime factors of $\binom{2n}{n}$ lie between 2 and $2n/3$. (It is easy to see that $\binom{2n}{n}$ can have no prime factor greater than $2n$, since such a factor would have to divide $(2n)!$ and so divide some $k \leq 2n$.) We will show that each of these factors appears only to a small exponent, forcing $\binom{2n}{n}$ to be small. The following is the claim we need in this direction.

Claim 2.1 *If* $p|\binom{2n}{n}$ *then*

$$p^{o_p\left(\binom{2n}{n}\right)} \leq 2n.$$

Proof: Let $r(p)$ be such that $p^{r(p)} \leq 2n < p^{r(p)+1}$. We have

$$\begin{aligned} o_p\left(\binom{2n}{n}\right) &= o_p((2n)!) - 2o_p(n!) \\ &= \sum_{i=1}^{r(p)} \left[\frac{2n}{p^i}\right] - 2\sum_{i=1}^{r(p)} \left[\frac{n}{p^i}\right] \\ &= \sum_{i=1}^{r(p)} \left(\left[\frac{2n}{p^i}\right] - 2\left[\frac{n}{p^i}\right]\right) \\ &\leq r(p), \end{aligned} \tag{6}$$

and so

$$p^{o_p\left(\binom{2n}{n}\right)} \leq p^{r(p)} \leq 2n.$$

In (6) we use the easily verified fact that for all real x, $0 \leq [2x] - 2[x] \leq 1$; indeed, this quantity is 0 if $x - [x] < 1/2$ and is 1 otherwise. □

Corollary 2.2 *The number of prime numbers dividing* $\binom{2n}{n}$ *is at least* $\log_2 \binom{2n}{n} / \log_2(2n)$.

Proof: Let $p_1, \ldots, p_\ell$ be the distinct primes dividing $\binom{2n}{n}$. We have

$$\begin{aligned} \binom{2n}{n} &= \prod_{i=1}^{\ell} p_i^{o_{p_i}\left(\binom{2n}{n}\right)} \\ &\leq (2n)^\ell \end{aligned} \tag{7}$$

with (7) using Claim 2.1. Taking logarithms, the result follows. □

Before writing down the estimates that upper bound $\binom{2n}{n}$, we need one more simple result.

Claim 2.3 *For all $n \geq 2$, $\prod_{p \leq n} p \leq 4^n$, where the product is over primes.*

Proof: We proceed by induction on n. For small values of n, the claim is easily verified. For larger even n, we have

$$\prod_{p \leq n} p = \prod_{p \leq n-1} p \leq 4^{n-1} \leq 4^n,$$

the equality following from the fact that n is even and so not a prime, and the first inequality following from the inductive hypothesis. For larger odd n, say $n = 2m+1$, we have

$$\begin{aligned}
\prod_{p \leq n} p &= \prod_{p \leq m+1} p \prod_{m+2 \leq p \leq 2m+1} p \\
&\leq 4^{m+1} \binom{2m+1}{m} && (8) \\
&\leq 4^{m+1} 2^{2m} && (9) \\
&= 4^{2m+1} = 4^n.
\end{aligned}$$

In (8) we use the induction hypothesis to bound $\prod_{p \leq m+1} p$ and we bound $\prod_{m+2 \leq p \leq 2m+1} p$ by observing that all primes between $m+2$ and $2m+1$ divide $\binom{2m+1}{m}$. In (9) we bound $\binom{2m+1}{m} \leq 2^{2m}$ by noting that $\sum_{i=0}^{2m+1} \binom{2m+1}{i} = 2^{2m+1}$ and $\binom{2m+1}{m} = \binom{2m+1}{m+1}$ and so the contribution to the sum from $\binom{2m+1}{m}$ is at most 2^{2m}. $\square$

We are now ready to prove Bertrand's postulate. Let $n \geq 3$ be such that there is no prime p with $n < p \leq 2n$. Then we have

$$\begin{aligned}
\binom{2n}{n} &\leq (2n)^{\sqrt{2n}} \prod_{\sqrt{2n} < p \leq 2n/3} p && (10) \\
&\leq (2n)^{\sqrt{2n}} \prod_{p \leq 2n/3} p \\
&\leq (2n)^{\sqrt{2n}} 4^{2n/3}. && (11)
\end{aligned}$$

The main point is (10). We have first used the simple fact that $\binom{2n}{n}$ has at most $\sqrt{2n}$ prime factors that do not exceed $\sqrt{2n}$, and, by Claim 2.1, each of these prime factors contributes at most $2n$ to $\binom{2n}{n}$; this accounts for the factor $(2n)^{\sqrt{2n}}$. Next, we have used that by hypothesis and by (5) all of the prime factors p of $\binom{2n}{n}$ satisfy $p \leq 2n/3$, and the fact that each such p with $p > \sqrt{2n}$ appears in $\binom{2n}{n}$ with exponent 1 (this is again by Claim 2.1); these two observations together account for the factor $\prod_{\sqrt{2n} < p \leq 2n/3} p$. In (11) we have used Claim 2.3.

Combining (11) with (4) we obtain the inequality

$$\frac{4^n}{2n+1} \leq (2n)^{\sqrt{2n}} 4^{2n/3}. \tag{12}$$

This inequality can hold only for small values of n. Indeed, for any $\epsilon > 0$ the left-hand side of (12) grows faster than $(4-\epsilon)^n$ whereas the right-hand side grows more slowly than $(4^{2/3}+\epsilon)^n$. We may check that in fact (12) fails for all $n \geq 468$ (Mathematica calculation), verifying Bertrand's postulate for all n in this range. To verify Bertrand's postulate for all $n < 468$, it suffices to check that

$$2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 631. \tag{13}$$

is a sequence of primes, each term of which is less than twice the term preceding it; it follows that every interval $\{n+1, \ldots, 2n\}$ with $n < 468$ contains one of these 11 primes. This concludes the proof of Theorem 1.1.

3 Comments, conjectures and consequences

A stronger result than (2) is known (due to Lou and Yao [7]): for all $\epsilon > 0$ there exists $n(\epsilon) > 0$ such that for all $n \geq n(\epsilon)$, there is always a prime in the interval $(n, n + n^{\frac{1}{2}+\frac{1}{22}+\epsilon}]$. The Riemann hypothesis would imply that we could shorten this interval to $(n, n+n^{\frac{1}{2}+\epsilon}]$, and a very strong conjecture of Cramér [2] would imply we could shorten it to $(n, n+(1+\epsilon)\ln^2 n]$.

Here is a very lovely open question much in the spirit of Bertrand's postulate.

Question 3.1 *Is it true that for all $n \geq 2$, there is always a prime p with $n^2 < p < (n+1)^2$?*

As mentioned in the introduction, a consequence of Bertrand's postulate is the appealing Theorem 1.2. We give the proof here.

Proof of Theorem 1.2: We proceed by induction on n. For $n = 1$ the result is trivial. For $n > 1$, let p be a prime satisfying $2n < p \leq 4n$. Since $4n$ is not prime we have $p = 2n+m$ for $1 \leq m < 2n$. Pair $2n$ with m, $2n-1$ with $m+1$, and continue up to $n + \lceil m \rceil$ with $n + \lfloor m \rfloor$ (this last a valid pair since m is odd). This deals with all of the numbers in $\{m, \ldots, 2n\}$; the inductive hypothesis deals with $\{1, \ldots, m-1\}$ (again since m is odd). $\square$

Finally, we turn to the proof of Theorem 1.3. We begin with the upper bound. We have

$$\sqrt{x}^{\pi(x)-\pi(\sqrt{x})} \leq \prod_{p \leq x} p \leq 4^x.$$

The second inequality is Claim 2.3. For the first inequality, there are $\pi(x) - \pi(\sqrt{x})$ primes in the interval $(\sqrt{x}, x]$, and each of them is at least $\sqrt{x}$. It follows that

$$\pi(x) \leq \frac{4x}{\log_2 x} + \pi(\sqrt{x}) \leq \frac{4x}{\log_2 x} + \sqrt{x} \leq \frac{6x}{\log_2 x},$$

the last inequality valid for all $x \geq 2$ (a Mathematica calculation). This gives the right-hand side of the first inequality in Theorem 1.3 (with $C = 6$); noting that $\sqrt{x} = o(x/\log_2 x)$ as $x \to \infty$, we also get the right-hand side of (3).

For the lower bound on $\pi(x)$, $x \geq 2$, let $2n$ be the least even integer that is not smaller than x. Note that $2n - 2 < x \leq 2n$, and that $\pi(2n) \leq \pi(x) + 1$. Since the primes dividing $\binom{2n}{n}$ are all at most $2n$, Corollary 2.2 gives

$$\pi(2n) \geq \frac{\log_2 \binom{2n}{n}}{\log_2(2n)},$$

and so, using $\binom{2n}{n} \geq 4^n/(2n+1)$ together with our observations on relations between x and n, we get

$$\pi(x) \geq \frac{\log_2\left(\frac{2^x}{x+3}\right)}{\log_2(x+2)} = \frac{x - \log_2(x+3)}{\log_2(x+2)}.$$

A Mathematica calculation shows that this latter quantity is at least $x/2\log_2 x$ for all $x \geq 8$. This gives the left-hand side of the first inequality in Theorem 1.3 (with c perhaps smaller that 1/2 to deal with $x \in [2, 8)$); noting that

$$\frac{x - \log_2(x+3)}{\log_2(x+2)} \sim \frac{x}{\log_2 x}$$

as $x \to \infty$, we also get the left-hand side of (3).

Acknowledgements

Iriet Danon, Pierre-Yves Gaillard and Michal Pokorny pointed out some errors in an earlier version of this note. Mark Sapir pointed out an error in the proof of the lower bound in Theorem 1.3, and found a fix; my presentation follows his treatment in [9].

References

[1] M. Aigner and G. M. Ziegler, *Proofs from the book* (4th ed.), Springer, 2010.

[2] Cramér, H., On the Order of Magnitude of the Difference Between Consecutive Prime Numbers, *Acta Arith.* **2** (1936) 23–46. 1936.

[3] Erdős, P., Beweis eines Satzes von Tschebyschef, *Acta Sci. Math. (Szeged)* **5** (1930–1932), 194–198.

[4] Greenfield, L. and Greenfield, S., Some problems of combinatorial number theory related to Bertrand's postulate, *J. Integer Seq.* **1** (1998), Article 98.1.2.

[5] Hadamard, J., Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques, *Bull. Soc. math. France* **24** (1896), 199–220.

[6] Havil, J., *Gamma: Exploring Euler's Constant*, Princeton University Press (2003).

[7] Lou, S. and Yau, Q., A Chebyshev's Type of Prime Number Theorem in a Short Interval (II), *Hardy-Ramanujan J.* **15** (1992), 1–33.

[8] Poussin, C. de la Vallée, Recherces analytiques sur la théorie des nombres premiers, *Ann. Soc. Sci. Bruxells* (1897).

[9] Sapir, M., Combinatorial algebra: syntax and semantics, Springer, 2014.