Beyond Popular Science

David H. Silver

BEYOND POPULAR SCIENCE
David H. Silver

https://www.openbookpublishers.com

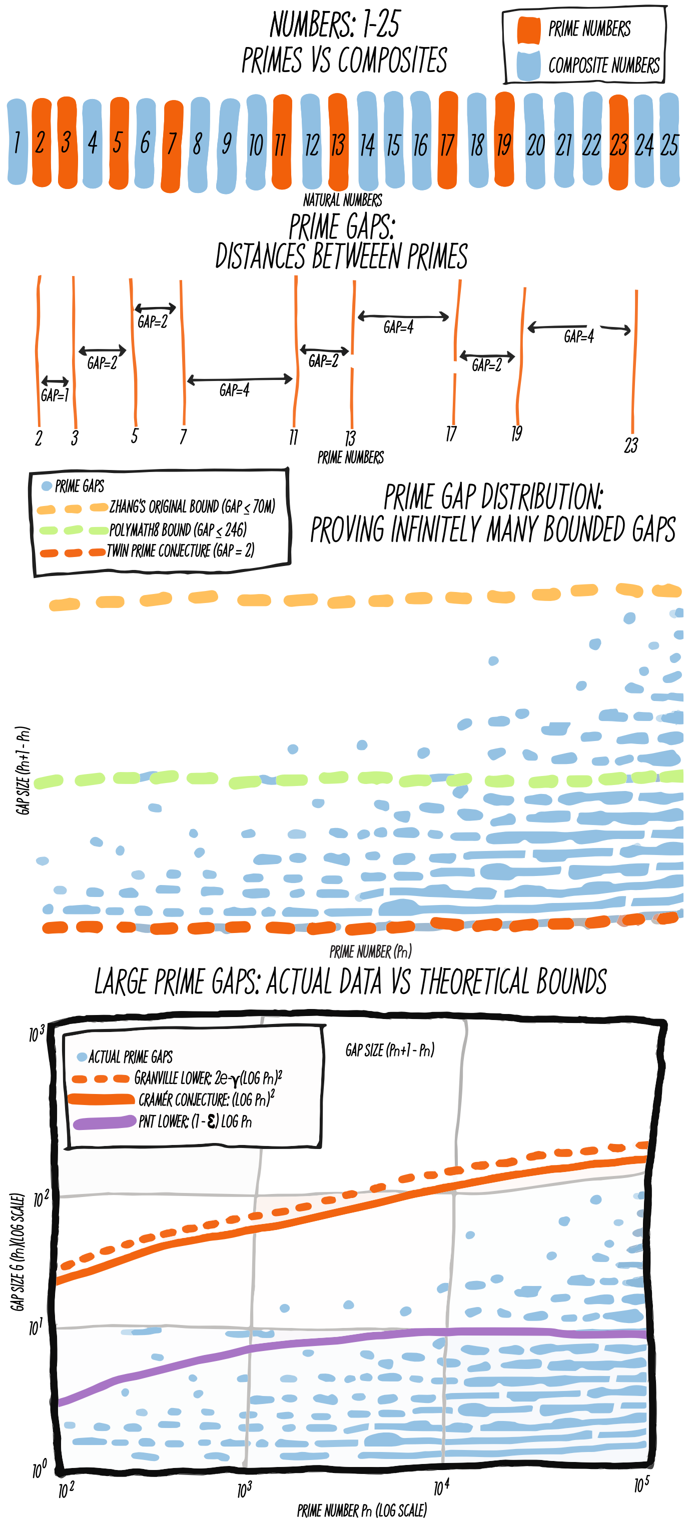
Digital material and resources associated with this volume are available at https://doi.org/10.11647/OBP.0526#resources

ISBN Paperback:	978-1-80511-877-0
ISBN Hardback:	978-1-80511-878-7
ISBN Digital (PDF):	978-1-80511-879-4
ISBN HTML:	978-1-80511-881-7
ISBN Digital ebook (epub):	978-1-80511-880-0
DOI:	10.11647/OBP.0526

Cover image by Enny Silver and David H. Silver
Cover design by Jeevanjot Kaur Nagpal

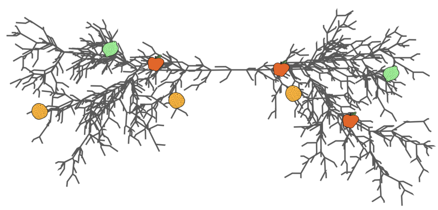
Mind the Gap

Top (Prime Distribution): Natural numbers 1–25 with primes (red) and composites (teal), showing that twin prime pairs like (3,5), (5,7), (11,13), (17,19).
Second (Gap Sizes): Consecutive primes with gap sizes marked by blue arrows, from the minimal gap of 2 to larger separations such as 8 between 23 and 31.
Third (Gap Analysis): Scatter plot of prime gaps $p_{n+1} - p_n$ versus p_n for the first 10,000 primes, showing that most gaps are small with rare large exceptions. Reference lines mark key bounds: twin primes (gap = 2), Polymath8's refined bound (246), and Zhang's original breakthrough (70 million). Infinitely many gaps stay below some fixed bound, guaranteeing points always appear below these lines.
Note the key distinction: for small gaps we seek absolute bounds (fixed numbers such as 246), while for large gaps the bounds are functions of p_n that grow as primes get larger.
Bottom (Maximum Gaps): How large can prime gaps become? This shows the largest observed gaps (light blue) compared to theoretical predictions. The red lines show upper bound conjectures for maximum gap size, while coloured lower bounds prove that gaps must occasionally be large. The scatter demonstrates that while most gaps are small relative to the local prime density, some gaps are much larger than the typical spacing in their region.



Mind the Gap

In 2013, an unaffiliated Yitang Zhang proved there exists a finite bound B (initially 70,000,000) such that infinitely many prime pairs differ by at most B. While prime gaps can grow arbitrarily large, this breakthrough showed they cannot drift apart arbitrarily far. The Polymath8 collaboration subsequently reduced this bound to a few hundred. Zhang's approach combined distribution properties of primes in arithmetic progressions with an advanced sieving technique, resolving a fundamental question about number patterns while falling short of proving the Twin Prime Conjecture that infinitely many primes differ by exactly 2.



Prime Gaps ∘ Twin Prime Conjecture ∘ Zhang's Bounded Gaps ∘ 70 Million Bound ∘ Weighted Sieve Method ∘ Polymath8 Collaboration ∘ Maynard Simplification ∘ Green-Tao Progressions ∘ Cramér's Conjecture ∘ Riemann Hypothesis ∘ Zhang's Trajectory

"If I were to awaken after having slept
for a thousand years,
my first question would be:
has the Riemann Hypothesis been proven?"

— David Hilbert, 1900

Mind the Gap

In the third century BCE, Euclid proved that there are infinitely many prime numbers. His argument, based on contradiction, became one of the earliest and most enduring examples of a general mathematical method. The search for primes—and for patterns among them—soon followed. Eratosthenes introduced a sieve procedure (an algorithm that filters out composites by eliminating multiples) for enumerating primes. By the time of Diophantus primes were already recognised as foundational to arithmetic.

In the eighteenth century, Euler showed that the sum of reciprocals of primes diverges (a stronger quantitative refinement of Euclid's theorem that, in particular, implies infinitude), and he introduced analytic tools that connected primes to infinite products and logarithmic identities. This initiated the study of prime distribution through analytic functions.

In 1859, Bernhard Riemann introduced the zeta function into number theory (a complex analytic function encoding prime information via its Euler product) and conjectured that all its nontrivial zeros lie on the critical line. This hypothesis remains unproven. Riemann's formulation marked the beginning of analytic number theory—a field that uses tools from complex analysis to study the distribution and density of primes. G. H. Hardy and others developed this perspective further in the early twentieth century.

The study of prime gaps took a more technical turn when Viggo Brun introduced sieve methods in the 1910s (combinatorial procedures for bounding the count of integers with prescribed divisibility). Brun proved that the sum of reciprocals of twin primes converges, implying their overall scarcity, even if they might be infinite in number. Later refinements by Selberg and Bombieri led to the Bombieri–Vinogradov theorem (an average-case version of the Generalised Riemann Hypothesis for arithmetic progressions), which became central to modern sieve theory.

In the early 2000s, Goldston, Pintz, and Yıldırım (GPY) introduced a method for bounding small gaps between primes using weighted sums over admissible tuples (integer patterns that avoid local divisibility obstructions—for example, $0, 2, 4$ is not admissible, since modulo 3 it covers all residue classes, whereas $0, 2, 6$ is admissible). Their work showed that if primes are sufficiently regular in arithmetic progressions, then bounded gaps should follow. The approach relied on conjectural input—notably the Elliott–Halberstam conjecture (a proposed uniformity result for primes in arithmetic sequences).

In 2013, Yitang Zhang proved that there are infinitely many pairs of primes separated by at most 70 million. His argument used a modified version of the GPY sieve. This was the first proof that bounded prime gaps occur infinitely often, without relying on unproven conjectures.

Prime numbers are integers greater than 1 that have no positive divisors other than 1 and themselves. They form the multiplicative building blocks of arithmetic. The first few primes, 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, occur without any obvious pattern. Their spacing varies.

As numbers increase, primes become less frequent, because they have more possible factors. The Prime Number Theorem formalises this observation (Hadamard, 1896; de la Vallée Poussin, 1896): the number of primes less than x grows like $x/\log x$. This gives an average spacing between primes near x of about $\log x$, but does not constrain individual gaps.

The differences between consecutive primes can be small or large. Pairs such as (3, 5), (5, 7), (11, 13), (17, 19), and (29, 31) each differ by 2. The smallest prime gap of 2 recurs often at the start of the number line. By contrast, the primes 370,261 and 370,373 are separated by 112, with no primes between them. For any given n, there exist consecutive primes with a gap larger than n. One construction uses the sequence $(n+1)!+2, (n+1)!+3, \ldots, (n+1)!+(n+1)$, which yields n consecutive composite numbers, hence a gap of at least n between the bounding primes.

What remains unknown is whether small gaps, such as a fixed difference of 2, occur infinitely often. The Twin Prime Conjecture asserts that there are infinitely many primes p such that $p+2$ is also prime. This remains one of the major conjectures in number theory.

In 2013, Yitang Zhang proved that there exists a constant B such that infinitely many pairs of primes differ by at most B (as opposed to exactly a fixed gap such as 2 in the twin prime conjecture). His original bound was $B < 70{,}000{,}000$—while this does not resolve the twin prime conjecture, it proves that small prime gaps occur infinitely often.

Zhang's method extended work by Goldston, Pintz, and Yıldırım. He combined improved estimates on the distribution of primes in arithmetic progressions with a weighted sieve construction that amplified configurations where primes appear close together. This yielded a finite bound on the gap size that recurs infinitely often.

Following Zhang's proof, the Polymath8 collaboration reduced the bound from 70 million to below 250 through analytic refinements. Later on, James Maynard introduced a simplified sieve method that removed the need for strong distributional estimates and extended the technique to detect many primes within bounded intervals.

Zhang's result drew attention for its mathematical content and the circumstances of its discovery—after completing his doctorate, he spent years outside academic mathematics, with no permanent university position and limited research output. His proof was written and submitted independently, lacking collaborators or institutional support. The publication of his result led to rapid follow-up work, large-scale collaboration, and the re-entry of a long-standing problem into the mathematical mainstream.

While Zhang's work addressed bounded gaps, the opposite question—how large prime gaps can become—has also attracted intense study. Let $G(x)$ be the largest gap between consecutive primes less than x. It is known that $G(x)$ increases faster than $\log x$, which is the average spacing predicted by the Prime Number Theorem. A classical result due to Paul Erdős shows that $G(x)$ exceeds a constant multiple of $\log x$ times another slowly growing function. This means that although most prime gaps are relatively small, unusually large gaps must still occur infinitely often. The best known upper bounds on $G(x)$ remain far from matching the lower bounds. Some of the strongest predictions, such as Cramér's conjecture, suggest that the maximal gap should grow no faster than $\log^2 x$, but this has not been proven.

Alongside these increasingly long gaps, primes can also appear in patterns. In 2004, Ben Green and Terence Tao proved that the primes contain arbitrarily long arithmetic progressions. For any integer k, there exists a sequence of the form $p, p + d, p + 2d, \ldots, p + (k - 1)d$ in which all terms are prime. The length k can be taken as large as desired. Although such progressions become rarer as k increases, the result shows that they never stop appearing.

The Green–Tao theorem uses tools from ergodic theory and additive combinatorics—it begins by approximating the set of primes using related sequences whose behaviour is easier to control. A transference principle then carries results from these surrogate sequences back to the primes themselves. The original result was later extended by Green and Ziegler to cover polynomial progressions, such as $p, p + q, p + 4q, p + 9q, p + 16q$ in which the differences between terms follow a fixed polynomial pattern and all terms are again required to be prime.

Many questions regarding the distribution of primes, including the spacing between them and the occurrence of patterned arrangements, remain unresolved. Some of these questions cannot be settled with current methods because their answers depend on an open conjecture in complex analysis and number theory. This conjecture is known as the Riemann Hypothesis (Riemann, 1859).

The Riemann Hypothesis (RH) concerns a function called the Riemann zeta function. This function is initially defined as a sum over positive integers, $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$, which converges when the complex number s has real part greater than 1. Through a process known as analytic continuation, the function is extended to other values of s in the complex plane. The hypothesis asserts that all nontrivial zeros of $\zeta(s)$—that is, all values of s for which $\zeta(s) = 0$ and which are not negative even integers—lie on the vertical line $\mathrm{Re}(s) = \frac{1}{2}$.

Assuming RH, bounds on the error terms in prime-counting functions become significantly tighter. For example, the difference between the actual number of primes up to x and the estimate $x/\log x$ can be bounded more sharply. The hypothesis also leads to improved estimates on how often primes occur in short intervals and how large the gaps between consecutive primes can be. Without a proof, many of these refinements remain conditional. The Riemann Hypothesis has been verified for many individual zeros through numerical computation, and no counterexamples have been found. Nevertheless, the general statement remains unproven.

Transparent Statements, Resistant Proofs

The central claim of this chapter can be stated in one line and requires no definitions beyond the integers. This is typical of many problems in number theory. Simplicity of formulation does not imply tractability.

Consider these easy-to-state problems that remain unsolved:

- **Twin Prime Conjecture**: Are there infinitely many primes p such that $p + 2$ is also prime?
- **Goldbach's Conjecture**: Can every even integer greater than 2 be written as the sum of two primes?
- **Odd Perfect Numbers**: Does there exist an odd perfect number (a number equal to the sum of its proper divisors)?

Each can be explained to a child, yet they have resisted centuries of mathematical effort. They can be tested on billions of examples, but no general proof exists.

Zhang's proof is concise and intricate. Its validity depends on a balance between distributional estimates and the sieve framework. Subsequent refinements by the Polymath8 project, led by Terence Tao, and by Maynard's independent method reduced the bound but did not simplify the analytic core.

This chapter is included because of Zhang's personal trajectory and because it illustrates this broader principle: many problems in number theory are easy to state and test numerically, yet remain inaccessible to current methods. Let's demonstrate it further.

The following questions of existence, stated in a single line, can fall anywhere along the spectrum of difficulty:

- **Impossible to answer**: Hilbert's Tenth Problem asked for a general procedure to decide whether any Diophantine equation (equation in integer variables with polynomial coefficients) has a solution. Matiyasevich's work, building on Davis–Putnam–Robinson, showed (Matiyasevich, 1970) that no such algorithm can exist.
- **Unknown**: The *Collatz Conjecture* asks whether repeated iteration of the rule $n \mapsto n/2$ if n is even and $3n + 1$ if n is odd always reaches 1. No proof is known.
- **Hard No**: Fermat's Last Theorem—that $x^n + y^n = z^n$ has no nontrivial integer solutions for $n > 2$—resisted proof until Wiles.

Proofs: Infinitude of Primes

Two nice proofs of the infinitude of primes—one using topology and another via a trigonometric product identity—.

1. **Furstenberg's Topological Proof**
   Define a topology on $\mathbb{Z}$ by taking as a basis of open sets all arithmetic sequences of the form:
   $$S_{a,b} = \{a + bn \mid n \in \mathbb{Z}\}.$$
   These sets mimic modular residue classes and are closed under finite intersections, forming a well-defined topology. Suppose there are only finitely many primes, $p_1, p_2, \dots, p_k$. The union of their corresponding arithmetic sequences,
   $$S = \bigcup_{i=1}^{k} S_{0,p_i} = \{n \mid n \equiv 0 \mod p_i \text{ for some } i\},$$
   consists of all integers divisible by at least one prime and would be a closed set. Its complement, the set of integers not divisible by any p_i, must then be open. However, every basic open set $S_{a,b}$ is infinite, meaning an open set cannot be finite. Since the complement of S consists of finitely many integers (the units ± 1 modulo $p_1 p_2 \dots p_k$), we obtain a contradiction. Thus, the assumption that the set of primes is finite must be false.
2. **Trigonometric Product Proof**
   Assume for contradiction that the set $\mathcal{P}$ of primes is finite. Then consider the product:
   $$0 < \prod_{p \in \mathcal{P}} \sin(\pi/p).$$
   Since each term is positive, the product itself remains positive.
   Now define $N = 2\prod_{p'} p'$, the product of all assumed primes. By construction, every prime p divides N, so the term $1 + N$ must be divisible by some prime q. That is, for some integer k, $1 + N = kq$. Then, evaluating the sine function:
   $$\sin(\pi(1 + N)/q) = \sin(k\pi) = 0.$$
   This forces the right-hand side of the original product identity to be zero:
   $$\prod_{p \in \mathcal{P}} \sin(\pi(1 + N)/p) = 0.$$
   But this contradicts the assumption that the left-hand side was positive. Hence, the set of primes must be infinite.

Remark: Both proofs subtly rely on Euclid's key step: that some prime p must divide the product of primes plus one, making these proofs disguised versions of the classic argument.

References: H. Furstenberg, *Bull. Amer. Math. Soc.*, 62 (1955), 353.& *A One-Line Proof of the Infinitude of Primes*, Amer. Math. Monthly, 122 (2015), 466.

Bounded Gaps Between Primes

Statement of Result and Outline of Method

In 2013, Yitang Zhang proved that there exists a constant N such that infinitely many prime pairs (p, q) satisfy $q - p \leq N$. The initial bound was $N < 7 \times 10^7$. Zhang's approach refined the Goldston–Pintz–Yıldırım (GPY) method through two components:

- **Distribution in Arithmetic Progressions:** Primes remain evenly distributed across residue classes beyond the Bombieri–Vinogradov range.
- **Weighted Sieve:** Modified sieve to detect multiple primes within admissible tuples $n + h_i$.

Sieve Setup and Admissibility

A tuple $\mathcal{H} = \{h_1, \ldots, h_k\}$ is admissible if for every prime p, the set $\mathcal{H} \mod p$ does not cover all residue classes modulo p.

The GPY strategy constructs a weighted sum:

$$S(n) := \left(\sum_{i=1}^{k} \Lambda(n + h_i) \right) w(n),$$

where Λ is the von Mangoldt function and $w(n)$ is a smooth function supported on $n \in [x, 2x]$. The weights emphasise values where multiple $n + h_i$ are likely prime:

$$\sum_n S(n) = \sum_n \left(\sum_i \Lambda(n + h_i) \right) w(n).$$

If this exceeds the random baseline, then for some n, at least two $n + h_i$ are prime.

Example

The admissible set $\{0, 2, 6\}$ avoids covering all residue classes modulo any prime. For $n = 5$, we get $\{5, 7, 11\}$—three primes. The method proves such cases occur infinitely often.

Zhang's Level of Distribution

Zhang proved that primes remain equidistributed up to moduli $q \leq x^\theta$ for $\theta > 1/2$, surpassing the Bombieri–Vinogradov barrier ($\theta = 1/2$).

Define:

$$\theta(x; q, a) = \sum_{\substack{p \leq x \\ p \equiv a \pmod{q}}} \log p.$$

The deviation of $\theta(x; q, a)$ from $x/\phi(q)$ remains small across a wide range of moduli, enabling uniform error control. Zhang bypassed the Elliott–Halberstam conjecture by achieving a weaker but sufficient level of distribution.

Maynard's Modification and Polymath Refinements

Maynard introduced new sieve weights detecting primes in admissible tuples without requiring $\theta > 1/2$, simplifying the construction.

The Polymath8 project refined and extended both approaches:

- **Polymath8a:** Improved error analysis, reduced N to 4,680.
- **Maynard's Variant:** Lowered N further, generalised to m primes in bounded intervals.
- **Polymath8b:** Reduced bound below 250.

References:

Zhang, Y. (2014). Bounded gaps between primes. *Annals of Mathematics*, 179(3), 1121–1174.

Maynard, J. (2015). Small gaps between primes. *Annals of Mathematics*, 181(1), 383–413.