AI-Assisted Exploration in Algebra and Number Theory: A New Paradigm for Mathematical Discovery

[1]Dr.Sharad Kumar Agarwal,sharad.agarwal.aimt@gmail.com
[2]Dr.Vijay Vir Singh,singh_vijayvir@yahoo.com
[3]Dr.Yazdani Hasan,yazhassid@gmail.com
[1]Noida International University, Department of Mathematics
[2]Noida International University, Department of Mathematics
[3]Noida International University, Department of Computer Science

Abstract

The integration of artificial intelligence into mathematical research represents a transformative shift in how mathematicians explore, conjecture, and prove results. This paper examines the emerging paradigm of AI-assisted exploration in algebra and number theory, focusing on the development of evolutionary coding agents and large language models (LLMs) as collaborators in mathematical discovery. We analyze the foundational work of Georgiev, Gomez-Serrano, Tao, Wagner, and Google DeepMind on AlphaEvolve, which demonstrates autonomous discovery of mathematical constructions across 67 problems spanning analysis, combinatorics, geometry, and number theory. We then explore a systematic three-stage pipeline—computational exploration, informal reasoning, and formal verification—as proposed by recent research initiatives at Emory University. Additional case studies, including the resolution of an Erdős prize conjecture using LLM-assisted formal verification and applications of neural networks to Dirichlet L-function analysis, illustrate the breadth and power of these methods. We conclude by discussing open problems in semi-global fields as promising targets for future AI-assisted investigation and consider the implications of this methodology for the future of mathematical practice.

Keywords: artificial intelligence, number theory, algebraic exploration, large language models, evolutionary algorithms, formal verification, Lean theorem prover

1. Introduction

For centuries, mathematical discovery has been a uniquely human endeavor—a creative process of pattern recognition, conjecture formation, and rigorous proof. The mathematician's toolkit has expanded dramatically in recent decades with the advent of computational assistance, yet the core activities of exploration and discovery have remained fundamentally human-guided. This landscape is now undergoing a profound transformation.

Recent breakthroughs in artificial intelligence have demonstrated that machine learning systems can not only assist with routine calculations but can actively participate in the creative process of mathematical research. From solving Olympiad geometry problems to discovering novel constructions in extremal combinatorics, AI systems are evolving from passive computational tools into active collaborators in mathematical exploration.

This paper examines one particularly promising direction: AI-assisted exploration in algebra and number theory . These fields, with their rich structures and deep open problems, present ideal testing grounds for AI-driven discovery methods. We focus on three interconnected themes:

First, we analyze the AlphaEvolve framework, an evolutionary coding agent that combines the generative capabilities of large language models with automated evaluation to discover and refine mathematical constructions at unprecedented scale. Second, we examine a systematic three-stage pipeline that integrates computational exploration, informal reasoning, and formal verification—a methodology being developed to address classical problems in number theory and extend them to less-understood settings such as semi-global fields. Third, we explore complementary approaches, including LLM-assisted formal verification that recently contributed to resolving a long-standing Erdős conjecture and neural network applications for conjecture verification in analytic number theory.

The convergence of these methods suggests a future in which AI systems serve as genuine intellectual partners in mathematics—generating conjectures, exploring vast search spaces, and even contributing to rigorous proofs through formal verification. This paper aims to provide both a survey of current capabilities and a roadmap for future investigations.

2. The AlphaEvolve Framework: Evolutionary Discovery at Scale

2.1 Architecture and Methodology

The AlphaEvolve system, introduced by Georgiev, Gomez-Serrano, Tao, Wagner, and their collaborators at Google DeepMind, represents a significant advance in AI-driven mathematical discovery. At its core, AlphaEvolve is an evolutionary coding agent that combines the generative flexibility of large language models with automated evaluation in an iterative evolutionary framework.

The system operates through a cycle of three phases: propose, test, and refine . Starting from a problem specification—which can be as simple as a description of a mathematical object to be constructed or an inequality to be optimized—AlphaEvolve generates candidate solutions in the form of algorithmic code. These candidates are then evaluated automatically against objective criteria. Successful solutions are retained and serve as parents for the next generation, with LLMs guiding the mutation and recombination operations that produce new candidates.

A key innovation lies in AlphaEvolve's ability to operate at multiple levels of abstraction. The system evolves not only the parameters of a mathematical construction but also the algorithmic strategy used to discover it—a "meta-level evolution" that mirrors how human mathematicians adapt their approaches as a problem unfolds. Initial heuristics might focus on broad exploration of the solution space, while later-stage heuristics refine near-optimal solutions with greater precision.

2.2 Performance Across Mathematical Domains

To demonstrate the breadth of AlphaEvolve's capabilities, the researchers assembled a test suite of 67 problems spanning mathematical analysis, combinatorics, geometry, and number theory . The results were striking: AlphaEvolve rediscovered the best-known solutions in most cases and, crucially, discovered improved solutions in several instances.

Particularly noteworthy is the system's ability to generalize from finite examples to universal formulas

. In some test cases, AlphaEvolve was able to take solutions valid for a finite set of input values and extrapolate them into formulas applicable for all inputs—a capacity that moves beyond mere numerical optimization toward genuine mathematical insight.

The efficiency gains are equally impressive. Problem setup with AlphaEvolve typically requires only a few hours of preparation time, compared to the weeks or months often needed for traditional computational approaches to similar problems. This dramatic reduction in setup overhead enables "constructive mathematics at scale"—the systematic exploration of entire problem classes rather than isolated questions.

2.3 Integration with Proof Assistants

AlphaEvolve does not operate in isolation. The framework is designed for integration with complementary AI systems, particularly proof assistants like Deep Think and AlphaProof . This integration creates a powerful pipeline: AlphaEvolve generates candidate mathematical constructions through evolutionary search; Deep Think provides extended reasoning capabilities to develop informal insights; and AlphaProof translates successful constructions into formally verified theorems.

This layered approach addresses a fundamental challenge in AI-assisted mathematics: the tension between exploration and verification. Evolutionary methods excel at exploring vast search spaces and discovering novel constructions, but they cannot guarantee correctness. Proof assistants provide rigorous verification but require precisely specified problems and cannot generate novel candidates on their own. By combining these capabilities, the integrated framework offers both discovery power and mathematical certainty.

3. A Systematic Pipeline for Algebraic and Number-Theoretic Exploration

3.1 The Three-Stage Framework

Building directly on the AlphaEvolve methodology, a research initiative at Emory University has proposed a systematic pipeline for applying AI-assisted exploration specifically to problems in algebra and number theory. This framework, developed under the mentorship of Dr. Deependra Singh, organizes the discovery process into three complementary stages:

Stage 1: Computational Exploration. Using evolutionary and code-based agents in the spirit of AlphaEvolve, researchers conduct large-scale numerical and symbolic experiments. The goal is to identify patterns, generate plausible conjectures, and discover novel mathematical constructions. This stage leverages the ability of LLM-guided evolution to explore vast search spaces that would be impractical to examine manually.

Stage 2: Informal Mathematical Reasoning. The patterns and conjectures emerging from computational exploration are then developed into coherent, human-readable arguments. This stage draws on extended reasoning systems similar to DeepThink, which can articulate the intuitive insights behind discovered constructions and suggest potential proof strategies.

Stage 3: Formal Verification. Finally, successful arguments are translated into fully rigorous, machine-checked proofs using proof assistants such as Lean, guided by tools similar to AlphaProof. This stage ensures that discoveries meet the highest standards of mathematical certainty and contributes to the growing library of formalized mathematics.

3.2 Calibration on Classical Problems

As an initial calibration, the Emory project plans to apply this pipeline to a set of classical problems in number theory, including results on sums of squares, quadratic forms, and the u-invariant of local and global fields . These problems serve dual purposes: they provide training grounds for deploying the methodology, and they enable evaluation of its strengths and limitations since the underlying results are well understood.

The choice of these classical problems is strategic. The u-invariant of a field—the maximal dimension of anisotropic quadratic forms over that field—has been extensively studied for local fields (such as $\mathbb{Q}_p$) and global fields (such as $\mathbb{Q}$). By applying AI-assisted exploration to these well-understood settings, researchers can verify that the pipeline reproduces known results before moving to more challenging territory.

3.3 Extending to Semi-Global Fields

Having calibrated the approach, the project aims to investigate open problems in less-explored settings , particularly semi-global fields such as $\mathbb{C}((t))(x)$ and $\mathbb{Q}_p(x)$. These fields—function fields over complete discretely valued fields—present rich arithmetic structures that are less thoroughly understood than their local and global counterparts.

For semi-global fields, even basic questions about quadratic forms and u-invariants remain open in many cases. The vast search spaces involved make these problems ideal candidates for AI-assisted exploration. An evolutionary agent could, for example, search for anisotropic forms of unusually high dimension, or explore the behavior of u-invariants under field extensions. Any novel constructions discovered could then be developed into conjectures and, potentially, formally verified theorems.

This extension from classical to frontier problems illustrates the broader promise of AI-assisted mathematics: the ability to systematically explore mathematical territories that have remained inaccessible due to their complexity or the sheer scale of possibilities they present.

4. Complementary Approaches: LLMs in Conjecture Generation and Formal Verification

4.1 Human-Assisted Proof and the Erdős Conjecture

A striking demonstration of AI's potential in mathematical research emerged recently with the resolution of a long-standing conjecture of Paul Erdős concerning Sidon sets and perfect difference sets. The conjecture asked whether every finite Sidon set can be extended to a finite perfect difference set—a problem that had remained open for decades, with a $1000 prize attached.

The research team, led by Boris Alexeev and Dustin G. Mixon at The Ohio State University, discovered that the set $\{1, 2, 4, 8, 13\}$ serves as a counterexample to Erdős's conjecture. Interestingly, they also uncovered a previously overlooked counterexample published by Marshall Hall, Jr. decades earlier. To ensure the validity of both findings, the team turned to artificial intelligence.

Using ChatGPT 5 , they generated Lean code to produce formal, machine-verified proofs of both counterexamples. The resulting proof exceeded 6000 lines of code and established with mathematical certainty that Erdős's conjecture is false. The researchers adopted a "just-do-it" proof style, focusing on directly implementing mathematical logic in Lean rather than meticulously planning every step—an approach inspired by the recent "vibe coding" technique where code is generated based on intuition and a general understanding of the problem.

This case study illustrates several important themes: the power of LLMs to assist with formal verification, the value of combining human mathematical insight with machine precision, and the potential for AI tools to uncover and verify results that might otherwise remain overlooked.

4.2 Neural Networks and Dirichlet L-Functions

A different but complementary approach appears in recent work by Saraeb, who applied neural networks and ensemble methods to investigate conjectures in analytic number theory. The folklore conjecture under investigation concerned Dirichlet characters and their L-functions : specifically, whether the modulus q of a Dirichlet character χ is uniquely determined by its initial nontrivial zeros.

Using data from the LMFDB (L-functions and Modular Forms Database), Saraeb formulated a classification problem with feature vectors derived from initial zeros and labels given by the modulus q. A feed-forward neural network and random forest classifier, combined via a meta-ensemble, achieved perfect test accuracy of 1.0 when appropriate zero statistics were included.

Based on these results, Saraeb proposed two new conjectures: (i) that hidden statistical patterns exist in the nontrivial zeros of each Dirichlet L-function, and (ii) that an underlying statistical connection links zeros of L-functions of characters sharing the same modulus. This work demonstrates how machine learning techniques can not only verify existing conjectures but also generate new ones based on detected patterns—a form of AI-assisted mathematical intuition.

4.3 Automated Formula Discovery and Unification

A third thread of AI-assisted mathematics concerns the discovery and unification of formulas for mathematical constants . Researchers at the Technion have developed a hybrid LLM-symbolic tool system that automatically harvests formulas from hundreds of thousands of arXiv publications, validates them numerically, and feeds them into novel symbolic algorithms.

This system has demonstrated that many formulas for π, including famous ones by Euler, Gauss, and the Ramanujan Machine, are generated by a common mechanism. The work represents one of the first successful applications of LLMs to number theory, a domain that had remained largely elusive to AI due to its reliance on deep structural understanding rather than pattern matching alone.

Similarly, the EPFL AI Center has launched a project on "Document-Level Autoformalization," aiming to develop open-source tools that can automatically translate mathematical research papers into computer-readable format for verification in Lean. This initiative, supported by the $18 million AI for Math Fund, seeks to bridge the gap between human and machine understanding of mathematics, making high-level mathematics more transparent and accessible.

5. A Research Agenda for Semi-Global Fields

5.1 Why Semi-Global Fields?

Semi-global fields—function fields over complete discretely valued fields—occupy a fascinating middle ground in arithmetic geometry. They share properties with both local fields (such as $\mathbb{Q}_p$) and global fields (such as $\mathbb{Q}$), but their additional structure from the function field component creates rich new phenomena. Fields of the form $\mathbb{C}((t))(x)$ and $\mathbb{Q}_p(x)$ are prime examples.

For these fields, fundamental questions about quadratic forms remain open. The u-invariant, which measures the maximum dimension of anisotropic quadratic forms, is well understood for local fields (u = 4 for non-dyadic local fields) and for global fields (u = 4 for number fields, u = 2 for function fields over finite fields). For semi-global fields, however, the situation is more complex and less thoroughly mapped.

5.2 Applying the AI-Assisted Pipeline

The three-stage pipeline described in Section 3 offers a systematic approach to investigating these questions. An evolutionary agent could:

1. Explore the space of quadratic forms over a given semi-global field, searching for anisotropic forms of progressively higher dimension.
2. Identify patterns in the forms discovered—perhaps relating to ramification behavior, residue field characteristics, or interactions between the local and function field components.
3. Generate conjectures about the u-invariant and its behavior under field extensions or completions.
4. Attempt formal verification of the most promising conjectures using proof assistants.

The search space for such an investigation is vast, encompassing infinitely many forms with potentially subtle anisotropy conditions. This is precisely the kind of problem where AI-guided exploration could complement human intuition, suggesting constructions or counterexamples that might otherwise remain hidden.

5.3 Potential Outcomes

What might such an investigation yield? Several outcomes are plausible:

- Discovery of new anisotropic forms with unexpectedly high dimension, suggesting lower bounds on the u-invariant.

Identification of structural constraints that limit anisotropy, suggesting upper bounds.

Patterns linking u-invariants to other arithmetic invariants of the field, such as its Brauer group or cohomological dimension.

- Counterexamples to plausible conjectures, analogous to the Sidon set discovery, that reshape our understanding of the terrain.

Any of these outcomes would constitute a genuine contribution to arithmetic geometry—and would demonstrate the power of AI-assisted exploration to advance research in pure mathematics.

6. Conclusion: The Future of Mathematical Practice

The convergence of evolutionary algorithms, large language models, and formal verification tools is reshaping the landscape of mathematical research. The AlphaEvolve framework demonstrates that AI systems can autonomously discover novel mathematical constructions, often matching or improving upon the best-known results. Systematic pipelines for AI-assisted exploration promise to extend these capabilities to new domains, including the rich arithmetic territory of semi-global fields. Complementary approaches—from LLM-assisted formal verification to neural network conjecture generation—illustrate the breadth of AI's potential contributions.

Several themes emerge from these developments:

Scale and efficiency. AI systems can explore mathematical search spaces at a scale impossible for human mathematicians alone, and with dramatically reduced preparation time. This enables systematic investigation of entire problem classes rather than isolated questions.

Collaboration, not replacement. In each case study, AI serves as a complement to human intelligence—generating candidates for human evaluation, assisting with formal verification, or suggesting patterns for human interpretation. The most powerful paradigm is human-AI collaboration, not automation.

Formal verification as a bridge. The growing integration of proof assistants like Lean into the discovery pipeline offers a path from computational exploration to mathematical certainty. Formal verification ensures that discoveries meet the highest standards of rigor and contributes to the cumulative enterprise of mathematics.

New questions, new methods. As AI tools become more sophisticated, they will not only help answer existing questions but will also suggest new ones. The patterns detected by neural networks in Dirichlet L-functions, for example, point toward conjectures that might never have occurred to human mathematicians.

The research program outlined in this paper—applying AI-assisted exploration to open problems in algebra and number theory—represents a natural next step in this evolving collaboration between human and machine intelligence. By combining the pattern-recognition capabilities of AI with the structural understanding and creativity of human mathematicians, we may hope to make progress on questions that have resisted solution for decades, and perhaps to discover entirely new territories of mathematical truth.

References

[1] Georgiev, B., Gomez-Serrano, J., Tao, T., Wagner, A. Z., & Google DeepMind. (2025). Mathematical exploration and discovery at scale. arXiv preprint arXiv:2511.02864 .

[2] Singh, D. (2026). AI-Assisted Exploration in Algebra and Number Theory. Emory REU Computational Mathematics for Data Science. https://www.math.emory.edu/site/cmds-reuret/projects/2026-algebra-nt/

[3] Hugging Face. (2025). Paper page - Mathematical exploration and discovery at scale. https://huggingface.co/papers/2511.02864

[4] Quantum Zeitgeist. (2025). DeepMind AlphaEvolve AI System Advances Math, Rediscovering & Improving Solutions. https://quantumzeitgeist.com/ai-advanced-math/

[5] Emory REU. (2026). Summer 2026. https://www.math.emory.edu/site/cmds-reuret/tag/summer-2026/

[6] Saraeb, A. (2025). Artificial Intelligence in Number Theory: LLMs for Algorithm Generation and Neural Networks for Conjecture Verification. arXiv preprint arXiv:2504.19451 .

[7] Quantum Zeitgeist. (2025). Human-assisted Proof Identifies {1, 2, 4, 8, 13} As Counterexample To $1000 Erdős Prize Conjecture On Sidon Subsets. https://quantumzeitgeist.com/1000-human-assisted-proof-identifies-counterexample-erdos-prize-conjecture/

[8] Raz, T. (2025). From Euler to AI: Unifying Formulas for Mathematical Constants. Technion Faculty of Data and Decision Sciences. https://dds.technion.ac.il/seminars/graduate-student-seminar/from-euler-to-ai-unifying-formulas-for-mathematical-constants/

[9] EPFL AI Center. (2025). Advancing Mathematics with AI. https://ai.epfl.ch/advancing-mathematics-with-ai/

[10] Yeungnam University Library. (2025). Generative Modeling for Mathematical Discovery. arXiv preprint arXiv:2503.11061 .

[11] Novikov, G., et al. (2025). AlphaEvolve: Evolutionary coding agents for scientific discovery. In preparation .

[12] DeepMind. (2025). AlphaProof: Formal mathematical reasoning with reinforcement learning. Technical report .

[13] LMFDB Collaboration. (2025). The L-functions and Modular Forms Database. https://www.lmfdb.org

[14] Buzzard, K., Commelin, J., & Massot, P. (2025). Formalizing mathematics in Lean. Notices of the AMS .

[15] Alexeev, B., & Mixon, D. G. (2025). Forbidden Sidon subsets of perfect difference sets. arXiv preprint arXiv:2510.19804 .

[16] Romera-Paredes, B., et al. (2024). Mathematical discoveries from program search with large language models. Nature , 625, 468-475.

[17] Trinh, T. H., et al. (2024). Solving olympiad geometry without human demonstrations. Nature , 625, 476-482.

[18] Kaminer, I., et al. (2025). The Ramanujan Machine: Automated generation of fundamental mathematical constants. Proceedings of the National Academy of Sciences .

[19] Kunčak, V., & Viazovska, M. (2025). Document-Level Autoformalization. AI for Math Fund project description.

[20] Tao, T. (2025). Machine-assisted proof. Notices of the AMS , forthcoming.