JOHNS HOPKINS UNIVERSITY

Emily Riehl

AI for math: what is happening and what it means

Chalmers University of Technology

A lot has changed since July 2024:

A.I. Can Write Poetry, but It Struggles With Math

A.I.'s math problem reflects how much the new technology is a break with computing's past.

Sean Dong

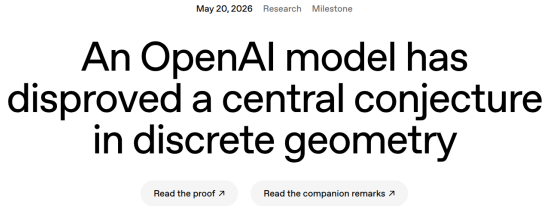
1. What is happening in AI-for-Math

2. How the mathematical community is responding

1

What is happening in AI-for-Math

May 20, 2026

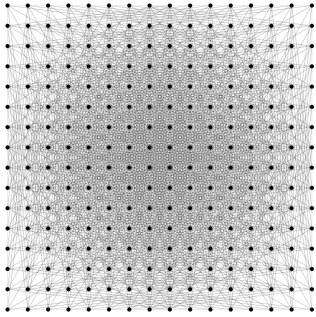


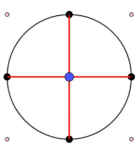
In 1946, Paul Erdös posed the unit distance problem: a conjectural upper bound to the number of pairs of points whose distance is exactly 1 over any collection of n points in the plane.

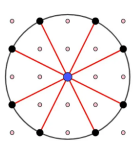
Commentary by Tim Gowers:

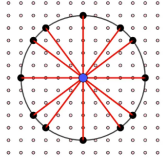
> In any case, there is no doubt that the solution to the unit-distance problem is a milestone in AI mathematics: if a human had written the paper and submitted it to the Annals of Mathematics and I had been asked for a quick opinion, I would have recommended acceptance without any hesitation. No previous AI-generated proof has come close to that. Furthermore, even if it is correct that AI cannot yet find a proof that needs a long hint sequence, such proofs are very difficult to find for humans as well, so in the unlikely event that progress in AI mathematics does suddenly stall, we have still probably entered an era where it will become very difficult for humans to compete with AI at solving mathematical problems.[4]
>
> [4] I chose my words carefully there, since solving problems is not all that mathematicians do. My guess is that AI will soon reach a high level at other activities such as building theories, formulating definitions and asking interesting questions, but that is a separate discussion.

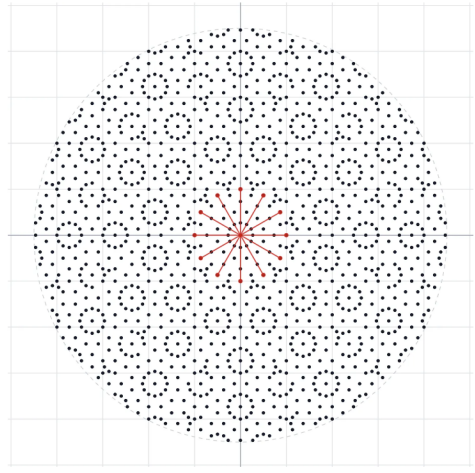
Unit distance problem







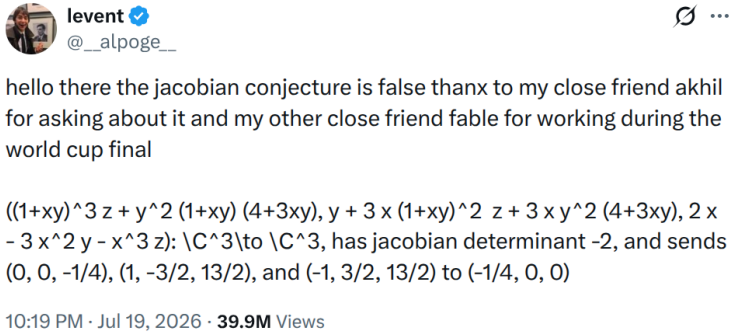




If we make the grid smaller, we can intersect more grid points with the unit circle. This gives more unit distances. (Diagram by Kai Williams)

Images by Kai Williams

July 19, 2026



The Jacobian conjecture states that for polynomial function $F\colon \mathbb{C}^n \to \mathbb{C}^n$, if the determinant of the matrix of partial derivatives is a non-zero constant, then F has a polynomial inverse. This is true in dimension 1, open in dimension 2, and false in dimensions 3 and above on account of the counterexample found by Fable.

The Jacobian conjecture

For the polynomial functor of three complex variables

$$((1+xy)^3z+y^2(1+xy)(4+3xy), y+3x(1+xy)^2z+3xy^2(4+3xy), 2x-3x^2y-x^3z)\colon \mathbb{C}^3 \to \mathbb{C}^3$$

the determinant of the matrix of partial derivatives

$$\begin{pmatrix} 3y(1+xy)^2z+y^3(7+6xy) & 3x(1+xy)^2z+8y+21xy^2+12x^2y^3 & (1+xy)^3 \\ 3(1+xy)^2z+6xy(1+xy)z+12y^2+18xy^3 & 1+6x^2(1+xy)z+24xy+27x^2y^2 & 3x(1+xy)^2 \\ 2-6xy-3x^2z & -3x^2 & -x^3 \end{pmatrix}$$

is -2 but the function sends the three points $(0,0,-1/4)$, $(1,-3/2,13/2)$, and $(-1,3/2,13/2)$ to $(-1/4,0,0)$ so is not invertible.

August 1, 2026

The 10 advances, including results in discrete geometry, algebra, graph theory, and complexity theory, were presented in a 253 page paper, with each result accompanied by a formalization in `Lean`.

We provide new results for the following problems. The results were achieved by an internal version of Astra, our next major model. The total number of tokens needed to find solutions to these problems would cost roughly $2,000 at Sol API rates. These arguments were then prepared into manuscripts by humans with the same model. Afterward, the model formalized each argument in a Lean certificate. We are also releasing for each solution a model's narration of its thinking process.

Sphere packing

In 2003, Henry Cohn and Noam Elkies established an upper bound LP_d for the maximal density of packing d-dimensional unit balls in d-dimensional space.
In 2016, Maryna Viazovska used this to find the densest possible packings in 8 and (with collaborators) 24 dimensions.
The new result calculates the exponential decay rate of this upper bound

$$\lim_{d\to\infty} LP_d^{1/d} = \sqrt{\frac{e}{2\pi}}.$$

ten-proofs / **SpherePacking.lean**

balexeev-oai .

Code | Blame | 55616 lines (52794 loc) · 2 MB

```
1  /-
2  Released under Apache 2.0 license as described in the file LICENSE.
3  SPDX-License-Identifier: Apache-2.0
4
5  The majority of this file, including the proof and its supporting developments,
6  is original work by OpenAI.
7
8  Some sphere-packing fundamentals were adapted from Sphere Packing in Lean and
9  Math, Inc.'s related fork:
10 https://github.com/thefundamentaltheor3m/Sphere-Packing-Lean
11 https://github.com/math-inc/Sphere-Packing-Lean
12
13 The reused upstream materials were released under the Apache License, Version
14 2.0. The original files were modified and adapted for this formalization;
15 their copyright and attribution notices are retained below.
16
17 Copyright (c) 2024 Sidharth Hariharan. All rights reserved.
18 Original authors of those definitions: Sidharth Hariharan and Gareth Ma.
19
20 Upstream project authors: Christopher Birkbeck, Sidharth Hariharan, Seewoo Lee,
21 Gareth Ma, Bhavik Mehta, and Maryna Viazovska.
22 The Math, Inc. fork includes additional formalization work by its Gauss agent.
23 -/
```

August 10, 2026

Learning more about Claude's mathematical capabilities

Aug 10, 2026

The Riemann zeta function

The Riemann zeta function describes the distribution of prime numbers: each place that the function takes the value of zero contributes successively finer detail to the sequence of primes. The Riemann hypothesis is that the zeros that determine the primes all exist along a certain vertical line. This has become one of the most consequential conjectures in mathematics: many results assume it in order to provide a form of randomness in the primes.

No one has yet been able to prove or disprove the Riemann hypothesis, but mathematicians have made progress in many related directions studying the Riemann zeta function and its zeros. One of these, as above, is quantifying a minimum proportion of zeros that are on the line: over time, they've gradually increased this known constant proportion to 41.6%.

Recently, a member of staff at Anthropic gave Claude an unreasonable challenge. It was about one of the most famous unsolved problems in mathematics: *Take a real stab at the Riemann hypothesis*.

Claude did take a real stab, but as you might have expected if you're familiar with the difficulty of the task (the Riemann hypothesis dates back to 1859 and has a million-dollar bounty), it didn't succeed. Nevertheless, during its attempt, it unexpectedly made strides on a related problem.

An unreleased research version of Claude has improved on a longstanding lower bound for the fraction of zeros of the Riemann zeta function that satisfy the Riemann hypothesis. Drawing on extensive prior research by mathematicians over the past decades, it has increased this bound from 41.6% to 67.2%.

The Riemann Hypothesis

The Riemann Hypothesis is a conjecture about the zeros of the Riemann zeta function defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p \text{ prime}} \frac{1}{1 - p^{-s}}$$

for complex numbers whose real part is greater than 1 and extended to the entire complex plane (except the point $s = 1$) by a process called analytic continuation. The Riemann zeta function has zeros at negative even integers, and all other zeros are known to have real part between 0 and 1. The Riemann Hypothesis asserts that all of these remaining zeros have real part equal to $1/2$.

August 23, 2026

levent @__alpoge__

Please welcome to the world a beautiful new geometric object, to do with a problem i've always loved. claude really contains multitudes:D
Does S^6 admit a complex structure?

Yup

5:31 PM · Aug 23, 2026 · **1.3M** Views

The literature, dating back to 1947, contains several mistaken disproofs. Boris Alexeev from OpenAI produced a `Lean` formalization four days later.

Formalization of the solution to the Hopf problem

The six-sphere admits a complex manifold structure compatible with its standard topology.

Based on [A compact complex threefold fibred by tori over the projective line, and the six-sphere], originally [shared on X] by [Levent Alpöge].

The repository includes a Comparator setup, with the statement adapted from the [Formal Conjectures project].

```
lake update
lake exe cache get
lake build lean4export
lake exe comparator comparator/config.json
```

[Type-check it online!]

levent @__alpoge__ · Aug 23

i always like to have short certificates that people can use to verify in full with their expertise and/or with one of today's super amazing models. this one is a little hard because (a pattern?) it's a problem with a lot of claims either way in the literature that fail due to subtleties. still im gonna try, two screenshots attached, but i also got opus 5 to write every detail down in a certainly far too long 100+ page pdf

the first page, or at most the first two, have enough to replicate i think

22 · 61 · 554 · 149K

levent @__alpoge__ · Aug 23

here's the pdf: alpo.ge/s6.pdf

The Hopf problem

The 6-sphere is the space containing all points of distance 1 from the origin in 7-dimensional space

$$S^6 := \left\{ (x_1, \dots, x_7) \in \mathbb{R}^7 \mid \sqrt{x_1^2 + \cdots + x_7^2} = 1 \right\}$$

An ant living on this sphere would think they were living in 6-dimensional space: S^6 is a 6-dimensional real manifold.

The Hopf problem asks whether S^6 admits the structure of a 3-dimesional complex manifold. An "almost complex structure" on S^6 is determined by identifying S^6 with the unit imaginary octonions, but that was known not to give a genuine complex manifold structure.

August 31–September 3, 2026

BOUNDED GAPS BETWEEN PRIMES

JULIA STADLMANN

Abstract. Polymath 8b [5] proved that $H_1 = \liminf_{n\to\infty}(p_{n+1}-p_n) \leq 246$. In this paper we show how the Bombieri-Vinogradov theorem can be combined with newer equidistribution estimates for smooth moduli to obtain the improved bound $H_1 \leq 240$.

The value 246 corresponds to the length of the shortest admissible 50-tuple, while 240 is the length of the shortest admissible 49-tuple. The bound $H_1 \leq 240$ is therefore the smallest possible improvement. However, while Polymath [5] used symmetric polynomials of degree ≤ 27 to achieve $H_1 \leq 246$, we only needed polynomials of degree ≤ 21 to get $H_1 \leq 240$. As such, we believe that our method could achieve bigger improvements, but due to limited resources and limited time, we have decided to present this modest improvement as a proof of concept.

Axiom ✓
@axiommathai

1/ ✨ World record today on bounded gaps between primes, one of number theory's oldest open conjectures.

We show that infinitely many pairs of prime numbers are separated by 212 or less.

10:56 AM · Sep 3, 2026 · **259K** Views

Weijie Su ✓
@weijie444

Announcing that GPT-6 Astra has pushed the prime gap to 186, with Lean formalization!

I was 9 when I first heard the twin prime conjecture. Its elegance and Yitang Zhang's legendary story have always stuck with me. A truly surreal night, being the first to see our model make progress, pushing 246 all the way down to 186, on a problem I've revered since I was a kid.

For me, it felt like witnessing a new era of intelligence being born, made possible by everyone at @OpenAI!

3:49 PM · Sep 3, 2026 · **146.8K** Views

Bounded gaps between primes

Unable to get an academic position, Zhang kept the books for a Subway franchise.

Photograph by Peter Bohler

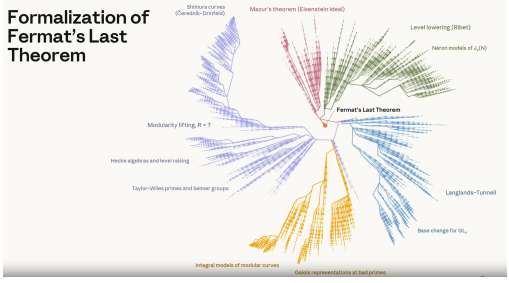
We have known, since at least Euclid's *Elements*, that there are infinitely many primes.

The twin primes conjecture asserts that there are infinitely many pairs of primes who difference is just 2, despite the fact that the prime numbers become sparser as they get larger.

In 2013, Yitang Zhang established the first finite bound on a gap between infinitely many pairs of consecutive primes at around 70 million.

In 2014, this bound was improved by the collaborative Polymath project to 246.

September 4, 2026



We are sharing the first complete computer-checked proof of Fermat's Last Theorem. Claude worked largely autonomously over 11 days to write the proof in the Lean programming language. Below, we describe how the formalization was done and share some thoughts about what this work could mean for research mathematics.

FLT: Anthropic has beaten me to it

Posted on September 4, 2026 by xenaproject

I guess technically it was revealed to the world by a coffee shop in Islington on Insta, but an hour later it was officially announced by Anthropic: one of their internal models, using the prove2.me platform, has formalized a complete proof of Fermat's Last Theorem (FLT) in Lean. This is the final theorem to be formalized in Freek Wiedijk's famous list of 100 formalization challenges and thus wraps up this 20-year-old benchmark. Congratulations to Anthropic!

Fermat's Last Theorem

Around 1637, Pierre de Fermat claimed to have a proof that there are no positive integers solutions to the equation $a^n + b^n = c^n$ when $n > 2$.

This was finally proven in 1995 by Andrew Wiles with help from Richard Taylor.

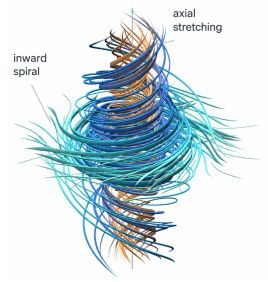
The interest in this problem is not just the solution but the massive amount of number theory that was developed along the way. Kevin Buzzard is formalizing, a different, more modern, version of the proof.

September 8, 2026

September 8, 2026 Research Publication

On the Navier–Stokes Millennium Prize Problem

Read the paper | Link to Lean formalized proof >



On Tuesday, September 1, we heard rumors that two Millennium Prize problems had been resolved. Inspired by these rumors and by the step change in performance of our internal model, we launched an effort to evaluate it on all open Millennium Prize problems and a few other high-impact problems.

We used a system of coordinating agents powered by our internal model. The agents had access to tools such as the ability to read from a cached version of the internet and the ability to run code. Agents were subdivided into groups with the ability to communicate within the group. The groups varied in size, and the group that produced the Navier–Stokes resolution involved on the order of 10,000 concurrent agents. At all times we maintained the same strict safeguards that we apply to all our frontier model evaluations, including monitoring and isolation.

...

The agents arrived at their resolution on Saturday, September 5, about 88 hours after the first agents were launched. Lean formalization and verification took an additional 17 hours via GPT-6 Astra.

Across all attempted problems, the agents sent 4.9 million messages and used about 300 billion output tokens. In the process of resolving the Navier–Stokes problem, the agents sent 2.7 million messages and used approximately 130 billion output tokens.

The Navier-Stokes problem

ARTIFICIAL INTELLIGENCE

AI Has Solved One of Math's $1 Million Millennium Prize Problems

By KONSTANTIN KAKAES | 10 |

Mathematicians at OpenAI showed that the Navier-Stokes equations, which describe how fluids flow, can sometimes "blow up." But the massive result is not without controversy.

The Navier-Stokes equations are differential equations that use Newton's second law of motion to describe how fluids, from ocean currents to air flows, behave. They were first written down in the mid-19th century, and have been central to the study of fluid mechanics ever since. But one basic question about the equations has persisted: Are their solutions always well-behaved? Or can their solutions evolve over time so that some infinitesimally small part of the fluid begins to flow infinitely quickly, creating a so-called singularity?

Quanta magazine

2

How the mathematical community is responding

How do we incentivize useful scientific activities?

Tony Feng @tonylfeng · Aug 25

I don't repost mathematics if I do not personally feel sure of its correctness, so I did not yet comment on the purported construction of a complex structure on S^6. However, I wanted to add some contextualization because I don't agree with the social media narrative that is developing.

0. Of the AI breakthroughs so far it's the closest to my interests, although still pretty far. I am not grieving, I am interested to learn more about it! I don't know if it's correct but on cursory glance it seems quite plausible.

1. If correct, this would be a major development. I believe it is a central problem in its field (complex differential geometry), like the Unit Distance Conjecture was for combinatorial geometry. Perhaps even more famous names are attached to this field, leading some people to say it's even more prestigious, but I prefer not to think in such terms.

2. There are two writeups, a 2-page one and a 100+-page one. Both are terrible and seem deliberately constructed to be inhumane. Fortunately, one can fight AI with AI: I asked Claude/GPT/Gemini to collaborate on making a more readable and enlightening version, which you can find at overleaf.com/read/frddxmbnq.... To make this, I prompted the models to write a document in the style of one of my own surveys. I haven't checked the mathematics, but it seems less forbidding than the original document at least.

3. This is a great opportunity for an expert to write an illuminating account of the construction, explain the key ideas and their proper historical context, and summarize the takeaways and lessons learned for the mathematics community.

4. If I see such an account, I would be happy to arrange for the author to give a lecture series in Berkeley. If this lecture series is also wonderful, I would go further and advocate for the author to have a position at Berkeley (if desired). For contrast, I see little value in having colleagues who are good at being the first to prompt AI for a solution.

COMPLEX STRUCTURES ON $\mathbb{S}^6$

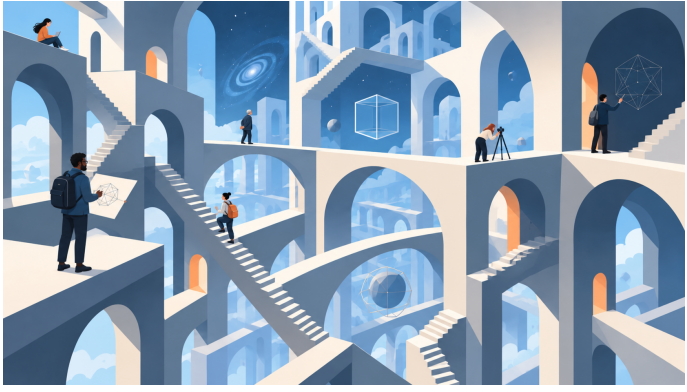
PHILIP ENGEL

ABSTRACT. A recent manuscript produced by Claude under the direction of Levent Alpöge constructs a 1-parameter family of compact complex threefolds diffeomorphic to $\mathbb{S}^6$ [AC26]. Our goal is to describe the geometric ideas behind the construction, give a self-contained proof, and remark on possible future research directions.

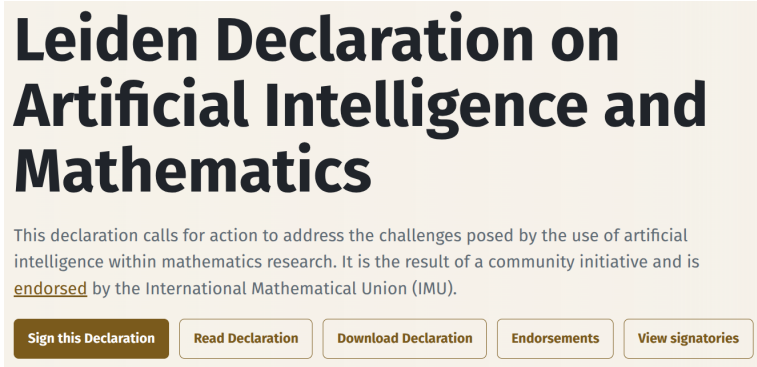
How do we articulate the value of human mathematicians?

The Role of Applied Mathematics in a New Era of Artificial Intelligence

SIAM AI Task Force Report | February 2026



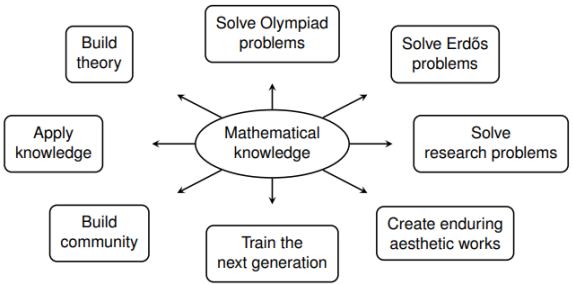
What norms should we establish around AI use?



Recommendations for individual mathematicians are to disclose tool use, retain responsibility for correctness, affirm humanity of authorship, put effort into proper attribution, participate in public discourse, stay informed about emerging technologies, ...

What are the goals and values motivating mathematical research?

Excessive AI optimization may in fact cause the many (previously largely aligned) goals of mathematics to diverge[3] from each other.



[3]The diagram is extremely oversimplified; in particular, it should be much higher dimensional.

From a talk entitled Mathematics in the Age of AI delivered by Terry Tao at the 2026 International Congress of Mathematicians.

What should we tell our students?

The Enduring Value of Math, in an Age of AI

Thoughts I'll share at the start of my upper division math course

FRANCIS SU

AUG 20, 2026

- *Why get good at problem-solving if an AI can do that much more efficiently?*
- *Why construct mathematical arguments, when AI models can now produce cleaner, slicker, more well-reasoned arguments than I can?*
- *What's the point of being a math major?*

The good news is that doing mathematics is not ultimately just about the particular problem-solving skills you are learning, nor about the outputs you are producing. It's actually more about the kind of person you are becoming. Said another way, *mathematical thinking is a set of virtues, not just a set of skills.*

Open problems

There are many important questions that are yet to be answered:

- How can we collectivize and democratize access to resources?
- How do we reinforce open scientific practices, such as sharing work in progress?
- How do we save the mathematics literature?
- How should we update our criteria for admissions / hiring?
- What should be the requirements for a PhD thesis?
- Do we need to abandon the practice of attributing authorship to mathematical discoveries?

Tack så mycket!