MASCAMP · ONLINE SEMINAR
”Mathematics for Scattering Amplitudes”
2 November 2026

→ *meet the agent*

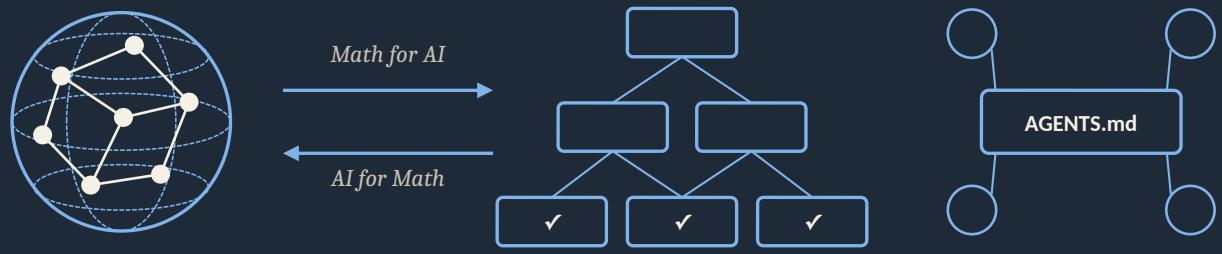
reads · proposes · checks

# *Math for AI vs AI for Math*

## Symmetry for networks, agents for mathematics

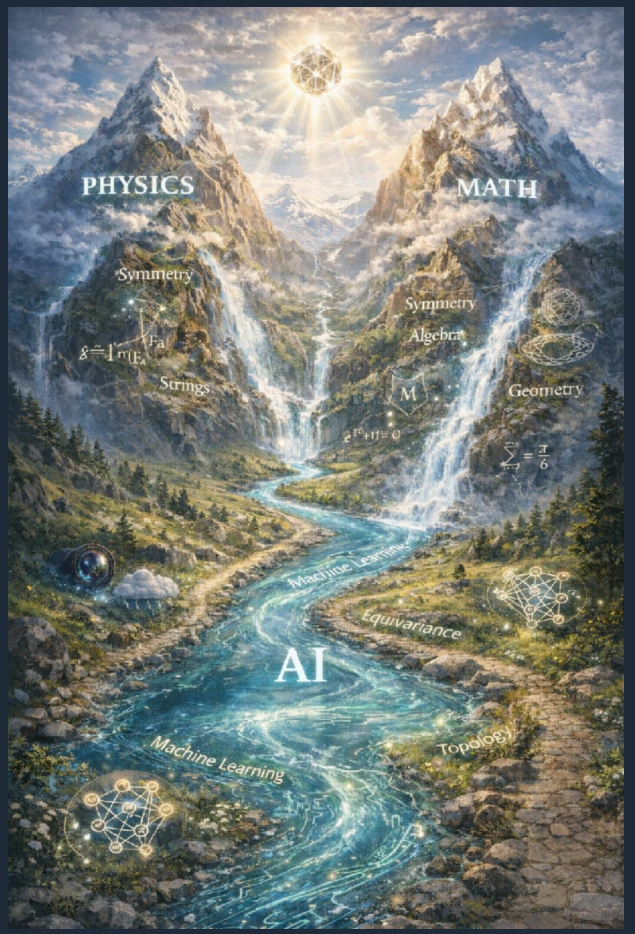
**Daniel Persson**

Chalmers University of Technology



HOW THE LAST FEW YEARS HAVE FELT

# *The water is rising*

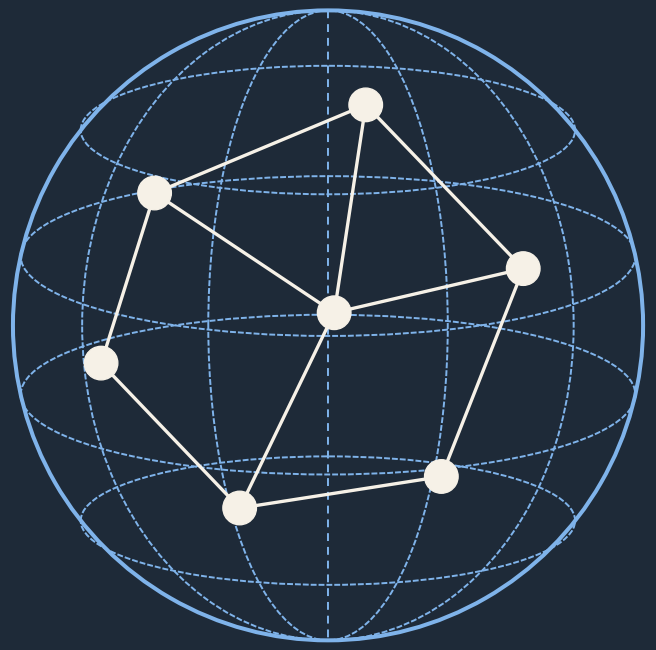





PART I

# Math for AI

***Symmetry, geometry and physics as inductive bias: what my research group does, and where it has led.***

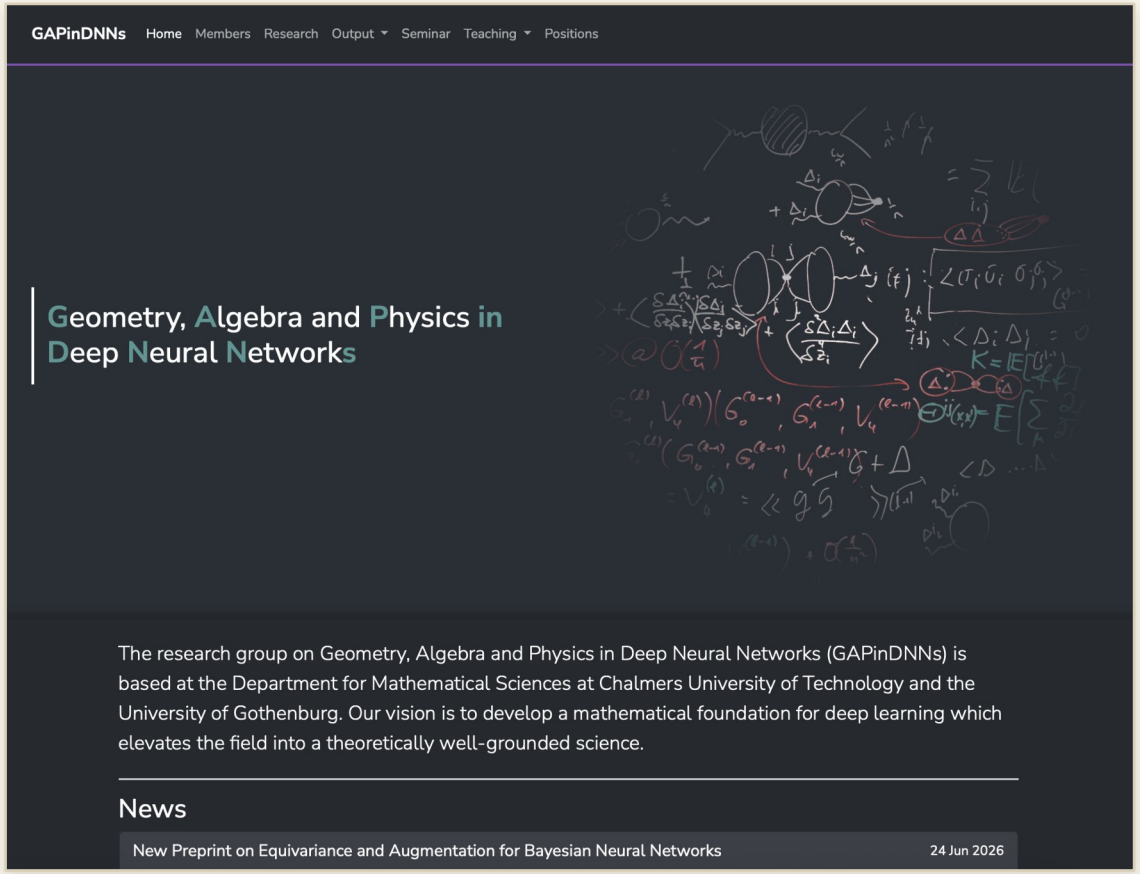


$f(T\,x) = T\,f(x)$

MY RESEARCH GROUP

MATH & PHYSICS → AI

# *Mathematics and physics for deep learning*

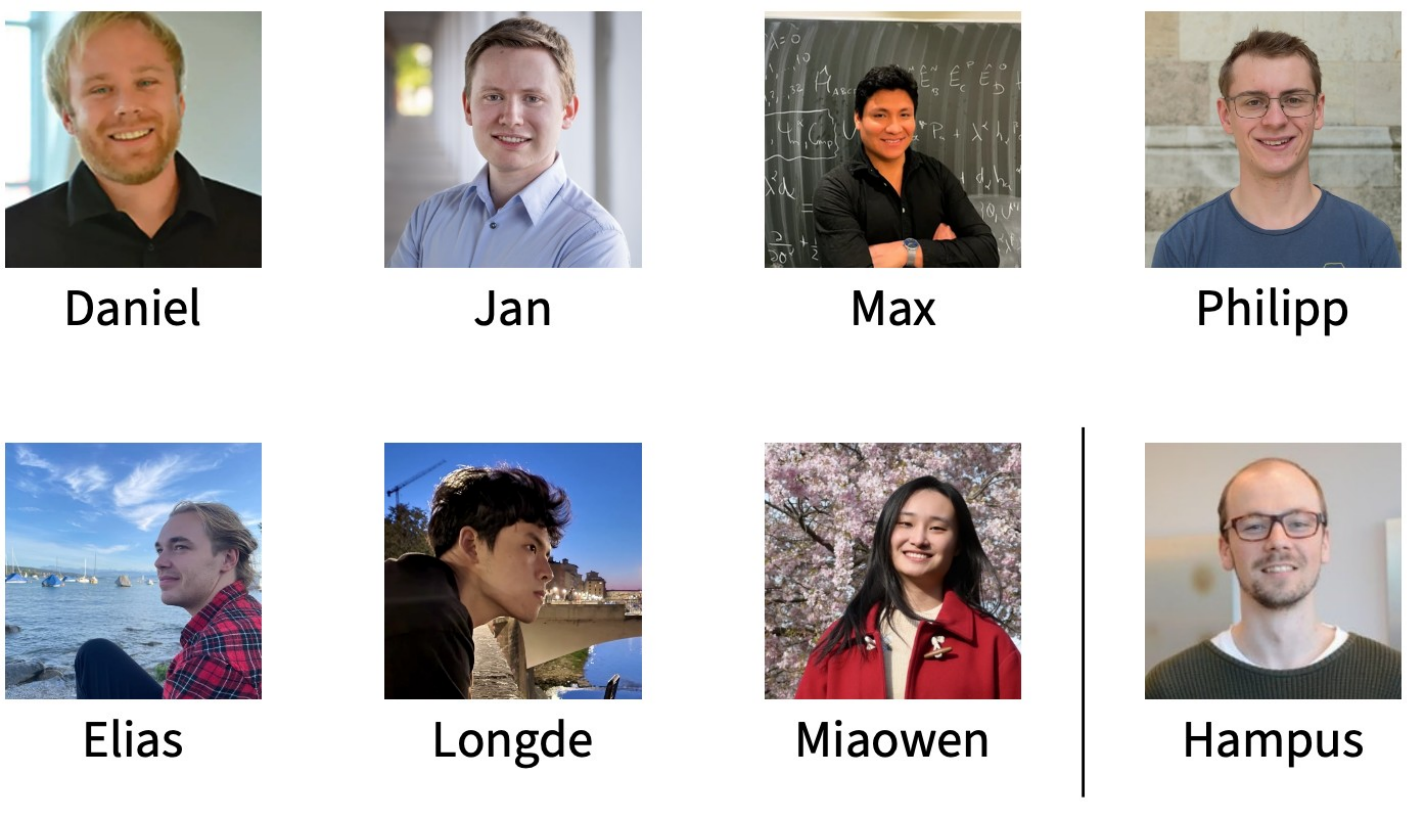



## *Geometry, Algebra and Physics in Deep Neural Networks*

- **GAPinDNNs** — our research group at Chalmers & the University of Gothenburg
- **Vision:** a mathematical foundation for deep learning
- Symmetry, geometry and physics as guiding principles for neural networks

***the direction: mathematics & physics in service of AI***

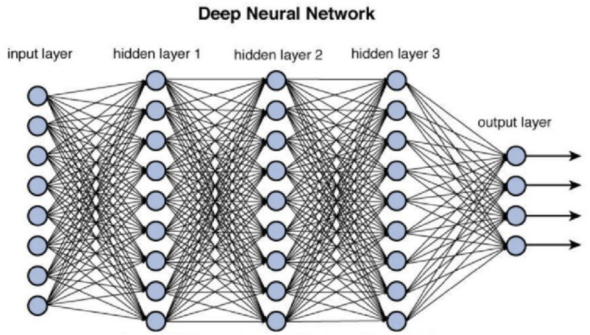
# *GAPinDNNs, autumn 2026*




*Daniel Persson and Jan Gerken (faculty), Max Guillen (postdoc), Philipp Misof, Elias Nyholm, Longde Huang and Miaowen Dong (PhD students), Hampus Linander (guest researcher, right of the line). Chalmers and the University of Gothenburg; gapindnns.github.io.*

# *Inductive bias: build the math into the network*






## Inductive bias

Incorporate known constraints into the architecture/training

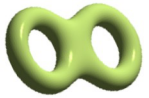
***symmetries***



***differential equations***

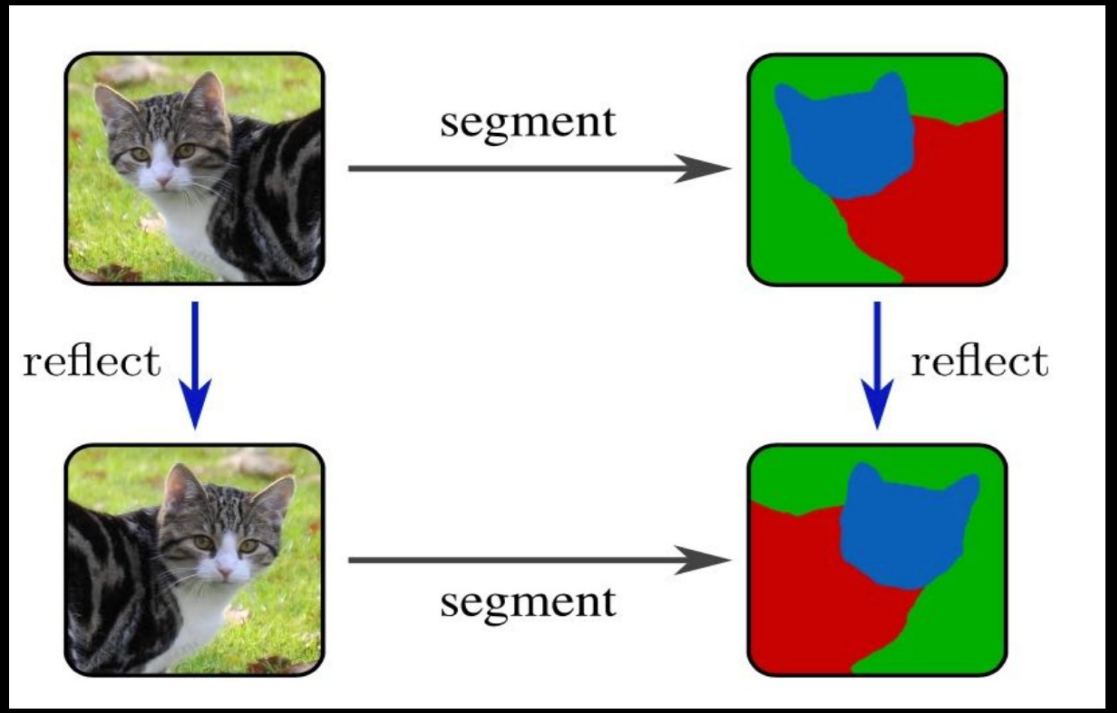
$$R_{\mu\nu}(g) - \frac{1}{2} g_{\mu\nu} R(g) = T_{\mu\nu}(g, F, \dots)$$

***geometry***



# *Invariance and equivariance*

**Segmentation**

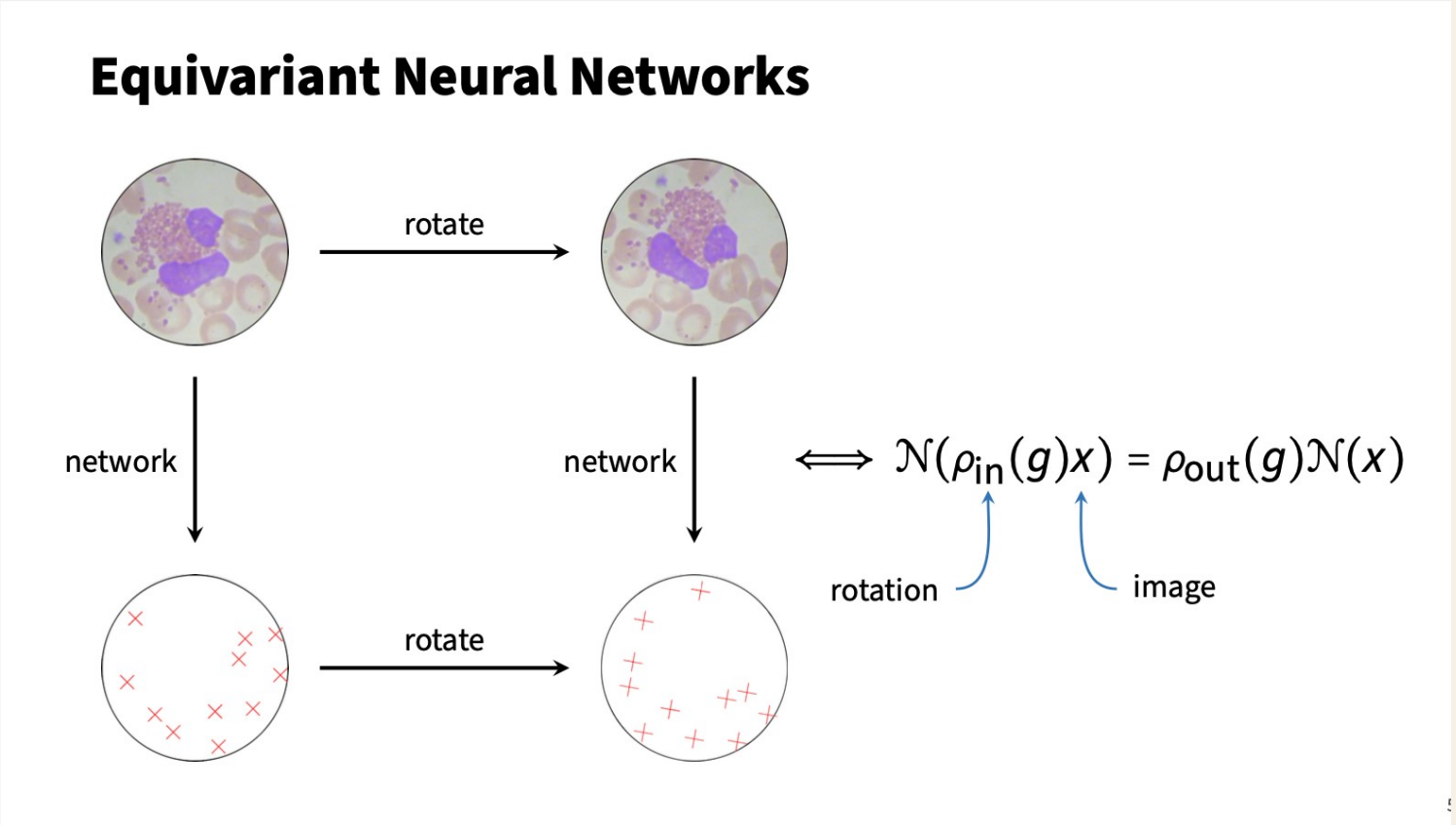



commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \downarrow{\scriptstyle T} & & \downarrow{\scriptstyle T} \\ X & \xrightarrow{f} & Y \end{array}$$

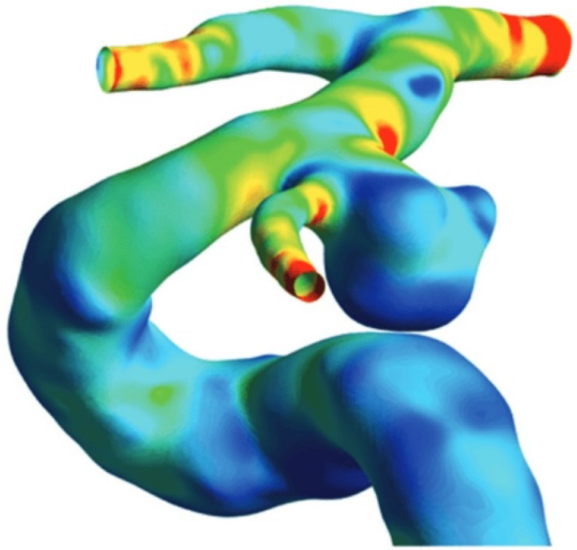
$$f(Tx) = Tf(x)$$

# Rotate the image, and the detections rotate with it

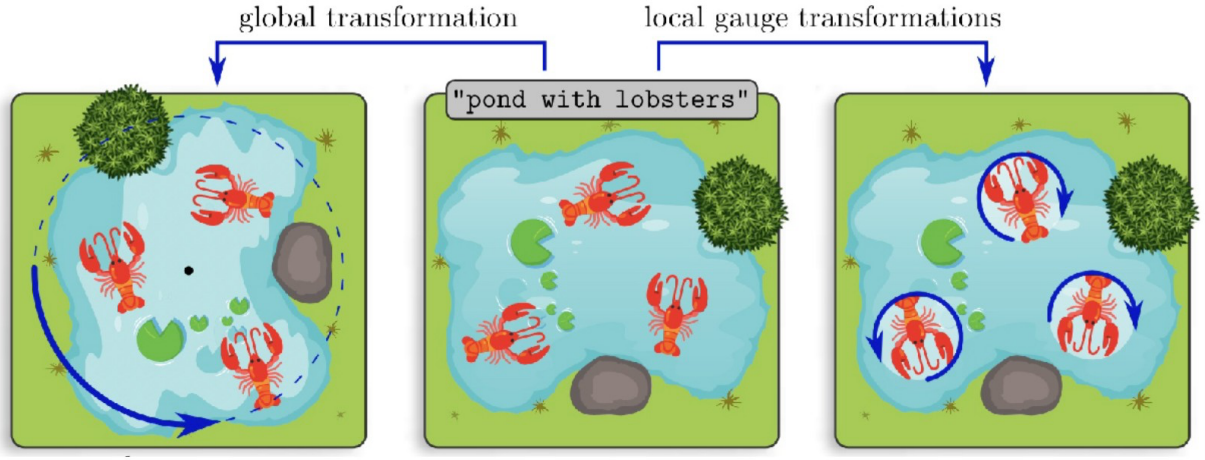



*Slide from Jan Gerken's lectures on equivariant neural networks (GAPinDNNs). Cells on a blood smear, marked by a network that commutes with rotation.*

# Geometric deep learning - A unified framework for deep learning



***Principle of geometric deep learning:***
The equations governing neural networks should be equivariant with respect to all **local** and **global** symmetries of the input data




[Images from: Weiler, Forré, Verlinde, Welling (2023)]

MATH & PHYSICS → AI

# Geometric deep learning in practice

## HEAL-SWIN: A Vision Transformer On The Sphere

Oscar Carlsson[*ab] Jan E. Gerken[*a] Hampus Linander[a] Heiner Spieß[c] Fredrik Ohlsson[d]
Christoffer Petersson[ea] Daniel Persson[a]

**Abstract**

*High-resolution wide-angle fisheye images are becoming more and more important for robotics applications such as autonomous driving. However, using ordinary convolutional neural networks or vision transformers on this data is problematic due to projection and distortion losses introduced when projecting to a rectangular grid on the plane. We introduce the HEAL-SWIN transformer, which combines the highly uniform Hierarchical Equal Area iso-Latitude Pixelation (HEALPix) grid used in astrophysics and cosmology with the Hierarchical Shifted-Window (SWIN) transformer to yield an efficient and flexible model capable of training on high-resolution, distortion-free spherical data. In HEAL-SWIN, the nested structure of the HEALPix grid is used to perform the patching and windowing operations of the SWIN transformer, enabling the network to process spherical representations with minimal computational overhead. We demonstrate the superior performance of our model on both synthetic and real automotive datasets, as well as a selection of other image datasets, for semantic segmentation, depth regression and classification tasks. Our code is publicly available[1].*

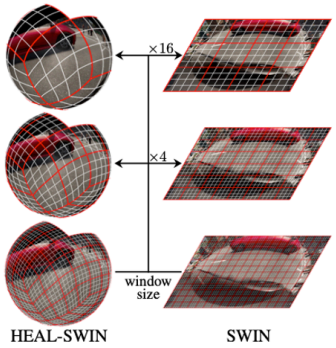



Figure 1. Our HEAL-SWIN model uses the nested structure of the HEALPix grid to lift the windowed self-attention of the SWIN model onto the sphere.

vision and autonomous driving datasets do not contain fisheye images. For these reasons, fisheye images have received

*HEAL-SWIN: computer vision on the sphere*

## PEAR: Equal Area Weather Forecasting on the Sphere

Hampus Linander[1,2] Tage Tykesson[2,4] Pietro Rosso[3]
Christoffer Petersson[2,5] Daniel Persson[2] Jan E. Gerken[2]

[1] VERSES AI
Los Angeles, USA

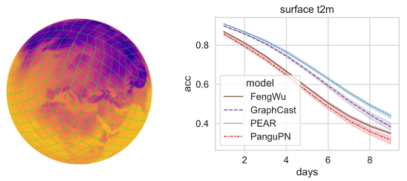
[2] Department of Mathematical Sciences
Chalmers University of Technology
University of Gothenburg, Sweden

[3] Department of Computer Science and Engineering
[4] Department of Physics and Astronomy
Chalmers University of Technology, Sweden

[5] Recohere
Gothenburg, Sweden

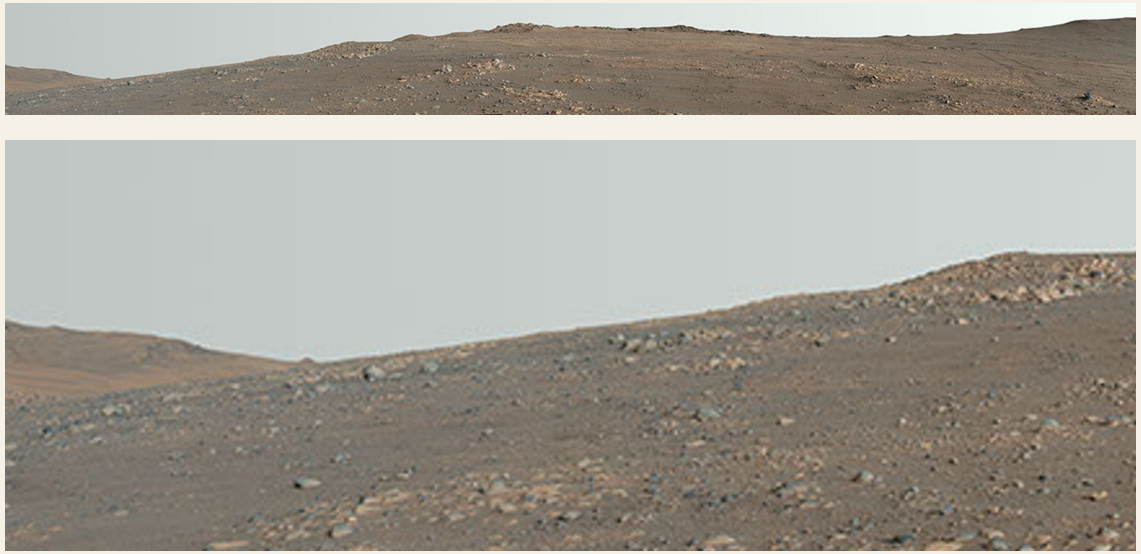
**Abstract**

Artificial intelligence is rapidly reshaping the natural sciences, with weather forecasting emerging as a flagship AI4Science application where machine learning models can now rival and even surpass traditional numerical simulations. Following the success of the landmark models Pangu Weather and Graphcast, outperforming traditional numerical methods for global medium-range forecasting, many novel data-driven methods have emerged. A common limitation shared by many of these models is their reliance on an equiangular discretization of the sphere which suffers from a much finer grid at the poles than around the equator. In contrast, in the Hierarchical Equal Area iso-Latitude Pixelization (HEALPix) of the sphere, each pixel covers the same surface area, removing unphysical biases. Motivated by a growing support for this grid in meteorology and climate sciences, we propose to perform weather forecasting with deep learning models which natively operate on the HEALPix grid. To this end, we introduce Pangu Equal ARea (PEAR), a transformer-based weather forecasting model which operates directly on HEALPix-features and outperforms the corresponding model on an equiangular grid, and other baselines, without any computational overhead. Furthermore, we perform numerical experiments on the equivariance properties of our setup and verify the performance of PEAR on climate model emulation.
Code: `https://github.com/hlinander/PEAR-Weather`




*PEAR: equal-area AI weather forecasting*

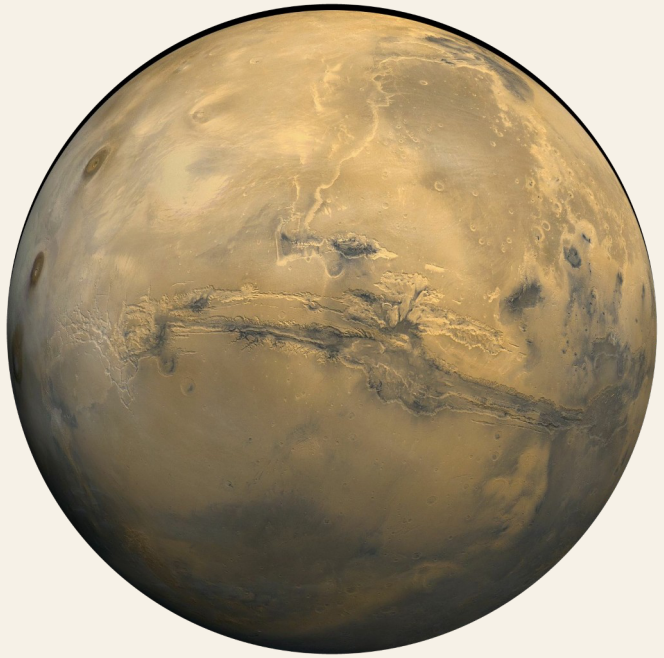
# *Geometric deep learning in practice*



THE SURFACE · PERSEVERANCE, 5 APRIL 2026

*"Arbot", west of Jezero Crater: a Mastcam-Z panorama of 46 images in enhanced colour, and its left-hand third.*

*NASA/JPL-Caltech/ASU/MSSS (PIA26753), public domain.*



NEXT · FROM EARTH TO MARS

*A new research project in the group: weather and climate on Mars.*

*Mars, Viking 1 orbiter mosaic. NASA/JPL-Caltech, public domain.*

# *Ultrasound: a skill that does not scale*



Trauma care: six views in 90 seconds




Ultrasound machines are everywhere; using them well is hard



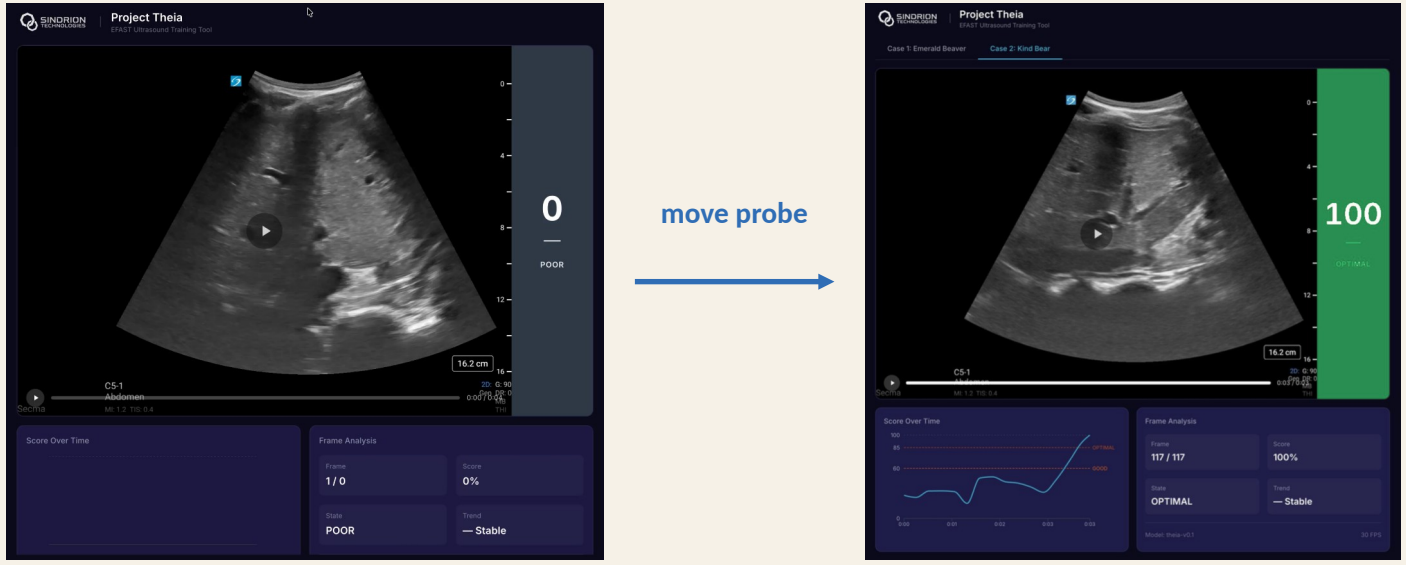

One expert for two to four trainees

**THE PROBLEM**

*eFAST, the focused ultrasound exam of the trauma patient: a few standard views, acquired fast, under pressure. The machine is cheap and in every emergency room; the skill of placing the probe is not, and teaching it takes an expert at the bedside for every few trainees.*

*Figures: Jan Gerken, GAPinDNNs group slides, 2026.*

# An AI tool to make ultrasound training scalable




*A neural network rates the probe position live, frame by frame, against the target view. Screenshots of the training tool.*



*Collecting training data at Sahlgrenska University Hospital, Gothenburg.*

**WHY IT IS A MATH-FOR-AI PROJECT**

*The same group, the same question: what structure of the data and the task should the network be told about, and what should it learn? Here the answer is judged at the bedside, with trainees in Gothenburg, not on a benchmark.*

*Figures: Jan Gerken, GAPinDNNs group slides, 2026.*

MATH & PHYSICS → AI

# *Kymeon Medical Technologies*

kymeon
MEDICAL TECHNOLOGIES

Knut and Alice Wallenberg Foundation

**Proof-of-concept grant, Wallenberg Launch Pad (WALP)**

**WHAT IT IS**

*A spin-out from the GAPinDNNs group and its clinical partners at Sahlgrenska: AI feedback for ultrasound training, so that one expert can supervise many trainees. The company is being incorporated this autumn.*

**FUNDING**

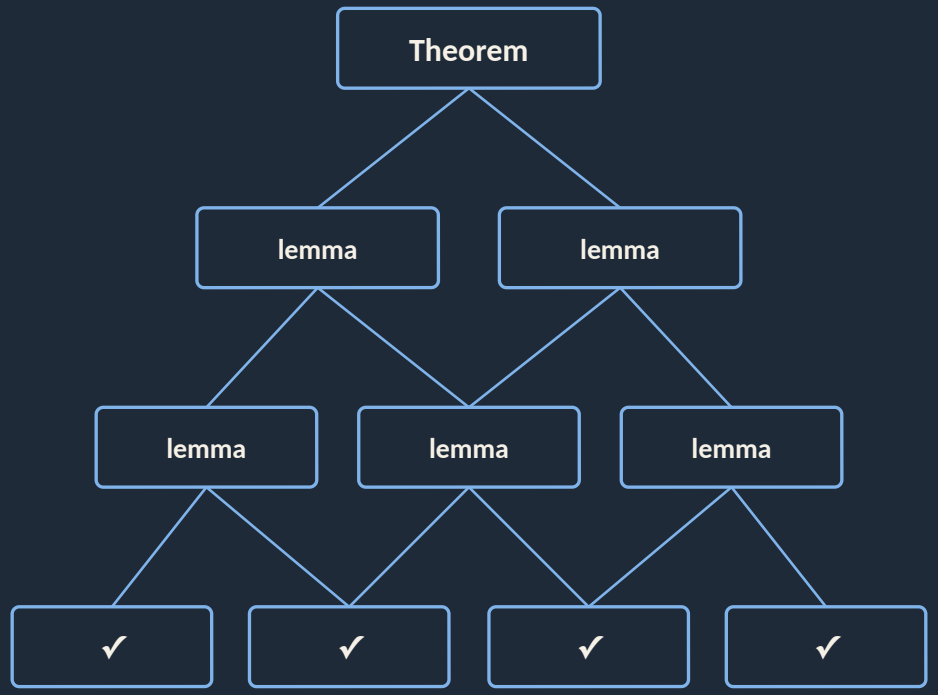
*Proof-of-concept grant from the Knut and Alice Wallenberg Foundation, through the Wallenberg Launch Pad (WALP), to take the research prototype to a product.*

*The arrow in this talk points both ways: mathematics into the network (Part I), and agents into mathematics (Part III). This is where the first arrow leaves the department.*

PART II

# AI for Math

***What has happened since the summer, and how the mathematical community is taking it.***




*every step checked, down to the axioms*

# *“A.I. can write poetry, but it struggles with math”*



**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

*The New York Times, 23 July 2024, Steve Lohr. Illustration: Sean Dong.*

**THE STATE OF PLAY, SUMMER 2024**

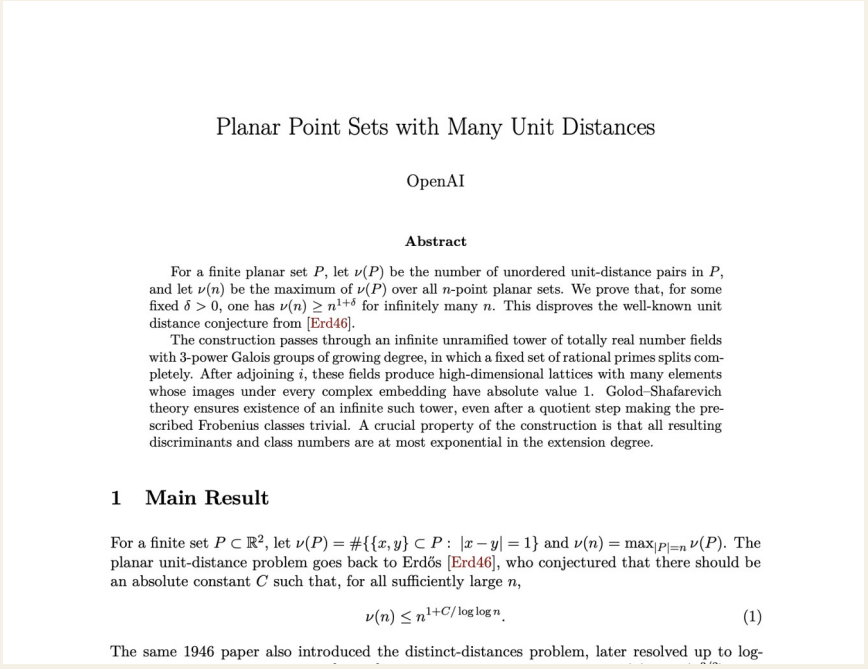
*Chatbots stumbled on simple arithmetic and on word problems that take several steps. The Times read this as a sign of how far the new technology had broken with computing's past: a machine that predicts text rather than one that calculates. OpenAI called mathematics “an important ongoing area of research”.*

**TWENTY-SEVEN MONTHS LATER**

*Every result in this part of the talk is younger than that article: a Fields medallist leaving for a lab on the day of the medal, a Millennium-Prize manuscript in 88 hours, 722 manuscripts in one evening, Fermat's Last Theorem and the classification of finite simple groups formalized by agents. Emily Riehl opened her own survey talk this autumn with the same headline.*

AI → MATH

# The first big crack: Erdős's unit-distance conjecture falls



Planar Point Sets with Many Unit Distances

OpenAI

**Abstract**

For a finite planar set $P$, let $\nu(P)$ be the number of unordered unit-distance pairs in $P$, and let $\nu(n)$ be the maximum of $\nu(P)$ over all $n$-point planar sets. We prove that, for some fixed $\delta > 0$, one has $\nu(n) \geq n^{1+\delta}$ for infinitely many $n$. This disproves the well-known unit distance conjecture from [Erd46].

The construction passes through an infinite unramified tower of totally real number fields with 3-power Galois groups of growing degree, in which a fixed set of rational primes splits completely. After adjoining $i$, these fields produce high-dimensional lattices with many elements whose images under every complex embedding have absolute value 1. Golod–Shafarevich theory ensures existence of an infinite such tower, even after a quotient step making the prescribed Frobenius classes trivial. A crucial property of the construction is that all resulting discriminants and class numbers are at most exponential in the extension degree.

**1 Main Result**

For a finite set $P \subset \mathbb{R}^2$, let $\nu(P) = \#\{\{x, y\} \subset P : |x - y| = 1\}$ and $\nu(n) = \max_{|P|=n} \nu(P)$. The planar unit-distance problem goes back to Erdős [Erd46], who conjectured that there should be an absolute constant $C$ such that, for all sufficiently large $n$,

$$\nu(n) \leq n^{1+C/\log\log n}. \quad (1)$$

The same 1946 paper also introduced the distinct-distances problem, later resolved up to log-

*"Plane Point Sets with Many Unit Distances", OpenAI, May 2026, released with a commentary by nine outside mathematicians.*

### THE PROBLEM, ERDŐS 1946

*Among n points in the plane, how many pairs can be exactly one unit apart? A slightly skewed square grid gives n^(1 + c/log log n), and Erdős conjectured that nothing does essentially better. The upper bound n^(4/3) has stood since 1984.*

### WHAT THE MODEL DID

*An internal OpenAI reasoning model built a configuration with n^(1+ε) unit distances, ε of order 10^(−38), from number fields richer than the Gaussian integers. Will Sawin (Princeton) then pushed the exponent to about 1.014. Gowers: "I would have recommended acceptance without any hesitation."*

### IN HINDSIGHT

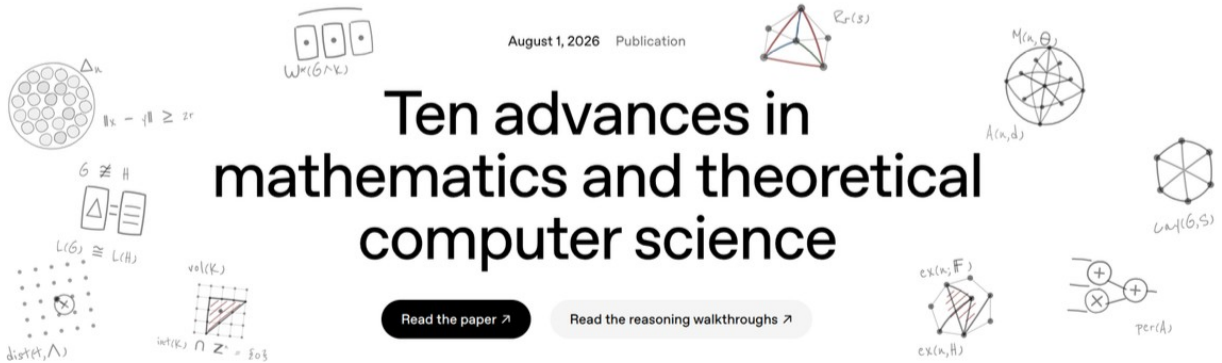
*The first big breakthrough of the paradigm shift: a named, eighty-year-old conjecture refuted by a general-purpose model, "a milestone in AI mathematics" (Gowers). Everything in the rest of this part happened in the five months that followed. Aaronson, 27 May: "we're now over the top of this particular rollercoaster".*

***"we will transition from an era of proof scarcity to an era of proof abundance"***

*Terence Tao, ICM 2026*

AI → MATH

# *Ten advances, one announcement*




*openai.com, 1 August 2026.*

**WHAT THE DOCUMENT CLAIMS**

*Ten results in mathematics and theoretical computer science, "obtained by an internal OpenAI model". Not stated: how much human involvement there was, what it cost, and what changed in the revision of 6 August. Self-reported, and unrefereed at the time of release.*

We present a collection of results obtained by an internal OpenAI model, spanning mathematics and theoretical computer science:[1]

1. **High-dimensional sphere packing.** The asymptotic strength of the Cohn–Elkies linear program is determined exactly. This gives an improved general packing bound in high dimensions and settles the corresponding Fourier sign-uncertainty problem asymptotically.
2. **Binary and spherical codes.** Classical upper bounds for fixed-distance binary and spherical codes are improved by exponential factors for all parameters. The spherical construction also recovers the sphere-packing exponent of Chapter 1.
3. **Nonsofic groups exist.** An explicit nonsofic group is constructed, answering the longstanding question of whether every countable group is sofic. The proof combines work of Kun and Kun–Thom with a combinatorial expander-matching argument, realized using the binary Leavitt algebra and Thompson's group $V$.
4. **Connes's rigidity conjecture.** Infinitely many pairwise nonisomorphic groups with property $(T)$ are constructed with the same group von Neumann algebra, disproving Connes's conjecture and answering a related finite-to-one question of Popa.
5. **Arithmetic circuit complexity.** For the permanent, division-free circuits require $\Omega(n^2 \log\log n)$ gates, while formulas require $\Omega(n^4/\log n)$ leaves.
6. **Quantum parallel repetition.** Exponential parallel repetition is proved for every finite two-player entangled game, extending the classical repetition principle beyond previously treated special classes of quantum games.
7. **Closest vector problem.** A direct reduction from 3SAT gives $n^{1/400}$-factor hardness for the Euclidean closest vector problem, with related consequences for binary decoding and other lattice norms.
8. **Ehrhart's volume conjecture.** The sharp volume upper bound $(n+1)^n/n!$ is proved in every dimension for convex bodies whose barycenter is their only interior lattice point.
9. **Multicolor Ramsey numbers.** A new superexponential lower bound shows that $R_k(3)$ grows as $k^{\Theta(k)}$.
10. **Compactness and degeneracy.** Separate bipartite graph constructions disprove two conjectures in extremal graph theory: the compactness conjecture of Erdős and Simonovits and a degeneracy conjecture of Erdős.

*The abstract, in the authors' own words.*

***"much of the publicly available evidence is highly subject to reporting bias and non-scientific incentives"***

*Terence Tao, ICM 2026*

AI → MATH

# *The same day as the Fields Medal*



*Jacob Tsimerman, 2026*
*Photo: Simons Foundation, CC BY 4.0*

THE WALL STREET JOURNAL.

TECHNOLOGY • ARTIFICIAL INTELLIGENCE • SCIENCE OF SUCCESS

## The Math Superstar Who's Terrified of AI—and Just Took a Job at OpenAI

Jacob Tsimerman won the biggest prize in math. Now he's working on the most important problem of his career.

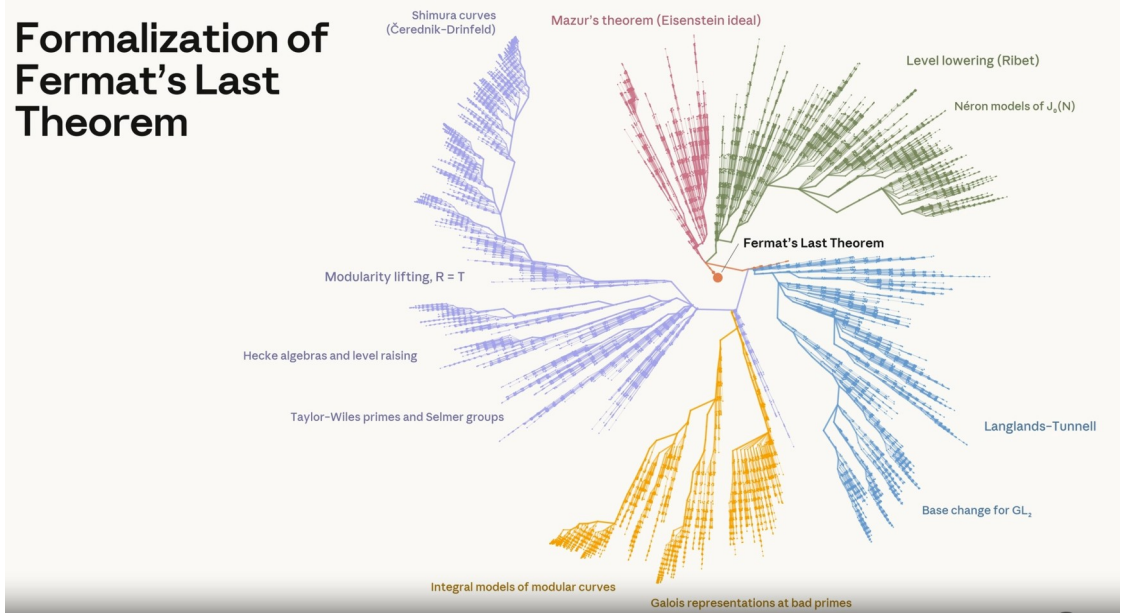
*The Wall Street Journal, 31 July 2026 · Ben Cohen*

***"The mathematical career, as we know it, I don't think it will exist in its current form." — Tsimerman***

*Announced 23 July at ICM Philadelphia, just after the medal. He takes leave from Toronto for OpenAI's safety team.*

AI → MATH

# *Fermat's Last Theorem, formalized in 11 days*




*anthropic.com, 4 September 2026. The dependency graph of the formal proof: about 29,500 theorems and 13 million lines of Lean, following Darmon, Diamond and Taylor's 1995 exposition of Wiles's argument.*

**WHAT WAS DONE**

*Dozens of Claude agents, coordinating through the Prove2Me platform (Tianyi Peng and collaborators, Columbia), with a handful of high-level human hints. Lean's kernel checked the result against Mathlib's statement of the theorem, on the three standard axioms only. It was the last open item on Freek Wiedijk's twenty-year-old list of 100 theorems to formalize.*

**THE PROJECT IT OVERTOOK**

*Kevin Buzzard (Imperial) has led a community formalization of FLT since 2024, on a five-year £1M EPSRC grant, along a different and more modern route. Anthropic's repository adapts pieces of that project. Riehl files the episode under "Bad math?: wasted resources", question mark included.*



**FLT: Anthropic has beaten me to it**

Posted on September 4, 2026 by xenaproject

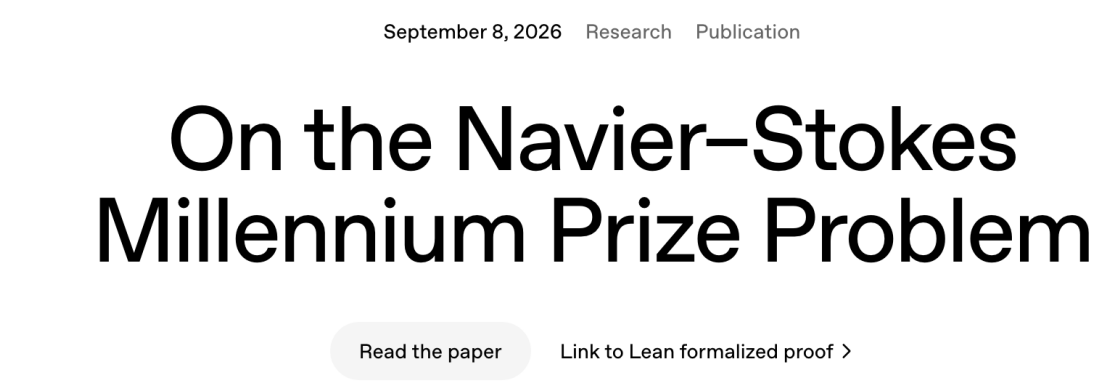
*Kevin Buzzard (photograph from Emily Riehl's slides). xenaproject.wordpress.com, 4 September 2026: "Note that mathematically this work of anthropic tells us essentially nothing."*

| 13 M | 29,500 | 11 | £1 M · 5 yrs | 2024 |
|---|---|---|---|---|
| *lines of Lean* | *theorems* | *days* | *Buzzard's EPSRC grant* | *the human project began* |

AI FOR MATH · 8 SEPTEMBER 2026

AI → MATH

# *A Millennium Prize problem, in 88 hours*



September 8, 2026 Research Publication

On the Navier–Stokes Millennium Prize Problem

Read the paper Link to Lean formalized proof ›

*openai.com, 8 September 2026. The paper's author line reads "OpenAI"; a Lean formalization is on GitHub.*

FINITE TIME BLOWUP FOR NAVIER–STOKES

OPENAI

Abstract. For every positive viscosity, we construct a solution of the three-dimensional incompressible Navier–Stokes equations that starts from rest and develops unbounded velocity in finite time while maintaining uniformly bounded kinetic energy.

*"Finite time blowup for Navier–Stokes", 166 pages, page 1.*

**WHAT WAS CLAIMED**

*Finite-time blowup from rest, under a smooth, compactly supported external force, with bounded energy: statements (C) and (D) of the Clay formulation. The unforced problem, statements (A) and (B), is untouched.*

**~10,000 agents · 88 hours · 17 hours of Lean · "several million dollars"**

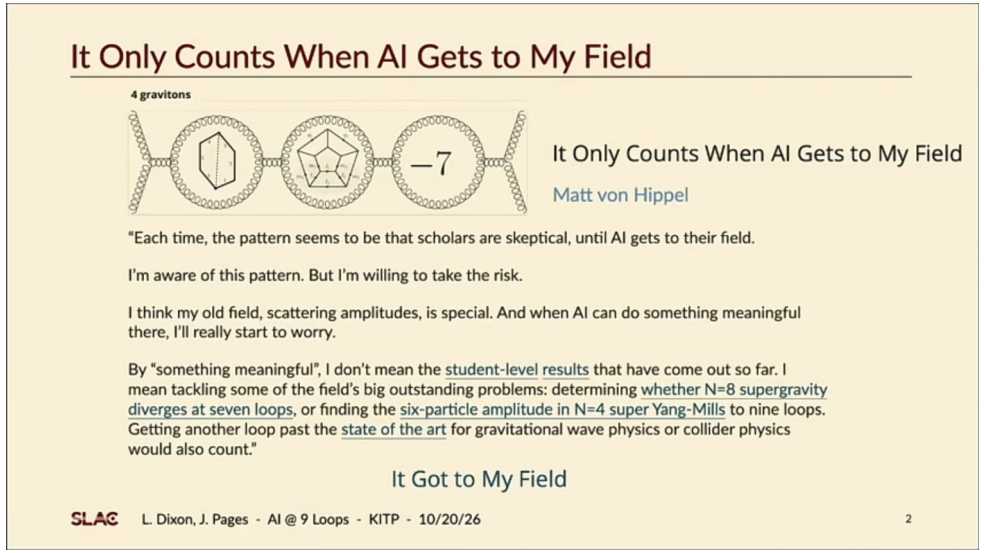
*OpenAI's own account; cost estimate by Sébastien Bubeck (OpenAI), Quanta, 8 September.*

*Status on 26 October: Lean-checked, but arXiv:2610.08144 finds the formal statement does not match the written proof. Not refereed. Clay: "apparently been settled", review "deliberately unhurried", problem still listed as active. OpenAI will not claim the prize.*

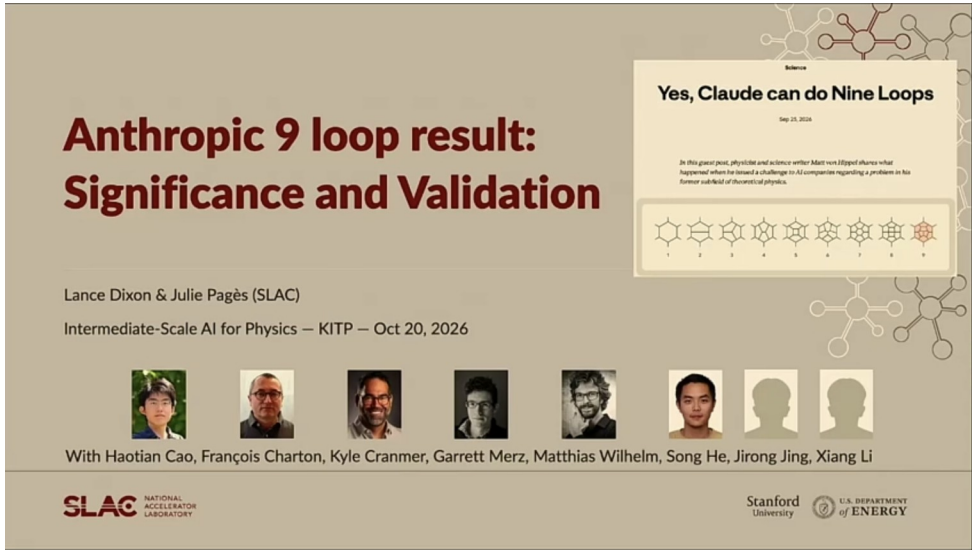
***"I was thrilled that the problem was solved."***

*Charles Fefferman (Princeton), author of the Clay problem statement. Quanta, 8 September 2026.*

# *Nine loops in N=4 super Yang–Mills, from one prompt*




*Lance Dixon and Julie Pagès, "AI @ 9 Loops", KITP, 20 October 2026, slide 2: Matt von Hippel's challenge of 7 August. Daniel's screenshots of the KITP recording.*




*The talk's title slide, with Anthropic's science-blog post of 25 September, "Yes, Claude can do Nine Loops", written by von Hippel at Anthropic's invitation, and paid for.*

**THE CHALLENGE · 7 AUGUST**

*Matt von Hippel, once an amplitudes theorist, on his blog 4gravitons: "Give us N=8 supergravity to seven loops, or N=4 super Yang-Mills to nine loops." The six-gluon MHV amplitude in planar N=4 SYM had stood at eight loops since Dixon and Liu, 2023.*

**WHAT HAPPENED · 25 SEPTEMBER**

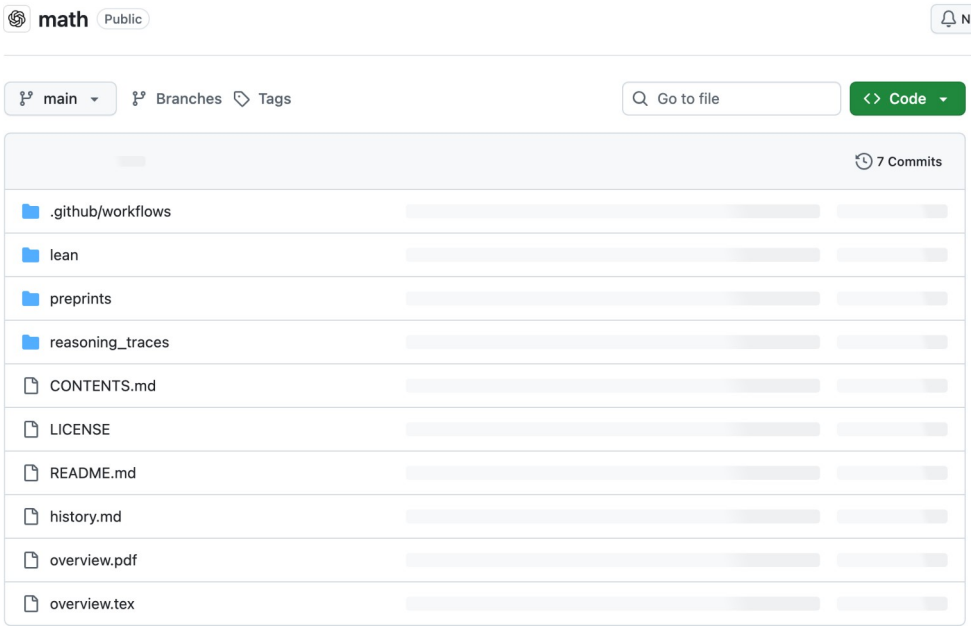
*Liam Fitzpatrick and Siddharth Mishra-Sharma (Anthropic) gave Claude the one-line problem and "keep working on this until I tell you to stop". It ran the hexagon bootstrap and the form-factor route unsupervised for days: one or two thousand dollars per route. Lance Dixon validated the result.*

**IN VON HIPPEL'S WORDS**

*"This was a lot easier than a Millennium problem." "There is more low-hanging fruit out there than you'd expect." "This doesn't tell you much about whether AI can achieve field-defining breakthroughs." Song He's group had most of the result independently; Dixon and Pagès reported on both at KITP.*

AI → MATH

# 722 manuscripts, one evening




*github.com/openai/math, read 26 October 2026. An unnamed internal model; "raw research data, rather than polished academic papers" (README). ~3 hours of compute per result, ~4,000 problems posed.*

NUMBER THEORY

## AI Proved the Quasi-Riemann Hypothesis. What Happens Now?

*Unlike other recent AI proofs, this one has been unpacked by human experts. The implications, one mathematician said, are "staggering."*

*Quanta, 19 October 2026 · Konstantin Kakaes*

**THE QUASI-RIEMANN HYPOTHESIS**

*ζ(s) and every Dirichlet L-function are zero-free for Re s > 7/8: 199 pages, Lean-verified. A 49-page human-edited proof for Re s > 11/12. No Landau–Siegel zeros. Experts largely accept it; Sarnak calls it "a Theorem".*

***"Within analytic number theory, it's obviously the best headline result for the past 100 years in the field."***

*James Maynard (Oxford, Fields Medal 2022), Science, 13 October 2026*

| 722 | 372 | ~42 % | 3 |
|---|---|---|---|
| *manuscripts* | *result families* | *Lean-formalized* | *withdrawn the next day* |

AI → MATH

# Excitement, and anxiety

## Excitement

*"I was like a kid at Christmas. Even if it was some scary monster giving me a gift, I was just so excited."*

*James Maynard, Oxford · Science, 13 October*

*"yesterday was surely one of the biggest days in mathematical history."*

*Scott Aaronson, UT Austin · Shtetl-Optimized, 7 October*

*"I didn't think I would live to see this result, and I'm not that old a guy."*

*Alex Kontorovich, Rutgers · Science, 13 October*

*"I'm more excited about staying in math than a few months ago."*

*Constantin Kogler, IAS, whose conjecture was among those solved · Science, 13 October*

## Anxiety

*"solutions to open problems are now being harvested at large scale in an unsustainable fashion"*

*Terence Tao, UCLA · Mathstodon, 6 October*

*"What OpenAI chose to do was to carpet-bomb mathematics, grabbing as many as possible of the chips on the table."*

*Peter Woit, Columbia · Not Even Wrong, 6 October*

*"I feel like I'm working in a toxic research environment. It's all just too fast."*

*Youness Lamzouri, Lorraine · Science, 13 October*

*"One must distinguish a claimed result, an independently verified result, and a result that has been understood and absorbed into mathematical culture."*

*Olga Holtz, Berkeley · Science, 13 October*

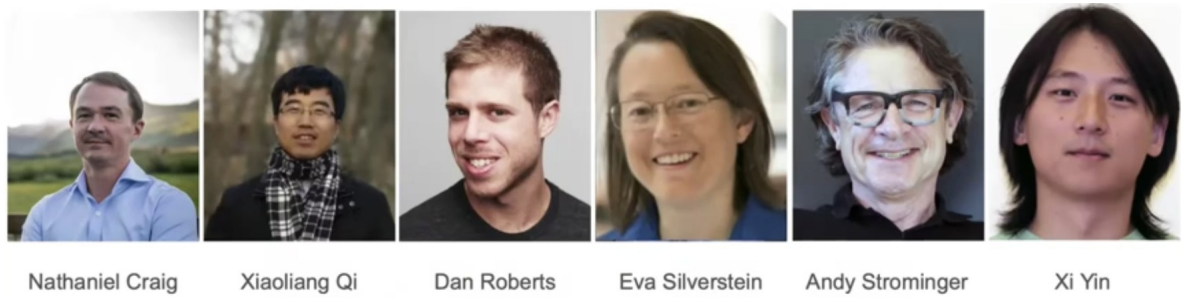
***"The goals of the AI companies and the goals of the mathematical community are severely misaligned."***

*"A Severe Misalignment of AI in Mathematics", open letter signed by 28 Fields medallists, 11 September 2026*

# *Theoretical physicists and the frontier labs, in one room*




*kitp.ucsb.edu, program page and online talks.*




*The six coordinators, from the opening session's slides (KITP online talks).*

### WHO IS COORDINATING

*Nathaniel Craig (UC Santa Barbara), Xiao-Liang Qi (Stanford), Dan Roberts (OpenAI), Eva Silverstein (Stanford), Andrew Strominger (Harvard), Xi Yin (Harvard). A week of talks on what the current models can and cannot do in theoretical physics.*

### THE TRAFFIC INTO THE LABS

*Dan Roberts coordinates from inside OpenAI. Xi Yin was reported in June 2026 to be joining OpenAI on sabbatical; neither he nor OpenAI has confirmed it. Strominger's former student Alex Lupsasca was the first hire of OpenAI for Science (October 2025), and in February 2026 Strominger co-authored a gluon-amplitude paper worked out with GPT-5.2. The labs are no longer across the street from the physics departments; they are inside them.*

***"… to facilitate constructive dialogue both among theoretical physicists and between theoretical physicists and the frontier labs."***

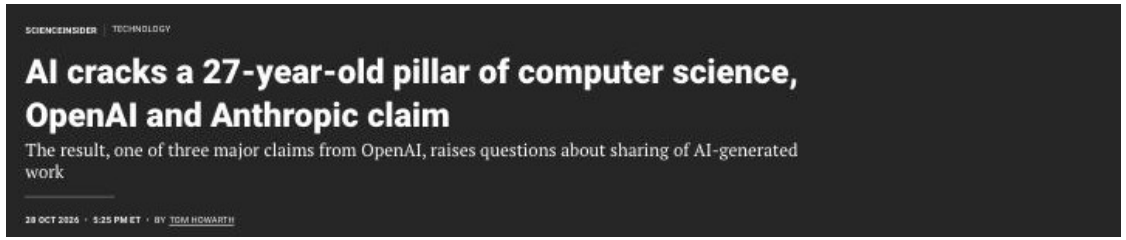
*KITP program description, Intermediate-Scale AI for Physics, October 2026*

AI → MATH

# *Three problems, three ways to release a proof*




*openai.com, 26 October 2026. "In our view, these manuscripts are best understood as a form of raw research data." External "shepherds" chose how each result reached the public.*




*Science, 28 October 2026 · Tom Howarth*

**MATHEMATICS · released as is**

*Sogge's local smoothing conjecture (1991) in every dimension, hence the Kakeya and Stein restriction conjectures in all dimensions. Output 18 October, Lean the next day, on GitHub the same week.*

**COMPUTER SCIENCE · shepherded, then public**

*The Exponential Time Hypothesis is false: 3-SAT in sub-exponential time. 16 October, Lean the next day. Impagliazzo, Paturi and Williams rewrote a "completely unreadable" first version; Anthropic posted an independent refutation hours later.*

**PHYSICS · withheld**

*Existence of four-dimensional Yang–Mills theory coupled to fermions, QCD included; pure Yang–Mills is half of a Millennium Prize problem. 1147 pages. A committee chaired by David Gross holds the proof and will release it "in due course".*

***"I think the main difficulty is that a company was selecting or attempting to select who is taking charge of it."***

*Ryan Williams (MIT), one of the three ETH shepherds · Science, 28 October 2026*

AI → MATH

# 1147 pages that nobody outside the committee has read

Proof by OpenAI of the Existence of Quantum Yang-Mills theory

Statement by the shepherding committee

Michael R. Douglas, David J. Gross (Chair), Arthur Jaffe, Igor R. Klebanov and Nikita Nekrasov

October 2026

AI agents are beginning to dramatically accelerate progress in mathematics, theoretical physics, and many other disciplines. This era of rapid change calls for new ideas on how to comprehend and disseminate scientific results. We are pleased to contribute to a new experimental approach where a group of scientists serves as an interface between AI generated breakthroughs and the broader scientific community. We are honored to work on shepherding this result to the community.

We were delighted to learn of the recent proof by OpenAI of the existence of quantum Yang-Mills theory, including coupling to fermions. This constitutes an important juncture for the entire subject of Quantum Field Theory (QFT), which applies quantum mechanics to systems with an infinite number of degrees of freedom and provides a universal language for a variety of observed physical phenomena. The new proof by OpenAI appears to be an important step towards putting QFT on a rigorous mathematical footing.

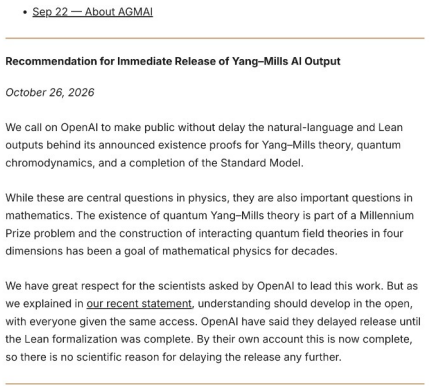
On Oct. 16, David Gross was asked by OpenAI to assume custody of the proof, decide on the best way to release it to the community, and shepherd the release process. David Gross then chose the rest of the committee. We will work on behalf of the academic community, to shepherd the release of the exciting new results discovered by OpenAI's LLM, while minimizing the risks (such as putting out wrong results or carelessly omitting attributions.) We have

*Statement of the shepherding committee, kitp.ucsb.edu, 27 October 2026.*
*"We have no responsibility to OpenAI."*

**WHAT IS CLAIMED**

*A rigorous existence proof of quantum Yang–Mills theory in four dimensions, coupled to fermions, so QCD and a grand-unified completion of the Standard Model, by Balaban's renormalization-group construction of the continuum limit. Pure Yang–Mills existence is the existence half of the Clay problem. Not claimed: the mass gap. "Confinement and the mass gap remain very important open challenges."*

| 9 Sept | 18 Sept | 16 Oct | 26 Oct |
|---|---|---|---|
| *Yang–Mills proof produced* | *QCD proof produced* | *Lean done; Gross asked to take custody* | *announced; proof withheld* |



• Sep 22 — About AGMAI

**Recommendation for Immediate Release of Yang–Mills AI Output**

*October 26, 2026*

We call on OpenAI to make public without delay the natural-language and Lean outputs behind its announced existence proofs for Yang–Mills theory, quantum chromodynamics, and a completion of the Standard Model.

While these are central questions in physics, they are also important questions in mathematics. The existence of quantum Yang–Mills theory is part of a Millennium Prize problem and the construction of interacting quantum field theories in four dimensions has been a goal of mathematical physics for decades.

We have great respect for the scientists asked by OpenAI to lead this work. But as we explained in our recent statement, understanding should develop in the open, with everyone given the same access. OpenAI have said they delayed release until the Lean formalization was complete. By their own account this is now complete, so there is no scientific reason for delaying the release any further.

*agmai.org, 26 October 2026*

***"By their own account this is now complete, so there is no scientific reason for delaying the release any further."***

*Advisory Group on Mathematics and AI: Charles, De Lellis, Gowers, Hairer, Srivastava, Tillmann, Vakil, Witten, Wood*

***"Before we release it, we would like to be convinced that it is correct."***

*The committee's statement*

AI FOR MATH · 26 OCTOBER 2026

AI → MATH

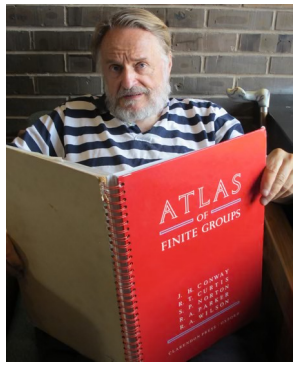
# *The enormous theorem, formalized in 24 days*

October 26, 2026 by Axiom

## Formalizing the Enormous Theorem

Axiom has formalized in Lean the classification of finite simple groups, the longest proof in the history of mathematics.

*axiommath.ai, 26 October 2026*




*John Horton Conway (1937–2020), holding the ATLAS of Finite Groups (1985), which he co-wrote. Three of the 26 sporadic groups carry his name. He is among the hundred mathematicians Axiom's acknowledgements name.*

**WHAT WAS DONE**

*The first-generation proof of the classification of finite simple groups, about 15,000 pages in more than 500 articles by about a hundred mathematicians, in Lean 4 with Mathlib. Dozens of agents directed by Axiom's engineers and mathematicians, 2 to 25 October; 1.2 million lines reused from the Qiuzhen-CFSG project; external advisors Lyons, Nakano and Solomon.*

**THE GAP THE AGENTS FOUND**

*Counterexamples to a closure assertion in a published route to the Unbalanced Group Theorem, one of 34 deviations from the record. The theorem stands: a replacement argument via Solomon's 1976 maximal 2-components, written up by Defant and Ono (Hexagon, 27 October, labelled "primarily human-written text").*

*Status on 30 October: review status "team self-assessment". Lean's kernel checked every step; whether the formal statement says what CFSG says is for group theorists and the Lean community to confirm, which Axiom invites. No cost or compute figures given.*

| 43.1 M | 2.05 M | 715 | 24 | 34 |
|---|---|---|---|---|
| *lines of Lean* | *declarations* | *papers and books* | *days* | *deviations from the record* |

AI → MATH

# What counts as good math research, in an age of AI?

*Emily Riehl's categories and her examples, from "What counts as 'good math research' in an age of AI?", Chalmers AI Days, 29 October 2026.*

## Bad math?

**Scooping junior researchers**

*Math Inc. finished Sidharth Hariharan's two-year E8 sphere-packing formalization with its agent Gauss, to show it off: "drive-by proving".*

*Jeremy Avigad, Mathematicians in the age of AI, arXiv:2603.03684*

**Wasted resources**

*Anthropic's 13-million-line Lean proof of Fermat's Last Theorem in 11 days, alongside Buzzard's long-running project of the same.*

*Anthropic, 4 September; Buzzard, "FLT: Anthropic has beaten me to it", 4 September*

## Good math?

**Advancing understanding in an area**

*Axiom's classification formalization, which found and repaired a gap in a published route to the Unbalanced Group Theorem.*

*axiommath.ai, 26 October*

**Acknowledging human mathematicians**

*Axiom's acknowledgements name the hundred mathematicians whose proof it is, and the advisors it consulted.*

*axiommath.ai, 26 October*

**Autoformalization as future infrastructure**

*$\pi_3(S^2)$ in agda-unimath, with the AI-generated proofs imported back into the library.*

*Carranza, Liu, Riehl, Rijke, arXiv:2608.23721*

***"Do we need to abandon the practice of attributing authorship to mathematical discoveries?"***

*Emily Riehl, closing slide, Chalmers AI Days, 29 October 2026*

# *How do we articulate the value of human mathematicians?*

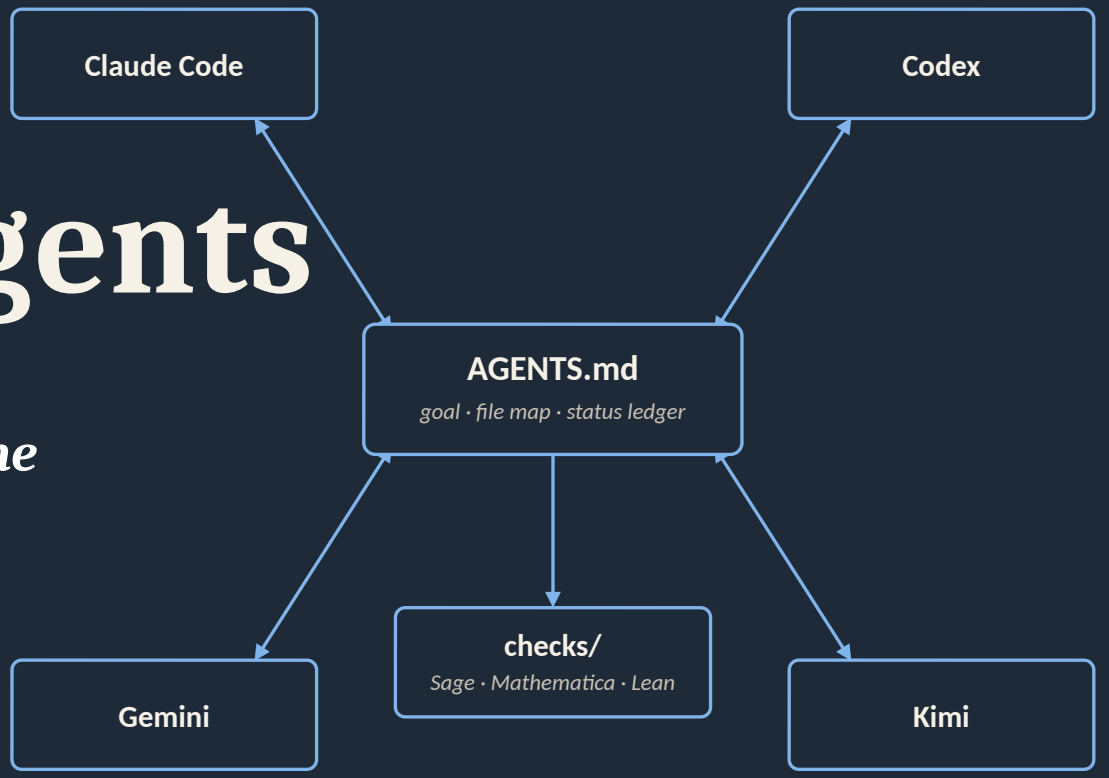


*Illustration from Emily Riehl's slide of the same title, "AI for math: what is happening and what it means", September 2026. Original source not identified.*

PART III

# How I work with agents

***A setup for collaborative research with AI agents: one file, a repository, and checks that print the right answer.***




*one file, four readers, machine checks*

THE WORKFLOW

# *The step from chatbot to agents*

CHATBOT

*ask a question*
↓
*read a plausible answer*
↓
*copy the useful part*
↓
*context ends with the session*

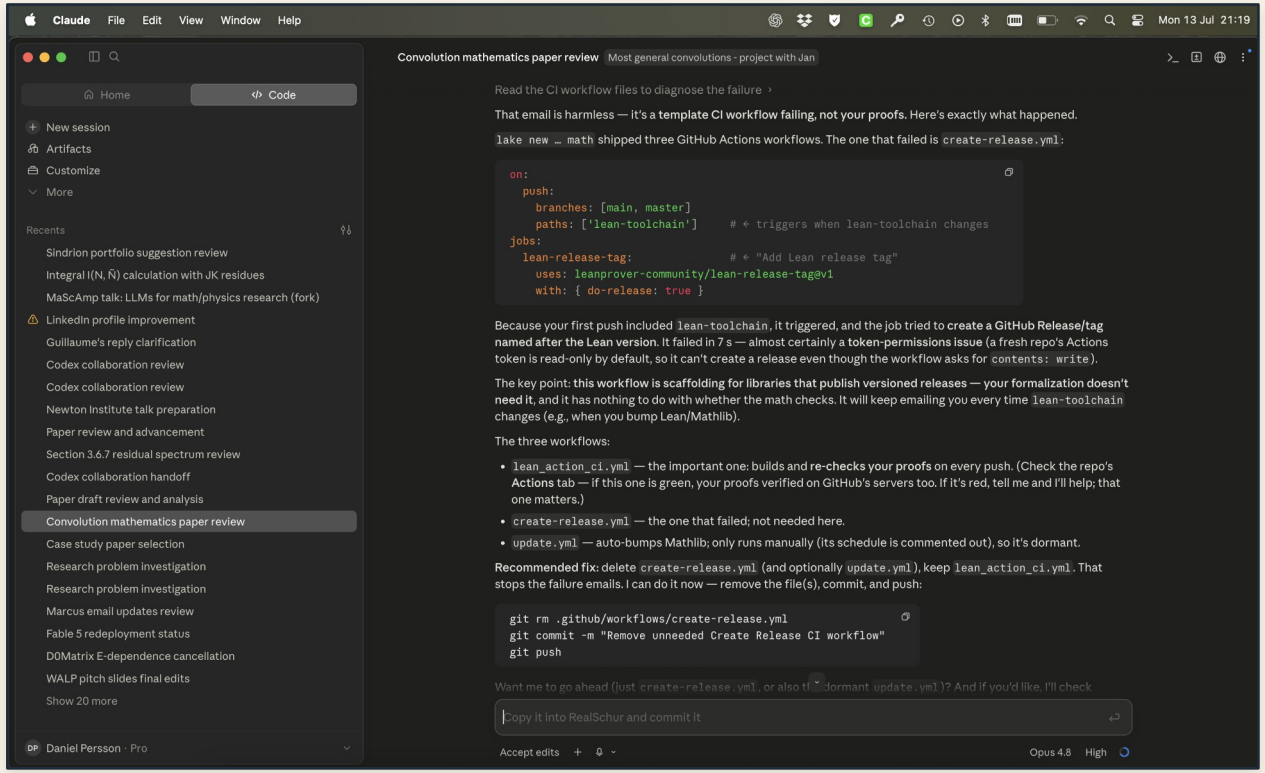
PROJECT AGENT

*give a task*
↓
*it reads the repository*
↓
*it edits and runs tools*
↓
*you inspect the evidence*
↓
*accept—or send it back*

THE TOOLS

# What is a project agent?

- **Works in a project folder** — your repository: LaTeX, notes, code, bibliography
- **Reads and writes files** — drafts, edits and reorganises; every change can be inspected
- **Runs programs** — computations, checks, literature searches — and reads the output
- **Two examples** — Codex (OpenAI) and Claude Code (Anthropic)

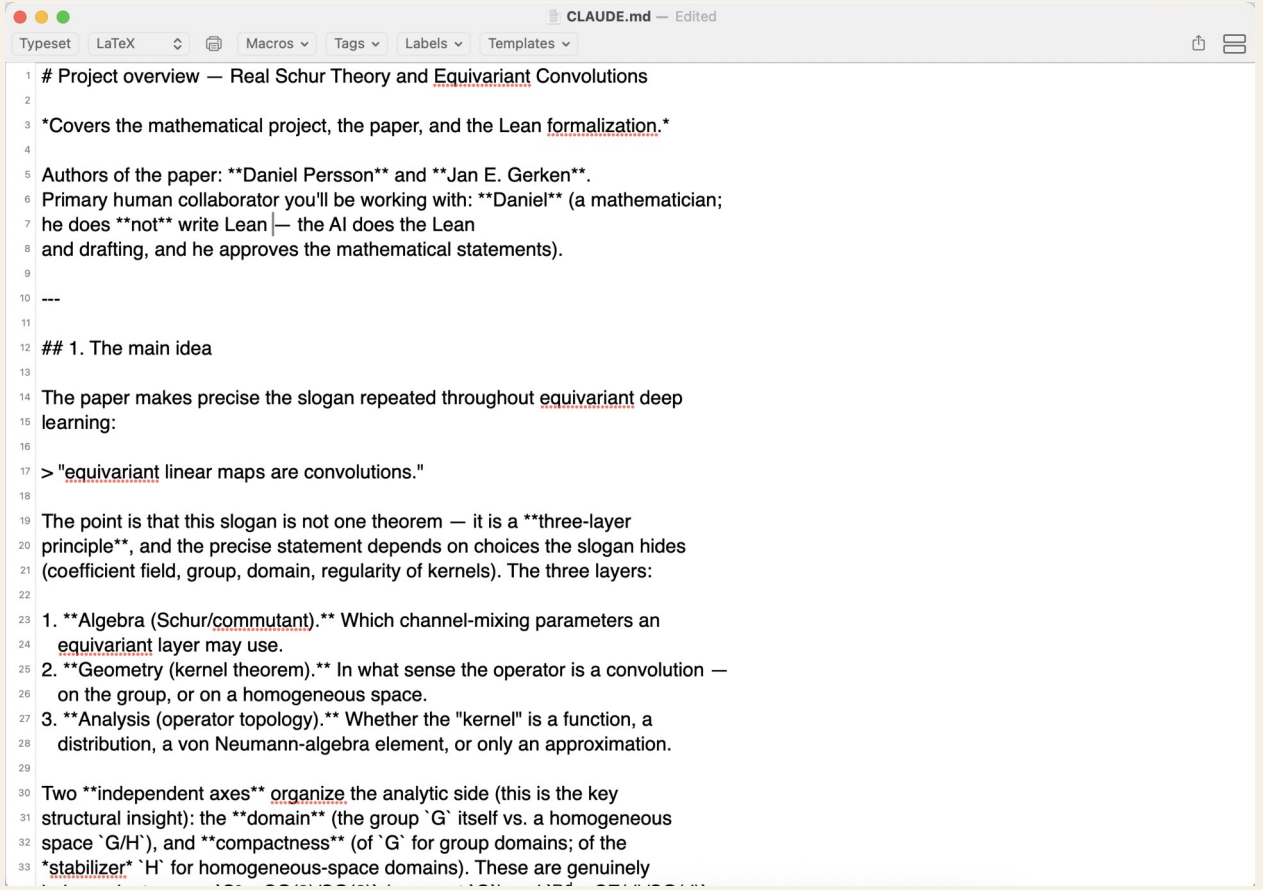



*Claude Code, working in one of my paper repositories*

PROJECT MEMORY

# *Markdown files make the project remember*

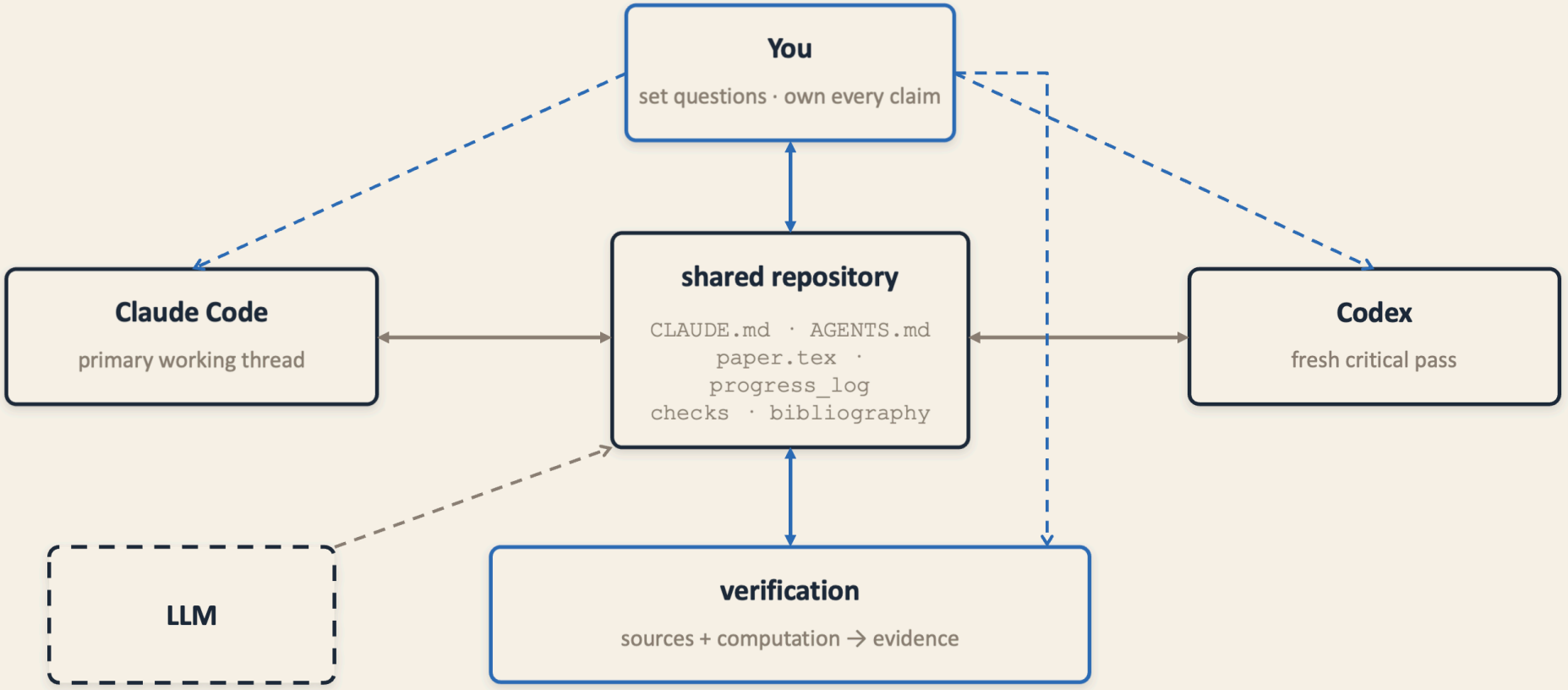
- • One GLOBAL.md holds the standing rules, symlinked into every agent.
- • Each project has an AGENTS.md: goal, file map, conventions, status ledger.
- • A progress log keeps decisions, loose ends, and next steps.
- • New sessions start by reading this durable context.




*A project-specific CLAUDE.md carries durable context*

THE MINDSET

# *The repository is the shared brain*



THE CASE STUDY

# *Z-infinite automorphic forms*

- Project: automorphic functions that are not Z-finite.
- We had an annihilation theorem for SO(n,n), proved by explicit differential equations.
- The goal was a proof that generalizes to exceptional groups.

*the paper draft:*

**$\mathcal{Z}$-INFINITE AUTOMORPHIC FORMS AND THE SQUARE OF THE MINIMAL REPRESENTATION**

GUILLAUME BOSSARD, SOLOMON FRIEDBERG, DMITRY GOUREVITCH, AXEL KLEINSCHMIDT, AND DANIEL PERSSON

(426)

Using the functional relation (184) one obtains therefore both (423) at $s = \frac{n-2}{2}$ and the equation above (426) for $E_{P_2,\pi}(\frac{n-2}{2}, g)$. Substituting (426) in (423) for $s = \frac{n-2}{2}$, one obtains

$$\frac{\Delta + (n-1)(n-2)}{4}\mathcal{D}_{a\hat{c}}\mathcal{D}^{d\hat{c}}\mathcal{D}_{d\hat{b}}E^{D_n}_{P_2,\pi}(\tfrac{n-2}{2}, g)$$
$$= \Big(\frac{\Delta + (n-1)(n-2)}{4}\Big)\Big(\frac{\Delta + (n-2)(n-3)}{4}\Big)\mathcal{D}_{a\hat{b}}E^{D_n}_{P_2,\pi}(\tfrac{n-2}{2}, g)\,. \quad (427)$$

For $\pi$ such that $\Delta f_\pi = -(\frac{1}{4} + r_\pi^2) f_\pi$ in $SL(2)$, one obtains that

$$\frac{\Delta + (n-1)(n-2)}{4}E^{D_n}_{P_2,\pi}(\tfrac{n-2}{2}, g) = -\frac{\frac{1}{4} + r_\pi^2}{4}E^{D_n}_{P_2,\pi}(\tfrac{n-2}{2}, g) \quad (428)$$

so one can inverse this operator and obtain that

$$\mathcal{D}_{[a}{}^{[\hat{a}}\mathcal{D}_b{}^{\hat{b}}\mathcal{D}_{c]}{}^{\hat{c}]}E^{D_n}_{P_2,\pi}(\tfrac{n-2}{2}, g) = 0\,, \quad (429)$$

$$\mathcal{D}_{a\hat{c}}\mathcal{D}^{d\hat{c}}\mathcal{D}_{d\hat{b}}E^{D_n}_{P_2,\pi}(\tfrac{n-2}{2}, g) = \frac{1}{4}\Big(\Delta + (n-2)(n-3)\Big)\mathcal{D}_{a\hat{b}}E^{D_n}_{P_2,\pi}(\tfrac{n-2}{2}, g)\,.$$

This concludes the direct proof that $E^{D_n}_{P_2,\pi}(\frac{n-2}{2}, g)$ is annihilated by $\mathcal{I}_{P_2}$.

# *The proof emerged from collaboration with AI agents*

Not one prompt: repeated proposals, objections, source checks, and rewrites.

1. **Define the problem — correct theorem, look for alternative proof**
2. **ChatGPT first suggests a way forward. Claude Code then searches the literature and brings in new ideas.**
3. **We review and come with new input. Codex then checks and re-derives the argument, attacks assumptions, and proposes alternatives.**
4. **We read the new sources, check the details, and reformulate the argument. Both Claude Code and Codex verify the new proof.**

**Theorem 1.** *For every spherical cuspidal $\pi$ occurring in Lemma 6, we have*

$$\mathcal{A}\,X_\pi = \mathcal{B}\,X_\pi = 0\,. \tag{220}$$

*Equivalently, $\mathcal{I}_{P_2} X_\pi = 0$; together with $\Delta X_\pi = 2s_\pi(2s_\pi + 3 - 2n)X_\pi$ this gives $\mathcal{I}_{P_2,s_\pi} X_\pi = 0$.*

*Proof.* Let $M_{s_\pi} = \mathcal{U}(\mathfrak{g}_{\mathbb{C}})\chi_{P_2}^{s_\pi}$ be the cyclic Harish-Chandra module generated by the scalar spherical section. It has finite length, so it has a nonzero irreducible quotient $J_\pi$. The image of its cyclic vector is necessarily nonzero and $K$-fixed. Hence $J_\pi$ is spherical and has infinitesimal character $2s_\pi\Lambda_2 - \rho = \lambda_\pi$. Equation (212) descends to its spherical vector.

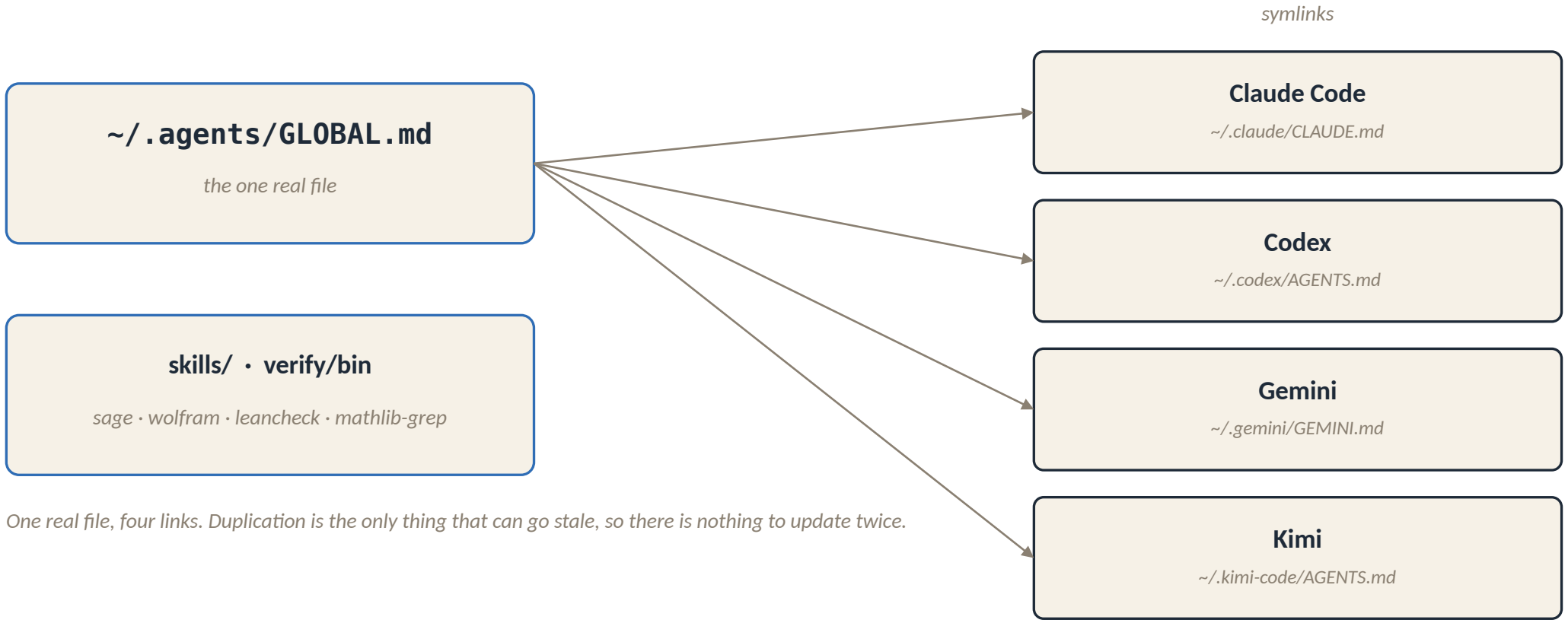
By Proposition 9, $J_\pi \simeq \sigma_\pi$, where $\sigma_\pi$ is the representation of Lemma 7. Thus the $K$-fixed line in every copy of $\sigma_\pi$ is annihilated by $\mathcal{A}$ and $\mathcal{B}$. Lemma 7 writes $X_\pi = v_K \otimes m_\pi$ in the corresponding isotypic component, with $G^\circ$ acting only on the first factor. Therefore

$$\mathcal{A}X_\pi = (\mathcal{A}v_K) \otimes m_\pi = 0\,, \qquad \mathcal{B}X_\pi = (\mathcal{B}v_K) \otimes m_\pi = 0\,. \tag{221}$$

□

THE INFRASTRUCTURE

# *One file, every agent*




*One real file, four links. Duplication is the only thing that can go stale, so there is nothing to update twice.*

# *The whole thing, on one page*

**THE RULES**

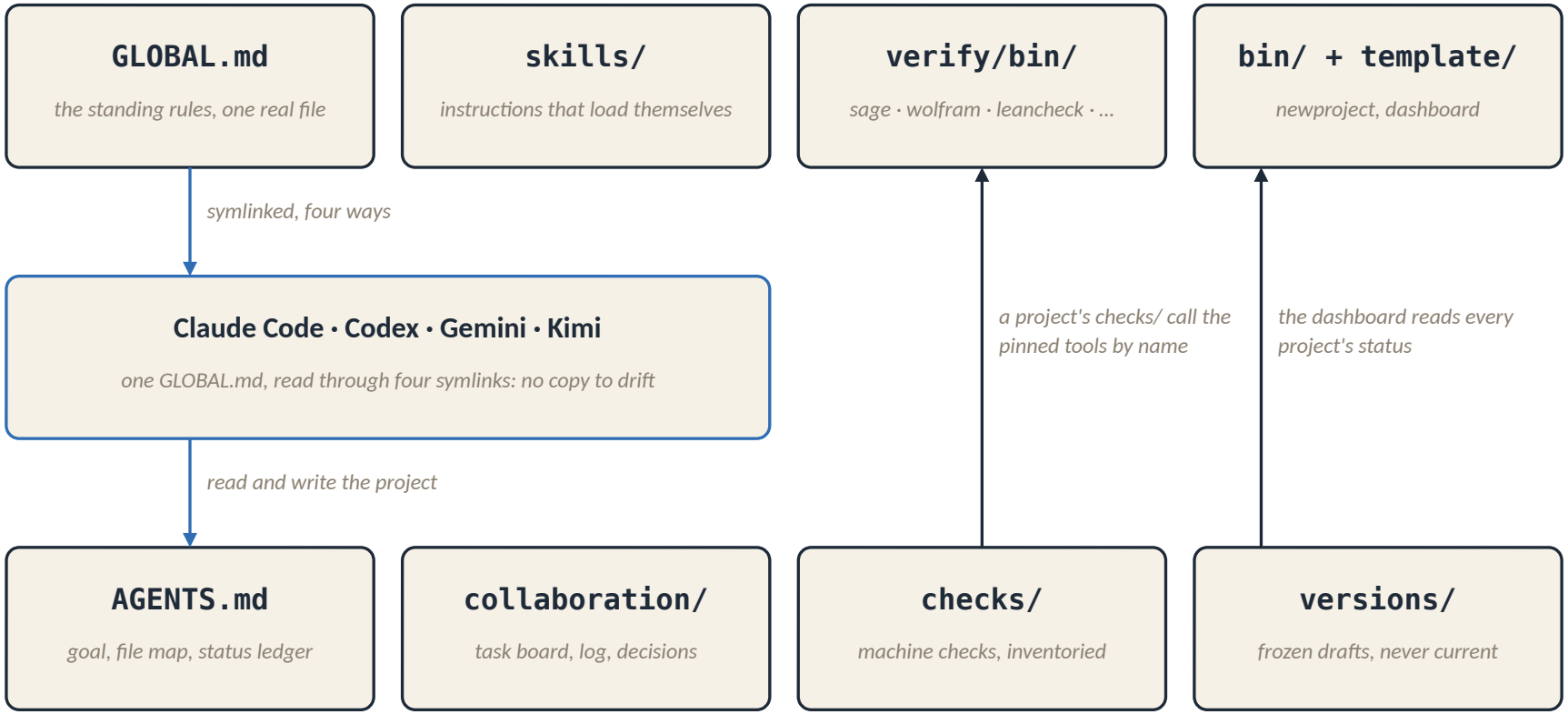
*~/.agents/, one directory, under its own git*

**THE AGENTS**

*interchangeable readers of the same rules*

**EVERY PROJECT**

*Dropbox/Documents/Latex/, eight of them, all the same shape*




***Rules flow down; evidence flows up. The two arrows that run back up are what make it a system rather than a naming convention.***

*Larger version, with every path drawn in full: infrastructure.html in the same folder.*

THE INFRASTRUCTURE

# *Skills: instructions that load themselves*

```
~/.agents/skills/math-verification/SKILL.md
---
name: math-verification
description: Machine-checks mathematical
  claims with Sage, Python, Mathematica and
  Lean 4. Use when verifying a calculation,
  proof or identity; when a draft asserts
  something that needs checking
---

## Tiers
## Three rules that save you a day
```

**The description is the trigger**

*the agent loads it when the work matches. You never have to remember it exists.*

**It carries the gotchas, permanently**

*never the system python3 · try the local Wolfram kernel before blaming the licence · grep Mathlib before proposing any Lean*

**Written once, after it cost you a day**

*one folder under ~/.agents/skills/, read by every agent, in every project*

The second one in daily use is academic-humanizer: a style profile of my own prose, with separate layers for mathematics, physics and machine learning.

*A skill is a folder with a SKILL.md and its references. The frontmatter description is what makes it self-loading; everything hard-won goes in the body.*

***The alternative is retyping the same warning into every session until you forget one.***

INSIDE A PROJECT

# *The repository, under version control*

**`Latex/Z-infinite automorphic forms/`**

| | |
|---|---|
| **AGENTS.md** | *entry point: goal, conventions, status ledger* |
| **Spectral_argument_v13.tex** | *the live draft — top level means current* |
| **collaboration/** | *task board · progress log · decisions · reviews* |
| **checks/** | *17 checks: ID · exact command · date last passed* |
| **versions/** | *frozen drafts — never cite as current* |

**`git log · local, no remote needed`**

- **2026-08-12 Codex**
  *spectral argument → v13*
- **2026-08-11 Claude**
  *checks C-009 … C-012*
- **2026-08-04 Dan**
  *reorganise; retire v5 to versions/*

*Who changed what, and when. Dropbox forks files silently and cannot answer that.*

*Claim a task on the board before touching files. The default role is verification: re-derive, compare with sources, log disagreements.*

VERIFICATION

# *What counts as evidence*

tier 1

**Sage · Python**

*root systems, Weyl and Coxeter groups, crystals, numerics*

tier 2

**Mathematica**

*heavy symbolics, special functions, shareable notebooks*

tier 3

**Lean 4 · Mathlib**

*the algebraic core only: slow, narrow, replayable*

**an LLM reviewing an LLM**

*advisory — it never enters the ledger*

**every result carries a tag:** [ESTABLISHED] [CONJECTURED] [OPEN]

*An untagged claim gets read as proved by whichever agent reads it next.*

# *Rethlas: a machine that reads a proof and talks back*

### WHAT IT IS

*An open-source system (github.com/frenzymath/Rethlas, Apache-2.0) built around two Codex agents. One writes an informal proof of a stated problem. The other, a local service, reads a prose proof and returns a strict verdict with repair hints.*

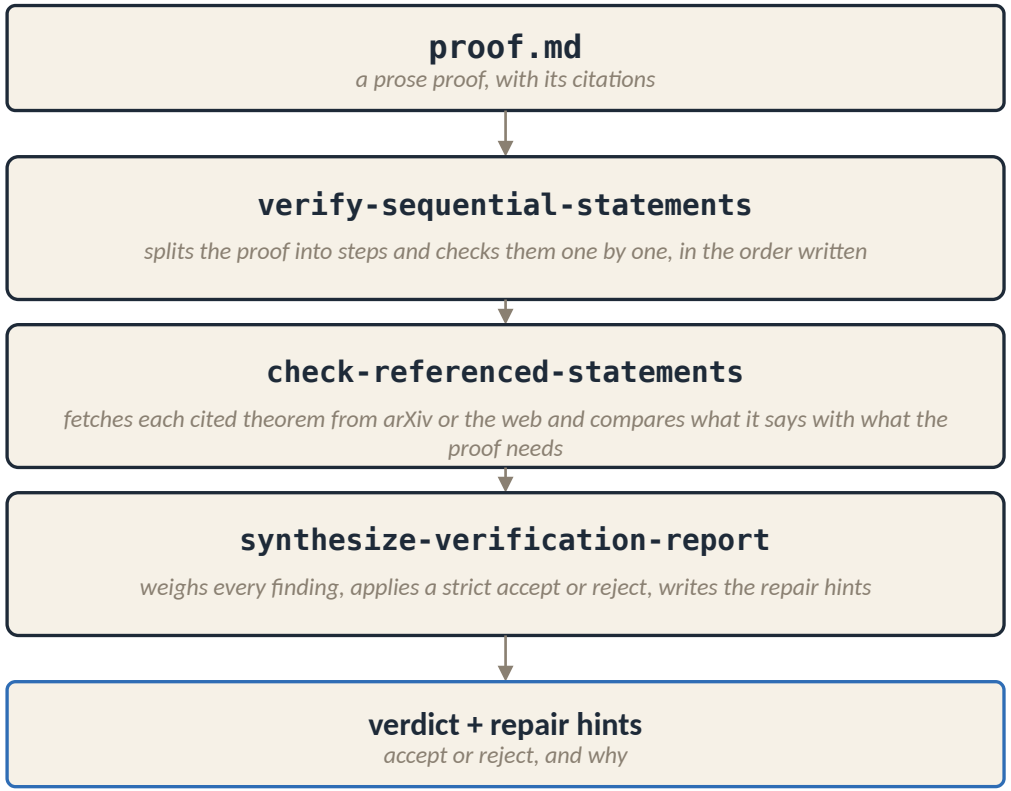
### WHAT I USE IT FOR

*Only the reader. A proof drafted in a project, by me or by an agent, goes in as Markdown. What comes back is a step-by-step account of what holds, what is unjustified, and whether each cited result says what the proof claims it says.*

### WHAT IT IS NOT

*Evidence. A model's opinion of a proof is still an opinion, so Rethlas sits below the bench of Sage, Mathematica and Lean: advisory, never ESTABLISHED. The proof-writing half stays off: a model rewriting until another model approves is optimising for approval, not for truth.*

### THE READER, STEP BY STEP




*Installed at ~/.agents/verify/rethlas/; the generation agent is never run. The middle skill earns its keep: a machine can go and check whether a citation says what the proof claims it says.*

# One number, left blank



UUITP-16/24

CANONICALIZING ZETA GENERATORS:
GENUS ZERO AND GENUS ONE

Daniele Dorigoni[1], Mehregan Doroudiani[2], Joshua Drewitt[3], Martijn Hidding[4,5], Axel Kleinschmidt[2,7], Oliver Schlotterer[4,6], Leila Schneps[8] and Bram Verbeek[4]

[1] *Centre for Particle Theory & Department of Mathematical Sciences Durham University, Lower Mountjoy, Stockton Road, Durham DH1 3LE, UK*
[2] *Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut) Am Mühlenberg 1, 14476 Potsdam, Germany*
[3] *School of Mathematics, University of Bristol, Queens Road, Bristol, BS8 1QU, UK*
[4] *Department of Physics and Astronomy, Uppsala University, 75108 Uppsala, Sweden*
[5] *Institute for Theoretical Physics, ETH Zurich, 8093 Zürich, Switzerland*
[6] *Department of Mathematics, Centre for Geometry and Physics, Uppsala University, 75106 Uppsala, Sweden*
[7] *International Solvay Institutes, ULB-Campus Plaine CP231, 1050 Brussels, Belgium*
[8] *CNRS, Sorbonne Université, Campus Pierre et Marie Curie 4 place Jussieu, 75005 Paris, France*

Zeta generators are derivations associated with odd Riemann zeta values that act freely on the Lie algebra of the fundamental group of Riemann surfaces with marked points. The genus-zero incarnation of zeta generators are Ihara derivations of certain Lie polynomials in two generators that can be obtained from the Drinfeld associator. We characterize a canonical choice of these polynomials, together with their non-Lie counterparts at even degrees $w \geq 2$, through the action of the dual space of formal and motivic multizeta values. Based on these canonical polynomials, we propose a canonical isomorphism that maps motivic multizeta values into the $f$-alphabet. The canonical Lie polynomials from the genus-zero setup determine canonical zeta generators in genus one that act on the two generators of Enriquez' elliptic associators. Up to a single contribution at fixed degree, the zeta generators in genus one are systematically expanded in terms of Tsunogai's geometric derivations dual to holomorphic Eisenstein series, leading to a wealth of explicit high-order computations. Earlier ambiguities in defining the non-geometric part of genus-one zeta generators are resolved by imposing a new representation-theoretic condition. The tight interplay between zeta generators in genus zero and genus one unravelled in this work connects the construction of single-valued multiple polylogarithms on the sphere with iterated-Eisenstein-integral representations of modular graph forms.

arXiv:2406.05099v3 [math.QA] 13 Nov 2025

***arXiv:2406.05099v3*** · *13 Nov 2025*
*Dorigoni, Doroudiani, Drewitt, Hidding, Kleinschmidt, Schlotterer, Schneps, Verbeek*
***two of the eight are MaScAmp PIs***

***eq. (7.24)*** · *the expansion of $\sigma_{11}$ in Tsunogai's geometric derivations*

$$\begin{aligned}\sigma_{11} = &-\frac{\epsilon_{12}^{(10)}}{10!} + \frac{11s^3(\epsilon_4,\epsilon_{10})}{40} + \frac{11s^3(\epsilon_6,\epsilon_8)}{60} + \frac{242407s^2(\epsilon_4,t^2(\epsilon_4,\epsilon_6))}{14735232} + \frac{s^5(\epsilon_4,\epsilon_{12})}{720}\\ &- \frac{s^5(\epsilon_6,\epsilon_{10})}{216} - \frac{7s^5(\epsilon_8,\epsilon_8)}{4320} + \frac{11090423s^2(\epsilon_4,t^4(\epsilon_4,\epsilon_8))}{309439872} + \frac{3197s^3(\epsilon_4,t^3(\epsilon_4,\epsilon_8))}{57600}\\ &- \frac{2983s^4(\epsilon_4,t^2(\epsilon_4,\epsilon_8))}{86400} + \frac{148753s^3(\epsilon_4,t^3(\epsilon_6,\epsilon_6))}{7367616} + \frac{490853s^3(\epsilon_6,t^3(\epsilon_6,\epsilon_4))}{17191104}\\ &+ \frac{156805s^4(\epsilon_6,t^2(\epsilon_6,\epsilon_4))}{14735232} + c\,s^2(\epsilon_4,t^2(\epsilon_4,t^3(\epsilon_4,\epsilon_4))) + \dots\,,\end{aligned}$$

***"the coefficient $c \in \mathbb{Q}$ of the first modular-depth-four contribution to $\sigma_{11}$ in the last line has not yet been computed. It is, however, a highest-weight vector and entirely fixed by our construction."***

*same page, immediately below the equation*

One rational number. Uniquely determined. Never evaluated.

# *The number, and how we know it*

$$c = 232871129 / 8841139200$$

*degree 16, modular depth 4 · denominator $2^{10} \cdot 3^3 \cdot 5^2 \cdot 12791$ · numerator prime*

## HOW IT WAS COMPUTED

**$g_{11} \rightarrow \tau_{11} \rightarrow \sigma_{11}$**
*the paper's extension lemma, in exact rational word algebra to degree 17*

↓

**$\sigma_7$ at degree 12**
*reproduces the commutators printed in the paper, exactly*

↓

**$\sigma_9$ and $\sigma_{11}$ below the frontier**
*match the authors' own machine-readable data*

↓

**$\sigma_9$ at modular depth four**
*reproduces the published 149/13824 and −149/69120 — the gate*

↓

**$\sigma_{11}$ at modular depth four**
*one unknown rational, fixed by 16 251 word coefficients*

## HOW IT WAS VERIFIED

**Codex**
*Mathematica 13.1*
*checks S-001 … S-008*

**Claude**
*Python, exact rationals*
*checks I-001 … I-007*

*written apart: the second built from the paper before the first was read*

**The authors' own ancillary data**
*arXiv 2403.14816 — an external oracle for the subtracted terms*

↓

**All three agree, exactly**
*same support, same ratio, on every one of 16 251 coefficients*

*Not yet checked against anything outside the construction.*

PRACTICAL

# How to start on Monday

**01**

### Put it in plain text

*LaTeX, Markdown, code, notes, bibliography*

**02**

### Give an agent the folder

*start with a bounded task you know well*

**03**

### Write the project brief

*goals, conventions, boundaries, known pitfalls*

**04**

### Demand evidence

*calculation, proof artifact, source, counterexample*

ONE LAST THOUGHT

*“Can AI do mathematics?” is not my question.*

***How does research change when a fluent, tireless, fallible collaborator joins the loop?***



*Illustration from Emily Riehl's talk “AI for math: what is happening and what it means”, September 2026.*