



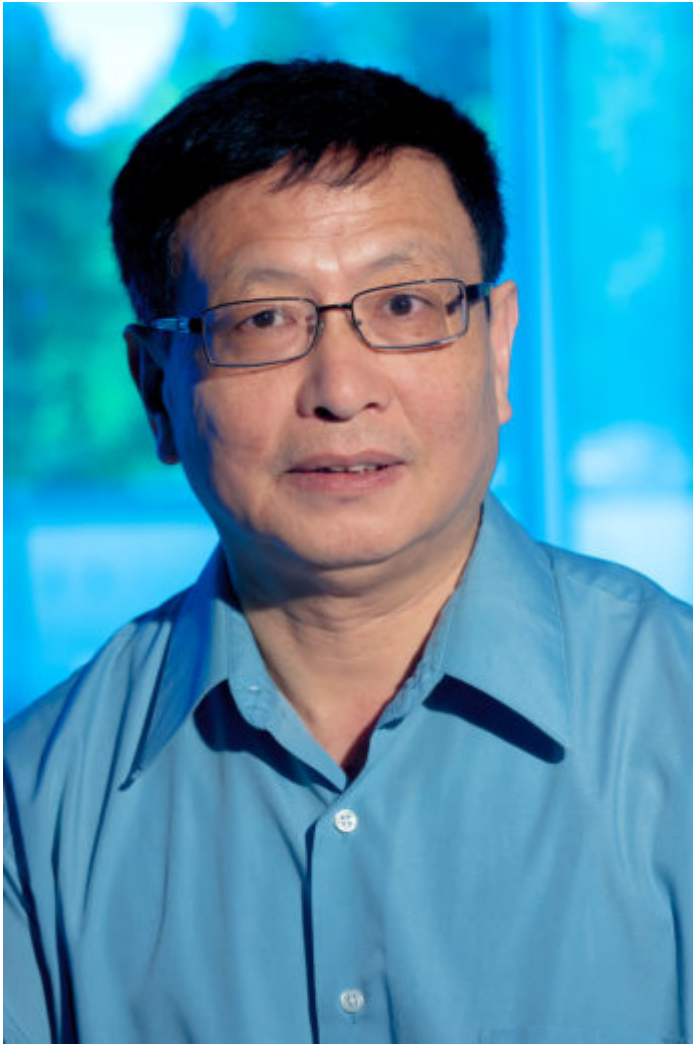
Unheralded Mathematician Bridges the Prime Gap

By Erica Klarreich

On April 17, a paper arrived in the inbox of *Annals of Mathematics*, one of the discipline's preeminent journals. Written by a mathematician virtually unknown to the experts in his field — a 50-something lecturer at the University of New Hampshire named Yitang Zhang — the paper claimed to have taken a huge step forward in understanding one of mathematics' oldest problems, the twin primes conjecture.

Editors of prominent mathematics journals are used to fielding grandiose claims from obscure authors, but this paper was different. Written with crystalline clarity and a total command of the topic's current state of the art, it was evidently a serious piece of work, and the *Annals* editors decided to put it on the fast track.

Just three weeks later — a blink of an eye compared to the usual pace of mathematics journals — Zhang received the referee report on his paper.



Yitang Zhang

“The main results are of the first rank,” one of the referees wrote. The author had proved “a landmark theorem in the distribution of prime numbers.”

Rumors swept through the mathematics community that a great advance had been made by a researcher no one seemed to know — someone whose talents had been so overlooked after he earned his doctorate in 1991 that he had found it difficult to get an academic job, working for several years as an accountant and even in a Subway sandwich shop.

“Basically, no one knows him,” said Andrew Granville, a number theorist at the Université de Montréal. “Now, suddenly, he has proved one of the great results in the history of number theory.”

Mathematicians at Harvard University hastily arranged for Zhang to present his work to a packed audience there on May 13. As details of his work have emerged, it has become clear that Zhang achieved his result not via a radically new approach to the problem, but by applying existing methods with great perseverance.

“The big experts in the field had already tried to make this approach work,” Granville said. “He’s not a known expert, but he succeeded where all the experts had failed.”

The Problem of Pairs

Prime numbers — those that have no factors other than 1 and themselves — are the atoms of arithmetic and have fascinated mathematicians since the time of Euclid, who proved more than 2,000 years ago that there are infinitely many of them.

Because prime numbers are fundamentally connected with multiplication, understanding their additive properties can be tricky. Some of the oldest unsolved problems in mathematics concern basic questions about primes and addition, such as the twin primes conjecture, which proposes that there are infinitely many pairs of primes that differ by only 2, and the Goldbach conjecture, which proposes that every even number is the sum of two primes. (By an astonishing coincidence, a weaker version of this latter question was settled in a [paper posted online](#) by Harald Helfgott of École Normale Supérieure in Paris while Zhang was delivering his Harvard lecture.)

Related Articles:

- [Together and Alone, Closing the Prime Gap](#)
- [After Prime Proof, an Unlikely Star Rises](#)
- [Prime Gap Grows After Decades-Long Lull](#)

Prime numbers are abundant at the beginning of the number line, but they grow much sparser among large numbers. Of the first 10 numbers, for example, 40 percent are prime — 2, 3, 5 and 7 — but among 10-digit numbers, only about 4 percent are prime. For over a century, mathematicians have understood how the primes taper off on average: Among large numbers, the expected gap between prime numbers is approximately 2.3 times the number of digits; so, for example, among 100-digit numbers, the expected gap between primes is about 230.

But that's just on average. Primes are often much closer together than the average predicts, or much farther apart. In particular, "twin" primes often crop up — pairs such as 3 and 5, or 11 and 13, that differ by only 2. And while such pairs get rarer among larger numbers, twin primes never seem to disappear completely (the largest pair discovered so far is $3,756,801,695,685 \times 2^{666,669} - 1$ and $3,756,801,695,685 \times 2^{666,669} + 1$).

For hundreds of years, mathematicians have speculated that there are infinitely many twin prime pairs. In 1849, French mathematician Alphonse de Polignac extended this conjecture to the idea that there should be infinitely many prime pairs for any possible finite gap, not just 2.

Since that time, the intrinsic appeal of these conjectures has given them the status of a mathematical holy grail, even though they have no known applications. But despite many efforts at proving them, mathematicians weren't able to rule out the possibility that the gaps between primes grow and grow, eventually exceeding any particular bound.

Now Zhang has broken through this barrier. His paper shows that there is some number N smaller than 70 million such that there are infinitely many pairs of primes that differ by N . No matter how far you go into the deserts of the truly gargantuan prime numbers — no matter how sparse the primes become — you will keep finding prime pairs that differ by less than 70 million.

The result is "astounding," said Daniel Goldston, a number theorist at San Jose State University. "It's one of those problems you weren't sure people would ever be able to solve."

A Prime Sieve

The seeds of Zhang's result lie in a [paper from eight years ago](#) that number theorists refer to as GPY, after its three authors — Goldston, János Pintz of the Alfréd Rényi Institute of Mathematics in Budapest, and Cem Yıldırım of Boğaziçi University in Istanbul. That paper came tantalizingly close but was ultimately unable to prove that there are infinitely many pairs of primes with some finite gap.

Instead, it showed that there will always be pairs of primes much closer together than the average spacing predicts. More precisely, GPY showed that for any fraction you choose, no matter how tiny, there will always be a pair of primes closer together than that fraction of the average gap, if you go out far enough along the number line. But the researchers couldn't prove that the gaps between these prime pairs are always less than some particular finite number.

GPY uses a method called "sieving" to filter out pairs of primes that are closer together than average. Sieves have long been used in the study of prime numbers, starting with the 2,000-year-old Sieve of Eratosthenes, a technique for finding prime numbers.

To use the Sieve of Eratosthenes to find, say, all the primes up to 100, start with the number two, and cross out any higher number on the list that is divisible by two. Next move on to three, and cross out all the numbers divisible by three. Four is already crossed out, so you move on to five, and cross out all the numbers divisible by five, and so on. The numbers that survive this crossing-out process are the primes.

	2	3	4	5	6	7	8	9	10	Prime numbers
11	12	13	14	15	16	17	18	19	20	
21	22	23	24	25	26	27	28	29	30	
31	32	33	34	35	36	37	38	39	40	
41	42	43	44	45	46	47	48	49	50	
51	52	53	54	55	56	57	58	59	60	
61	62	63	64	65	66	67	68	69	70	
71	72	73	74	75	76	77	78	79	80	
81	82	83	84	85	86	87	88	89	90	
91	92	93	94	95	96	97	98	99	100	
101	102	103	104	105	106	107	108	109	110	
111	112	113	114	115	116	117	118	119	120	

The Sieve of Eratosthenes This

procedure, which dates back to the ancient Greeks, identifies all the primes less than a given number, in this case 121. It starts with the first prime — two, colored bright red — and eliminates all numbers divisible by two (colored dull red). Then it moves on to three (bright green) and eliminates all multiples of three (dull green). Four has already been eliminated, so next comes five (bright blue); the sieve eliminates all multiples of five (dull blue). It moves on to the next uncolored number, seven, and eliminates its multiples (dull yellow). The sieve would go on to 11 — the square root of 121 — but it can stop here, because all the non-primes bigger than 11 have already been filtered out. All the remaining numbers (colored purple) are primes.

The Sieve of Eratosthenes works perfectly to identify primes, but it is too cumbersome and inefficient to be used to answer theoretical questions. Over the past century, number theorists have developed a collection of methods that provide useful approximate answers to such questions.

"The Sieve of Eratosthenes does too good a job," Goldston said. "Modern sieve methods give up on trying to sieve perfectly."

GPY developed a sieve that filters out lists of numbers that are plausible candidates for having prime

pairs in them. To get from there to actual prime pairs, the researchers combined their sieving tool with a function whose effectiveness is based on a parameter called the level of distribution that measures how quickly the prime numbers start to display certain regularities.

The [level of distribution is known to be at least \$\frac{1}{2}\$](#) . This is exactly the right value to prove the GPY result, but it falls just short of proving that there are always pairs of primes with a bounded gap. The sieve in GPY could establish that result, the researchers showed, but only if the level of distribution of the primes could be shown to be more than $\frac{1}{2}$. Any amount more would be enough.

The theorem in GPY “would appear to be within a hair’s breadth of obtaining this result,” the researchers wrote.

But the more researchers tried to overcome this obstacle, the thicker the hair seemed to become. During the late 1980s, three researchers — Enrico Bombieri, a Fields medalist at the Institute for Advanced Study in Princeton, John Friedlander of the University of Toronto, and Henryk Iwaniec of Rutgers University — had developed a way to tweak the definition of the level of distribution to [bring the value of this adjusted parameter up to \$\frac{4}{7}\$](#) . After the GPY paper was circulated in 2005, researchers worked feverishly to incorporate this tweaked level of distribution into GPY’s sieving framework, but to no avail.

“The big experts in the area tried and failed,” Granville said. “I personally didn’t think anyone was going to be able to do it any time soon.”

Closing the Gap

Meanwhile, Zhang was working in solitude to try to bridge the gap between the GPY result and the bounded prime gaps conjecture. A Chinese immigrant who received his doctorate from Purdue University, he had always been interested in number theory, even though it wasn’t the subject of his dissertation. During the difficult years in which he was unable to get an academic job, he continued to follow developments in the field.

“There are a lot of chances in your career, but the important thing is to keep thinking,” he said.

Zhang read the GPY paper, and in particular the sentence referring to the hair’s breadth between GPY and bounded prime gaps. “That sentence impressed me so much,” he said.

Without communicating with the field’s experts, Zhang started thinking about the problem. After three years, however, he had made no progress. “I was so tired,” he said.

To take a break, Zhang visited a friend in Colorado last summer. There, on July 3, during a half-hour lull in his friend’s backyard before leaving for a concert, the solution suddenly came to him. “I immediately realized that it would work,” he said.

Zhang’s idea was to use not the GPY sieve but a modified version of it, in which the sieve filters not by every number, but only by numbers that have no large prime factors.

“His sieve doesn’t do as good a job because you’re not using everything you can sieve with,” Goldston said. “But it turns out that while it’s a little less effective, it gives him the flexibility that allows the argument to work.”

While the new sieve allowed Zhang to prove that there are infinitely many prime pairs closer together than 70 million, it is unlikely that his methods can be pushed as far as the twin primes

conjecture, Goldston said. Even with the strongest possible assumptions about the value of the level of distribution, he said, the best result likely to emerge from the GPY method would be that there are infinitely many prime pairs that differ by 16 or less.

But Granville said that mathematicians shouldn't prematurely rule out the possibility of reaching the twin primes conjecture by these methods.

"This work is a game changer, and sometimes after a new proof, what had previously appeared to be much harder turns out to be just a tiny extension," he said. "For now, we need to study the paper and see what's what."

It took Zhang several months to work through all the details, but the resulting paper is a model of clear exposition, Granville said. "He nailed down every detail so no one will doubt him. There's no waffling."

Once Zhang received the referee report, events unfolded with dizzying speed. Invitations to speak on his work poured in. "I think people are pretty thrilled that someone out of nowhere did this," Granville said.

For Zhang, who calls himself shy, the glare of the spotlight has been somewhat uncomfortable. "I said, 'Why is this so quick?'" he said. "It was confusing, sometimes."

Zhang was not shy, though, during his Harvard talk, which attendees praised for its clarity. "When I'm giving a talk and concentrating on the math, I forget my shyness," he said.

Zhang said he feels no resentment about the relative obscurity of his career thus far. "My mind is very peaceful. I don't care so much about the money, or the honor," he said. "I like to be very quiet and keep working by myself."

Meanwhile, Zhang has already started work on his next project, which he declined to describe. "Hopefully it will be a good result," he said.

Correction: This article was revised on May 21, 2013, to reflect that Yitang Zhang's doctorate was issued by Purdue University in 1991, not 1992.

This article was [reprinted on Wired.com](#) and translated into Chinese on [Wired.tw](#).