The Beauty of Bounded Gaps
A huge discovery about prime numbers
—and what it means for the future of mathematics

Jordan Ellenberg

Last May, Yitang "Tom" Zhang, a popular math professor at the University of New Hampshire, stunned the world of pure mathematics when he announced that he had proven the "bounded gaps" conjecture about the distribution of prime numbers—a crucial milestone on the way to the even more elusive twin primes conjecture, and a major achievement in itself.

The stereotype, outmoded though it is, is that new mathematical discoveries emerge from the minds of dewy young geniuses. But Zhang is over 50. What's more, he hasn't published a paper since 2001. Some of the world's most prominent number theorists have been hammering on the bounded gaps problem for decades now, so the sudden resolution of the problem by a seemingly inactive mathematician far from the action at Harvard, Princeton, and Stanford came as a tremendous surprise.

But the fact that the conjecture is *true* was no surprise at all. Mathematicians have a reputation of being no-bullshit hard cases who don't believe a thing until it's locked down and proved. That's not quite true. All of us believed the bounded gaps conjecture before Zhang's big reveal, and we all believe the twin primes conjecture even though it remains unproven. Why?

Let's start with what the conjectures say. The prime numbers

Lisa Nugent, University of New Hampshire Photographic Services

Tom Zhang, lecturer in mathematics at the University of New Hampshire.

are those numbers greater than 1 that aren't multiples of any number smaller than themselves and greater than 1; so 7 is a prime, but 9 is not, because it's divisible by 3. The first few primes are: 2, 3, 5, 7, 11, and 13.

Every positive number can be expressed in just one way as a product of prime numbers. For instance, 60 is made up of two 2s, one 3, and one 5. (This is why we don't take 1 to be a prime, though some mathematicians have done so in the past; it breaks the uniqueness, because if 1 counts as prime, 60 could be written as $2 \times 2 \times 3 \times 5$ and $1 \times 2 \times 2 \times 3 \times 5$ and $1 \times 1 \times 2 \times 2 \times 3 \times 5$ ...)

The primes are the atoms of number theory, the basic indivisible entities of which all numbers are made. As such, they've been the objects of intense study ever since number theory started. One of the very first theorems in number theory is that of Euclid, which tells us that the primes are infinite in number; we will never run out, no matter how far along the number line we let our minds range.

But mathematicians are greedy types, not inclined to be satisfied with mere assertion of infinitude. After all, there's infinite and then there's *infinite*. There are infinitely many powers of 2, but they're very rare. Among the first 1,000 numbers, there are only 10 powers of 2: 1, 2, 4, 8, 16, 32, 64, 128, 256, and 512.

There are infinitely many even numbers, too, but they're much more common: exactly 500 out of the first 1,000. In fact, it's pretty apparent that out of the first x numbers, just about $(1/2)x$ will be even.

Primes, it turns out, are intermediate—more common than the powers of 2 but rarer than even numbers. Among the first x numbers, about $x/\log(x)$ are prime; this is the prime number theorem, proven at the end of the 19th century by Hadamard and de la Vallée Poussin. This means, in particular, that prime numbers get less and less common as the numbers get bigger, though the decrease is very slow;

a random number with 20 digits is half as likely to be prime as a random number with 10 digits.

Naturally, one imagines that the more common a certain type of number, the smaller the gaps between instances of that type of number. If you're looking at an even number, you never have to travel farther than 2 numbers forward to encounter the next even; in fact, the gaps between the even numbers are always *exactly* of size 2. For the powers of 2, it's a different story. The gaps between successive powers of 2 grow exponentially, and there are finitely many gaps of any given size; once you get past 16, for instance, you will never again see two powers of 2 separated by a gap of size 15 or less.

Those two problems are easy, but the question of gaps between consecutive primes is harder. It's so hard that, even after Zhang's breakthrough, it remains a mystery in many respects.

And yet we think we know what to expect, thanks to a remarkably fruitful point of view—we think of primes as *random numbers.* The reason the fruitfulness of this viewpoint is so remarkable is that the viewpoint is so very, very false. Primes are not random! Nothing about them is arbitrary or subject to chance. Quite the opposite—we take them as immutable features of the universe and carve them on the golden records we shoot out into interstellar space to prove to the ETs that we're no dopes.

If you start thinking really hard about what "random" *really* means, first you get a little nauseated, and a little after that you find you're doing analytic philosophy. So let's not go down that road.

Instead, take the mathematician's path. The primes are not random, but it turns out that in many ways they *act as if they were.* For
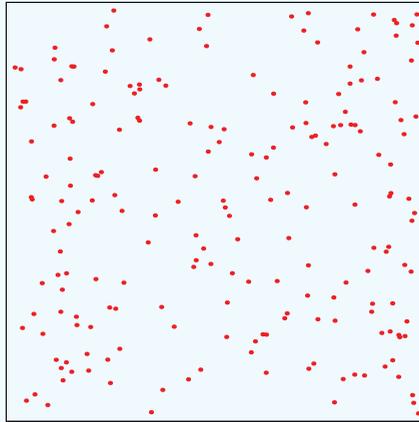
A random collection of points exhibits clumps and clusters.

example, when you divide a random number by 3, the remainder is 0, 1, or 2, and each case arises equally often. When you divide a big prime number by 3, the quotient can't come out even; otherwise, the so-called prime would be divisible by 3, which would mean it wasn't really a prime at all. But an old theorem of Dirichlet tells us that remainder 1 shows up about equally often as remainder 2, just as is the case for random numbers. So as far as "remainder modulo 3" goes, prime numbers, apart from not being multiples of 3, look random.

What about the gaps between consecutive primes? You might think because prime numbers get rarer and rarer as numbers get bigger, that they also get farther and farther apart. On average, that's indeed the case. But what Yitang Zhang just proved is that there are infinitely many pairs of primes that differ by at most 70,000,000. In other words, that the gap between one prime and the next is bounded by 70,000,000 infinitely often—thus, the "bounded gaps" conjecture.

On first glance, this might seem a miraculous phenomenon. If the primes are tending to be farther and farther apart, what's causing there to be so many pairs that are close together? Is it some kind of prime gravity?

Nothing of the kind. If you strew numbers at random, it's very likely that some pairs will, by chance, land very close together. (The picture at left is a nice illustration of how this works in the plane; the points are chosen independently and completely randomly, but you see some clumps and clusters all the same.)

It's not hard to compute that, if prime numbers behaved like random numbers, you'd see precisely the behavior that Zhang demonstrated. Even more: You'd expect to see infinitely many pairs of primes that are separated by only 2, as the twin primes conjecture claims.

Among the first N numbers, about $N / \log N$ of them are primes. If these were distributed randomly, each number n would have a $1 / \log N$ chance of being prime. The chance that n and $n + 2$ are *both* prime should thus be about $(1 / \log N)^2$. So how many pairs of primes separated by 2 should we expect to see? There are about N pairs $(n, n + 2)$ in the range of interest, and each one has a $(1 / \log N)^2$ chance of being a twin prime, so one should expect to find about $(N / \log N)^2$ twin primes in the interval.

There are some deviations from pure randomness whose small effects number theorists know how to handle; a more refined analysis taking these into account suggests that the number of twin primes should in fact be about 32 percent greater than $(N / \log N)^2$. This better approximation gives a prediction that the number of twin primes less than a quadrillion should be about 1.1 trillion; the actual figure is 1,177,209,242,304. That's a lot of twin primes.

And a lot of twin primes is exactly what number theorists expect to find no matter how big the numbers get—not because we think there's a

est theorems of modern mathematics, like Pierre Deligne's results relating averages of number-theoretic functions with the geometry of high-dimensional spaces. (More clas-

deep, miraculous structure hidden in the primes, but *precisely because we don't think so.* We expect the primes to be tossed around at random like dirt. If the twin primes conjecture were false, *that* would be a miracle, requiring that some hitherto unknown force be pushing the primes apart.

Not to pull back the curtain too much, but a lot of famous conjectures in number theory are like this. The Goldbach conjecture that every even number is the sum of two primes? The ABC conjecture, for which Shin Mochizuki controversially claimed a proof last fall? The conjecture that the primes contain arbitrarily long arithmetic progressions, whose resolution by Ben Green and Terry Tao in 2004 helped win Tao a Fields Medal? All are immensely difficult, but they are all exactly what one is guided to believe by the example of random numbers.

It's one thing to know what to expect and quite another to prove one's expectation is correct. Despite the apparent simplicity of the bounded gaps conjecture, Zhang's proof requires some of the deep-

sically minded analytic number theorists are already wondering whether Zhang's proof can be modified to avoid such abstruse stuff.)

Building on the work of many predecessors, Zhang is able to show in a precise sense that the prime numbers look random in the first way we mentioned, concerning the remainders obtained after division by many different integers. From this (following a path laid out by Goldston, Pintz, and Yıldırım, the last people to make any progress on prime gaps) he can show that the prime numbers look random in a

totally different sense, having to do with the sizes of the gaps between them. Random is random!

Zhang's success (along with the work of Green and Tao) points to a prospect even more exciting than any individual result about primes—that we might, in the end, be on our way to developing a richer theory of randomness. How wonderfully paradoxical: What helps us break down the final mysteries about prime numbers may be new mathematical ideas that structure the concept of structurelessness itself. ■

Jordan Ellenberg is a professor of mathematics at the University of Wisconsin. His book, How Not to Be Wrong, *is scheduled to appear in 2014.*

Email: ellenber@math.wisc.edu

This article first appeared in Slate Magazine (www.slate.com).

· ·

Further Reading

Number theorist Emmanuel Kowalski offers a first report on Zhang's paper: *http://bit.ly/10KAyvf.*

Here's Terry Tao on the dichotomy between structure and randomness: *http://bit.ly/11ZWyd0.*

Since Zhang's result appeared in May, several researchers have been working to decrease the size of the bounded gap. Zhang's gap of 70,000,000 is currently down to 5,414. Progress can be tracked here: *http://bit.ly/15yECDq.*