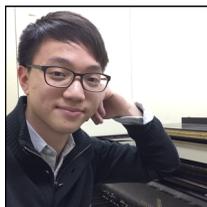


Goldbach's Pigeonhole

Edward Early, Patrick Kim, and Michael Proulx



Edward Early (eduarde@stedwards.edu) received his Ph.D. from the Massachusetts Institute of Technology. He is an associate professor of mathematics at St. Edward's University in Austin, TX. His research is mostly in combinatorics, but he enjoys dabbling in number theory. He also likes spending time with his family and training in martial arts. **Patrick Kim** (musicpkim@gmail.com) is a senior at Stanford University majoring in materials science and music. His interests lie in multifunctional nanoparticles along with phase-change materials. He is an avid pianist and violinist, spending many hours playing chamber music. In his free time, Patrick enjoys swimming, playing Ultimate, and hungrily watching the Food Network late at night. **Michael Proulx** (michaelproulx@college.harvard.edu) is a senior at Harvard College whose research interests include analytic number theory and low-dimensional topology. Aside from mathematics, Michael enjoys backpacking and mountaineering; he is also a lifelong student of Go.

The Goldbach conjecture was first proposed in a 1742 letter from German mathematician Christian Goldbach to Swiss mathematician Leonhard Euler [1]. In the letter, Goldbach proposed to his colleague that every even integer greater than 2 can be expressed as the sum of two primes, a hypothesis known today as *Goldbach's strong conjecture* or simply *Goldbach's conjecture*. There are many variations and related problems, including connections to the Riemann hypothesis, which have attracted the attention of numerous mathematicians over the years [2, 4].

Is it possible that the pigeonhole principle could give an easy proof of Goldbach's conjecture? It turns out there are hundreds of integers for which we can prove the existence of two such primes nonconstructively via the pigeonhole principle. It does not always work, but how do we know when we have found all the cases for which it does? The answer turns out to be a nice application of asymptotic formulas, or very scary-looking formulas for complicated functions that actually turn out to be easy to use with a little modification. A little computer programming then finishes off the argument.

Some number theory and bounds

A *Goldbach partition* is an expression of a given even integer n as the sum of two primes. We will use three number-theoretic functions. Euler's totient function $\varphi(n)$ denotes the number of positive integers less than or equal to and relatively prime to n . It can be computed via the formula $\varphi(n) = n \cdot \prod_{p|n} (p - 1)/p$ [3]. The prime counting

<http://dx.doi.org/10.4169/college.math.j.46.2.131>
MSC: 11P32

function $\pi(n)$ denotes the number of primes less than n , and $\omega(n)$ denotes the number of distinct prime factors of n . Throughout, we shall use n to denote a positive even integer and p to denote a prime. Our work focuses on relating $\varphi(n)$ and $\pi(n)$ to satisfy the conditions for a particular counting argument.

In 1962, Rosser and Schoenfeld improved the asymptotic estimates and bounds of several functions that are pertinent to our work [5]. Specifically, they derived the following upper bound for $\pi(n)$ and lower bound for $\varphi(n)$.

Theorem 1. For $n > 1$,

$$\pi(n) < \frac{n}{\ln(n)} \left(1 + \frac{1.5}{\ln(n)} \right). \quad (1)$$

Theorem 2. For $n > 3$,

$$\varphi(n) > \frac{n}{e^\gamma \ln(\ln(n)) + \frac{2.5}{\ln(\ln(n))}} \quad (2)$$

except when $n = 223,092,870$, in which case

$$\varphi(n) > \frac{n}{e^\gamma \ln(\ln(n)) + \frac{2.50637}{\ln(\ln(n))}}, \quad (3)$$

where $\gamma \approx 0.577216$ is Euler's constant.

The pigeonhole principle is a standard counting argument that states that if n items (pigeons) are placed into m containers (holes) where $n > m$, then at least one container must contain more than one item.

We do not purport to prove Goldbach's conjecture; rather, we explore an interesting approach to the conjecture using the pigeonhole principle. We use the term *pair* to refer to a set of two positive integers that sum to a given even integer n . By expressing each n as a list of distinct pairs, we can prove the existence of a Goldbach partition if we can demonstrate that there are more primes (pigeons) than pairs (pigeonholes). We would like to remove all pairs composed of two composites, but this seems too difficult to count rigorously. However, pairs can be removed when the two numbers are not relatively prime, which will be related to $\varphi(n)$, and there are $\pi(n) - \omega(n)$ primes to consider.

After demonstrating that we can verify the conjecture for certain values of n using the pigeonhole principle, we prove that $n = 90,090$ is the greatest value that this method can support.

Placing the primes into the pairs

We begin with the observation that any n satisfying $\varphi(n)/2 < \pi(n) - \omega(n)$ must have a Goldbach partition.

Consider $n = 20$; the pairs are $\{1, 19\}$, $\{2, 18\}$, $\{3, 17\}$, $\{4, 16\}$, \dots , $\{10, 10\}$.

The length of this list can then be reduced to $\varphi(n)/2$ by removing pairs that are not relatively prime to n . The key observation here is that if a and b are integers such that $a + b = n$, then a and b are relatively prime if and only if they are each relatively prime to n . This is true because any common factor of a and b will also divide $a + b = n$, and any common factor of a and n will also divide $n - a = b$. In our example, we see that there are four remaining pairs, $\{1, 19\}$, $\{3, 17\}$, $\{7, 13\}$, $\{9, 11\}$.

Next, we count the number of primes less than and relatively prime to $n = 20$, given by $\pi(n) - \omega(n)$. There are six such primes, 3, 7, 11, 13, 17, and 19. Therefore, by the pigeonhole principle, at least one pair must comprise two primes, thus there is at least one Goldbach partition of 20 (in fact, there are two).

More generally, n has at least $\pi(n) - \omega(n) - \varphi(n)/2$ Goldbach partitions.

The inequality can be computed easily using Mathematica, Maple, Sage, etc. We leave the details to the reader. This could also be the basis for an exercise in any class involving discrete mathematics and some computer programming.

For $n < 100$ we can thus verify Goldbach's conjecture for all n other than 4, 6, 10, 58, 82, 86, 92, and 94. Even though the method fails for some small values of n , it works for some larger numbers such as 53,130, 60,060, 78,540, and 90,090. Because $\pi(n) - \omega(n)$ is a relatively predictable function, whether this pigeonhole argument will work when applied to a given n is largely dependent on the totient function.

More specifically, values of n with many distinct prime factors will generally have lower values of $\varphi(n)$ than nearby values of n with few distinct prime factors. Because $\pi(n)$ is monotonically increasing and $\omega(n)$ is small relative to n , numbers with lower values of $\varphi(n)$ are more likely to work with the pigeonhole argument and numbers with higher values of $\varphi(n)$ are more likely to fail the pigeonhole argument.

When the Goldbach pigeonhole method stops working

There are some special cases where the list of pigeonholes may be reduced, extending the argument to more values of n . For example, if $n - 1$ is composite, then $\{1, n - 1\}$ can be removed from the list of pairs without decreasing the number of primes appearing in the pairs. While it may be possible to efficiently implement more such refinements to extend the applicability of the argument, our focus for now is on analyzing the pigeonhole process without any such enhancements. As it is, the method can be used to verify Goldbach's conjecture nonconstructively for 586 values of n ranging from 12 to 90,090. Our main result is that the method does not work for any larger value.

Theorem 3. *If $n > 90,090$, then $\varphi(n)/2 > \pi(n) - \omega(n)$, thus the above approach with the pigeonhole principle cannot be used to verify Goldbach's conjecture for n .*

A graph provides insight into how to proceed. Figure 1 shows the graph of $\varphi(n)/2$ for even values of n as a jagged area, which is actually a very rapidly and irregularly oscillating curve (though the function is discrete, consecutive points have been connected for the purpose of graphing). There are three other curves shown. The one that starts the lowest and ends the highest is the lower bound on this function from (2). The curve that ends the lowest is $\pi(n)$, and the remaining curve is the upper bound on $\pi(n)$ from (1). We can see the lower bound of $\varphi(n)/2$ surpassing the upper bound of $\pi(n)$ at $n \approx 210,000$ (the precise value is 210,909).

Proof. Comparing inequalities (1) and (2), we want to prove that, for n sufficiently large,

$$\frac{n}{2 \left(e^\gamma \ln(\ln(n)) + \frac{2.5}{\ln(\ln(n))} \right)} > \frac{n}{\ln(n)} \left(1 + \frac{1.5}{\ln(n)} \right),$$

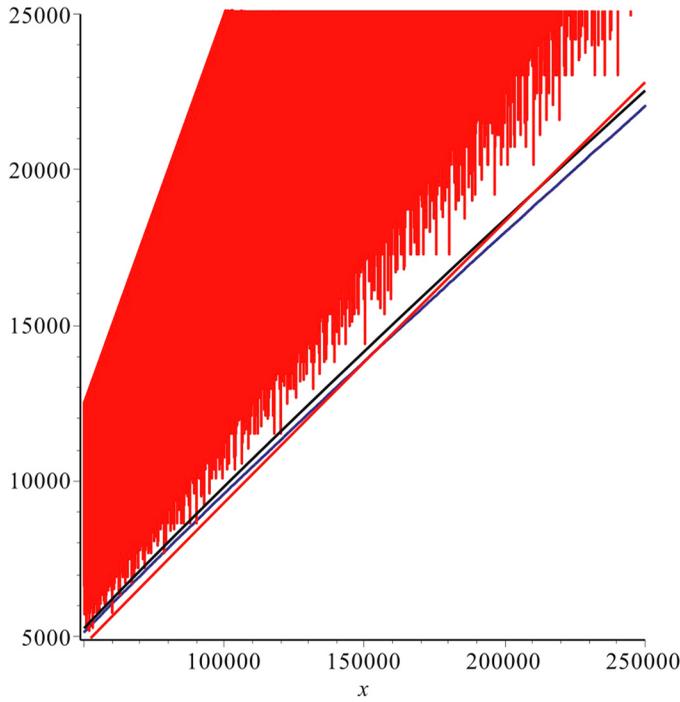


Figure 1. Comparison of the bounds

equivalently

$$\ln(n) > 2 \left(1 + \frac{1.5}{\ln(n)} \right) \left(e^\gamma \ln(\ln(n)) + \frac{2.5}{\ln(\ln(n))} \right). \quad (4)$$

Assume $n > 210,000$ and $n \neq 223,092,870$. Since the right-hand side of (4) is difficult to work with, we sacrifice some precision for the sake of simplicity. It is easily verified that

$$\frac{1.5}{\ln(n)} < \frac{1}{8}, \quad \frac{2.5}{\ln(\ln(n))} < 1, \quad e^\gamma < \frac{9}{5}.$$

These substitutions yield

$$\left(1 + \frac{1.5}{\ln(n)} \right) \cdot 2 \left(e^\gamma \ln(\ln(n)) + \frac{2.5}{\ln(\ln(n))} \right) < \frac{81}{20} \ln(\ln(n)) + \frac{9}{4},$$

so it will suffice to prove that

$$\ln(n) > \frac{81}{20} \ln(\ln(n)) + \frac{9}{4}. \quad (5)$$

Let $f(n) = \ln(n)$ and $g(n)$ the right-hand side of (5). Both $f(n)$ and $g(n)$ are continuous and monotonically increasing functions on the positive reals. The two functions appear to cross twice, once at $n \approx 2$ and again at $n \approx 260,202$. We therefore want to confirm that the two functions do not intersect anywhere after $n = 260,202$. To this end, it will suffice to show that $f'(n) > g'(n)$ for $n > 260,202$ since f is then both greater than and increasing faster than g .

Now, $f'(n) = 1/n$ and $g'(n) = 81/(n \ln(n))$, thus $f'(n) > g'(n)$ whenever $\ln(n) > 81/20$, or $n > e^{81/20} \approx 57.397457$. This establishes (5) and hence (4) for $n > 260,202$.

Since inequality (2) does not apply when $n = 223,092,870$, we could modify the argument using (3), or simply observe that $\varphi(223,092,870)/2 = 18,247,680$ is greater than $\pi(223,092,870) = 12,283,531$.

To complete the proof, we ran a computer search to verify that there are no values of n between 90,092 and 260,202, inclusive for which the method works. ■

Acknowledgments. The authors thank the Mathworks program at Texas State University for providing the opportunity to work together on this project at the Honors Summer Math Camp.

Summary. Goldbach's conjecture states that every even integer greater than two can be written as the sum of two primes. For some even integers, we can prove the existence of two such primes nonconstructively via the pigeonhole principle. Using a computer search and asymptotic bounds on classic number-theoretic functions, we determine the greatest value for which this approach works.

References

1. C. Goldbach, Letter to L. Euler, June 7, 1742, *Oper. Omnia* 765, available at <http://eulerarchive.maa.org/correspondence/letters/000765.pdf>.
2. R. K. Guy, *Unsolved Problems in Number Theory*. Third edition. Springer, New York, 2004.
3. G. H. Hardy, E. M. Wright, *An Introduction to the Theory of Numbers*. Oxford Univ. Press, Oxford, 1979.
4. R. Munroe, Goldbach Conjectures, 2013, <http://xkcd.com/1310/>.
5. J. B. Rosser, L. Schoenfeld, Approximate formulas for some functions of prime numbers, *Illinois J. Math.* **6** (1962) 62–94.