Erdős's proof of Bertrand's postulate

David Galvin*

April 21, 2006

Abstract

In 1845 Bertrand postulated that there is always a prime between n and $2n$, and he verified this for $n < 3,000,000$. 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 $\{1, \ldots, 2n\}$ 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 [4] and de la Vallée-Poussin [7] in 1896, is the statement that

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

A consequence of the PNT is that

$$\forall \epsilon > 0 \; \exists n(\epsilon) > 0 : \; n > n(\epsilon) \Rightarrow \exists p \text{ prime}, \; n < p \leq (1+\epsilon)n. \tag{2}$$

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 $n > 0$ there is a prime p such that $n < p \leq 2n$.*

*Dept of Mathematics, University of Pennsylvania, Philadelphia, PA 19104; dgalvin@math.upenn.edu.

This is Bertrand's postulate, conjectured in the 1845, verified by Bertrand for all $N < 3\,000\,000$, and first proved by Tchebychev in 1850. (See [5, p. 25] for a discussion of the original references).

In his first paper Erdős [2] 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 a consequence in Section 3. One consequence is the following lovely theorem of Greenfield and Greenfield [3].

Theorem 1.2 *For $n > 0$, 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.

Another is an approximate version of (1).

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

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

2 Erdős's proof

We consider the middle binomial coefficient $\binom{2n}{n} = (2n)!/(n!)^2$. An easy lower bound is

$$\binom{2n}{n} \geq \frac{4^n}{2n+1}. \tag{3}$$

Indeed, $\binom{2n}{n}$ is the largest term in the $2n+1$-term sum $\sum_{i=0}^{2n} \binom{2n}{n} = (1+1)^{2n} = 4^n$. 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 $o_p(a/b) = o_p(a) - o_p(n)$. The heart of the whole proof is the following simple observation.

$$\text{If } \tfrac{2}{3}n < p \leq n \text{ then } o_p\left(\binom{2n}{n}\right) = 0 \text{ (i.e., } p \not| \binom{2n}{n}). \tag{4}$$

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 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 $2n/3$. 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*

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

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

$$\begin{aligned} o_p\left(\binom{2n}{n}\right) &= o_p((2n)!) - 2o_p(n!) \\ &= \sum_{i=1}^{r(p)} \left[\frac{2n}{p^i}\right] - 2\sum_{i=1}^{r(p)} \left[\frac{n}{p^i}\right] \\ &= \sum_{i=1}^{r(p)} \left(\left[\frac{2n}{p^i}\right] - 2\left[\frac{n}{p^i}\right]\right) \\ &\leq r(p), \qquad (5) \end{aligned}$$

and so

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

In (5) we use the easily verified fact that for integers a and b, $0 \leq [2a/b] - 2[a/b] \leq 1$. □

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

Claim 2.2 $\forall n$ $\prod_{p \leq n} p \leq 4^n$ *(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 an so not a prime and the first inequality following from the inductive hypothesis. For larger odd n, say $n = 2m+1$, we have

$$\begin{aligned} \prod_{p \leq n} p &= \prod_{p \leq m+1} p \prod_{m+2 \leq p \leq 2m+1} p \\ &\leq 4^{m+1}\binom{2m+1}{m} \qquad (6) \\ &\leq 4^{m+1}2^{2m} \qquad (7) \\ &= 4^{2m+1} = 4^n. \end{aligned}$$

In (6) 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 (7) 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 be such that there is no prime p with $n < p \leq 2n$. Then we have

$$\begin{aligned} \binom{2n}{n} &\leq (2n)^{\sqrt{2n}} \prod_{\sqrt{2n}<p\leq 2n/3} p && (8) \\ &\leq (2n)^{\sqrt{2n}} \prod_{p\leq 2n/3} p \\ &\leq (2n)^{\sqrt{2n}} 4^{2n/3}. && (9) \end{aligned}$$

The main point is (8). We have first used the simple fact that $\binom{2n}{n}$ has at most $\sqrt{2n}$ prime factors that are smaller than $\sqrt{2n}$, 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 (4) 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 (9) we have used Claim 2.2.

Combining (9) with (3) we obtain the inequality

$$\frac{4^n}{2n+1} \leq (2n)^{\sqrt{2n}} 4^{2n/3}. \quad (10)$$

This inequality can hold only for small values of n. Indeed, for any $\epsilon > 0$ the left-hand side of (10) 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 (10) fails for all $n \geq 468$ (Maple 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

$$2, 3, 5, 7, 13, 23, 43, 83, 163, 317, 631. \quad (11)$$

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 $n < 486$ contains one of these 11 primes. This concludes the proof of Theorem 1.1.

(If a Maple calculation is not satisfactory, it is easy to check that (10) reduces to $n/3 \leq \log_2(2n+1) + \sqrt{2n}\log_2 2n$. The left hand side of this inequality is increasing faster than the right, and the inequality is easily seen to fail for $n = 2^{10} = 1024$, so to complete the proof in this case we need only add the prime 1259 to the list in (11)).

3 Comments, conjectures and consequences

A stronger result than (2) is known (due to Lou and Yao [6]):

$$\forall\epsilon > 0\ \exists n(\epsilon) > 0 :\ n > n(\epsilon) \Rightarrow \exists p \text{ prime},\ n < p \leq n + n^{\frac{1}{2}+\frac{1}{22}+\epsilon}.$$

The Riemann hypothesis would imply

$$\forall\epsilon > 0\ \exists n(\epsilon) > 0:\ n > n(\epsilon) \Rightarrow \exists p \text{ prime},\ n < p \le n + n^{\frac{1}{2}+\epsilon}.$$

There is a very strong conjecture of Cramér [1] that would imply

$$\forall\epsilon > 0\ \exists n_0 > 0 : n > n_0 \Rightarrow \exists p \text{ prime},\ n < p \le n + (1+\epsilon)\ln^2 n.$$

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

Question 3.1 *Is it true that for all $n > 1$, 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 $2n < p \le 4n$. Since $4n$ is not prime we have $p = 2n+m$ for $1 \le m < 2k$. Pair $2n$ with m, $2n-1$ with $m+1$, and continue up to $n+\lceil k\rceil$ with $n+\lfloor k\rfloor$ (this last a valid pair since m is odd). This deals with all of the numbers in $\{m, \dots, 2n\}$; the inductive hypothesis deals with $\{1, \dots, m-1\}$ (again since m is odd). $\square$

Finally, we turn to the proof of Theorem 1.3. The upper bound will follow from Claim 2.2 while the lower bound will follow from Claim 2.1.

Proof of Theorem 1.3: For the lower bound on $\pi(x)$ choose n such that

$$\binom{2n}{n} \le x < \binom{2n+2}{n+1}.$$

For sufficiently large n we have $\ln\binom{2n}{n} > n$ (from (3)) and for all n we have $\binom{2n}{n}/\binom{2n+2}{n+1} \ge 1/4$ and so

$$\frac{\pi(x)\ln x}{x} \ge \frac{\pi\left(\binom{2n}{n}\right)\ln\binom{2n}{n}}{\binom{2n+2}{n+1}} \ge \frac{n\pi\left(\binom{2n}{n}\right)}{4\binom{2n}{n}} \tag{12}$$

We lower bound the number of primes at most $\binom{2n}{n}$ by counting those which divide $\binom{2n}{n}$. By Claim 2.1 each such prime contributes at most $2n$ to $\binom{2n}{n}$ and so $\pi\left(\binom{2n}{n}\right) \ge \binom{2n}{n}/2n$. Combining this with (12) we obtain (for sufficiently large x)

$$\pi(x) \ge \frac{x}{8\ln x}.$$

For the upper bound we use Claim 2.2 to get (for $x \ge 4$)

$$4^x \ge \prod_{p\le x} p \ge \sqrt{x}^{\pi(x)-\pi(x/2)}$$

and so

$$\pi(x) \le \frac{4x\ln 2}{\log x} + \pi(x/2).$$

Repeating this procedure $\lfloor \log_2 x \rfloor$ times we reach (for sufficiently large x)

$$\begin{aligned} \pi(x) &\leq \frac{8x \ln 2}{\log x} + \pi(2) \\ &\leq \frac{9x \ln 2}{\log x}. \end{aligned}$$

□

References

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

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

[3] 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.

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

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

[6] 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.

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