On mathematical superintelligence

Kyler Siegel[*][†]

November 12, 2025

Abstract

In this essay, we consider the near-future possibility of artificial intelligence with superhuman mathematical reasoning capabilities, and we attempt to flesh out some of the implications for the enterprise of mathematics research. We find it useful to subdivide our conjured future into three distinct "epochs". In Epoch I (arguably already in progress), AI emerges as a powerful productivity booster for human mathematicians, potentially ushering in a new golden age of discovery and creative fulfillment, albeit with many pitfalls which must be carefully navigated. In Epoch II, AI begins offloading progressively more technical heavy lifting while human mathematicians engage in high level prompt engineering (or "vibe mathing"), with the combined human + AI system largely more effective than either humans or AI alone. In Epoch III, AI reaches a level of dominance such that human mathematicians can no longer substantively contribute to the mathematical discovery process, with perspectives shifting instead towards appreciation, critique, personal enrichment, and so on. We begin the essay by detailing the context and motivation for performing this thought experiment at the present moment, and we end with some reflections on how mathematicians may positively influence the future of their endeavor. In order to keep our scope manageable and focused, we choose to avoid (however unnatural) any discussion of the broader societal or existential implications of superhuman artificial intelligence.

Contents

[*]Department of Mathematics, University of Southern California. Email: kyler.siegel@usc.edu, web: kylersiegel.xyz.

[†]K.S. is partially supported by NSF grant DMS-2506639 and a Simon Foundation Travel Support for Mathematicians Award. We also benefited greatly from conversations with many people, including Mohammed Abouzaid, Aravind Asok, Dan Cristofaro-Gardiner, Yoel Groman, Steve Heilman, Oleg Lazarev, Greta Panova, Dan Romik, Tasos Moulinos, Umut Varolgunes, the participants of the Spring 2025 FrontierMath Symposium, and the participants of the Spring 2025 Discussion on Artificial Intelligence in Mathematics at the University of Southern California.

1 Introduction

1.1 Prelude

We begin by recalling a brief excerpt from the 2015 Breakthrough Prize panel discussion which took place at Stanford University in November 2014[1] between distinguished laureates Maxim Kontsevich, Yuri Milner, Terrance Tao, Richard Taylor and moderator Yuri Milnor.

Excerpt from 2015 Breakthrough Prize Panel Discussion (highly paraphrased)[2]:

Yuri Milner: Could any of you imagine computers surpassing humans in mathematics like they did for chess, in a hundred or maybe a thousand years?

Terry Tao: Well computers could act like a really powerful search engine, but it will still be humans driving the show, I'm pretty sure.

Maxim Kontsevich: Actually, I don't think creating artificial intelligence is that hard. I think we'll have it pretty soon (nervous laughter from panel).

Yuri Milner: You are sort of a contrarian here saying it will happen so quickly. What makes you so optimistic?

Maxim Kontsevich: Actually it's pessimistic (more nervous laughter). I thought a little about it myself and I don't see any fundamental difficulty.

Yuri Milner: Then why don't you just work on that?

Maxim Kontsevich: I think it would be immoral (more nervous laughter).

Richard Taylor: But isn't the way a computer plays chess just a big combinatorial check? I would be very surprised to see a computer win a Fields medal in our lifetime.

Terry Tao: Well I think a computer could discover just by brute force a connection between two fields of mathematics, and then the person using the computer could actually flesh it out and maybe collect the medal.

[1]Note that this was approximately 3 years before the introduction of the transformer [Vas+17] and 8 years before the first public release of ChatGPT.

[2]See [Bre14] starting around minute 38.

Yuri Milner: So Jacob, are you terrified at all by what Maxim is saying?

Jacob Lurie: I'm not too worried...

1.2 What is this essay about?

In brief, our goal here is to record some thoughts and speculations on the future of mathematics research in light of recent rapidly unfolding advances in artificial intelligence.

1.3 What is meant by "mathematics research"?

It can be a little tricky to articulate exactly what is mathematics research, or even to convey what it looks like in practice to someone who has not participated in it firsthand. Very roughly, what we have in mind is simply those activities typically performed by "research mathematicians", who exist primarily in universities around the world as professors, postdocs and graduate students (and sometimes undergraduate students or younger), and who spend a sizeable amount of time trying to do things like

- prove novel theorems, or shed new light on old ones
- discover new compelling mathematical objects (e.g. definitions, formulas, algorithms, etc)
- make progress (even if modest) on important problems, or initiate new promising directions
- disseminate the results of these efforts through various outlets (e.g. journal articles and preprint servers, conference and seminar talks, textbooks, blog posts, etc).

We will not try to give a more precise definition here, partly because it would take us too far afield and deserves a more thoughtful treatment elsewhere, and mostly because it will not needed in order to get our main ideas across. In fact, we expect that many of the issues at play here are largely relevant for other areas of scientific research and the human experience more generally, but for concreteness and to keep our scope manageable (and closer to the author's domain of experience) we will try to mostly maintain this narrower focus.

1.4 Is there anything special about mathematics research?

As we mentioned above, many of the issues under consideration in this essay are quite broad and likely to be relevant for a variety of human endeavors under the influence of rapidly progressing artificial intelligence. In that regard, we encourage the reader to view this essay as a kind of hybrid case study / thought experiment. That said, it is worth reflecting on whether there are some unique aspects of mathematics research which may distinguish it from other activities as they pertain to AI. Here are a few standout features of mathematics research which seem worth highlighting:

(1) **rigor**: the potential for very high levels of rigor, both in the traditional mathematical sense and also in a more modern sense of formal verification (see §2.1g)

(2) **entry**: the lack of necessity (at least in principle) for any strong physical, financial, or social resources (e.g. hardware or lab equipment, proprietary data sets, etc)

(3) **safety**: the lack of any obvious direct safety concerns (e.g. injurious accidents, lab leaks, environmental destruction, etc).

(4) **purity**: takes place in an intellectually clean environment sheltered from messy real-world complications (e.g. social dynamics, human psychology, unpredictable politics, weather patterns, etc).

Let us make a few preliminary comments on the significance of these features. Regarding (1), it is natural to wonder whether rigor might end up being a useful tool for mitigating thorny issues around reliability and trustworthiness in contemporary stochastic AI systems such as large language models. For (2), we should add that this perceived low barrier to entry may only hold at a superficial level, since for example the environment where one is educated is often considered crucial for success, not to mention various other financial and social prerequisites to be able to focus one's time and energy on abstract research. Furthermore, the situation may change drastically if computational power and other resources become central factors underlying mathematical progress (see e.g. §2.2i). For (3), we note that the apparent lack of safety concerns could become quite significant if these become a major source a friction in other areas, together with various associated downstream effects such as bureaucratic red tape. Note that we are sidestepping any general large scale safety concerns around the development of advanced artificial intelligence, as these would lie outside the scope of this essay (c.f. §1.9). Finally, regarding (4), we note that mathematics research may be particularly attractive to current developers of artificial intelligence partly because it naturally offers high quality training data and clean objective notions of correctness (although measuring progress on longer timelines may be much more subjective or nebulous).

1.5 Why write about this topic now?

The main relevant context at the time of writing is that the mathematical capabilities of large language models (e.g. ChatGPT, Gemini, Claude, DeepSeek etc) have rapidly progressed in the last few years, as has awareness of their existence and capabilities. Indeed, prior to the release of ChatGPT in late 2022, very few mathematicians were aware, interested, or concerned about artificial intelligence in the context of mathematics research. Now, less than three years later, most mathematicians have gotten some exposure to these systems and have started to take their mathematical capabilities more seriously, although there is still remarkable heterogeneity of opinions about how far these systems have progressed, to what degree they exhibit "true understanding", how much further they are likely to progress in the next few years, and so on.

Before addressing this heterogeneity, let us further set the stage by sampling a few recent[3] developments of particular relevance to mathematics research:

- the introduction of reasoning language models, starting with OpenAI's o1 in 2024 and followed shortly thereafter by Open AI's o3, DeepSeek's R1, and so on, which "think" by exploiting additional time and computational resources at inference time and are particularly designed to tackle multistep reasoning tasks like mathematics problems
- at least two AI models, by Google DeepMind and OpenAI respectively, achieved gold medal scores on the 2025 International mathematical olympiad exam (see e.g. [CS25]), both solving five out of six problems and outscoring all but 26 human high school student participants (note that Google's "Deep Think" mathematics model is now available to subscribers of the $250 per month Google AI Ultra plan)
- the introduction of advanced mathematics AI benchmarks such as [Gla+24; Stu25; Sch+25], for instance FrontierMath Tier IV consists of highly challenging problems involving research-level concepts written by professional mathematicians in their areas of expertise, with already 9 out of 48 problems solved at some point by at least one AI model as of October 2025 (see also [Aso25])
- steadily increasing discourse on the future of mathematics research, by way of articles (e.g. [Ven; He24; Har24; Cep; Avi25; Sev24; Hen25]) and workshops (e.g. [UCI25; Lor25; KIAS25; Sim25; ICML25; Aug25; JMM25]), not to mention recorded lectures, social media posts, and so on
- the trickling emergence of recent preprints such as [Van25; DMN25; AM25; JR25; IX25] whose authors attribute some or all of their main ideas to ChatGPT or some other AI model (or to interactions therewith)
- Google DeepMind's introduction of AlphaEvolve [Nov+25; Geo+25] (now with some open source cousins such as [Sha25; LIC25]), which combines evolutionary techniques with large language models to perform optimization on spaces of algorithms, achieving new state-of-the-art results on an assortment of constructive mathematics problems (e.g. 4×4 matrix multiplication, circle packing, the "kissing number" problem, etc)
- the expanding popularity of the Lean proof assistant and its mathematics library mathlib, along with various ongoing attempts to automatically formalize natural language proofs into Lean and to deformalize Lean proofs back into natural language (see e.g. [Aze+23; Yan+23; Pol+22; Dee24])

[3] This list should be viewed only as a narrow snapshot of the perceived situation at the time of writing (fall 2025), keeping in mind that any specific items may quickly become outdated or overshadowed by subsequent progress. We are also ignoring many other major developments in different parts of artificial intelligence such as computer vision, game playing, computational biology, and so on.

- a growing community of researchers applying techniques from artificial intelligence in order to generate new conjectures or constructions in pure mathematics (see e.g. [He+25; Dav+21; Cha+24; ACH24; Vel+24])
- DeepMind's recently announced partnership with five top mathematics institutes to accelerate the advancement of artificial intelligence in mathematics research [KR25], along with known or rumored ongoing efforts to solve major mathematics problems such as Navier–Stokes (see e.g. [Ans25]).

For anyone paying close attention, these developments immediately raise a number of pressing questions, including:

- What exactly is the current status of artificial intelligence in mathematics research? Is it still just a fancy parlor trick, or are we witnessing the beginnings of a true transformation? How is one supposed to keep track of all the different models being released in rapid succession, and to differentiate between hype and reality? To what extent have mathematicians already started to integrate these tools into their workflows?
- Looking at the most recent milestones or those which can be plausibly inferred in the immediate future, what are their implications for mathematics research in the coming years? Will the landscape be changed dramatically once the dust settles, or essentially the same as before save for a few new nifty tools? Will it still be possible to opt out and perform "good old-fashioned mathematics research" as it was done until just a few years ago? If the field is truly revolutionized, will it be for better or for worse?
- Can we make any explicit forecasts for the next say 1–3 years? Are the mathematical capabilities of AI systems likely to hit a stumbling block and meet diminishing returns, or should we brace for something more extreme? Can we expect to see mathematical superintelligence in the not-too-distant future? If so, where will that take us?
- Why does there seem to be so little consensus on the answers to the above questions?

1.6 What is "mathematical superintelligence"?

In this essay, we use the term "mathematical (artificial) superintelligence" (MASI) to mean (somewhat vaguely) any machine which can perform essentially any aspect of mathematics (including research) in a substantively and unambiguously better way than any human being. Note that we will not make much distinction between human-level and superhuman-level intelligence, because it seems likely that once an AI reaches human-level performance it will then surpass it shortly thereafter via (if nothing else) routine incremental improvements in processing power. However, as we stress below, there is a big and important distinction between

(i) AI which outperforms human mathematicians unaided by AI, and

(ii) AI which meets the level of human mathematicians even when aided by AI.

These might be called *type (i)* and *type (ii)* mathematical superintelligence respectively, and can be summarized as

$$\text{(i)} \quad \text{humans} < \text{AI} < \text{humans} + \text{AI} \qquad\qquad \text{(ii)} \quad \text{humans} < \text{AI} = \text{humans} + \text{AI}.$$

Loosely speaking, in our narrative type (i) mathematical superintelligence occurs in Epoch II below (see §3) and type (ii) occurs in Epoch III (see §4). To first approximation, the existence of type (i) MASI means that artificial intelligence has become a critical tool for mathematicians but humans are still driving or meaningfully contributing to mathematics research, while type (ii) MASI means that human mathematicians have become more or less obsolete.

It is also important to keep in mind that the transition from type (i) to type (ii) mathematical superintelligence could be surprisingly lengthy, even if AI continues to improve at a steady or exponential rate. This is because as long as artificial intelligence operates *differently* from human intelligence, it may still be complemented by the latter in a positive and nontrivial way. In this situation, human mathematicians will naturally focus their efforts on those activities which have not been strictly subsumed by artificial intelligence. In terms of subjective experience, it is unclear whether this remaining region of human utility will feel like a small island soon to be submerged in water, or an expanding continent large enough to productively occupy the minds of countless human mathematicians.

1.7 Where do mathematicians currently stand on these issues?

While there seems to be a wide range of stances on the above questions amongst mathematicians, let us caricature these into a few broad types:

- **the unengaged**: has not been paying much attention to developments in mathematical artificial intelligence, perhaps due to indifference, ignorance, or simply not having the time or mental bandwidth
- **the skeptic**: believes that most of the more aggressive claims about the current or future AI capabilities are pure hype, and that AI will not have a dramatic effect on mathematics research within the foreseeable future
- **the techno-optimist**: is very excited about using ever-improving AI tools to supercharge their research and productivity and is keen on rapidly integrating these into their workflow
- **the doomsayer**: expects that mathematical superintelligence is not far off and will likely make human mathematicians largely obsolete.

Of course these categories are not entirely mutually exclusive. For example, there exist skeptical doomsayers, who believe that there is too much hype in the current climate but nevertheless anticipate core AI models to continue improving and for this to have profound implications. Incidentally, at least for the purposes of this article, we should focus on the capabilities of core state-of-the-art AI models and try not to let our impressions be clouded by potentially dubious AI-related noise (for instance the AI-selected watermelons on sale at the author's local market).[4]

It seems fair to say that, at least until quite recently, most mathematicians were rather conservative in their predictions of AI's impacts on mathematics research, i.e. generally some combination of unengaged and skeptical, perhaps with a hint cautious optimism.[5] In particular, the prospect of a superhuman mathematical artificial intelligence within the next few years or even decades would typically be dismissed offhand. It also seems fair to posit that most mathematicians, if queried a few years ago, would have drastically underpredicted the capabilities of today's frontier AI models.

Indeed, in the panel discussion recalled in §1.1, with the exception of Kontsevich (who is clearly an outlier here), most of the panelists imagine an advanced AI as a kind of brute force mechanistic machine which excels at calculating and searching but does not exhibit its own creativity, and certainly cannot supplant humans from the driver's seat of mathematics research. Even today, a typical stance amongst mathematicians is to draw a line (either explicitly or subconsciously) which artificial intelligence shall never cross, e.g. AI cannot exhibit true understanding or define new mathematical objects or devise its own creative research directions. Naturally, many of these lines have been periodically redrawn to keep up with the latest technological advancements.

It seems worth reflecting on why mathematicians have not been more aggressive in their predictions of AI advancements in their own field. Of course, some of this may be simply attributable to the difficulty in keeping up with rapidfire developments in real time while trying to filter out the hype. There is also a general hypothesis (not specific to mathematics) that humans struggle to intuit exponential growth and tend to overestimate technology progress in the short term but underestimate it in the long term. Of course this thesis is somewhat controversial, with its more extreme proponents (e.g. Ray Kurzeil [Kur05]) predicting a coming "technological singularity" and others dismissing this is as science fiction or an excuse to ignore the more pressing mundane issues of the present. Rather than wading into this general topic, which lies beyond the scope of the present essay, let us just enumerate a few commonly encountered dismissive attitudes which are more specific to mathematicians and may contribute to the aforementioned conservatism:

(1) **too mechanical**: the perception of AI as fundamentally just a brute force machine which follows algorithms devised by humans and may excel at computations or structured pattern matching but cannot exhibit true creativity or understanding

(2) **too esoteric**: the (somewhat cynical) sense that mathematics research is too esoteric for artificial intelligence developers to bother working on

[4] In the recent discourse around whether we are in an AI bubble, a natural comparison point is the dot-com bubble of the late 1990's, whose subsequent economic recession disposed of companies like Pets.com and Webvan but overall did not do much to temper the rise of the internet, at least when viewed from today's vantage point.

[5] Of course, opinions are rapidly evolving and may have already diverged from our characteriziation between the time of writing and the time of reading.

(3) **too unreliable**: pointing out examples of AI models performing poorly or unreliably and taking these as evidence that the whole premise is fundamentally misguided

(4) **general intelligence**: the presumption that AI which excels at advanced mathematics must necessarily exhibit broad general intelligence, the existence of which would cause such profound changes in society that it seems silly to contemplate the fate of something so niche as mathematics research.

Let us address each of these sentiments in turn and argue why from our current vantage point they appear at least partially short-sighted. Firstly, (1) now seems fundamentally at odds with the modern paradigm of stochastic AI models like large language models (LLMs), which if anything have complementary shortcomings. Indeed, by the nature of the way they are trained (or "grown") rather than programmed, they behave in unpredictable ways, often surprising even their own designers with both their capabilities and their shortcomings. While there is something undeniably remarkable and mysterious about the human brain's ability to perform abstract mathematical thoughts and generate new insights (without requiring a data center), it is by now difficult to maintain the stance that the outcome of this process is fundamentally unachievable by a computer.

Regarding (2), as it happens many AI developers are currently exhibiting keen interest in cultivating the mathematical reasoning capabilities and problem solving prowess of their models, and even consulting professional mathematicians to assist them (recall e.g. [Ans25; KR25]). This could be in part due to the existence of challenging mathematical benchmarks as concrete metrics for measuring iterative progress, and subsequent bragging rights when these benchmarks are shattered (not to mention the potential for attracting investments and/or paid subscribers). It could also be because of some of the other special features listed in §1.4, or perhaps just the personal tastes and interests of some current AI developers.

As for (3), these arguments run the danger of veering into wishful thinking or denialism. It is clear that large language models can have strange and surprising failure modes,[6] but this does not a priori obviate their success modes, especially in mathematics, where creative ideas can be much more valuable than precision. Plus, in contemplating the coming years, the particular quirks or weaknesses of current models should be less relevant than their rate of progress and perceived future trajectories. Of course, in general it can be quite difficult to holistically assess the reasoning capabilities of AI models and devise robust benchmarks, and individual mileage may vary greatly from person to person based on different use cases, subfields, model access, etc.

Finally, while point (4) feels plausible (if a little hubristic[7]), it too is not guaranteed to be valid, for reasons related to the list in §1.4. Namely, while it is likely that mathematics cannot be "solved" without very powerful general purpose machines, it is also possible

[6] For example, when queried on August 4th, 2025, ChatGPT 4o gave the output to $9.9 - 9.10$ to the input -0.2.

[7] Until a few years ago, many software engineers similarly expected their field to be safe from AI, at least in the absence of very powerful general-purpose artificial intelligence. Today, a few minutes spent using LLMs to write code makes it clear that computer programming has already undergone a fundamental and irreversible transformation.

that certain major sources of friction for the development, deployment, and adoption of AI in other areas turn out to be less relevant for mathematics research. For example, progress in other parts of science may be slowed down by safety concerns and regulations, lab equipment limitations, fundamental time constraints on natural processes, licensing issues, and so on. Similarly, in the technology sector, people may be slow to adopt new products due to ingrained habits, privacy concerns, matters of taste, etc. There are also artificial barriers to entry in many fields (e.g. law, medicine, or real estate) enforced by powerful unions, trade associations, or monopolies. Thus, it is conceivable that, for largely practical reasons, mathematics research gets revolutionized by artificial intelligence to a much greater and/or faster extent than other parts of science and technology. Of course, while most mathematicians would likely welcome seeing their field leap ahead to become far more internally compelling and externally useful, on the flip side it could imply that human mathematicians become obsolete sooner than their counterparts in other fields.

1.8 What are the goals of this essay?

We have tried to make the case above that there is a surprising lack of consensus amongst mathematicians regarding the future of artificial intelligence in mathematics research, and also a concerning overinclination towards conservatism, or at least a lack of serious discourse around the possibility of more transformative outcomes on shorter timelines. The risk is that these tendencies may inhibit the undertaking of deliberate concerted actions within the community. In particular, while established mathematicians may have the luxury of reacting to developments in real time, the recruitment and education of newer generations of mathematicians requires thoughtful long-term planning. Moreover, by default and without thoughtful planning and posturing, the ultimate influence on mathematics research by AI developers and other powerful actors may not be entirely aligned with those outcomes desired by the mathematics research community itself.

With these points in mind, our main goal is to further provoke serious discussions on the forthcoming future of mathematics research, and to offer a possible framework for identifying various opportunities and risks. Relative to some other recent writings on similar topics (e.g [Ven; He24; Har24; Cep; Avi25; Sev24]), we try to focus mostly on specific practical issues which may arise in the foreseeable future, leaving more philosophical questions of meaning or purpose to be treated elsewhere. While our stated topic is the future of mathematics research and the implications for its primary stakeholders, we reiterate that many of the issues raised here are likely to resonate far beyond that context.

1.9 What this essay does *not* do

Firstly, we do not make any precise quantitative predictions regarding the developmental timeline for artificial intelligence in the context of mathematical reasoning (or more generally). We view such forecasting as an important exercise which deserves more attention, but it lies outside of our scope (see e.g. [Kok+25] for a recent viral and somewhat controversial example). In particular, we will try to maintain a fairly agnostic tone about if and when mathematical artificial superintelligence will emerge, leaving the reader to assign their own probabilities to different possible outcomes.

We also devote relatively few words specifically to mathematics education and how it may transform under artificial intelligence. This topic is clearly closely intertwined with mathematics research but is expansive enough to deserve its own separate treatment elsewhere.

Perhaps more glaringly, we will also skirt around the "big questions" which one inevitably falls into when contemplating artificial superintelligence, such as:

- Under what conditions will superintelligence arise? Will there be a "singularity" (i.e. rapid takeoff in a very short time window) or will it play out over a longer period of time? With what entities (individuals, corporations, countries) will it be aligned, if any? Will there continue to be close competition by several entities or will one achieve clear dominance?

- Will state-of-the-art models require enormous computational powers (e.g. entire massive data centers) to run, or will smaller models have some relevance? Who will have access to these models and their benefits?

- How will superintelligence affect the world power structure and economics? What about jobs and careers? Wealth distribution? Medicine and human health? Technology? Human flourishing? Sense of purpose?

In particular, in §4 we operate under the naive assumption that at some point in the not-too-distant future we will have superintelligent systems which are reasonably well aligned with their human creators (or at least have not killed all humans). Our scope is thus artificially narrow, and may seem frivilous if one expects major upheavals across all aspects of society and human existence, and yet we think there still some value in considering a more targeted purview.[8]

Lastly, we make no particular claims about originality. Many other mathematicians have recently written articles or given talks on a similar range of topics, including those referenced above and surely many more. We expect that most or all of the issues and ideas raised here have been contemplated elsewhere in some form or other, and perhaps only our global synthesis is somewhat novel.

1.10 The structure of this essay

To help structure our discussion, we divide our speculative future into three sequential "epochs" as follows:

1. Epoch I (§2): AI boosts productivity
2. Epoch II (§3): type (i) superintelligence
3. Epoch III (§4): type (ii) superintelligence.

[8] For a much wider-ranging view under similarly naive assumptions, see the influential essay [Amo24].

Of course, the distinction between these epochs will likely be blurred in reality, with certain facets of mathematics research progressing through epochs at different rates and so on. Moreover, many of the points first raised in earlier epochs will continue to apply (often in even stronger form) in subsequent epochs.

Let us now give a slightly more detailed preview. Firstly, in §2 on Epoch I, we enumerate many different opportunities and risks involved with artificial intelligence as a powerful new productivity tool. The positive side includes things like rapid education and literature search, new possibilities for coding experiments and computations, new standards of rigor, new unexpected connections between subfields, and so on, while the negative side involves things like content overload, quality control issues, overreliance, reduction of equity, and more.

Next, in §3 on Epoch II, we consider a more sophisticated form of artificial intelligence which can do a large portion of the technical heavy lifting in mathematics research, for example converting a roughly stated lemma into a precise formulation and providing a rigorous proof, or even producing a fully detailed paper based on only high level user guidance. In this epoch, we envision human mathematicians as still fundamentally driving the research process based on their tastes, big picture ideas, and complementary advantages, even though AI now largely exceeds humans in technical prowess and also plays an increasingly autonomous role in the ideation process. Among other things, we attempt to carefully tease out the role of human mathematicians in this new delicate balance.

Then, in §4, we finally arrive at a level of mathematical artificial intelligence which is capable of fully autonomous long-term mathematics research with little or no guidance from human mathematicians. We speculate around what human mathematics will look like in this post-MASI world, i.e. whether mathematics as a field will behave more like humanities, or perhaps exist primarily as an activity of recreation or personal enrichment. A key question here is whether and how humanity may still benefit significantly from our shared corpus of mathematical insights, even if humans no longer contribute substantively to the discovery process.

Overall, the picture which emerges in §2,§3,§4 is roughly as follows. In the short term, we expect a new "golden era" of mathematics, provided that we adequately neutralize various potential pitfalls. This golden era will be characterized by rapidly accelerated productivity and abundant mathematical discoveries, including resolution of big open problems and many new emerging directions. At the same time, mathematicians will find themselves working at increasingly high level of abstraction, eventually deferring most or all technical details to machines. As artificial intelligence continues to advance, this golden era will shift gradually (or suddenly) to a new order in which human mathematicians play more managerial or supervisorial roles. Ultimately, with human mathematicians adding diminishing practical value to the research process, they will need to negotiate entirely new relationships with mathematics.

Note that our narrative inevitably becomes more speculative as we move out from the present and near-term future to the medium or long-term future. We nevertheless think it is productive to take these more speculative developments seriously, both because they may require significant planning and because they may arrive sooner than widely expected.

We conclude the essay in §5 with some preliminary thoughts on outlook and what practical steps could be taken to positively influence the future, or at least plan for it. As should be already be evident by now, throughout this essay we raise many more questions than we are able to answer in any definitive way.

1.11 AI disclosures

While the bulk of this article was written in the "old-fashioned way" without artificial intelligence, the author did use some AI tools (primarily ChatGPT, and Claude Sonnet via Cursor) for things like:

- compiling lists of relevant references and formatting them as BibTeX entries
- looking up and/or summarizing simple facts or news events (e.g. the release timeline of a given AI model)
- checking for spelling and grammar issues and pointing out awkward or ambiguous phrasings and so on.

References

[ACH24] Alberto Alfarano, François Charton, and Amaury Hayat. "Global lyapunov functions: a long-standing open problem in mathematics, with symbolic transformers". *Advances in Neural Information Processing Systems* 37 (2024), pp. 93643–93670.

[AM25] Boris Alexeev and Dustin G Mixon. "Forbidden Sidon subsets of perfect difference sets, featuring a human-assisted proof". *arXiv:2510.19804* (2025).

[Amo24] Dario Amodei. *Machines of Loving Grace*. 2024. URL: https://www.darioamodei.com/essay/machines-of-loving-grace.

[Ans25] Manuel Ansede. "Spanish mathematician Javier Gómez Serrano and Google DeepMind team up to solve the Navier-Stokes million-dollar problem". *El País (English Edition)* (2025). URL: https://english.elpais.com/science-tech/2025-06-24/spanish-mathematician-javier-gomez-serrano-and-google-deepmind-team-up-to-solve-the-navier-stokes-million-dollar-problem.html.

[Aso25] Aravind Asok. *Mathematicians interact with AI, July 2025 update*. 2025. URL: https://www.math.columbia.edu/~woit/wordpress/?p=15039.

[Aug25] *Workshop: AI Transforms Math Research*. Workshop held August 25–29, 2025 at the University of Augsburg, Germany. University of Augsburg, Faculty of Mathematics and Natural Sciences, 2025. URL: https://www.uni-augsburg.de/en/fakultaet/mntf/math/prof/geom/workshop-aimath/.

[Avi25] Jeremy Avigad. "Is Mathematics Obsolete?" *arXiv:2502.14874* (2025).

[Aze+23] Zhangir Azerbayev, Bartosz Piotrowski, Hailey Schoelkopf, Edward W Ayers, Dragomir Radev, and Jeremy Avigad. "Proofnet: Autoformalizing and formally proving undergraduate-level mathematics". *arXiv:2302.12433* (2023).

[Bre14] Breakthrough. *2015 Math Panel with Donaldson, Kontsevich, Lurie, Tao, Taylor, Milner*. https://www.youtube.com/watch?v=eNgUQlpc1m0. YouTube video. Dec. 2014.

[Buz20] Kevin Buzzard. *The future of mathematics?* 2020. URL: https://www.andrew.cmu.edu/user/avigad/meetings/fomm2020/slides/fomm_buzzard.pdf.

[Cep] Jordana Cepelewicz. "Mathematical Beauty, Truth and Proof in the Age of AI". *Quanta Magazine* (). URL: https://www.quantamagazine.org/mathematical-beauty-truth-and-proof-in-the-age-of-ai-20250430/.

[CGH24] Vincent Colin, Paolo Ghiggini, and Ko Honda. "The equivalence of Heegaard Floer homology and embedded contact homology III: from hat to plus". *Publications mathématiques de l'IHÉS* 139.1 (2024), pp. 349–385.

[Cha+24] François Charton, Jordan S. Ellenberg, Adam Zsolt Wagner, and Geordie Williamson. "PatternBoost: Constructions in Mathematics with a Little Help from AI". *arXiv:2411.00566* (2024).

[CS25] Katie Cai and Jaspreet Singh. "Google clinches milestone gold at global math competition, while OpenAI also claims win". *Reuters* (July 2025). URL: https://www.reuters.com/world/asia-pacific/google-clinches-milestone-gold-global-math-competition-while-openai-also-claims-2025-07-22/.

[Dav+21] Alex Davies et al. "Advancing mathematics by guiding human intuition with AI". *Nature* 600 (2021), pp. 70–74.

[Dee24] DeepMind. *AlphaProof: A formal approach to reasoning (solving IMO-level problems in Lean)*. https://deepmind.google/discover/blog/ai-solves-imo-problems-at-silver-medal-level/. 2024.

[DMN25] Charles-Philippe Diez, Luis da Maia, and Ivan Nourdin. "Mathematical research with GPT-5: a Malliavin-Stein experiment". *arXiv:2509.03065* (2025).

[Geo+25] Bogdan Georgiev, Javier Gómez-Serrano, Terence Tao, and Adam Zsolt Wagner. "Mathematical exploration and discovery at scale". *arXiv:2511.02864* (2025).

[Gla+24] Elliot Glazer et al. "Frontiermath: A benchmark for evaluating advanced mathematical reasoning in AI". *arXiv:2411.04872* (2024).

[Gra04] Jeremy J Gray. "Anxiety and abstraction in nineteenth-century mathematics". *Science in Context* 17.1-2 (2004), pp. 23–47.

[Har24] Michael Harris. "Automation compels mathematicians to reflect on our values". *Bulletin of the American Mathematical Society* 61.2 (2024), pp. 331–342.

[He+25] Yang-Hui He, Kyu-Hwan Lee, Thomas Oliver, and Alexey Pozdnyakov. "Murmurations of elliptic curves". *Experimental Mathematics* 34.3 (2025), pp. 528–540.

[He24] Yang-Hui He. "AI-driven research in pure mathematics and theoretical physics". *Nature Reviews Physics* 6.9 (2024), pp. 546–553.

[Hen25] Jonas Henkel. "The mathematician's assistant: integrating AI into research practice". *Mathematische Semesterberichte* (2025), pp. 1–28.

[ICML25] *2nd AI for Math Workshop @ ICML 2025*. Workshop held July 18, 2025 (hybrid) at ICML 2025, Vancouver, Canada. ICML 2025 Workshop Organizers, 2025. URL: https://sites.google.com/view/ai4mathworkshopicml2025/.

[IX25] Paata Ivanisvili and Xinyuan Xie. "Counterexample to majority optimality in NICD with erasures". *arXiv:2510.20013* (2025).

[JMM25] *AI Sessions and Events at the 2025 Joint Mathematics Meetings*. Meeting held January 8–11, 2025 at the Seattle Convention Center & Sheraton Grand Seattle, Seattle, WA, USA. Joint Mathematics Meetings (AMS & MAA), 2025. URL: https://jointmathematicsmeetings.org/meetings/national/jmm2025/2314_aisessions.html.

[JR25] Uijeong Jang and Ernest K Ryu. "Point Convergence of Nesterov's Accelerated Gradient Method: An AI-Assisted Proof". *arXiv:2510.23513* (2025).

[KIAS25] *Workshop: Machine Learning and Mathematics*. Workshop held July 7–11, 2025 at the Korea Institute for Advanced Study (KIAS), Seoul, South Korea. Korea Institute for Advanced Study (KIAS), 2025. URL: http://events.kias.re.kr/h/MLM2025/?pageNo=5841.

[Kok+25] Daniel Kokotajlo, Scott Alexander, Thomas Larsen, Eli Lifland, and Romeo Dean. *AI2027*. 2025. URL: https://ai-2027.com/.

[Kos+25] Nataliya Kosmyna, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein, and Pattie Maes. "Your brain on chatgpt: Accumulation of cognitive debt when using an AI assistant for essay writing task". *arXiv:2506.08872* 4 (2025).

[KR25] Pushmeet Kohli and Eugénie Rives. *Accelerating discovery with the AI for Math Initiative*. Blog post on Google DeepMind website. Oct. 2025. URL: https://blog.google/technology/google-deepmind/ai-for-math/.

[Kur05] Ray Kurzweil. *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking, 2005.

[LIC25] Robert Tjarko Lange, Yuki Imajuku, and Edoardo Cetin. "Shinkaevolve: Towards open-ended and sample-efficient program evolution". *arXiv:2509.19349* (2025).

[Lor25] *Workshop: Mechanization and Mathematical Research*. Workshop held September 15–19, 2025 at the Lorentz Center, Leiden, Netherlands. Lorentz Center, 2025. URL: https://www.lorentzcenter.nl/mechanization-and-mathematical-research.html.

[Nov+25] Alexander Novikov, Ngân Vũ, Marvin Eisenberger, Emilien Dupont, Po-Sen Huang, Adam Zsolt Wagner, Sergey Shirobokov, Borislav Kozlovskii, Francisco JR Ruiz, Abbas Mehrabian, et al. "AlphaEvolve: A coding agent for scientific and algorithmic discovery". *arXiv:2506.13131* (2025).

[Pol+22] Stanislas Polu, Jesse Michael Han, Kunhao Zheng, Mantas Baksys, Igor Babuschkin, and Ilya Sutskever. "Formal mathematics statement curriculum learning". *arXiv:2202.01344* (2022).

[Roo25] Kevin Roose. "Not a Coder? With A.I., Just Having an Idea Can Be Enough". *The New York Times* (Feb. 27, 2025). URL: https://www.nytimes.com/2025/02/27/technology/personaltech/vibecoding-ai-software-programming.html.

[Sch+25] Johannes Schmitt, Gergely Bérczi, Jasper Dekoninck, Jeremy Feusi, Tim Gehrunger, Raphael Appenzeller, Jim Bryan, Niklas Canova, Timo de Wolff, Filippo Gaia, et al. "IMProofBench: Benchmarking AI on Research-Level Mathematical Proof Generation". *arXiv:2509.26076* (2025).

[Sev24] Simone Severini. *The Future of Mathematics in a World of Machines*. LinkedIn post. Feb. 2024. URL: https://www.linkedin.com/pulse/future-mathematics-world-machines-simone-severini-v6odc/.

[SFK17] Martin Schmidt, Benedikt Fecher, and Christian Kobsda. *The 20 Highest Performing Authors in Mathematics Publish Third Most Articles But Have Comparably Low Bibliometrics*. Dataset on Zenodo. 2017. URL: https://doi.org/10.5281/zenodo.845806.

[Sha25] Asankhaya Sharma. *OpenEvolve: an open-source evolutionary coding agent*. Open-source implementation of AlphaEvolve. July 2025. URL: https://github.com/codelion/openevolve.

[Sim25] *Joint Workshop: AI for Mathematics and Theoretical Computer Science*. Workshop held April 11–? 2025 at the Simons Institute for the Theory of Computing and SLMath, Berkeley, CA, USA. Simons Institute for the Theory of Computing & SLMath, 2025. URL: https://simons.berkeley.edu/workshops/simons-institute-theory-computing-slmath-joint-workshop-ai-mathematics-theoretical.

[Smi+24] David Smith, Joseph Samuel Myers, Craig S. Kaplan, and Chaim Goodman-Strauss. "An aperiodic monotile". *Comb. Theory* 4.1 (2024).

[Stu25] Christian Stump. *Science Bench: A New Public Benchmark of PhD-Level Mathematics Problems*. Press Release. 2025. URL: https://math.xn--sciencebench-m09f.ai/press/press%E2%80%90release%E2%80%902025%E2%80%9009%E2%80%9001%E2%80%90v1.pdf.

[UCI25] *Workshop: Humans, Math, and Machines*. Workshop held March 7–8, 2025 at the University of California, Irvine. Department of Logic and Philosophy of Science, University of California, Irvine, 2025. URL: https://www.lps.uci.edu/workshops/aimath2025.php.

[Van25] Mark Van Raamsdonk. "Finite entropy sums in quantum field theory". *arXiv:2508.21276* (2025).

[Vas+17] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention is all you need". *Advances in neural information processing systems* 30 (2017).

[Vel+24] Petar Veličković, Alex Vitvitskyi, Larisa Markeeva, Borja Ibarz, Lars Buesing, Matej Balog, and Alexander Novikov. "Amplifying human performance in combinatorial competitive programming". *arXiv:2411.19744* (2024).

[Ven] Akshay Venkatesh. *Human mathematics in the age of reasoning machines*. URL: https://www.math.ias.edu/~akshay/research/HumanMath.pdf.

[Wik25] Wikipedia contributors. *Vibe coding — Wikipedia, The Free Encyclopedia.* [Online; accessed 11-November-2025]. 2025. URL: https://en.wikipedia.org/w/index.php?title=Vibe_coding&oldid=1320779131.

[Yan+23] Kaiyu Yang, Aidan Swope, Alex Gu, Rahul Chalamala, Peiyang Song, Shixing Yu, Saad Godil, Ryan J Prenger, and Animashree Anandkumar. "Leandojo: Theorem proving with retrieval-augmented language models". *Advances in Neural Information Processing Systems* 36 (2023), pp. 21573–21612.