https://althofer.de/math+ai.pdf

Math and AI
Making Mathematics with the Help of AI
Some Random Thoughts and Memories

Ingo Althoefer
ingo.althoefer@uni-jena.de
Version: November 22, 2025

Using ChatGPT-5 and other LLMs in Mathematics for:

* Quick reading, structuring and abstracting of papers

* Error search in papers (own ones or others)

* Performing small or medium-size computations to collect data (without classical programming)

* "little proofs" (Tao's experience: to save some hours of own working time)

* Using an AI like GPT-5 for searching data bases such OEIS, finding structures and formulating conjectures

* LLM in teamwork with a theorem prover, for instance LEAN
  LLM = Large Language Model

* * * * * * *

A few random hints and remarks

* important when using GPT-5: use strict mathematical mode ("Thinking mode")   -   GPT-5 account for 23 Euro per month

* very nice with GPT-5: during the thinking process current thoughts are shown in grey script, changing every fraction of a second. These intermediate thoughts are not protocolled!

* Problems: sometimes (or often; depending on patience) LLMs have serious logical problems with statements which are elementary for human users.

*** Chess with Computer Help ***

Karsten Müller (born 1970; Chess Grandmaster and PhD in Mathematics) was one of the early adoptors of chess programs to analyse positions (in 1992/1993).

In the last few months Karsten exclaimed several times about the development in mathematics: "I'm so glad I'm still alive to see this day!" He has been living with and for computer chess for more than three decades, but never lost his interest in math.

Chess program **Shredder** in 4-best analysis mode, version Shredder-3 from 1998 (author **Stefan Meyer-Kahlen**). See this site for more information on Shredder.

Screenshot of 4-best analysis mode (chess program Shredder from 1998): iterative deepening mode; for depth 8 the two best candidates have been analysed already. Currently no. 3 of depth 7 is under investigation. More screenshots (for other 2-player games than chess) here: https://althofer.de/k-best-visualisations.html

The company **ChessBase.com** (seated in Hamburg) produces and sells software for computer-aided chess analysis since 1987. World Champion **Garry Kasparov** was one of the very first power users and supporters of ChessBase, already in late 1986. It took the chess scene almost 20 years from the early professional adopters until chess programs were generally accepted as tools for analysis.

My prediction: Progress in computer-aided proving in Mathematics will be faster from now on. **The floodgates are open.**

Answers of GPT-5 and other LLM's are not deterministic.
And sometimes GPT is hallucinating.

** How can LLM-Hallucinating help in Mathematics?

- Brainstorming (example: two abstracts from a conferenc book: "meditate about mixing the topics in these two abstracts"
- finding sketchy new names for objects
- . . .

Preprints
(from my current personal preference list; Nov 13, 2025)

https://arxive.org/pdf/2112.11598
What is the point of computers? A question for pure mathematicians
Kevin Buzzard. April 18, 2022.

https://arxiv.org/pdf/2207.04779
Mathematical proof between generations
Jonas Bayer, Cristoph Benzmueller, Kevin Buzzard, Marco David, Leslie Lamport, Yuri Matiyasevich, Lawrence Paulson, Dierk Schleicher, Benedict Stock, and Efim Zelmanov. July 2022.

https://arxiv.org/html/2510.12350v1
O-Forge: An LLM + Computer Algebra Framework for Asymptotic Analysis
Ayush Khaitan and Vijay Ganesh. October 14, 2025.

https://arxiv.org/abs/2510.19804
Forbidden Sidon subsets of perfect difference sets, featuring a human-assisted proof
Boris Alexeev and Dustin G. Mixon. October 22, 2025.
https://arxiv.org/pdf/2511.02864
Mathematical exploration and discovery at scale
Bogdan Georgiev, Javier Gomez-Serrano, Terence Tao, and Adam-Zsolt Wagner. November 03, 2025.

https://arxiv.org/pdf/2511.07721
New Nikodym set constructions over finite fields
Terence Tao. November 11, 2025

Disclaimer: This is my very personal list. It will change over time. Other mathematicans will have other favorites.

* * * * * * *

Some screenshots with marks from the arxiv preprint 2511.07721,
highlighting interactive use of AI tools by Terence Tao

11 Nov 2025

NEW NIKODYM SET CONSTRUCTIONS OVER FINITE FIELDS

TERENCE TAO

ABSTRACT. For any fixed dimension $d \geq 3$ we construct a Nikodym set in $\mathbf{F}_q^d$ of cardinality $q^d - (\frac{d-2}{\log 2} + 1 + o(1))q^{d-1} \log q$ in the limit $q \to \infty$, when q is an odd prime power. This improves upon the naive random construction, which gives a set of cardinality $q^d - (d-1+o(1))q^{d-1} \log q$, and is new in the regime where $\mathbf{F}_q$ has unbounded characteristic and q not a perfect square. While the final proofs are completely human generated, the initial ideas of the construction were inspired by output from the tools `AlphaEvolve` and `DeepThink`. We also give a new construction of Nikodym sets in $\mathbf{F}_q^2$ for q a perfect square that match the existing bounds of $q^2 - q^{3/2} + O(q \log q)$, assuming that q is not the square of a prime $p \equiv 3 \pmod 4$.

* * * * * * *

see Section 2.1. However, this bound was only heuristic.

- After several failed attempts (by both the human authors and `DeepThink`) to make this heuristic precise, a weaker bound

$$\text{Nikodym}(d, q) \leq q^d - \left(\frac{d-2}{\log 2} + 1 + o(1)\right) q^{d-1} \log q \tag{1.9}$$

was established for odd q, which improved upon (1.7) when $d \geq 3$. (For even q, the

* * * * * * *

This construction was initially found by `AlphaEvolve` and verified by `DeepThink`, but the proof we give in Section 4 is human-written.

* * * * * * *

As discussed below, a more complicated variant of this method was first discovered by `Deepthink`. After suggesting a purely random construction, `Deepthink` was also able to reconstruct most of the details of the above argument.

* * * * * * *

one could increase k to approximately $\log(q^{d-1})/log2$, thus predicting the bound (1.8).

An initial attempt to use `Deepthink` to reproduce these heuristics was unsuccessful, and identified a geometric flaw in the construction: as quadratic polynomials of one variable will usu-

* * * * * * * * * * *

```
Two screenshots with marks from the arxiv preprint 2510.12350v2 by
A. Khaitan and V. Ganesh.
```

O-Forge: An LLM + Computer Algebra Framework for Asymptotic Analysis

Ayush Khaitan
Rutgers University
ayush.khaitan@rutgers.edu

Vijay Ganesh
Georgia Institute of Technology
vganesh@gatech.edu

Abstract

Large language models have recently demonstrated advanced capabilities in solving IMO and Putnam problems; yet their role in research mathematics has remained fairly limited. The key difficulty is verification: suggested proofs may look plausible, but cannot be trusted without rigorous checking. We present a framework, called LLM+CAS, and an associated tool, O-Forge, that couples frontier LLMs with a computer algebra systems (CAS) in an In-Context Symbolic Feedback loop to produce proofs that are both creative and symbolically verified. Our focus is on asymptotic inequalities, a topic that often involves difficult proofs and appropriate decomposition of the domain into the "right" subdomains. Many mathematicians, including Terry Tao, have suggested that using AI tools to find the right decompositions can be very useful for research-level asymptotic analysis. In this paper, we show that our framework LLM+CAS turns out to be remarkably effective at proposing such decompositions via a combination of a frontier LLM and a CAS. More precisely, we use an LLM to suggest domain decomposition, and a CAS (such as Mathematica) that provides a verification of each piece axiomatically. Using this loop, we answer a question posed by Terence Tao: whether LLMs coupled with a verifier can be used to help prove intricate asymptotic inequalities. More broadly, we show how AI can move beyond contest math towards research-level tools for professional mathematicians.

* * * * * * *

Our primary novelty is in being able to automate proof completion for difficult research problems that should take most research mathematicians lots of time and effort. No existing AI tools are able to complete and symbolically verify proofs of this kind. Moreover, although frontier LLMs may be able to produce some of these proofs, these proofs are often incorrect, and need to be manually verified. Our tool does away with the need for manual verification.

2 Acknowledgements

We gratefully acknowledge Terence Tao for repeatedly stress-testing our tool and suggesting substantial improvements to our website, o-forge.com. We also thank Swarat Chaudhuri, Aiman Kohli, Amitayush Thakur, and George Tsoukalas for their valuable help and advice.

My website on Collatz Prizes:
https://althofer.de/collatz-prizes.html

From time to time I get proposals for solutions from hobby mathematicians. Often a quick check by GPT-5 together with a serious control look by me is enough to identify serious errors. This typically saves me a lot of time. Of course it would be better if afficionados “check” the proofs with the help of GPT-5.1 or Gemini 3 before submitting them.

Thanks to Karsten Müller, Dietmar Wolz, Ulrich Tamm, Torsten Sillke, Michael Taktikos, and Heinrich Burger (first top correspondence chess player to use intensive computer help in 1988) for collaboration, discussions and encouragement.