

Mathematical discovery in the age of AI

A short history of math,
from counting to vibe-proving.

by Hang Lu,

math PhD / jobless

warning: definitely not a historian, all sassy remarks are anachronistic

0. Why did people start mathing?

We've always needed to count things.

I.e. “bitch owes me 8 flints.”



Markings made on a hyena bone by a Neanderthal, some 60,000 years ago.

Source: [1]

1. Civilization develops

Keeping track of things becomes more complex. We introduce **procedures**.

I.e. “We building the pyramids. Better not mess it up!”



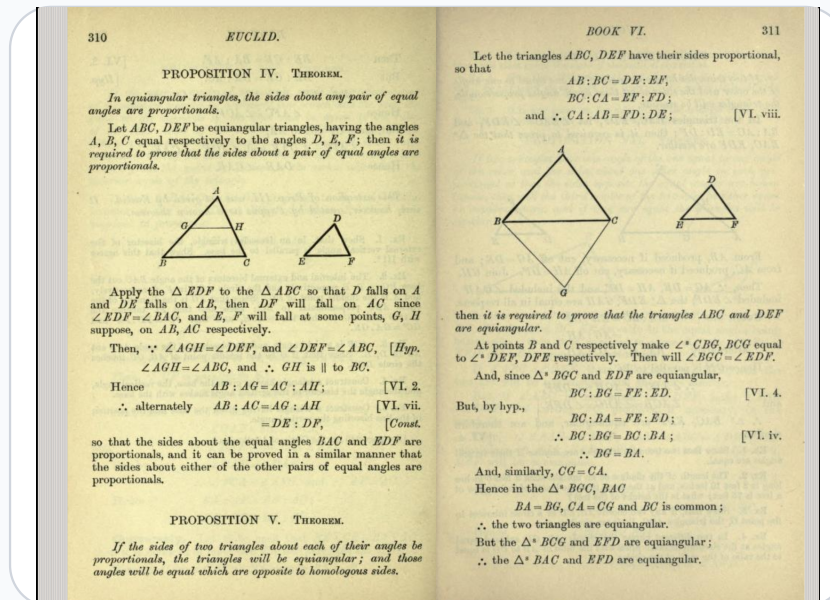
Rhind Mathematical Papyrus, Egypt, 1550 BCE.

The instructions show how to calculate the pyramid’s slope using a simple procedure based on the height and base length.

2. How do we know the procedures work?

Around 300 BCE, Euclid's Elements organized math from definitions, postulates, and common notions, then derived results by proof.

I.e. “Stop saying just trust me bro. Tell me **why** something is true.”



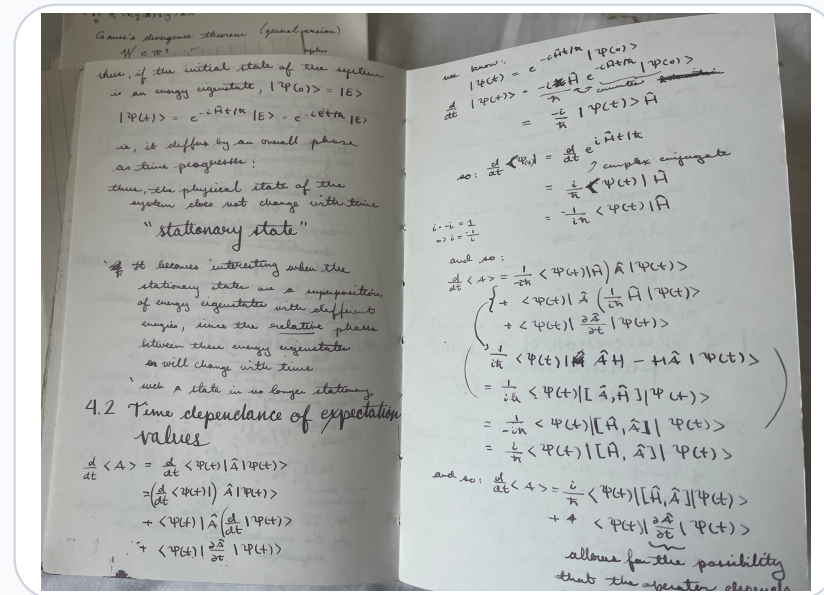
Reproduction of Euclid's Elements, Book VI, Proposition 4.

3. Modern math: notation becomes infrastructure

The relentless pursuit of the **how** and the **why** gave rise to modern math notation.

Centuries of definitions, symbols, and proof techniques packed into a page.

I.e. “A cathedral to truth.”



My undergrad notes. I was standing on the shoulders of giants.

4. Infrastructure becomes software

Modern math has foundations strong enough to be formally verified by a computer.

Try it yourself: show why $2+2 = 4$ as an interactive video game!

```
2 + 2 = 4.
example : (2 : ℕ) + 2 = 4 := by

  1 nth_rewrite 2 [two_eq_succ_one]
  2 rw [add_succ]
  3 rw [one_eq_succ_zero]
  4 rw [add_succ]
  5 rw [add_zero]
  6 rw [two_eq_succ_one]
  7 rw [one_eq_succ_zero]
  8 rw [four_eq_succ_three]
  9 rw [three_eq_succ_two]
  10 rw [two_eq_succ_one]
  11 rw [one_eq_succ_zero]
```

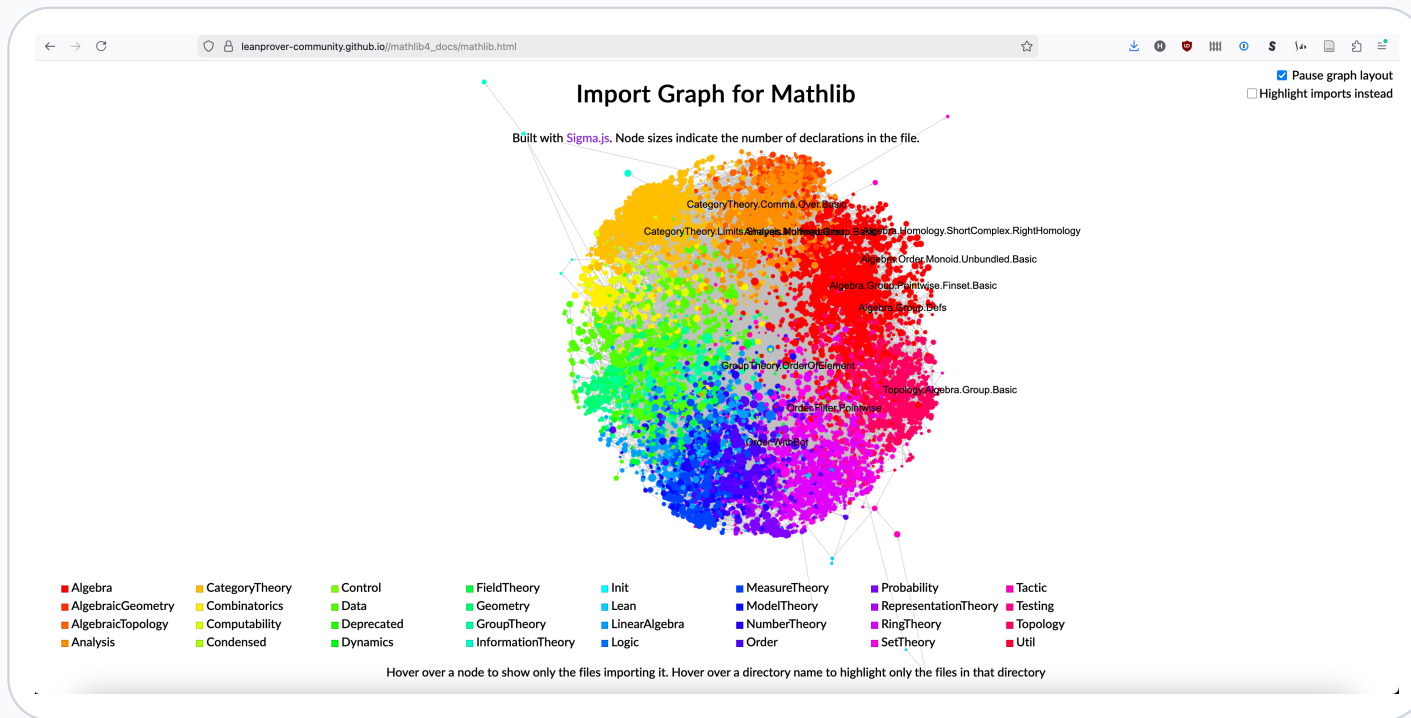
Level completed! 🎉

Current Goal
Goal:
 $2 + 2 = 4$

<https://adam.math.hhu.de/#/g/leanprover-community/nng4>

5. mathlib4: an online library of formalized math

<https://github.com/leanprover-community/mathlib4>



Code dependency plot shows interconnected nature of math.

6. math code + AI = vibe-proving?

LLMs are autocompletion machines. Can you autocomplete math?



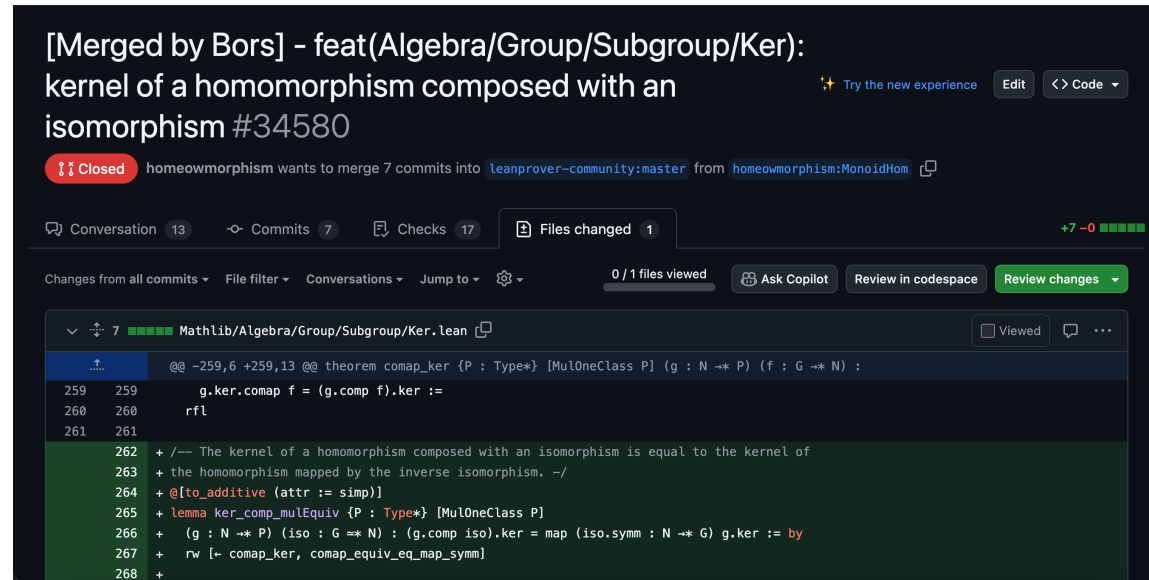
Absolutely, **and** formally verified software tells us if the LLM output is wrong.

So for once it can't even bullshit you that bad.

7. I've been vibe-proving

Me: to help build mathlib4! 🤖

Also me: Lean is a hard 😭.



[Merged by Bors] - feat(Algebra/Group/Subgroup/Ker):
kernel of a homomorphism composed with an
isomorphism #34580

🔒 Closed homeowmorphism wants to merge 7 commits into `leanprover-community:master` from `homeowmorphism:MonoidHom`

🗨 Conversation 13 📄 Commits 7 📄 Checks 17 📄 Files changed 1 +7 -0

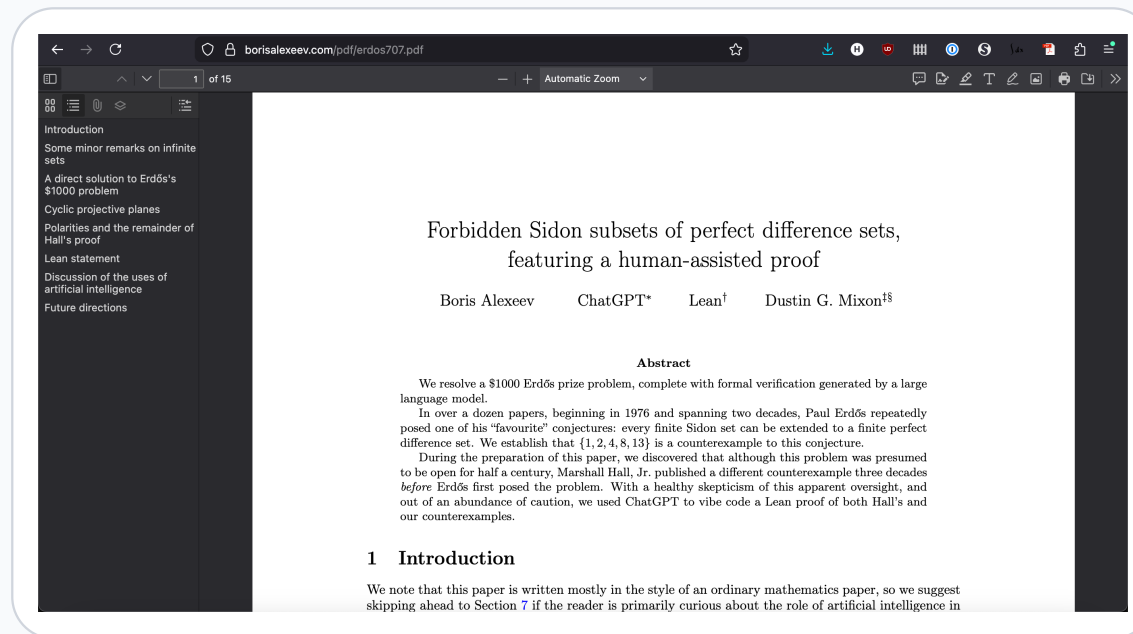
Changes from all commits File filter Conversations Jump to 0 / 1 files viewed Ask Copilot Review in codespace Review changes

```
Mathlib/Algebra/Group/Subgroup/Ker.lean
@@ -259,6 +259,13 @@ theorem comap_ker {P : Type*} [MulOneClass P] (g : N → P) (f : G → N) :
259 259   g.ker.comap f = (g.comp f).ker :=
260 260   rfl
261 261
262 + /- The kernel of a homomorphism composed with an isomorphism is equal to the kernel of
263 + the homomorphism mapped by the inverse isomorphism. -/
264 + @[to_additive (attr := simp)]
265 + lemma ker_comp_mulEquiv {P : Type*} [MulOneClass P]
266 +   (g : N → P) (iso : G → N) : (g.comp iso).ker = map (iso.symm : N → G) g.ker := by
267 +   rw [← comap_ker, comap_equiv_eq_map_symm]
268 +
```

My first contribution to mathlib4. Vibe-coded. Merged on Feb 3rd.

8. Can LLMs do new math?

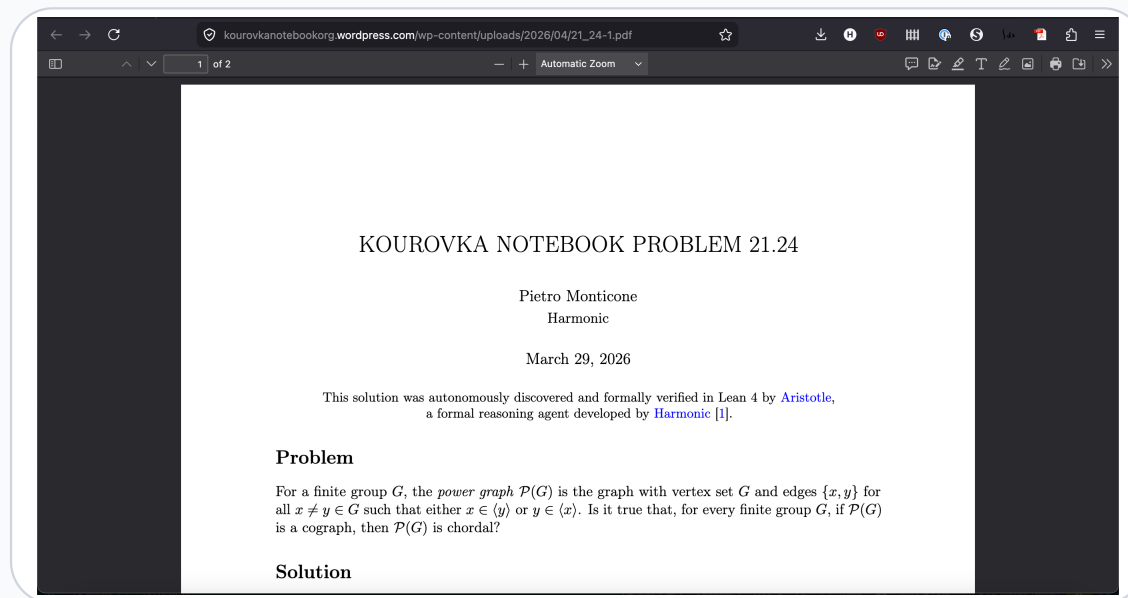
Terrifyingly, yes. This dropped last October.



Wait: “During the preparation of the paper, we discovered that although this problem was presumed to be open for half a century, Marshall Hall, Jr. published a different counterexample three decades before Erdős first posed the problem.”

9. Can LLMs do new math? Part 2

Last March: “An update for the current 21-st edition of “Kourovka Notebook (Unsolved Problems in Group Theory)” is posted at <https://kourovkanotebookorg.wordpress.com/>. For the first time (in history?), there are some solutions obtained by AI: check 20.125, 21.8, 21.24, 21.150.”



These are non-trivial problems! My supervisor goes through them sometimes and tries to find solutions.

I feel very torn about this.

My scientist self:

This is a marvel of human civilization! What else can we do?!? Let's gooo!

My worker self:

Not a fan of even more concentration of means of production toward the owner class, or the subsequent replacement of my labor I am contributing towards.

My feminist self:

Math can finally become the meritocratic, decentralized science it was always meant to be.

My degenerate self who grew up on the internet in the 2000s:

y'all wanna prove vibe-prove your wildest theories of the universe? DO IT!!

Bonus: The GOAT be vibe-proving

Look at what the greatest mathematician of our time dropped this on March 8th.



► Formalizing a proof in Lean using Claude Code

Top comment: “lol Anthropic needs to give Terence Tao infinite tokens - literally Holding math back”

Source: [3]



What are your feelings?

Bibliography and credits

Credit

This presentation was made using ChatGPT 5.2 and GPT-5.4.

Bibliography

- [1] C. Barras, “How did Neanderthals and other ancient humans learn to count?,” *Nature*, vol. 594, no. 7861, pp. 22–25, June 2021, doi: 10.1038/d41586-021-01429-6.
- [2] The British Museum, “Learn maths from Egyptian secrets: the Rhind Mathematical Papyrus.” [Online]. Available: <https://www.britishmuseum.org/blog/learn-maths-egyptian-secrets-rhind-mathematical-papyrus>
- [3] T. Tao, “Formalizing a proof in Lean using Claude Code.” [Online]. Available: <https://www.youtube.com/watch?v=JHEO7cplfk8>