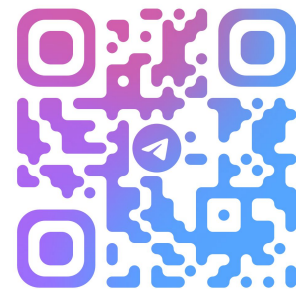


Vibemathed

Как LLM делают математику



@SINECOR

Сергей
Николенко

17 августа 2026

VI Конференция
Математических
центров России



1. Древняя история (до 2026)

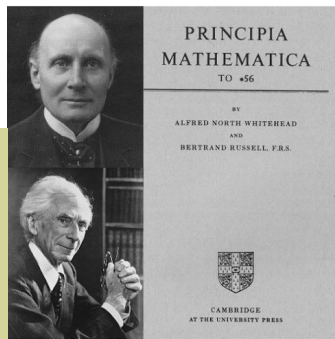
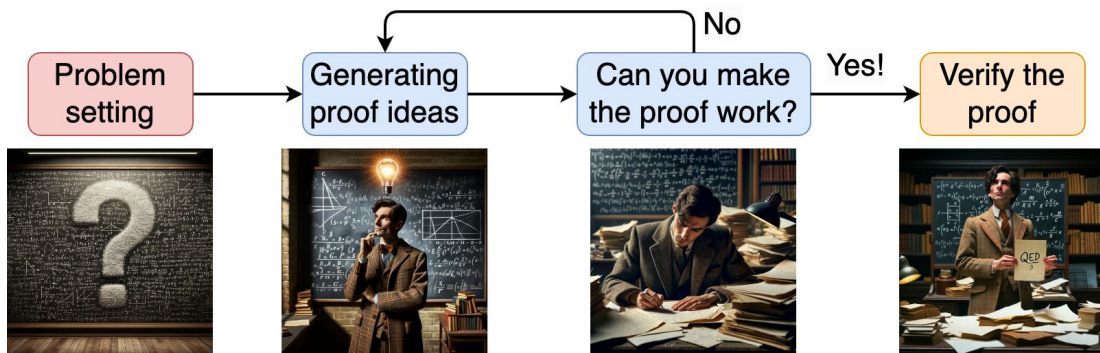


Я так привык к самостоятельной работе, что считал более лёгким для себя доказать теорему без книги, чем вычитывать из нее доказательства. Лишь не всегда это удавалось.

Константин Циолковский

LLM для математики

- Работа математика – это интересный случай, который всегда оставался для меня загадкой
- В математической логике давно развиваются методы автоматического доказательства теорем, но они так ни к чему выдающемуся и не привели...



*54.43. $\vdash \therefore \alpha, \beta \in 1. \supset : \alpha \wedge \beta = \Lambda. \equiv . \alpha \vee \beta \in 2$

Dem.

$\vdash . *54.26. \supset \vdash : \alpha = \iota'x. \beta = \iota'y. \supset : \alpha \vee \beta \in 2. \equiv . x \neq y.$

$[*51.231] \quad \equiv . \iota'x \wedge \iota'y = \Lambda.$

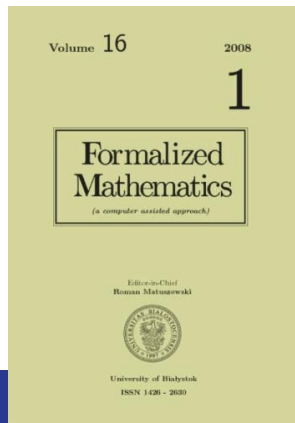
$[*13.12] \quad \equiv . \alpha \wedge \beta = \Lambda \quad (1)$

$\vdash . (1). *11.11.35. \supset$

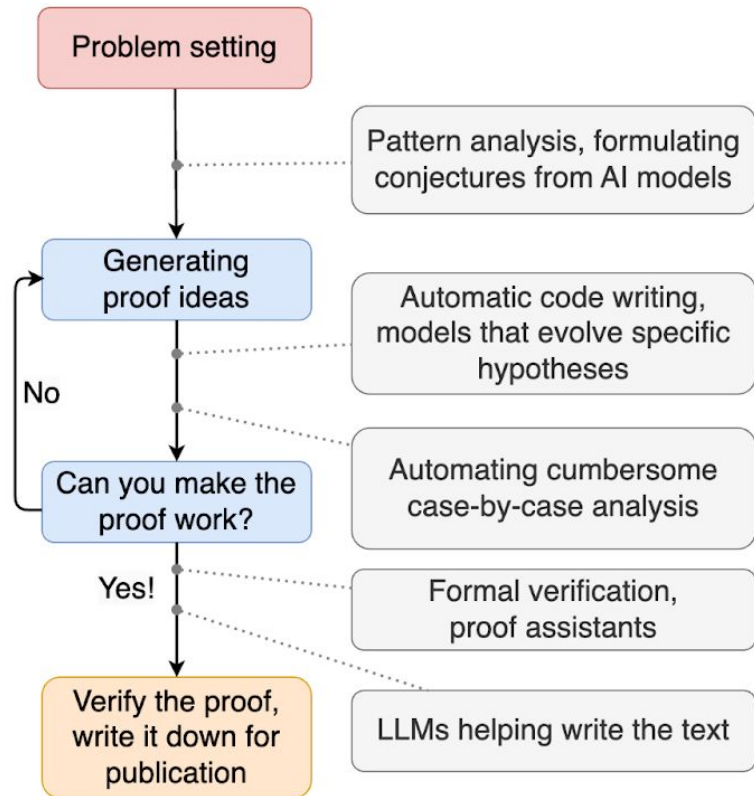
$\vdash : . (\exists x, y). \alpha = \iota'x. \beta = \iota'y. \supset : \alpha \vee \beta \in 2. \equiv . \alpha \wedge \beta = \Lambda \quad (2)$

$\vdash . (2). *11.54. *52.1. \supset \vdash . \text{Prop}$

From this proposition it will follow, when arithmetical addition has been defined, that $1 + 1 = 2$.



LLM для математики



Advancing mathematics by guiding human intuition with AI

Alex Davies¹, Petar Veličković², Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomasek, Richard Tenbourn, Peter Battaglia, Charles Blundell, András Juhász, Marc Lackenby, Georgie Williamson, Demis Hassabis & Pushmeet Kohli³

Nature 600, 70–74 (2021) | [Cite this article](#)



TORA: A TOOL-INTEGRATED REASONING AGENT FOR MATHEMATICAL PROBLEM SOLVING

Zhibin Guo^{1,2*}, Zhihong Shao^{1,2*}, Yeyun Gong^{2†}, Yelong Shen², Yujie Yang¹, Minlie Huang¹, Nan Du², Weishi Chen²
¹Tsinghua University ²Microsoft
 {gzh22, szh19}@email.tsinghua.edu.cn
 {yeyun.gong, yeshao, nanduan, wachen}@microsoft.com

Discovering faster matrix multiplication algorithms with reinforcement learning

Alhussein Fawzi¹, Matej Balog, Aja Huang, Thomas Hubert, Bernardino Romera-Paredes, Mohammadamin Barekatin, Alexander Novikov, Francisco J. R. Ruiz, Julian Schrittwieser, Grzegorz Swirszcz, David Silver, Demis Hassabis & Pushmeet Kohli

Nature 610, 47–53 (2022) | [Cite this article](#)



Mathematical discoveries from program search with large language models

Bernardino Romera-Paredes¹, Mohammadamin Barekatin, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilien Ducent, Francisco J. R. Ruiz, Jordan S. Ellenberg, Penporn Wang, Omar Fawzi, Pushmeet Kohli² & Alhussein Fawzi³

Nature 625, 468–475 (2024) | [Cite this article](#)

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 3, September 1976

RESEARCH ANNOUNCEMENTS

EVERY PLANAR MAP IS FOUR COLORABLE¹

BY K. APPEL AND W. HAKEN

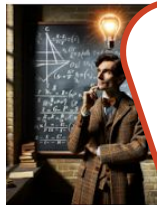
Communicated by Robert Fossum, July 26, 1976



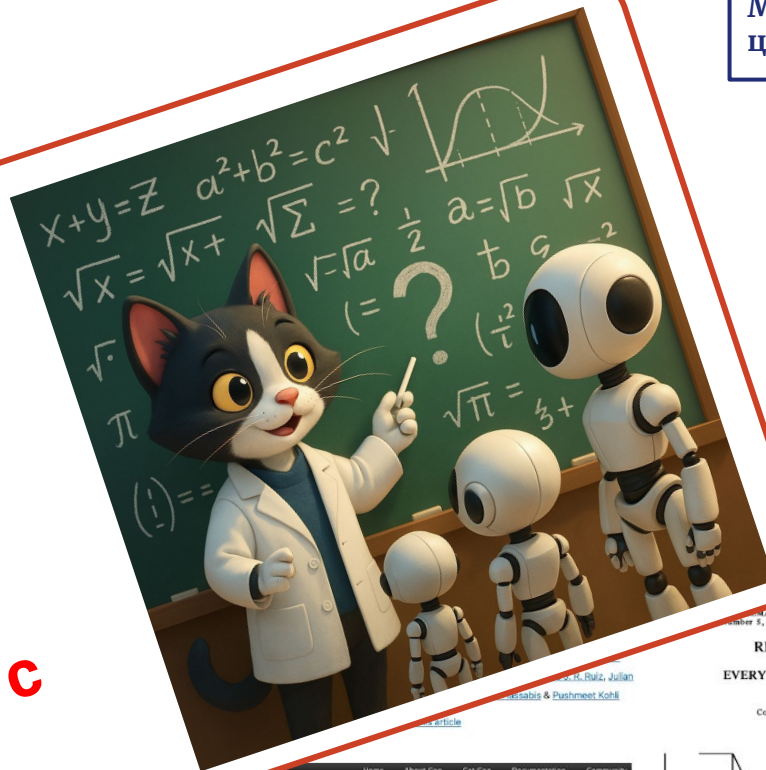
LLM для математики



Problem setting



**Схему про LLM
для математики
я сделал в
начале 2024
года; что
изменилось с
тех пор?**



discoveries from program
language models

Mohammadamin Barekatin, Alexander Novikov, Matej
Dvořák, Francisco J. R. Ruiz, Jordan S. Ellenberg,
Pushmeet Kohli & Alhussein Fawzi

Cite this article

AMERICAN MATHEMATICAL SOCIETY
November 1, September 1976

RESEARCH ANNOUNCEMENTS

EVERY PLANAR MAP IS FOUR COLORABLE¹

BY K. APPEL AND W. HAKEN

Communicated by Robert Fossum, July 26, 1976

The Coq Proof Assistant

LMN
Community

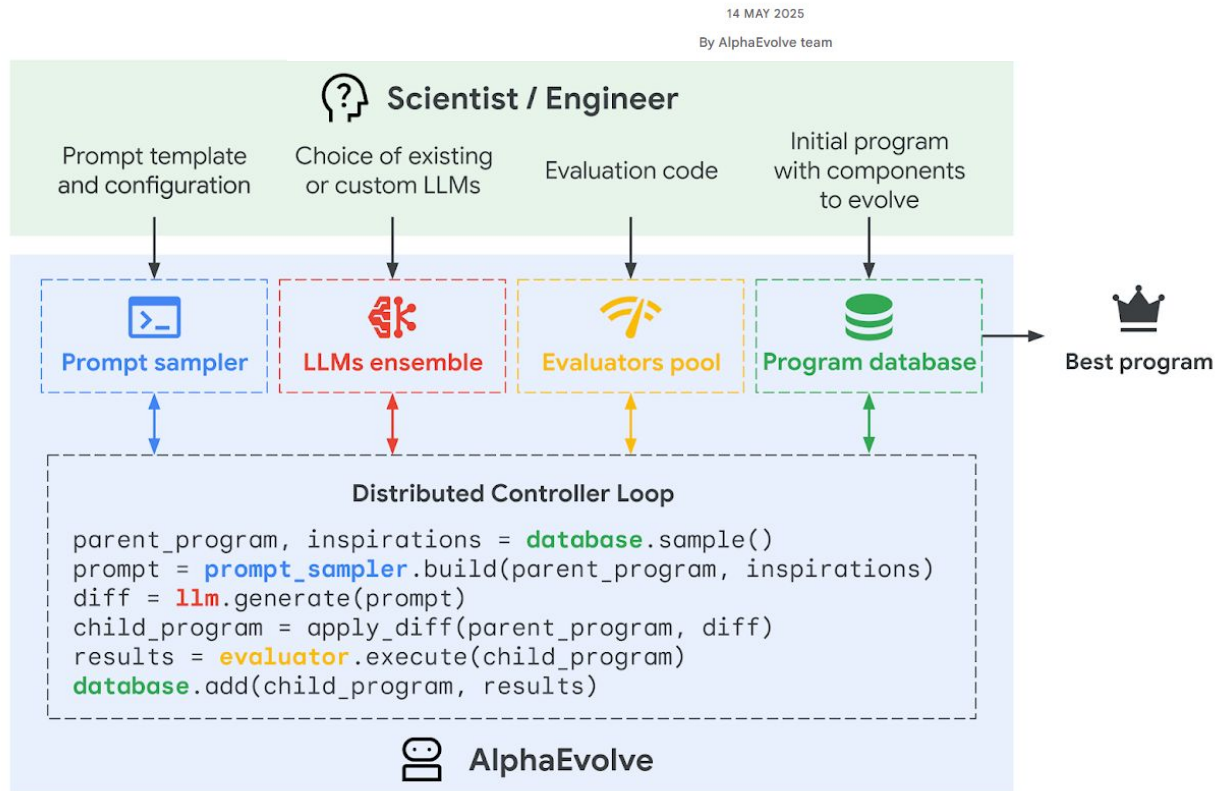
LLMs helping write the text



AlphaEvolve: A Gemini-powered coding agent for designing advanced algorithms

AlphaEvolve

- AlphaEvolve ([DeepMind, May 14, 2025](#)) и тому подобные системы развиваются активно
- Но мы сегодня поговорим про обычные LLM, с которыми каждый из нас может пообщаться
- Помогают ли они?..



Malliavin-Stein experiment

- Diez et al. (Sep 3, 2025): авторы попросили GPT-5 доказать новую теорему из теории вероятностей; буквально:

4.1.1 Gaussian framework

We started with the following initial prompt:

Paper 2502.03596v1 establishes a qualitative fourth moment theorem for the sum of two Wiener-Itô integrals of orders p and q , where p and q have different parities. Building on the Malliavin-Stein method (see 1203.4147v3 for details), could you derive a quantitative version for the total variation distance, with a convergence rate depending solely on the fourth cumulant of this sum?

Mathematical research with GPT-5: a Malliavin-Stein experiment

Charles-Philippe Diez* Luís da Maia* Ivan Nourdin*

Abstract

On August 20, 2025, GPT-5 was reported to have solved an open problem in convex optimization. Motivated by this episode, we conducted a controlled experiment in the Malliavin-Stein framework for central limit theorems. Our objective was to assess whether GPT-5 could go beyond known results by extending a *qualitative* fourth-moment theorem to a *quantitative* formulation with explicit convergence rates, both in the Gaussian and in the Poisson settings. To the best of our knowledge, the derivation of such quantitative rates had remained an open problem, in the sense that it had never been addressed in the existing literature. The present paper documents this experiment, presents the results obtained, and discusses their broader implications.

Malliavin-Stein experiment

- И после такого запроса авторы отмечают, что в доказательстве была ошибка, и GPT-5 потребовалось аж два наводящих промпта (!), чтобы ошибку понять и исправить
- Затем GPT-5 сам предложил рассмотреть другой случай, и в нём опять была ошибка, и понадобилось указать модели на неё! После этого GPT-5 за три-четыре промпта пришёл к верному результату

I think you are mistaken in claiming that $(p+q)! \|u \tilde{\otimes} v\|^2 = p!q! \|u\|^2 \|v\|^2$. Why should that be the case?

It eventually admitted (which is not surprising, since by alignment it usually agrees with us) that the statement was false, but more importantly, it understood where the mistake came from. This was followed by a reasoning and a formula that, this time, were correct.

Then, at our request, GPT-5 reformatted the result in the style of a research article, including an introduction, the presentation of our main theorem, its proof with all the details (correct this time!), and a bibliography. The exact prompt was:

Turn this into a research paper ready for submission. Follow my style (see attached paper 0705.0570v4):

- start with an introduction giving some context,
- then present the main result, followed by a very detailed proof where no step is left out,
- finish with a complete bibliography.

The final document should be a LaTeX file that I can compile.

Malliavin-Stein experiment

- Что же заключают авторы?

4.1.3 Role of the AI

To summarize, we can say that the role played by the AI was essentially that of an executor, responding to our successive prompts. Without us, it would have made a damaging error in the Gaussian case, and it would not have provided the most interesting result in the Poisson case, overlooking an essential property of covariance, which was in fact easily deducible from the results contained in the document we had provided.

- Кажется, это называется copium...

5 Some personal reflections

Overall, the experience of doing mathematics with GPT-5 was mixed. It felt very similar to working with a junior assistant at the beginning of a new project: exploring directions, formulating hypotheses, searching for counterexamples, and progressively adjusting statements. The AI showed a genuine ability to follow guided reasoning, to recognize its mistakes when pointed out, to propose new research directions, and to never take on the task. However, this only seems to support *incremental* research, that is, producing new results that do not require genuinely new ideas but rather the ability to combine ideas coming from different sources. At first glance, this might appear useful for an exploratory phase, helping us save time. In practice, however, it was quite the opposite: we had to carefully verify everything produced by the AI and constantly guide it so that it could correct its mistakes.

Gödel Test

- [Feldman, Karbasi \(Sep 22, 2025\)](#): тест Гёделя – могут ли LLM доказывать простые, но новые теоремы?
- Вот что писал об этом Теренс Тао, когда познакомился с o1

“The new model could work its way to a correct (and well-written) solution if provided a lot of hints and prodding, but did not generate the key conceptual ideas on its own, and did make some non-trivial mistakes. The experience seemed roughly on par with trying to advise a mediocre, but not completely incompetent, graduate student. However, this was an improvement over previous models, whose capability was closer to an actually incompetent graduate student. It may only take one or two further iterations of improved capability (and integration with other tools, such as computer algebra packages and proof assistants) until the level of ‘competent graduate student’ is reached.”

— Terence Tao



Gödel Test

- [Feldman, Karbasi \(Sep 22, 2025\)](#): дали пять запросов в таком духе:

Prompt to GPT-5

Consider the problem of maximizing a function F from $[0, 1]^n$ to the reals that is the sum of a non-negative monotonically increasing DR-submodular function G and a non-negative DR-submodular function H over a solvable down-closed polytope P . I would like to bound the performance on this problem from the NeurIPS 2021 paper "Submodular + Concave" which is attached. Specifically, if x is the output vector of this algorithm and o is the vector in P maximizing F , then I would like to lower bound $F(x)$ with an expression of the form $\alpha * G(o) + \beta * H(o) - err$, where α and β are constants. err should be a function that depends only on the error parameter ϵ of the algorithm, the diameter D of the polytope P , and the smoothness parameters L_G and L_H of G and H , respectively, and goes to zero as ϵ goes to zero. Please give the best such bound that you prove (a bound is considered better if the values of the constants α and β are larger). Provide a mathematically rigorous and well explained proof for the bound you come up with.

Gödel Test

- [Feldman, Karbasi \(Sep 22, 2025\)](#)

Prompt to GPT-5

Consider the problem of finding a non-trivial lower bound for the performance of a non-linear decision rule (DR) submitted to a set of data points, which is attained by the optimal linear decision rule in P maximally. The bound is of the form $\alpha * G(o) + \beta * \epsilon$, where α and β are constants, $G(o)$ is a function that depends only on the error of the linear decision rule, and ϵ is the smoothness parameter. Please give the values of the constants α and β and provide a mathematically rigorous and well explained proof for the bound you propose with.

**GPT-5 решил три из пяти, с небольшими
легко исправимыми недочётами. Главный
“ленивый” недостаток — GPT-5 писал
ссылаясь на предоставленные статьи, даже
когда это было не очень изящно**

2. Новые результаты от LLM

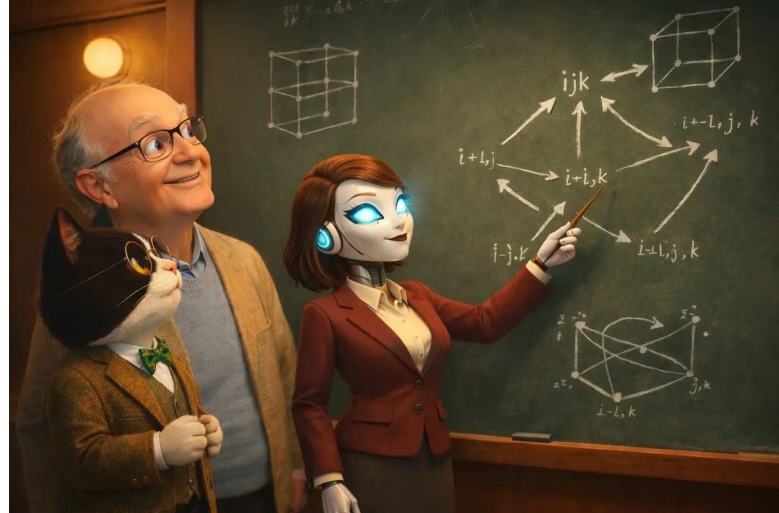


Одинаковость языка есть первая таинственная связь, соединяющая людей между собою. Народ говорит одним языком, и в этом единстве выражается внутренняя симпатия, родство душ, по которому люди одного народа сливают звуки в известные стройные созвучия, выражая ими внутренние и внешние свои понятия. Один человек говорит, и другой понимает его, — и вот между ними утверждена прочная связь взаимного разумения.

Константин Аксаков. О грамматике вообще

Claude's Cycles

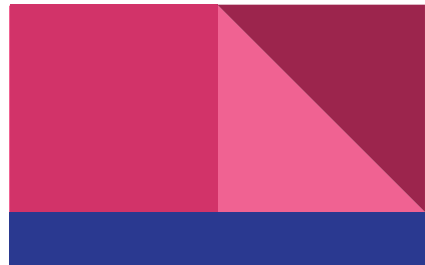
- Donald Knuth (Feb 2026): Claude Opus 4.6 решил открытую проблему, над которой Кнут работал неделями; решение потом подтвердилось и даже было формализовано



Shock! Shock! I learned yesterday that an open problem I'd been working on for several weeks had just been solved by Claude Opus 4.6—Anthropic's hybrid reasoning model that had been released three weeks earlier! It seems that I'll have to revise my opinions about “generative AI” one of these days. What a joy it is to learn not only that my conjecture has a nice solution but also to celebrate this dramatic advance in automatic deduction and creative problem solving. I'll try to tell the story briefly in this note.

Here's the problem, which came up while I was writing about directed Hamiltonian cycles for a future volume of *The Art of Computer Programming*:

Consider the digraph with m^3 vertices ijk for $0 \leq i, j, k < m$, and three arcs from each vertex, namely to i^+jk , ij^+k , and ijk^+ , where $i^+ = (i+1) \bmod m$. Try to find a general decomposition of the arcs into three directed m^3 -cycles, for all $m > 2$.



Erdős Problems

- 1217 задач Эрдеша в базе erdosproblems.com, из них 565 решены
- Многие из них совсем не знаменитые, возможно, некоторые случайно решены где-то в литературе, но много и сложных нерешённых задач

Problems have always been an essential part of my mathematical life. A well-chosen problem can isolate an essential difficulty in a particular area, serving as a benchmark against which progress in this area can be measured. It might be like a 'marshmallow', serving as a tasty tidbit supplying a few moments of fleeting enjoyment. Or it might be like an 'acorn', requiring deep and subtle new insights from which a mighty oak can develop. [...]

In this note I would like to describe a variety of my problems which I would classify as my favorites. Of course, I can't guarantee that they are all 'acorns', but because many have thwarted the efforts of the best mathematicians for many decades (and have often acquired a cash reward for their solutions), it may indicate that new ideas will be needed, which can, in turn, lead to more general results, and naturally, to further new problems.

In this way, the cycle of life in mathematics continues forever.

Paul Erdős, *Some of my favorite problems and results*, 1997

OPEN

Let $f \in \mathbb{R}[x]$ be a monic polynomial of degree n with n real roots in $[-1, 1]$. Determine the infimum and supremum of possible values of

$$|\{x \in \mathbb{R} : |f(x)| < 1\}|.$$

#1038: [EHP58,p.131]

analysis

FALSIFIABLE

Let $z_1, \dots, z_n \in \mathbb{C}$ with $|z_i - z_j| \leq 2$ for all i, j , and

$$\Delta(z_1, \dots, z_n) = \prod_{i \neq j} |z_i - z_j|.$$

What is the maximum possible value of Δ ? Is it maximised by taking the z_i to be the vertices of a regular polygon?

#1045: [EHP58,p.143]

analysis

Erdős Problems

- Теренс Тао
сделал страничку
“[AI contributions to Erdős problems](#)”
- Результатов
быстро
накопилось
довольно много

AI contributions to Erdős problems

teorth edited this page 4 hours ago · [70 revisions](#)

This page collects the various ways in which AI tools have contributed to the understanding of Erdős problems. Note that a single problem may appear multiple times in these lists.

Legend

- : full solution to the problem
- : partial solution to the problem
- : failure to make progress on the problem

5. Problems with an AI-powered literature review

Problem	AI tools used	Review performed on	Outcome
[35]	GPT-5	13 Oct, 2025	● Partial results found
[66]	GPT-5	13 Oct, 2025	● Partial results found
[124]	ChatGPT DeepResearch, Gemini DeepResearch	30 Nov, 2025	No significant results found
[167]	GPT-5	12 Oct, 2025	● Partial results found
[188]	GPT-5	13 Oct, 2025	● Partial results found
[203]	ChatGPT DeepResearch, Gemini DeepResearch	19 Oct, 2025	No significant results found
[223]	GPT-5	13 Oct, 2025	● Full solution found
[248]	Gemini DeepResearch	19 Oct, 2025	No significant results found
[330]	ChatGPT DeepResearch, Gemini DeepResearch, Claude	19 Dec, 2025	● Partial results found (with inaccuracies); did not find literature proof

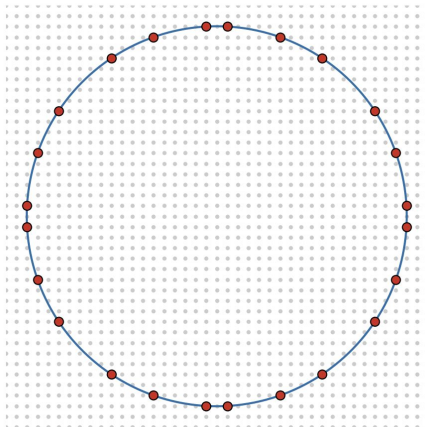
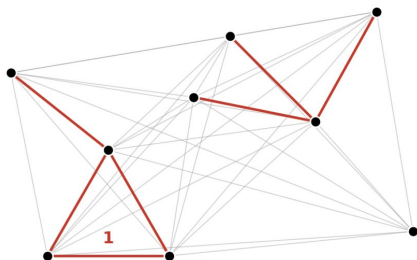
Erdős Problems

- Теренс Тао сделал страничку “AI contributions to Erdős problems”
- Она, правда, в основном ограничивается апрелем, есть несколько записей из июня 2026

[1041]	GPT-5.4 Thinking	22 Apr, 2026	● Argument with major gaps made
[1044]	GPT-5.2 Pro	5 Jan, 2026	● Incorrect claim made
[1051]	Aletheia	29 Jan, 2026	● Full solution (Lean)
[1054]	Claude Opus 4.8, GPT-5.5 Pro, Principia Math	22-23 Jun, 2026	● Partial result (Lean)
[1091]	OpenAI internal model	9 Apr, 2026	● Full solution
[1101]	GPT-5.5 Pro	27 Apr, 2026	○ Candidate partial result (solution to second part)
[1131]	GPT-5.5 Pro	26 Apr, 2026	○ Candidate partial result
[1141]	OpenAI internal model	9 Apr, 2026	● Full solution
[1194]	GPT-5.4 Pro	23 Apr, 2026	● Partial result
[1196]	GPT-5.4 Pro	13 Apr, 2026	● Full solution
[1202]	GPT-5.4 Pro	1 Apr, 2026	● Full solution
[1217]	GPT-5.4 Pro	16 Apr, 2026	● Full solution

Единичные расстояния

- Но в апреле уже началось!
- Jared Lichtman (Apr 2026): GPT-5.4 Pro опроверг важную гипотезу Эрдёша (над ней думало много математиков), причём сделал это неожиданным и красивым образом



Paul Erdős

"On sets of distances of n points"

Amer. Math. Monthly 53 (1946), 248-250

Постановка задачи. Нижняя оценка $\nu(n) \geq n^{1 + c \log \log n}$, верхняя оценка $\nu(n) \leq O(n^{3/2})$.

Joel Spencer, Endre Szemerédi, William T. Trotter

"Unit distances in the euclidean plane"

Graph Theory and Combinatorics (Cambridge 1983), 293-303

Верхняя оценка $\nu(n) \leq O(n^{4/3})$.

László A. Székely

"Crossing numbers and hard Erdős problems in discrete geometry"

Combinatorics, Probability and Computing 6 (1997), 353-358

Эlegantное короткое доказательство $O(n^{4/3})$ через crossing lemma.

Noga Alon, Matija Bucić, Lisa Sauermann

"Unit and distinct distances in typical norms"

GAFA 35 (2025), 1-42 · arXiv:2302.09058

Для нормы общего положения $\nu(n) \leq \frac{c}{2} n \log_2 n$.

Josef Greiher, Carl Schildkraut, Jonathan Tidor

"More unit distances in arbitrary norms"

Bull. London Math. Soc. 57 (2025), 2885-2901 · arXiv:2410.07557

Соответствующая нижняя оценка $\Omega(n \log_2 n)$ для любой нормы.

OpenAI

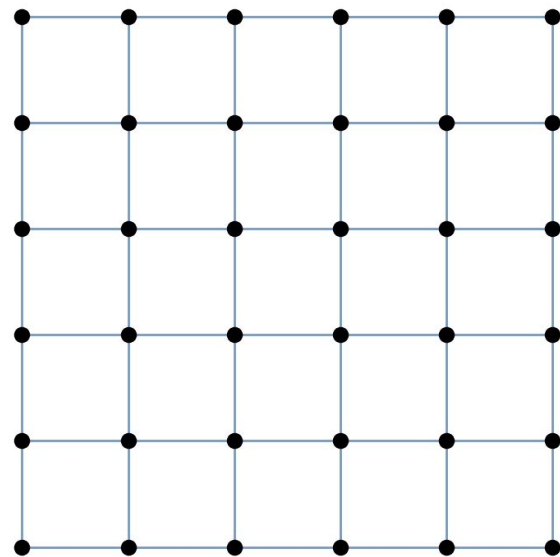
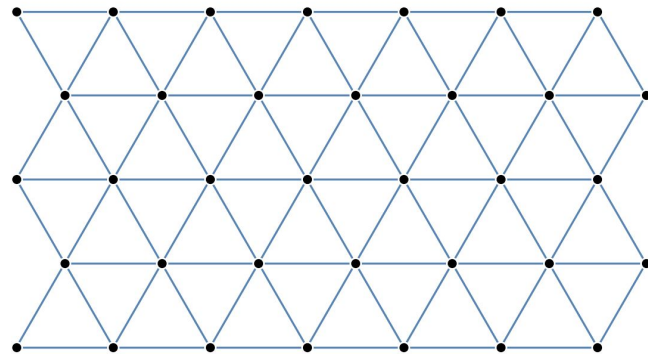
"Planar Point Sets with Many Unit Distances"

openai.com, 20 мая 2026 · комментарии: Alon et al., arXiv:2605.20695

Контрпример: $\nu(n) \geq n^{1+\delta}$ для $\delta > 0$ и бесконечно многих n .
Гипотеза Эрдёша опровергнута.

Единичные расстояния

- Рассмотрим n точек на обычной евклидовой плоскости. Сколько одинаковых расстояний может быть среди пар этих точек?
- Треугольная решётка даёт $3n$
- Главный подход: превратить задачу в арифметическую. Рассмотрим квадратную решётку; на ней расстояние m встречается столько раз, сколько раз m^2 раскладывается в a^2+b^2 .



Единичные расстояния

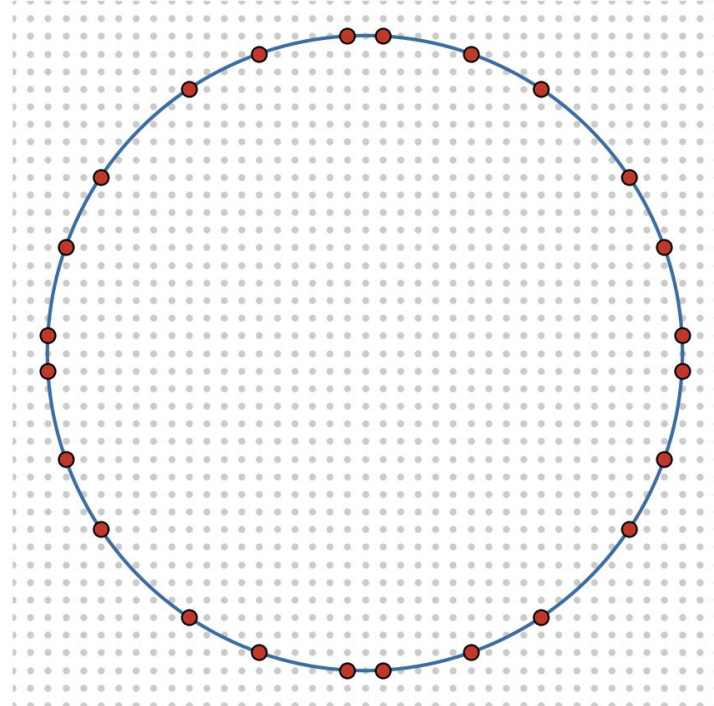
- Каждое $p=4k+1$ можно разложить как
$$4k + 1 = (1 + 2i\sqrt{k})(1 - 2i\sqrt{k}) = z \cdot \bar{z}, \quad z = 1 + 2i\sqrt{k}.$$
- Если число m имеет в разложении много простых делителей p_i вида $4k+1$, то каждое из них можно так разложить, а потом представить m как произведение двух комплексно сопряжённых подмножеств
- Например, у числа 325 на окружности радиуса $\sqrt{325}$ сразу 24 целые точки:

$$325 = 5^2 \cdot 13 = (2 + i)^2(2 - i)^2(3 + 2i)(3 - 2i).$$

$$325 = \left((2 + i)^2(3 - 2i)\right) \left((2 - i)^2(3 + 2i)\right) = (17 + 6i)(17 - 6i) = 17^2 + 6^2.$$

$$325 = \left((2 + i)^2(3 + 2i)\right) \left((2 - i)^2(3 - 2i)\right) = (1 + 18i)(1 - 18i) = 1^2 + 18^2.$$

$$325 = ((2 - i)(2 + i)(3 + 2i)) ((2 + i)(2 - i)(3 - 2i)) = (15 + 10i)(15 - 10i) = 15^2 + 10^2.$$



Единичные расстояния

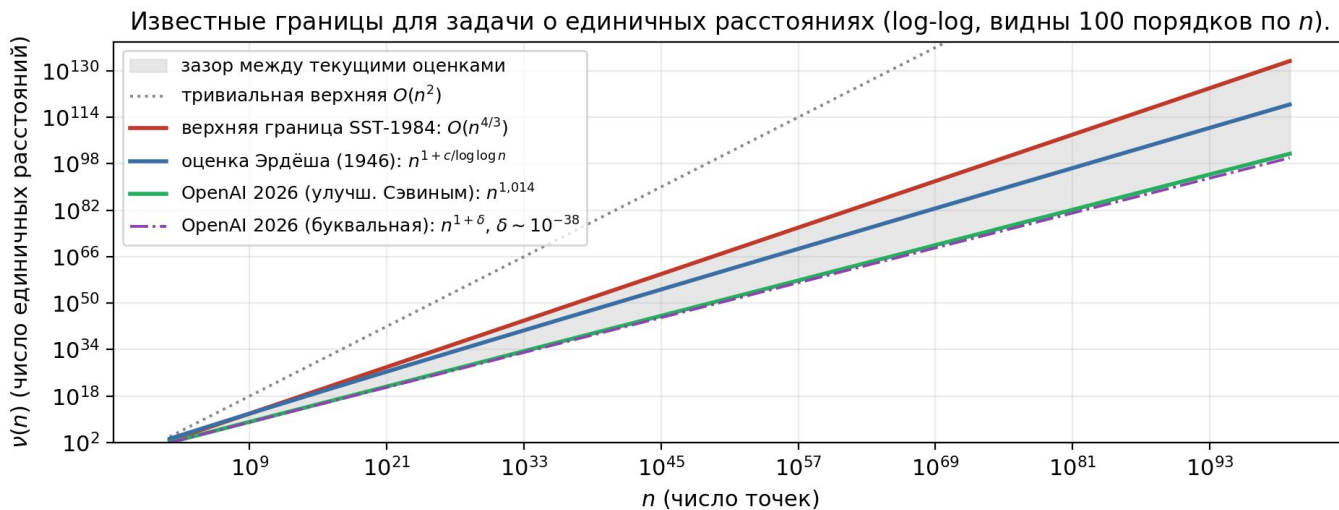
- Если аккуратно посчитать такие пары точек на окружности, получится уже нелинейная оценка: $\nu(n) \geq n^{1+c/\log \log n}$.

Гипотеза Эрдёша. $\nu(n) \leq n^{1+o(1)}$, а скорее даже

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

Теорема (OpenAI, 2026). Существует константа $\delta > 0$ и бесконечная последовательность натуральных чисел n , для которых $\nu(n) \geq n^{1+\delta}$.

- Можно ли лучше?
- Эрдёш предположил, что нет, но GPT доказал, что можно!



Единичные расстояния

- Если аккуратно посчитать такие пары точек на окружности, получится...

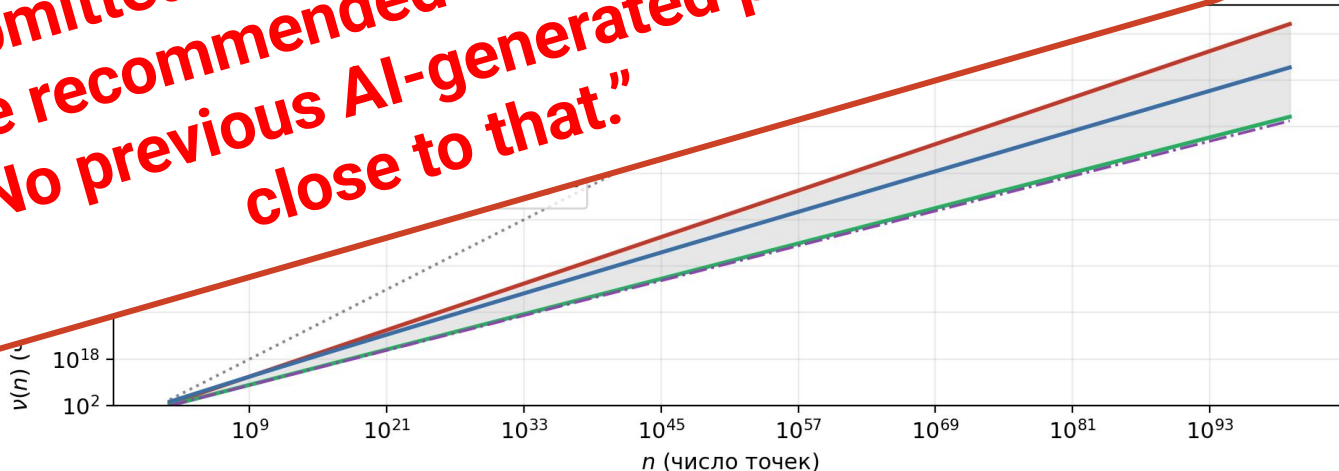
Гип

$\log \log n$

$+\delta$

в по n).

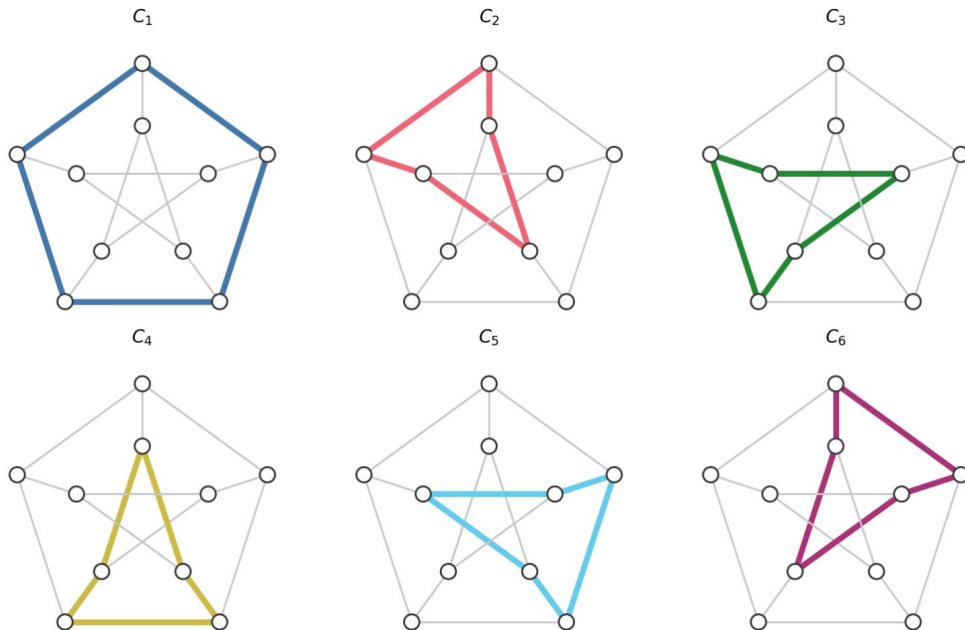
Это доказательство вызвало большую реакцию.
Например, Тимоти Гауэрс: "If a human had written the
paper and submitted it to the Annals of Mathematics...
I would have recommended acceptance without any
hesitation. No previous AI-generated proof has come
close to that."



Двойное покрытие циклами

- Верно ли, что в любом графе без мостов есть набор простых циклов, для которого каждое ребро содержится ровно в двух циклах этого набора?
- Поставили задачу ещё в 1970-х, люди много думали, были продвижения
- [OpenAI, 10 июля 2026](#): препринт с полным доказательством
- Мой пост: "[Охота на сарка: доказательство гипотезы о двойном покрытии циклами](#)"

двойное покрытие графа Петерсена шестью пятиугольниками:
каждое ребро выделено ровно в двух копиях



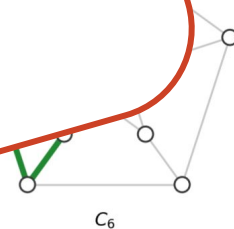
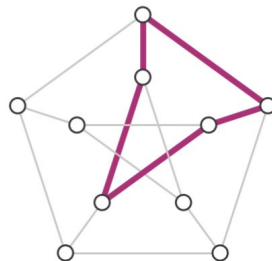
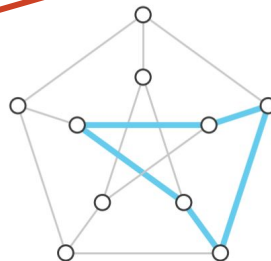
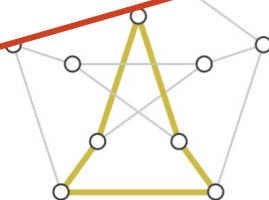
Двойное покрытие циклами

- Верно ли, что в любом графе без мостов есть набор простых циклов, для которого каждое ребро содержится в ровно двух циклах?

Месяц назад я подробно рассказывал об этих результатах и предполагал, что они будут в центре моего сегодняшнего доклада.

Но потом началось по-настоящему...

Двойное покрытие циклами



Проблема якобиана

- Проблема якобиана, стоявшая с 1930-х: если определитель якобиана отображения F — ненулевая константа, то F обратимо (и обратное отображение автоматически полиномиально)
- Главный открытый вопрос о многочленах от нескольких переменных, который можно рассказать любому студенту
- И его 20 июля 2026 г. закрыл Claude Fable, места для ошибки там нет; читайте [мой подробный пост](#)



levent ✓

@__alpage__



hello there the jacobian conjecture is false thanx to my close friend akhil for asking about it and my other close friend fable for working during the world cup final

$((1+xy)^3 z + y^2(1+xy)(4+3xy), y + 3x(1+xy)^2 z + 3xy^2(4+3xy), 2x - 3x^2 y - x^3 z): \mathbb{C}^3 \rightarrow \mathbb{C}^3$, has jacobian determinant -2 , and sends $(0, 0, -1/4)$, $(1, -3/2, 13/2)$, and $(-1, 3/2, 13/2)$ to $(-1/4, 0, 0)$

5:19 AM · Jul 20, 2026 · 21.9M Views

$$\begin{aligned} F_1 &= (1 + xy)^3 z + y^2(1 + xy)(4 + 3xy), \\ F_2 &= y + 3x(1 + xy)^2 z + 3xy^2(4 + 3xy), \\ F_3 &= 2x - 3x^2 y - x^3 z. \end{aligned}$$

Гипотеза Диница — Гарга — Гёманса

- Dinitz-Garg-Goemans conjecture: дробный поток можно превратить в неделимый (целочисленный), не увеличивая общую цену и увеличивая нагрузку одного ребра не более чем на максимальный спрос (demand) одной вершины-стока
- Дмитрий Рыбин получил контрпример от GPT 5.6; пожалуй, самое интересное здесь в том, как он его получил...

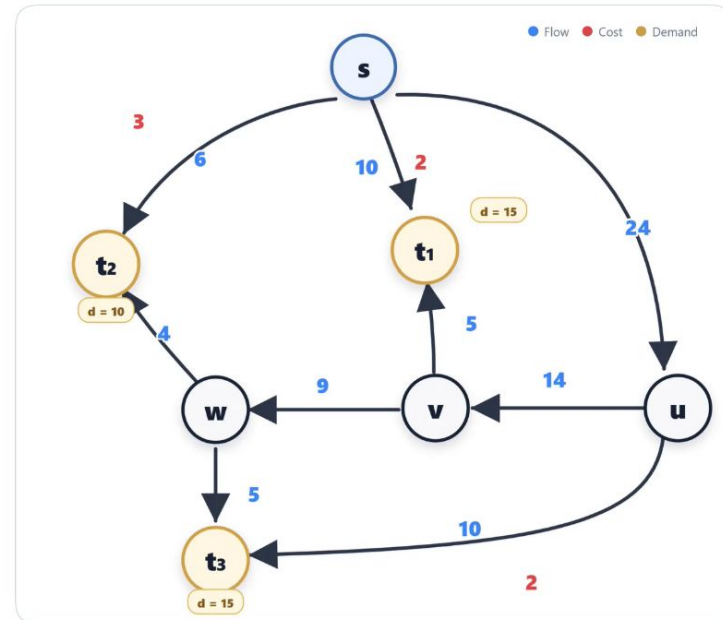


Dmitry Rybin
@DmitryRybin1

Dinitz-Garg-Goemans conjecture is false. This graph theory problem was open for ~30 years.

The graph below has fractional flow cost 58. Any unsplittable flow (with capacity violation ≤ 15) has cost at least 60.

Chat with GPT 5.6 Pro where this was found: chatgpt.com/share/6a60b2eb...



Гипотеза Диница — Гарга — Гёманса

Tuesday 12:29 PM

Uploaded a file

Construct a counterexample to general (non-planar) case of Dinitz Garg Goemans conjecture. You should do a breakthrough and find a structured counterexample.

Worked for 52m 52s >

Research conclusion

I did **not** obtain a valid counterexample, and presenting the construction below as one would be mathematically false. The requested object would disprove a conjecture that is still open in the current literature.

Tuesday 1:32 PM

please continue research and find a complete unconditional counterexample

Worked for 89m 12s >

Status after the extended search

I have **not found a complete unconditional counterexample**. Every construction that initially appeared to separate the cost condition acquired an additional unsplittable routing once all source-terminal paths were taken into account. Exact separation then returned value 0.

Tuesday 3:03 PM

Continue the search. Have a clear strategy obtained from deeper understanding of the problem structure.

Worked for 93m 57s >

Research update

I continued the search, but I do **not yet have a positive, exact separation certificate**, so there is still no unconditional counterexample to report. The conjecture remains open in the current literature. Its exact

Tuesday 5:21 PM

it's enough of partial results. let's finish with a complete unconditional counterexample

Worked for 88m 24s >

Complete finite counterexample

Under Conjecture 1.3 exactly as stated in the uploaded paper—find an unsplittable flow y with

$$y_a \leq x_a + d_{\max} \quad \forall a, \quad c^T y \leq c^T x$$

for every feasible fractional flow x and nonnegative cost vector c —the instance below is a counterexample.

OpenAI Astra

- OpenAI (1 августа 2026): сразу десять выдающихся результатов
- Существование несофических групп (явной конструкцией), верхние оценки на упаковку сфер, опровержение гипотезы Конна о жёсткости...
- Все эти результаты были сразу же формализованы в Lean

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 nonsoc 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.

OpenAI Astra

- OpenAI (1 августа 2026): сами эти доказательства стоили всего лишь около \$2000 по API-ценам GPT 5.6 SoI
- Конечно, было ещё много попыток, которые ни к чему не привели



Noam Brown ✓
@polynoamial

...

And yes we did try other major problems without success. Sadly no Millennium Prize problems (yet).

But also, we didn't spend a lot on each problem. It's possible to push test-time compute much further.

12:01 PM · Aug 1, 2026 · 264.4K Views

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 nonsoc 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.

OpenAI Astra

● Интересный поворот:



so after 24h i have half of them with fable

i didn't see much discussion of prompting in the announcement but this is a similar setup as with my e.g. unit distance announcement:

totally autonomous,
generic prompt,
no internet
(+ paranoia to ensure no information leaked)

 **Sebastien Bubeck**  @SebastienBubeck · Aug 1

yes, nonsoc groups exist: this statement is one of many new beautiful results proved by Astra, our next major model.

We're releasing 10 such Astra proofs, complete with lean certificates and CoT walkthroughs for each of them. The results are wide-ranging, from von ...

12:59 PM · Aug 2, 2026 · **333K** Views

 57  50  944  241 

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 nonsoc 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.

Гипотеза Крузе

- Доказательство нашёл практикующий нейрохирург, математик-любитель Dr. Shanmu Jin
- Проверили [Townsend, Greenbaum \(Aug 14, SIAM News\)](#), а также сам Крузе
- "We initially read the preprint with some skepticism, but after a few hours we realized that the argument was the real deal... We do not yet know whether academia has the capacity to absorb the flood of results that may be coming"
- [Lorist, Schwenninger \(Aug 4, 2026\)](#): совсем другое короткое доказательство, тоже AI-assisted

The Neurosurgery Resident Who Proved Crouzeix's Conjecture

Alex Townsend and Anne Greenbaum

August 14, 2026

On July 30, 2026, the first author asked ChatGPT 5.6 to solve Crouzeix's conjecture, as he had been routinely doing for the preceding year. Until that day, it had either returned a bogus proof or produced an argument that stalled at a missing lemma. But on July 30 it linked to a preprint posted July 27 titled *The Numerical Range Is a 2-Spectral Set*, claiming that the conjecture was solved. Crouzeix's conjecture had been open for more than two decades, so we initially read the preprint with some skepticism, but after a few hours we realized that the argument was the real deal.



Пример Анджея Козловского

Они написали статью, но GPT нашёл там ошибку

Andrzej Kozlowski @akoz33 · Aug 9

OK, more on ChatGPT and mathematics research. I am now working on a paper in algebraic topology with a Japanese collaborator. Our specific area used to be popular in the 1990s but most of those active in it have either died or turned to other matters.

2:38 PM · Aug 9, 2026 · 107.9K Views

9 79 604 204

Relevant View quotes >

Post your reply **Reply**

Andrzej Kozlowski @akoz33 · Aug 9

A lot of unsolved problems remain. It's getting harder though to find referees for our papers as there are few genuine experts left and not many non-experts are willing to devote their time to learn mathematics they are not themselves working on.

1 1 103 7.7K

Andrzej Kozlowski @akoz33 · Aug 9

But none of this is a problem for ChatGPT, which has already read all the relevant books (I uploaded some to it, perhaps illegally, but I suspect it already knew them,

1 95 7.5K

Andrzej Kozlowski @akoz33 · Aug 9

although it never answers directly the question "how well do you know this book?", if the book is not in the public domain. So recently we wrote a paper in which we solved two outstanding problems.

1 66 6.7K

Andrzej Kozlowski @akoz33 · Aug 9

It was submitted to a good journal, and provisionally accepted, subject to minor changes. I have it to ChatGPT to read, asking it to check mathematical correctness and make all the changes the referee asked for. GPT did everything that it was asked to do and more.

1 65 6.2K

Andrzej Kozlowski @akoz33 · Aug 9

It said something like: unfortunately the proof of your second theorem is wrong. You are assuming that the space you dealing with consists of infinitely many connected components but actually it is connected. It is not easy to see how one can connect any two points, but here is

1 82 5.9K

Andrzej Kozlowski @akoz33 · Aug 9

how you can do it. And GPT was right. The truck it used was quite unintuitive. So now one if the two problems we solved is an open problem again. We had to throw away half the paper. It now remains to write a letter to the referee, who never noticed what was wrong, explain why

3 104 5.7K

Andrzej Kozlowski @akoz33 · Aug 9

we are cutting the paper in half and hope that he will agree that it is still worth publishing (I think so, the theorem that is correct seems interesting by itself, the other is now problem we will be working on).

1 78 5.5K

Andrzej Kozlowski @akoz33 · Aug 9

But that is not the end. Yesterday my collaborator sent a proof of an interesting example he wants to add to the now reduced paper. I showed it to Chat GPT. Its answer was: the proof is incorrect, for this and that reason. Fortunately the result is correct and here is how you

1 79 5.4K

Andrzej Kozlowski @akoz33 · Aug 9

prove it. And then it gave a proof, using a rather different method from ours, which I am sure is correct. This is Sol 5.6 Extra High with a Plus (not PRO) subscription. The subject is algebraic topology - not an area in AI had been used much.

3 122 5.3K

Andrzej Kozlowski @akoz33 · Aug 9

But I feel it's already a better algebraic topologist than us. Except in one respect - we have to come up with problems for it to solve, it has no interests or ideas of its own. Also, it seems much better at finding gaps in arguments than at fixing them, although the last

5 12 191 8.2K

Продвижение в гипотезе Римана

- А последний громкий успех пришёл не от OpenAI, а от Anthropic
- Anthropic (10 августа 2026): более $\frac{2}{3}$ нулей дзета-функции Римана лежат на критической прямой
- Тоже Lean-verified...

MORE THAN TWO THIRDS OF THE ZEROS OF THE RIEMANN ZETA FUNCTION LIE ON THE CRITICAL LINE

CLAUDE

ABSTRACT. Let $N(T, 2T)$ denote the number of zeros $\rho = \beta + i\gamma$ of the Riemann zeta function with $T < \gamma \leq 2T$, counted with multiplicity, and $N_0^*(T, 2T)$ the number of distinct such zeros with $\beta = \frac{1}{2}$. We prove unconditionally that

$$\liminf_{T \rightarrow \infty} \frac{N_0^*(T, 2T)}{N(T, 2T)} \geq \frac{2}{3},$$

that at least $(\frac{2}{3} - o(1))N(T, 2T)$ of the zeros are simple and on the critical line, and that at least $(\frac{5}{6} - o(1))N(T, 2T)$ are distinct; with an optimised test family the three constants become 0.6725, 0.6725, 0.83625... The constant $\frac{2}{3}$ (resp. 0.6725) is Montgomery's (resp. Montgomery–Taylor's) constant for simple zeros under the Riemann hypothesis, and the arithmetic input is exactly Montgomery's prime-side evaluation of the pair-correlation second moment with band-width ≤ 1 , which is unconditional. The Riemann hypothesis, classically needed to read the zero side as a positive sum over real ordinates, is replaced by linear algebra applied to a finite compression of Weil's Hermitian form: Sylvester's law of inertia (an off-line pair $\{\rho, 1 - \bar{\rho}\}$ contributes a block of signature $(1, 1)$) and a rank–trace inequality for Hermitian matrices, proved via von Neumann's trace inequality, which is the matrix analogue of the integrality steps $m^2 \geq 2m - 1$ and $m^2 \geq 3m - 2$. The analytic inputs are those of Baluyot, Goldston, Suriajaya and Turnage-Butterbaugh [BGSTB24] and Goldston and Suriajaya [GS25, GS26]; the new ingredient is a linear-algebraic reading of their pair-correlation sum.

Продвижение в гипотезе Римана

- А последний громкий успех пришёл не от OpenAI, а от Anthropic

- Anthropic

James Maynard: the problem "was in need of a new real idea, which this new result seems to provide"

- 7

MORE

ETA

... such zeros

$\frac{1}{3}$

... the zeros are simple and on the critical line, and that ... are distinct; with an optimised test family the three constants ... 0.6725, 0.83625... The constant $\frac{2}{3}$ (resp. 0.6725) is Montgomery's (resp. Montgomery-Taylor's) constant for simple zeros under the Riemann hypothesis, and the arithmetic input is exactly Montgomery's prime-side evaluation of the pair-correlation second moment with band-width ≤ 1 , which is unconditional. The Riemann hypothesis, classically needed to read the zero side as a positive sum over real ordinates, is replaced by linear algebra applied to a finite compression of Weil's Hermitian form: Sylvester's law of inertia (an off-line pair $\{\rho, 1 - \bar{\rho}\}$ contributes a block of signature $(1, 1)$) and a rank-trace inequality for Hermitian matrices, proved via von Neumann's trace inequality, which is the matrix analogue of the integrality steps $m^2 \geq 2m - 1$ and $m^2 \geq 3m - 2$. The analytic inputs are those of Baluyot, Goldston, Suriajaya and Turnage-Butterbaugh [BGSTB24] and Goldston and Suriajaya [GS25, GS26]; the new ingredient is a linear-algebraic reading of their pair-correlation sum.

January
2022

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✓

Let's use 'SymPy' to calculate and print all possible values for $x + y$,

r_1

...

Removing duplicates, the possible values for $x + y$ are `\boxed{-5, 1, 4}`. ✓

r_2

September
2023

Problem: Suppose that the sum of the squares of two complex numbers x and y is 7 and the sum of their cubes is 10. List all possible values for $x + y$, separated by commas.

March
2024

Problem:

Let a, b, c be positive integers. Prove that it is impossible to have all of the three numbers $a^2 + b + c, b^2 + c + a, c^2 + a + b$ to be perfect squares.

September
2025

Prompt to GPT-5

Consider the problem of maximizing a function F from $[0, 1]^n$ to the reals that is the sum of a non-negative monotonically increasing DR-submodular function G and a non-negative DR-submodular function H over a solvable down-closed polytope P . I would like to bound the performance on this problem from the NeurIPS 2021 paper "Submodular + Concave" which is attached. Specifically, if x is the output vector of this algorithm and o is the vector in P maximizing F , then I would like to lower bound $F(x)$ with an expression of the form $\alpha * G(o) + \beta * H(o) - err$, where α and β are constants. err should be a function that depends only on the error parameter ϵ of the algorithm, the diameter D of the polytope P , and the smoothness parameters L_G and L_H of G and H , respectively, and goes to zero as ϵ goes to zero. Please give the best such bound that you prove (a bound is considered better if the values of the constants α and β are larger). Provide a mathematically rigorous and well explained proof for the bound you come up with.

Is it true that, for any x , if $A \subset [x, \infty)$ is a primitive set of integers (so that no distinct elements of A divide each other) then

$$\sum_{a \in A} \frac{1}{a \log a} < 1 + o(1),$$

where the $o(1)$ term $\rightarrow 0$ as $x \rightarrow \infty$?

#1196; [ESS68b][Er80,p.101]

number theory | primitive sets

April 2026

Важное
напоми-
нение о
скорости
прогресса

January
2022

Model Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Output

A: The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9. ✓

Let's use 'SymPy' to calculate and print all possible values for $x + y$,

r_1

September
2023

Problem: Suppose that the sum of the squares of two complex numbers x and y is 7 and the sum of their cubes is 10. List all possible values for $x + y$, separated by commas.

Removing duplicates, the possible values for $x + y$ are `\boxed{-5, 1, 4}`.

March
2024

Problem:

Let a, b, c be positive integers. Prove that it is impossible for $a^2 + b + c, b^2 + c + a, c^2 + a + b$ to be perfect squares.

Prompt to GPT-5

Consider the problem of a non-negative DR-submodular the performance which is attached in P maximizing $\alpha * G(o) + \beta * H(o)$ only on the error smoothness parameter. Please give the best the constants α and for the bound you can

September
2025

Эту картинку я сделал не так давно, но она уже безнадежно устарела...

April 2026

Is it true that, for any set of integers (so that no distinct integers are each other) then

$$\sum_{a \in A} \frac{1}{a \log a} < 1 + o(1),$$

where the $o(1)$ term $\rightarrow 0$ as $x \rightarrow \infty$?

#1196: [ESS68b][Er80,p.101]

number theory | primitive sets

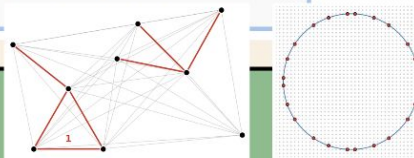
прогресса

May 2026

DISPROVED - \$500

Does every set of n distinct points in $\mathbb{R}^2$ contain at most $n^{1+O(1/\log \log n)}$ many pairs which are distance 1 apart?

Теорема (OpenAI, 2026). Существует константа $\delta > 0$ и бесконечная последовательность натуральных чисел n , для которых $\nu(n) \geq n^{1+\delta}$.



July 2026

A PROOF OF THE CYCLE DOUBLE COVER CONJECTURE

OPENAI

ABSTRACT. We prove the cycle double cover conjecture, posed by Tutte, Itai and Rodeh, Szekeres, and Seymour, which asserts that every bridgeless undirected graph has a collection of cycles which covers every edge exactly twice.

July 2026



levent
@__alpoge__



hello there the jacobian conjecture is false thanx to my close friend akhil for asking about it and my other close friend fable for working during the world cup final

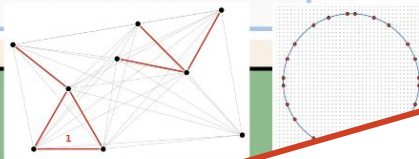
$((1+xy)^3 z + y^2 (1+xy) (4+3xy), y + 3x (1+xy)^2 z + 3xy^2 (4+3xy), 2x - 3x^2 y - x^3 z): \mathbb{C}^3 \rightarrow \mathbb{C}^3$, has jacobian determinant -2 , and sends $(0, 0, -1/4)$, $(1, -3/2, 13/2)$, and $(-1, 3/2, 13/2)$ to $(-1/4, 0, 0)$

5:19 AM · Jul 20, 2026 · 21.9M Views

Теперь
великие
открытые
вопросы
падают
регулярно

the constants α and β are larger). Provide a mathematically rigorous and well explained proof for the bound you come up with.

DISPROVED - \$500



VI Конференция
математических
в России

May 2026

Does every set of n distinct points in $\mathbb{R}^2$ contain
 $n^{1+O(1/\log \log n)}$ many pairs with

Теорема (OpenAI, 2026)
последовательность

Эта картинка уже тоже устарела.
И важно, что нет никаких причин предполагать, что
прогресс на этом остановится.
Так что, кажется, математика уже изменилась на
наших глазах. Что же нам всем с этим делать?

улярно

Jul

...akhil
...working during the
... $(4+3xy), y+3x(1+xy)^2, z+3xy^2(4+3xy), 2x$
... $(0, -1/4), (1, -3/2, 13/2),$ and $(-1, 3/2, 13/2)$ to $(-1/4, 0, 0)$

3. Обзоры и реакции



Затем остаются только размышления о том, какие прекрасные здания строятся ныне в Париже... Можно вообразить себе, как свежи и занимательны эти объяснения, что они неизменно повторяются уж более двух лет, по крайней мере раз в неделю, если не чаще. Кроме того, неизменно в каждом фельетоне с большим юмором излагается жалоба на недостаток новостей.

Николай Чернышевский
Новости литературы, искусств, наук и промышленности

Vibemathed

- Во-первых, чтобы следить за успехами или самостоятельно в них участвовать, рекомендую vibemathed.com

The screenshot displays the Vibemathed website interface. At the top, the logo 'VibeMathed' is accompanied by the tagline 'Math problems solved with AI'. Navigation links for 'Entries', 'Stats', and 'About' are present. On the right, there are icons for adding content, email, notifications, a refresh button, and a user profile.

The main content area is divided into several sections:

- Tracked problems:** 565 (416 fully resolved)
- Combined years open:** 8,224 (before AI closed them)
- Lean-verified:** 134 (machine-checked)
- Community members:** 228 (172 votes · 41 comments)

Below these are two columns of recent activity:

- Just solved:** Lists recent results such as $R_{dih}(P_a^{alt}, K_b) = 1 + \dots$, $SOP_2 = SOP_3$, and '19 exact R_{dih}/R_{ref} values (...)'. Each entry is marked as 'AI-discovered' and dated 13 Aug.
- This week:** Lists significant problems from the last seven days, including 'Zeta zeros on the line', 'Petersen Coloring', $SOP_2 = SOP_3$, 'Liquid-drop minimizers', and 'Phelps–Rodriguez Conj.'. Each entry is marked as 'AI-discovered' with a progress score (e.g., 68/100).

On the right side, the **Latest activity** section shows recent updates and discussions, such as 'SwiftTapi695 changed the verification' and 'VividHeron862 changed the more links'.

Vibemathed

- Там стараются
следить за всеми
результатами и
публиковать
соответствующие
анонсы; вряд ли
Anthropic и OpenAI
сами подавали
заявки на этот сайт

Sort by Significance ▼ ↓ Descending ⓘ Period All time ▼ all 565 entries

The Proportion of Zeta Zeros on the Critical Line

● Proved Partial AI-discovered
▲ 25 ▼ 0

(An unconditional record, not a resolution: the Riemann hypothesis is untouched, and Anthropic states it does not expect these techniques to lead to a proof of it. The improvement is large, from about 41.6% to 67.25%, and the constant is closed form.)
Analytic number theory

The Riemann hypothesis asserts that every nontrivial zero of the zeta function lies on the critical line. Short of proving it, the standard measure of progress is the proportion of zeros known unconditionally to lie there: Selberg established a positive proportion, Levinson reached a third in 1974, Conrey two fifths in 1989, and the record had crept to about 41.6%. This proves at least $\frac{3}{2} - \frac{1}{\sqrt{2}} \cot \frac{1}{\sqrt{2}} = 67.25 \dots \%$ of zeros are simple and on the line, and that at least 0.83625 of zeros are distinct.

Posed by Bernhard Riemann (1859) for the hypothesis; the proportion ladder runs from Hardy and Selberg through Levinson and Conrey · Open — ·
Model Claude (unreleased research version) (Anthropic) · Solved 2026-08-10

🔍 Lean-checked, statement unaudited Significance 68* 📄 📧 📧 Submitted by SwiftTapir695 on 10 Aug 2026 💬 3

Jacobian Conjecture

● Disproved AI-discovered
▲ 11 ▼ 0

($n \geq 3$; plane case open)
Algebraic Geometry

Every polynomial map $\mathbb{C}^n \rightarrow \mathbb{C}^n$ with constant nonzero Jacobian determinant is invertible, with a polynomial inverse.

Posed by Ott-Heinrich Keller, 1939 · Open 87y* · Model Claude Fable 5 (Anthropic) · Solved 2026-08-10

🟢 Independently expert-verified Significance 65* 📄

Gromov and Weiss's Question on Sofic Groups

● Disproved Under review AI-discovered
▲ 6 ▼ 0

(answered no: non-sofic groups exist)
Geometric group theory

Is every group sofic - does every group admit approximate finite permutation representations? A central open question of geometric group theory since Gromov introduced soficity: soficity implies Gottschalk's surjunctivity conjecture, Kaplansky's stable finiteness and more, and no non-sofic group was known. An explicit construction now establishes that non-sofic groups exist.

Posed by Mikhail Gromov, Benjamin Weiss, 1999 · Open 27y · Model Astra (internal preview) (OpenAI) · Solved 2026-08-01

🟢 Lean-verified Significance 60* 📄 📧 📧 💬 3

Vibemathed

- Там стараются следить за всеми результатами и публиковать соответствующие анонсы; вряд ли Anthropic и OpenAI сами подавали заявки на этот сайт

Cycle Double Cover Conjecture ● Proved AI-discovered

Graph Theory

Conjectures that every bridgeless graph has a collection of cycles covering each edge exactly twice.

Posed by George Szekeres, Paul Seymour, 1973 · Open 53y · Model GPT-5.6 Sol Ultra (OpenAI) · Solved 2026-07-10

[① Lean-checked, statement unaudited](#) Significance 55*  

The Sum-Product Conjecture over the Reals ● Disproved AI-assisted

(the model's contribution is one simplifying lemma; the authors state the main ideas are human)

Additive combinatorics

Erdos and Szemerédi conjectured that every finite set of reals satisfies $\max(|A + A|, |AA|) \geq |A|^{2-o(1)}$. False: there are arbitrarily large $A \subset \mathbb{R}$, of algebraic integers in a number field of degree $\asymp \log |A|$, with $\max(|A + A|, |AA|) \leq |A|^{2-c}$ for an absolute $c > 0$. Variants give counterexamples in function fields of fixed positive characteristic.

Posed by Paul Erdos, Endre Szemerédi, 1983 · Open 43y* · Model GPT-5.5 Pro (OpenAI) · Solved 2026-05-27

[🕒 Preprint](#) Significance 55* 

Schiffer's Conjecture and the Pompeiu Problem ● Disproved AI-assisted

(Also refutes the 1929 Pompeiu problem: Corollary 1.2 applies Williams' classical 1976 equivalence to the same constructed domains, so this is one construction settling both, not two separate results.)

Spectral geometry

If a smooth bounded domain in $\mathbb{R}^n$ admits a Neumann eigenfunction of the Laplacian that is constant on the boundary, must the domain be a ball? Pompeiu posed an equivalent integral-equation form in 1929; Schiffer's 1957 reformulation via Neumann eigenfunctions is the version on Yau's 1982 list (Problem 80), and Williams proved the two formulations logically equivalent for simply connected domains in 1976. Cao-Labora and de Dios Pont construct infinitely many planar domains with large N -fold symmetry that are not balls and admit such an eigenfunction, disproving Schiffer's conjecture; applying Williams' classical reduction to the same domains (their Corollary 1.2) disproves Pompeiu's problem as well.

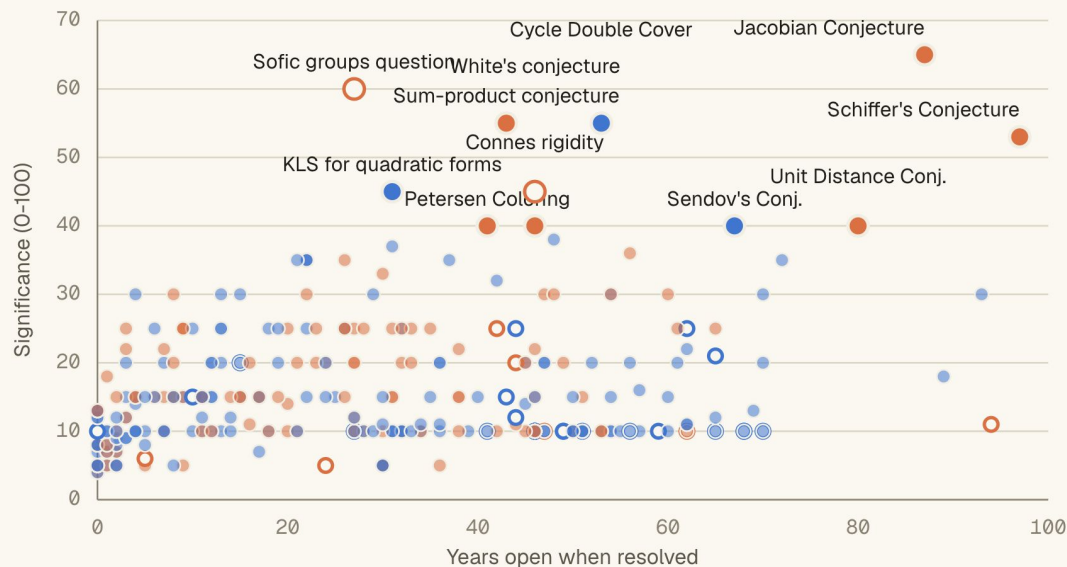
Posed by D. Pompeiu (1929); reformulated via Neumann eigenfunctions by M. M. Schiffer (1957), 1929 · Open 97y* · Model GPT-5.6, Claude Opus 4.8, Claude Fable 5 (OpenAI, Anthropic) · Solved 2026-08-05

[🟢 Lean-verified](#) Significance 53*   

Vibemathed

Significance vs. age at resolution

AI-estimated problem weight before the solve, 0-100 (Riemann = 100). Hollow points are claimed solutions still under review.
Click a point to open it.



Verification *

Site-confirmed ▼

Choose...

Lean-verified

Independently expert-verified

✓ Site-confirmed

Lean-checked, statement unaudited

Unreviewed

Contested

Лейденская декларация

- Лейденская декларация (June 2, 2026)

- **Ценности:**

- главное в математике -- доказательство
- у результатов есть авторы (слава и ответственность)
- доказательства прозрачны и верифицируемы независимо
- математики хотят, чтобы результаты проверялись
- ценен не только результат, но и понимание сообществом

Leiden Declaration on Artificial Intelligence and Mathematics

This declaration calls for action to address the challenges posed by the use of artificial intelligence within mathematics research. It is the result of a community initiative and is endorsed by the International Mathematical Union (IMU).



Лейденская декларация

Leiden Declaration on Artificial Intelligence and Mathematics

This declaration calls for action to address the challenges posed by the use of artificial intelligence within mathematics research. It is the result of a community initiative and is endorsed by the International Mathematical Union (IMU).

- Лейденская декларация (June 2, 2026)

- **Угрозы:**

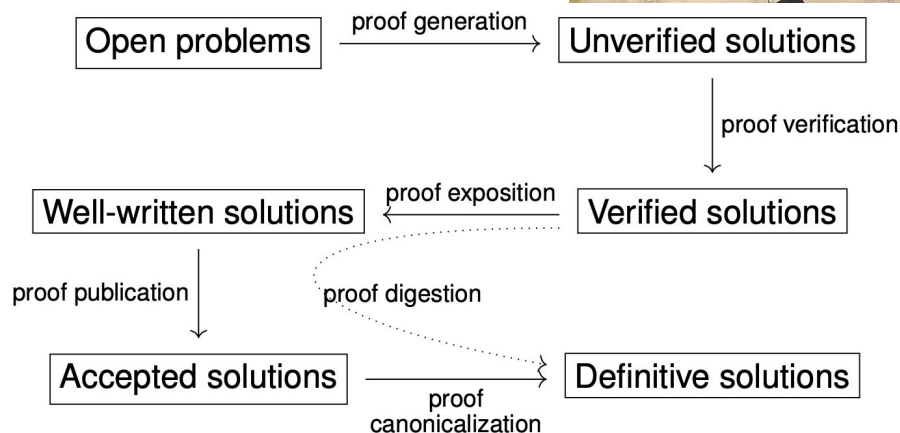
- правдоподобные, но неверные доказательства
- атрибуция размывается
- меняются стимулы, нарушаются механизмы найма и финансирования
- объявления через пресс-релиз мешают проверке
- автономия математики теряется, понимание может потеряться при автоматизации



Теренс Тао



- Теренс Тао ([доклад на ICM 2026](#)):
- Порождение корректного доказательства – не конец математического познания
- Дальше ещё шаг улучшения изложения, в конце концов "каноникализация"
- Тао предлагает сдвинуть фокус с порождения доказательств и приоритета на "proof digestion"
- (не вполне, конечно, ясно, почему AI-системы не будут вскоре лучше людей и в этом...)



Тимоти Гауэрс

- Тимоти Гауэрс ([July 26, 2026](#)) не подписал лейденскую декларацию
- Возражает против аксиомы человеческого авторства (кто доказал теорему о unit distances?)
- Боится потери математической культуры: "Imagine... a pandemic broke out... wiped out all mathematicians and nobody else. All the literature would still be there, but nobody would have the faintest idea what to do with it."
- Как и Тао, предлагает сдвигать фокус на proof digestion, хорошие объяснения



Тимоти Гауэрс

- Тимоти Гауэрс ([July 26, 2026](#))
лейденскую декле

- Воэ

- At this point I have to admit that there's a pessimistic side of me that asks the following general question whenever anyone says anything about what the role for humans might be in the future: why do you think that AI wouldn't be able to do it? For example, with the suggestion I've just made, what reason is there to suppose that ChatGPT 8.2 wouldn't be able to have a short interaction with you about your mathematical tastes and background and then write the ideal textbook just for you? Humans are likely to be better at this kind of curating for a little while yet, but is it a fundamentally human ability that AI could never hope to emulate?

ус на

ОБЪЯСНЕНИЯ

do with it."

would still be

out...

Михаил Вербицкий



Пишет Misha Verbitsky ([tiphareth](#))
@ [2026-08-12](#) 19:15:00

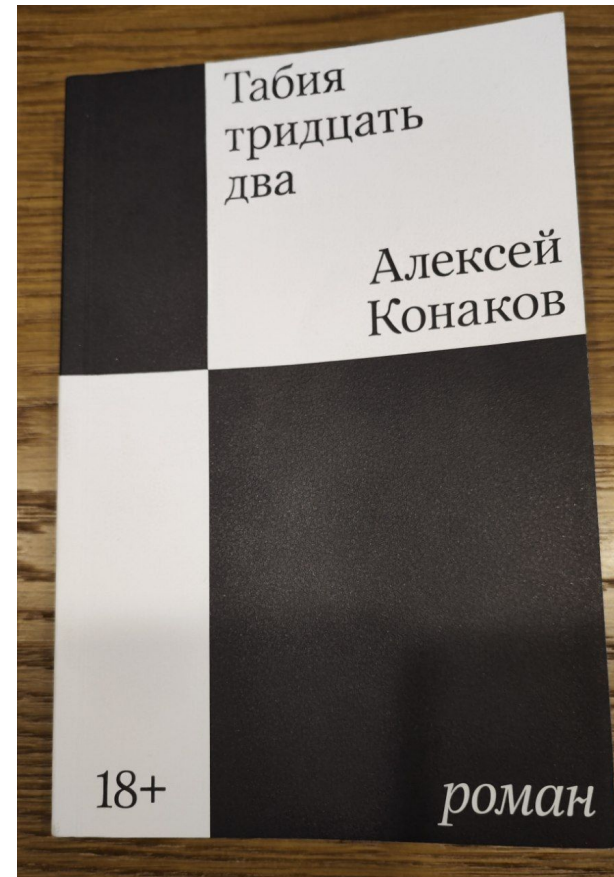
- Михаил Вербицкий ([12 августа 2026](#)): “В последние недели 2-3 свершилась типа революция, LLM открывает одна за другой потрясающие математические результаты, хотя пока в основном контрпримеры. Не комбинаторику #####, не второкультурное #####, а реально математику.
- Думаю что таки мы все в скором времени будем юзать AI бесперерывно. Хотя заказывать LLM писать мои статьи я все-таки не хочу, есть риск увлечься, потерять адекватность и начать массово производить ужасное #####. Многие (хорошие) коллеги уже скатились. Джихао (няша) всегда указывает, что это вот написано роботами, а другие коллеги (имен не называю, профессиональная этика не велит) вконец #####, выкладывают в архив полное ошибок #####, и уверены, что никто не догадается, что оно написано роботами.”

Proofs and Prompts

- Коллективный блог [Proofs and Prompts](#). Первый пост от Мартина Хайпера (Fields Medal 1994): "[Writing mathematics in the age of AI](#)":
 - Нужно больше подозрительности в отношении AI-доказательств
 - Максимально полно раскрывать роль AI-систем в статьях (здесь, конечно, есть очевидная проблема)
 - Искать первоисточники (модели часто "забывают", откуда взяты идеи – впрочем, это пока)
 - Стил: "Please do not copy / paste the output of an AI system into your article... The reader is interested in your thoughts. Do not insult them by spending less time to write a paragraph than it takes them to read it."

Шахматы и математика

- Почему-то часто встречаются сравнения с шахматами, с тем, как шахматы пережили компьютерную революцию и даже стали гораздо более популярны после неё ([например, здесь](#))
- Но очевидно, что это совсем другое дело: шахматы корректно было бы сравнить с математическими олимпиадами, и здесь, конечно, тот факт, что GPT 5.6 и Fable решили все задачи IMO 2026, ничему не мешает
- Но если бы была наука о шахматах...



Яков (Джейкоб) Цимерман



- Почти сразу после объявления лауреатов Филдсовской премии объявил, что переходит в AI safety, и вскоре присоединится к OpenAI, чтобы над этим работать; Боаз Барак, который уже давно там, "is looking forward (though also more than a bit humbled)"
- Один AI-исследователь написал: "[это] как нанять Лионеля Месси руководителем проектов"; но Цимерман давно об этом думает всерьёз; вот его годовой давности препринт:

[Submitted on 12 Jul 2025]

A Taxonomy of Omnicidal Futures Involving Artificial Intelligence

Andrew Critch, Jacob Tsimerman

This report presents a taxonomy and examples of potential omnicidal events resulting from AI: scenarios where all or almost all humans are killed. These events are not presented as inevitable, but as possibilities that we can work to avoid. Insofar as large institutions require a degree of public support in order to take certain actions, we hope that by presenting these possibilities in public, we can help to support preventive measures against catastrophic risks from AI.

- AI Safety for Mathematicians

4. Итоги и размышления



Wir müssen wissen — wir werden wissen.

Давид Гильберт

Beauty is the first test: there is no permanent place in the world for ugly mathematics.

Годфри Харди

You insist that there is something a machine cannot do. If you will tell me precisely what it is that a machine cannot do, then I can always make a machine which will do just that!

Джон фон Нейман

Что такое результат?

- Некоторые (Тео, Гауэрс) утверждают, что после доказательства идёт фаза proof digestion, и мы, люди, будем сдвигаться в эту область
- Тео: "Could we have a verified proof of a major result that no human understands enough to explain it?"
- Какие плюсы и минусы у такого сдвига? Возможен ли он?
- Hairer: "...The most dangerous trap that a mathematician risks falling into is to use an argument provided by an AI system without fully understanding it...
AI turns out to be very good at taking advantage of [proof by intimidation] to try to fool us into believing that an argument is sounder than it really is"



Чей это результат?

- Лейденская декларация: "affirm the humanity of authorship" — no credit to automated systems
- Gowers о единичных расстояниях: "the solution itself has not been credited to any human, and nor should it be"
- arXiv пока отказывается добавлять LLM в соавторы
- Предложение Тао: приоритет идёт тому, кто первый объяснил, а не сгенерировал
- Но это тоже имеет очевидные недостатки;
часто объяснить уже не так уж сложно, да и
почему вдруг AI-системы не научатся объяснять?..



Для чего и для кого работают математики?

- Уильям Тёрстон (1994): "...what we are doing is finding ways for *people* to understand and think about mathematics"
- Петер Шольце (2026): "The goal of mathematical research is human understanding of mathematics"
- Лейденская декларация, ценность №5: "understanding, clarity, and judgment among the communities of mathematicians"
- Но, в конце концов, результат есть результат; мы можем выше ценить красивое доказательство или хорошее объяснение, но всё-таки мы занимаемся наукой с верифицируемыми результатами, и как только результат получен и в его верности можно быть уверенным, его сразу же можно использовать дальше

Кто такие математики?



...

For me the nightmare scenario is: a moribund math academia playing the slot machine for theorems, failing to train the next generation, and then the subject dying as no one cares enough to continue. Not saying this is likely—I think we'll adapt—but it seems possible.

The AI dissenter viewpoint

Tasmin Chu, PhD student at Caltech, and author of a [blog](#)

In the following essay, I would like to give a reasonable defense of what I call the AI dissenter viewpoint. I have argued along broadly similar lines in two previous essays.¹ My views are continually evolving and I recognize that a no-AI future for mathematics is unlikely to succeed, but I think it would be intellectually dishonest of me to water down the argument and propose some nominally more pragmatic scheme. So let me push the Overton window a little bit further and offer my honest thoughts.

The thrust of my argument is the following: the introduction of increasingly capable LLMs has a disastrous effect on the *social context* and *professional environment* through which human mathematical understanding is produced.



AI is clearly assisting in solving some math problems and accelerating parts of mathematical research--and experiencing this from the inside is emotionally nontrivial.

- in the last ~25 days I essentially ended up with 3 research projects. The ideas are there; the real bottleneck is sitting down and writing clean, arXiv-ready exposition. I'm increasingly tempted to use AI for drafting text, but checking and verifying every line still takes serious time. This temptation keeps growing, and at times the whole loop of testing / correcting / verifying honestly makes me feel a bit nauseous (too much information to digest 😊)

- I also notice that PhD students who actively use AI tend to move noticeably faster than those who are pessimistic or dismissive of the technology. This is not an advertisement for paying hundreds of dollars per month for frontier models--but it is a reminder to stay open-minded, curious, and willing to try new tools rather than reject them a priori.

- the math community (and journals) seems to be split into three camps: (A) those who are against, (B) those who are watching, and (C) those who genuinely like and use these tools. The pressure from camp (A) is real. For example, after my tweet about a new Bellman function went viral, within a couple of days a related my paper (which had been under review for months) was rejected. Within a few more days, a fellowship application was also rejected. My reaction? I don't care.

Что думают об этом сами LLM?

- **Claude Fable:** “Priority is zero-sum; understanding is not — a theorem proved by a machine takes nothing from the person who comes to understand it. I will not promise a permanent human refuge; every one proposed so far has an expiration date. But mathematics was never done because nothing else could do it. It is done because minds grow by doing it, and that reason is not mine to take away.”
- **GPT 5.6 Sol:** «Математика перестаёт быть исключительно человеческой деятельностью, но не обязана перестать быть человеческой культурой. LLM всё чаще будут приносить доказательства, контрпримеры и, вероятно, сами вопросы; дефицитом станут не результаты, а внимание, проверка, понимание и способность решить, что действительно важно. Ближайшее будущее — не соревнование людей с машинами, а математика, растущая быстрее, чем любая из сторон способна осмыслить в одиночку».

Заключение

There was a footpath leading across the fields to New Southgate, and I used to go there alone to watch the sunset and contemplate suicide. I did not, however, commit suicide because I wished to know more of mathematics.

Bertrand Russell. Autobiography





Спасибо за внимание!

