Bertrand's Postulate

Michael Tang

October 2, 2015

1 Introduction

Bertrand's Postulate is a theorem in number theory on the existence of prime numbers:

Theorem 1 (Bertrand's Postulate)**.** *For all positive integers $n \geq 2$, there is a prime p such that $n < p < 2n$.*

In this article, we follow the proof of the theorem that uses the beautiful, elementary method found by Paul Erdős[1], and reflect on the implications of the ideas brought forth.

2 Prerequisites

For a clear reading, the reader should also be familiar with the following:

- Binomial coefficients,
- Factorials, $v_p(x)$ notation, and Legendre's formula[2]
- Product (Capital pi) notation,
- Proof by induction.

3 Proof

3.1 Main idea

The idea of the proof is the following: we consider the (central) binomial coefficient $\binom{2n}{n}$, and prove that it must have a prime factor $n < p < 2n$ in order to be large enough; that is, the central binomial coefficients are such that they cannot consist solely of powers of "small" prime numbers. Therefore, first, we analyze the prime factors and prime powers dividing $\binom{2n}{n}$, and second, we give bounds that force the existence of a prime p in the range $n < p < 2n$.

[1] Erdős (1913-1996) discovered this proof in 1932, at the age of nineteen!

[2] Legendre's formula states that the highest power of a prime p that divides $n!$ is

$$v_p(n!) = \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \dots$$

3.2 Prime factors

For the sake of contradiction, suppose that Theorem 1 is false: that for some $n \geq 2$, there is no prime p in the range $n < p < 2n$. Consider the prime factors of the integer

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2}.$$

This integer has no prime factors greater than $2n$, since if a prime divides $\binom{2n}{n}$, it must also divide its numerator, $(2n)!$. Also, since $n \geq 2$, the integer $2n$ is not a prime, so all of the prime factors of $\binom{2n}{n}$ are strictly less than $2n$. By our assumption, there are no primes $n < p < 2n$, so all prime factors of $\binom{2n}{n}$ are now at most n.

In fact, we can increase this restriction even further. It may be helpful to consider the following example, taken from the (first ever!) AIME contest:

Problem (AIME, 1983). Find the greatest two-digit prime factor of $\binom{200}{100}$.

Solution. We try the largest two-digit prime, 97. Sadly, 97 is not a factor of $\binom{200}{100} = \frac{200!}{(100!)^2}$, because the numerator has two factors of 97 (one each from 97 and $2 \cdot 97 = 194$), as does the denominator (from two 97s). The same reasoning holds for the next prime, 89, as well as the next primes $83, 79, \ldots$ However, when we get to the prime 61, we note that 200! has three factors of 61 (from $61, 2 \cdot 61$, and $3 \cdot 61$), and $(100!)^2$ has two factors of 61. Therefore, 61 divides $\binom{200}{100}$, so it is our answer. $\square$

Using the same methods as in this problem, we establish the following result.

Lemma 2 (No primes greater than $2n/3$)**.** *The integer $\binom{2n}{n}$ has no prime factors p in the range $\frac{2n}{3} < p \leq n$.*

Proof. If a prime lies in the range $\frac{2n}{3} < p \leq n$, then the numerator $(2n)!$ has two factors of p (from p and $2p$), as does the denominator $(n!)^2$ (from two factors of p). Therefore, p does not divide $\binom{2n}{n}$. $\square$

Hence, all of the prime factors of $\binom{2n}{n}$ are at most $\frac{2n}{3}$. We now try to bound the sizes of these powers of "smaller" primes:

Lemma 3 (Prime power bounding)**.** *If a prime power p^k divides $\binom{2n}{n}$, then $p^k \leq 2n$.*

Proof. It suffices to prove that k, the power of p, is at most $\log_p(2n)$. For this, we apply Legendre's formula:

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

Consider the summand here. If we let $x = n/p^i$, then the summand is $\lfloor 2x \rfloor - 2\lfloor x \rfloor$. This quantity is either 0 or 1, depending on the fractional part of x: if $\{x\} \geq 1/2$, then it equals 1, and otherwise it equals 0.

In addition, when $x < 1/2$, both terms equal zero, so the summand equals zero. Now, $x < 1/2$ when $n/p^i < 1/2$, or $p^i > 2n$, which gives $i < \log_p(2n)$. Therefore, the only terms of the sum that can be nonzero (i.e. equal to 1) are when $i \leq \log_p(2n)$; Hence,

$$\sum_{i=1}^{\infty} \left(\left\lfloor \frac{2n}{p^i} \right\rfloor - 2\left\lfloor \frac{n}{p^i} \right\rfloor \right) \leq \log_p(2n),$$

which is exactly what we wanted to prove. □

In addition, this result ($k \leq \log_p(2n)$) gives the following:

Lemma 4 (Larger primes only occur once)**.** *If a prime p is larger than $\sqrt{2n}$, then it is a factor of $\binom{2n}{n}$ at most once. That is, p^2 does not divide $\binom{2n}{n}$.*

Proof. This is a direct result of the previous Lemma: if p^2 divides $\binom{2n}{n}$, then by Lemma 3, $p^2 \leq 2n$, so $p \leq \sqrt{2n}$. □

This improves our result about prime powers: Lemma 3 gives us that all prime powers dividing $\binom{2n}{n}$ are at most $2n$. However, Lemma 4 gives that prime powers of a prime $p > \sqrt{2n}$ are at most p, which is in turn at most $2n/3$.

Let us present the "big picture" now. The binomial coefficient $\binom{2n}{n}$ has all prime factors at most $\frac{2n}{3}$. Furthermore, each prime power that divides it is at most $2n$, and each prime factor greater than $\sqrt{2n}$ (which is a relatively small number compared to n and $2n$) can only occur once. Having developed these strict bounds on the prime factors of $\binom{2n}{n}$, we seek now to derive some contradictory inequality on n (or an inequality true only for very certain, small values of n).

3.3 Bounding

We begin with a lower bound on $\binom{2n}{n}$:

Lemma 5 (Binomial coefficient lower bound)**.** *For all positive integers n, $\frac{4^n}{2n} \leq \binom{2n}{n}$.*

Proof. For this proof, we view

$$4^n = 2^{2n} = (1+1)^{2n}$$

and use the Binomial Theorem:

$$4^n = 1 + \binom{2n}{1} + \binom{2n}{2} + \ldots + \binom{2n}{2n-1} + 1.$$

Adding the two 1s separately, we obtain exactly $2n$ terms, one of which is 2 and the others being of the form $\binom{2n}{r}$, for some r. Since $\binom{2n}{n}$ is the largest combination in its row (and is always at least 2), we have

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

which is what we wanted to prove. □

Now we may employ our results from the first half of the proof. For a prime p, let $f(p)$ be the largest power of p that divides $\binom{2n}{n}$. Then

$$\binom{2n}{n} = \prod_{1 \leq p \leq 2n/3} f(p).$$

We know that $f(p) \leq 2n$ for all p, but for $p > \sqrt{2n}$, we have an even better bound: $f(p) \leq p^1 = p$. Therefore, we write

$$\prod_{1 \leq p \leq 2n/3} f(p) = \left(\prod_{1 \leq p \leq \sqrt{2n}} f(p)\right)\left(\prod_{\sqrt{2n} < p \leq 2n/3} f(p)\right).$$

There are at most $\sqrt{2n}$ primes from 1 to $\sqrt{2n}$, inclusive, and each prime power is at most $2n$, so $\prod_{1 \leq p \leq \sqrt{2n}} f(p) \leq (2n)^{\sqrt{2n}}$. We have $f(p) \leq p$ for $p > \sqrt{2n}$ as well, which gives the inequality

$$\left(\prod_{1 \leq p \leq \sqrt{2n}} f(p)\right)\left(\prod_{\sqrt{2n} < p \leq 2n/3} f(p)\right) \leq (2n)^{\sqrt{2n}} \cdot \prod_{\sqrt{2n} < p \leq 2n/3} p.$$

At this point, we introduce a new function to give a bound on the remaining product.

Definition 6. For a positive real number x, the **primorial**[3] of x, written $x\#$, is the product of all primes less than or equal to x. (If $x < 2$, $x\#$ is the empty product, which by convention equals 1.)

Then $\displaystyle\prod_{\sqrt{2n}<p\le 2n/3} p = \frac{(2n/3)\#}{(\sqrt{2n})\,\#}$. We bound the primorial with the following lemma:

Lemma 7 (Primorial bound)**.** *For all $x \ge 1$, the primorial $x\# \le 4^x$.*

Proof. First note that it suffices to prove the lemma when x is an integer, since if x is not an integer, then

$$x\# = (\lfloor x \rfloor)\,\# \le 4^{\lfloor x \rfloor} \le 4^x.$$

Also, isolate the base cases $x = 1, 2, 3$, which are easy to check by hand.

The critical idea is as follows. Note that $\binom{2x-1}{x} = \frac{(2x-1)!}{(x-1)!x!}$ is divisible by every prime p in the range $x + 1 \le p \le 2x - 1$, since the numerator is divisible by p but not the denominator. Since $(2x-1)\#/x\#$ is the product of all primes $x + 1 \le p \le 2x - 1$, we have

$$\frac{(2x-1)\#}{x\#} \le \binom{2x-1}{x}.$$

We also have that

$$2^{2x-1} = (1+1)^{2x-1} \ge \binom{2x-1}{x-1} + \binom{2x-1}{x} = 2\binom{2x-1}{x},$$

so $\binom{2x-1}{x} \le 2^{2x-2} = 4^{x-1}$. Thus,

$$(2x-1)\# \le 4^{x-1} \cdot x\#$$

for all x. This allows us to induct on the even and odd integers.

If x is odd, say $x = 2y - 1$, then

$$x\# = (2y-1)\# \le 4^{y-1} \cdot y\# \le 4^{y-1} \cdot 4^y = 4^{2y-1},$$

concluding the induction step. If x is even, then since $x \ge 4$ is not prime, we have

$$x\# = (x-1)\# \le 4^{x-1} \le 4^x$$

by applying the previous inequality to the odd integer $x - 1$. This completes the induction. $\square$

Returning to the problem, we have $\displaystyle\prod_{\sqrt{2n}<p\le 2n/3} p = \frac{(2n/3)\#}{(\sqrt{2n})\,\#} \le \left(\frac{2n}{3}\right)\# \le 4^{2n/3}$. Hence,

$$\binom{2n}{n} \le (2n)^{\sqrt{2n}} \cdot \prod_{\sqrt{2n}<p\le 2n/3} p \le (2n)^{\sqrt{2n}} \cdot 4^{2n/3}$$

[3]primorial, *not* primordial!

so we have the inequality

$$\frac{4^n}{2n} \leq (2n)^{\sqrt{2n}} \cdot 4^{2n/3}.$$

Rearranging, we have $4^{n/3} \leq (2n)^{1+\sqrt{2n}}$, or, taking logs,

$$\frac{n}{3} \ln 4 \leq (1 + \sqrt{2n}) \ln(2n).$$

Rewriting, we get

$$0 \leq \frac{(1 + \sqrt{2n}) \ln(2n)}{n} - \frac{\ln 4}{3}.$$

It turns out that (as a function over the real numbers), the right-hand side decreases steadily for $n \geq 3$, which can be seen by differentiating; $n = 467.4$ is the approximate location of a root of the right-hand side, so for $n \geq 468$, this inequality is false, and we have our contradiction.

For $n \leq 467$, it is easy to check manually. In fact, consider the sequence of primes

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

Each prime is less than twice the previous prime, so for $n \leq 467$, there must be at least one prime p in the range $n < p < 2n$. This gives a contradiction as well, so Theorem 1 is proven.

4 Reflections

Here are a few discussion points about the content and nature of this proof.

- The function we ended up with is something approximately

  $$\frac{\sqrt{n} \ln n}{n} = \frac{\ln n}{\sqrt{n}}$$

  (eliminating some small factors). The natural logarithm grows slower than any power of n, so as $n \to \infty$, $\ln n/\sqrt{n} \to 0$. Formally, by l'Hôpital's rule,

  $$\lim_{n \to \infty} \frac{\ln n}{\sqrt{n}} = \lim_{n \to \infty} \frac{n^{-1}}{1/2 \cdot n^{-1/2}} = \lim_{n \to \infty} 2n^{-1/2} = 0.$$

  This is why we expected the inequality to hold only for small n.

- Why is the theorem called a "postulate," a term reserved for basic assumptions and axioms only? The explanation is a historical relic: French mathematician Joseph Bertrand (1822-1900), who worked in group theory, found that the statement of Theorem 1 would be helpful to him in solving a group theory problem. He could not prove the theorem conclusively, but it seemed incredibly obvious, so he took it as a "postulate" for his purposes.

- Note that some of the bounds we placed were incredibly weak. We wrote, for example, "There are at most $\sqrt{2n}$ primes from 1 to $\sqrt{2n}$, inclusive," or, equivalently, that *all prime numbers are positive integers.* For the lower bound proved in Lemma 5, we used only the fact that the central binomial coefficient is the largest of the entries in its row (and a similar idea to prove Lemma 7), which becomes very, very weak as n grows. There are ways to improve the bounds, and as a result, decrease the number 467 for a smaller manual check (see [1] for an example), but the proofs are all the same in spirit.

- Even though some bounds were incredibly weak, it is quite subtle what was really necessary to establish a contradictory result. For instance, one might think, "why is Lemma 2 necessary, if it only eliminates one-third of the prime factors possible? Lemma 4 eliminates all but $\sqrt{2n}$ prime factors, which is much smaller in comparison. However, if we replace $2n/3$ with n, we get the inequality

$$\frac{4^n}{2n} \leq (2n)^{\sqrt{2n}} \cdot 4^n$$

  which is *always true*! So one has to be careful with the exact details of bounding.

- The proof really reveals how little we know about primes (how elusive prime numbers can be). Theorem 1 seems incredibly simple, almost trivial - we know from intuition that there ought to be not just one, but many primes between n and $2n$ - yet it takes very detailed and precise methods to prove the theorem via elementary means. The idea of considering the binomial coefficient on its own is quite groundbreaking, and not by any means easy to discover. Other proofs of Theorem 1 exist (a notable one by Ramanujan (1887-1920) in [3] is short but quite advanced), but Erdös' proof, as was so much of his other work, is truly a classic.

References

[1] Shigenori Tochiori *Stronger proof of Bertrand-Chebyshev's theorem*, available at `http://www.chart.co.jp/subject/sugaku/suken_tsushin/76/76-8.pdf`

[2] Wikipedia *Proof of Bertrand's postulate*, available at `https://en.wikipedia.org/wiki/Proof_of_Bertrand\%27s_postulate`

[3] Jaban Meher, M. Ram Murty *Ramanujan's Proof of Bertrand's Postulate*, available at `http://www.mast.queensu.ca/~murty/Meher-Murty-Monthly.pdf`