Michael Ruofan Zeng 曾若凡

Ph.D. candidate in Mathematics · Director, UW Math and AI Lab

zengrf@uw.edu · zengrf.github.io · github.com/zengrf

Department of Mathematics, University of Washington, Seattle, WA

Research interests: Algebraic geometry, algebraic combinatorics, motivic homotopy theory, and enumerative geometry; AI for math, interpretability, automated reasoning, and AI for humanities.

Education

University of Washington *September 2023 – present*
Ph.D. candidate in Mathematics; expected June 2028.
Advisors: Jarod Alper and Sara Billey.

University of California, Berkeley *Fall 2026*
Visiting Student Researcher, Department of Mathematics. Faculty sponsor: Hannah K. Larson.
SLMath semester: *Motivic Homotopy Theory: Connections and Applications*.

University of Pennsylvania
M.A. in Applied Mathematics and Computer Science *August 2022 – May 2023*
Concurrent submatriculation with the B.A.
B.A. in Mathematics and Linguistics, *summa cum laude* *August 2020 – May 2023*
Distinction in Mathematics; Dean's List, 2021–2023.

Honors and Research Funding

Simons Dissertation Fellowship in Mathematics *2026*
Two-year research fellowship for excellent graduate students in the final years of their mathematics Ph.D.

Schmidt Sciences HAVI Research Grant *2026*
Deep Learning for Jibo Extraction from Classical Japanese Manuscripts; led by Paul Atkins (UW) and Tomo Miyazaki (Tohoku). One of 19 projects selected from 501 proposals.

Jinmoncom Best Paper Award and Student Encouragement Award *2025*
For *Deriving Orthographic Data from Classical Japanese Texts with Machine-Learning Methods* [4].

Gerald B. Folland Endowed Graduate Fellowship (2026); UW Jane Street Fellowship (2026); McFarlan Fellowship; UW Provost's Fellowship; Thomas and Grace Tseng Fellowship (2023); Simpson Center Digital Humanities Summer Fellowship (2025); Freiwald Scholars Award (2020).

Appointments and Research Leadership

UW Math and AI Lab, University of Washington *January 2025 – present*
Director, June 2026 – present; Project Lead and Research Mentor, January 2025 – present.

- Lead teams in mathematical machine learning, interpretability, and automated reasoning; set research directions and design experimental pipelines from data generation through training and analysis.
- Secured grants and donations and led teams to ICLR, ICML, and NeurIPS workshop publications. Projects include PatternMiner, ComboFormer, learned circuit optimization, and mathematical formalization.

Math, Inc., San Francisco *June – July 2026*
Intern, Member of Technical Staff.

- Led mathematics and Lean 4 formalization of foundational results toward verifying a major number-theory proof, in collaboration with top mathematicians.
- Co-built the human interface connecting mathematicians to the proof-generation harness and codebase for expert review and guidance.

China Electric Power Research Institute, Beijing *July – August 2023*
Optimization Intern, Perceptive Intelligence Lab. Built a Python route optimizer for inspection of 7,000+ power towers; reduced total route distance by 7% through hub placement and dispatch optimization.

Publications and Preprints

Pure Mathematics

[20] **M. R. Zeng**. The integral K_0-ring of the stack of smooth hyperelliptic curves. Preprint. 2026. arXiv: 2610.27267.

[19] D. Ito, **M. R. Zeng**, and X. Zeng. Tensor-generating line bundles on toric varieties. Preprint. 2026. arXiv: 2610.22279.

[9] A. Debnath and **M. R. Zeng**. Oriented Cohomology Rings of Some Moduli Spaces via Blowups. 2026. arXiv: 2604.14536.

[5] **M. R. Zeng**. The Grothendieck group of the variety of spanning line configurations. 2025. arXiv: 2512.22769.

[3] T. Brazelton, D. Chan, B. Mudrak, B. Spitz, C. Vogeli, C. Wang, **M. R. Zeng**, and S. Zotine. C_p-Mackey functors in Macaulay2. Accepted by *Journal of Software for Algebra and Geometry*. 2025. arXiv: 2509.05456.

[2] D. Bidleman, T. Duff, J. Kendrick, and **M. R. Zeng**. Brackets and projective geometry in Macaulay2. 2025. arXiv: 2504.00889.

[1] M. Harada and M. Precup. Torus fixed point sets of Hessenberg Schubert varieties in regular semisimple Hessenberg varieties. In: *Osaka Journal of Mathematics* **60** (2023), pp. 637–652. arXiv: 2112.13250. Appendix by **M. R. Zeng**.

Machine Learning and AI for Mathematics

[18] E. X. Meng, X. Yang, J. Hasan, **M. R. Zeng**, and H. Kvinge. Studying the Effects of Task Diversity Using Small Mathematical Worlds. In: *NeurIPS 2026: The Sixth Workshop on Mathematical Reasoning and AI*. Accepted. 2026. arXiv: 2610.13750.

[17] **M. R. Zeng**. Anytime-Valid FDR Control for Fixed Formulaic Alpha Registries. In: *NeurIPS 2026 Workshop on E-values: From Statistics to ML*. Accepted. 2026.

[16] J. Hasan and **M. R. Zeng**. To Infinite and Beyond: How Neural Networks Learn Compact Lie Groups, a Mechanistic Study. In: *NeurIPS 2026 Workshop on Interpretability for Discovery*. Accepted. 2026.

[15] T. Gong, Y. Yang, and **M. R. Zeng**. Total 3-closure for projective special linear groups. 2026. arXiv: 2608.17878.

[14] T. Gong, Y. Yang, and **M. R. Zeng**. Alternating Groups and Embeddings into Groups Invariably Generated by Two Prime-Order Elements. 2026. arXiv: 2608.00703.

[13] T. Gong, **M. R. Zeng**, and Y. Yang. Albilich: Steerable Proof-State Orchestration for LLM-Based Mathematical Research with CAS Integration. 2026. arXiv: 2607.27705.

[12] R. Pandey, **M. R. Zeng**, W. K. Zhang, K. Jin, N. Morato, A. Ganapule, B. Mehta, and J. Alper. FactorLibrary: From Polynomials to Circuits via Recursive Subgoals. In: *3rd AI for Math Workshop (ICML 2026): Toward Self-Evolving Scientific Agents*. 2026. arXiv: 2606.25394.

[8] W. K. Zhang, R. Pandey, B. Mehta, K. Jin, N. Morato, A. Ganapule, **M. R. Zeng**, and J. Alper. CircuitBuilder: From Polynomials to Circuits via Reinforcement Learning. ICLR 2026 Workshop on AI with Recursive Self-Improvement. 2026. arXiv: 2603.17075.

AI and Digital Humanities

[11] P. S. Atkins and **M. R. Zeng**. Quantitative Orthographic Analysis of the Ogura *Shikishi*. In: *manuscript cultures* **26**.1 (June 2026), pp. 1–60. DOI: 10.15460/s8e93666.

[10] **M. R. Zeng**, H. Chau, and P. S. Atkins. Jibo-labeled classical hiragana image dataset. Version v1. Zenodo, 2026. DOI: 10.5281/zenodo.18765466.

[4] H. Chau, **M. R. Zeng**, and P. Atkins. Deriving orthographic data from classical Japanese texts with machine-learning methods. In: *Jinmoncom 2025: Computers and the Humanities Symposium*. Best Paper Award and Student Encouragement Award. 2025, pp. 19–24.

Submitted Manuscripts

[7] A. Ganapule, K. Jin, R. Pandey, N. Morato, W. Zhang, J. Alper, and **M. R. Zeng**. CircuitPilot: Learning Circuit Optimization through Imitation and Self-Play. Submitted to ICLR 2027. 2026.

[6] **M. R. Zeng**, J. Hasan, X. Yang, E. X. Meng, and H. Kvinge. PatternMiner: Automated Mathematical Discovery through LLM-Guided Training and Interpretation of Small Models. Submitted to ICLR 2027. 2026.

Manuscripts in Preparation

- T. Brazelton, A. Jaramillo-Puentes, H. Nuer, and **M. R. Zeng**. An Enriched Count of Lines in a Pencil of Quartic Surfaces and Related Problems. Manuscript in preparation. 2026.
- R. Gandhi, R. Liu, and **M. R. Zeng**. Exceptional Isomorphisms for GKM Varieties. Manuscript in preparation. 2026.
- E. Bishop, A. Debnath, M. Pernice, and **M. R. Zeng**. The K_0-Ring of $\overline{\mathscr{M}}_{1,3}$ via Patching. Manuscript in preparation. 2026.
- E. Bishop and **M. R. Zeng**. The order of the top Hodge class on $\mathscr{A}_g$ and $\mathscr{M}_g$. Manuscript, work in progress. 2026.

Research Programs

Algebraic Geometry, K-Theory, and Enriched Enumerative Geometry

Thesis work

- Develop GKM criteria for exceptional isomorphisms with Raj Gandhi and Ruizhen Liu, and homotopical criteria with Shiyue Li.
- Computed the integral K_0-ring of the stack $\mathscr{H}_g$ of smooth hyperelliptic curves for $g \geq 2$ in characteristic zero [20]. With Emma Bishop, determined the order of λ_g in integral cohomology of $\mathscr{A}_g$ and $\mathscr{M}_g$ and computed $K_0(\mathscr{M}_3)$. With Emma Bishop, Arkamouli Debnath, and Michele Pernice, compute $K_0(\mathscr{M}_{1,2})$ and $K_0(\mathscr{M}_{1,3})$.
- With Daigo Ito and Xiangru Zeng, characterized tensor-generating line bundles on proper toric varieties through combinatorial criteria and proved an analogue of Kleiman's criterion for tensor-generation [19].
- With Arkamouli Debnath, developed blowup formulas for oriented cohomology rings of moduli spaces, including an algebraic-cobordism computation of the 3,264-conic count [9]. Ongoing work studies enriched counts of lines and planes in pencils.
- Identified the Grothendieck ring of spanning line configurations with a generalized coinvariant algebra through K-theoretic localization; extended pipe-dream formulas to words [5].

Mathematical Discovery through Learning and Interpretation

Project lead: PatternMiner, ComboFormer, IntWorld/PermWorld, and Lie-group interpretability

- **PatternMiner** [6]: designed an LLM-guided discovery pipeline connecting exact symbolic data, model training, mechanistic interpretation, hard-negative mining, counterexample searches, and proof attempts. Led work yielding new results on major open problems in combinatorics, including proofs of two Schubert nonvanishing conjectures, a character-row evaluation algorithm, and Schur positivity for a family of immanant coefficients.
- **ComboFormer and IntWorld/PermWorld:** with Henry Kvinge and collaborators, develop transformers trained purely on algebraic combinatorics to extract conjectures and algorithms through interpretability. Design pretraining data, training, and post-training end to end; coordinate team experiments. IntWorld/PermWorld provides controlled studies of task diversity and transfer within exact mathematical worlds [18].
- **Lie-group interpretability** [16]: designed the pipeline, selected interpretation tools, and coordinated experiments across nine compact Lie groups and five seeds. Identified representations capturing 93–98% of encoder variance and tested their causal role with subspace ablations and activation patching.

Automated Reasoning, Formalization, and Circuit Optimization

Project lead: Albilich, mathematical benchmarks, CircuitBuilder, FactorLibrary, and CircuitPilot

- **Albilich** [13]: developed a mathematical research harness connecting persistent proof-state tracking, specialized reasoning agents, and computer algebra. With Ting Gong and Yong Yang, used the system in work resolving Kourovka Notebook problems 21.142 and 20.2 [15, 14].
- **Autoformalizing mathematical benchmarks:** with Theodore Meek, lead formalization of LemmaBench and OpenConjecture through an actor-verifier loop combining Lean checking with mathematician review of the statements.

- **CircuitBuilder** [8], **FactorLibrary** [12], and **CircuitPilot** [7]: designed reinforcement-learning architectures and implemented and trained models on UW Tillicum and AWS. Extended circuit synthesis to imitation and self-play for selecting symbolic optimization procedures, with independent circuit-equivalence checks.

Classical Manuscript Analysis and Quantitative Orthography

Researcher with Paul S. Atkins and Herman Chau; collaboration with Tohoku University

- Co-developed jibo-aware OCR combining character detection with vision-transformer classification; obtained 0.96 average F1 on the held-out test set [4]. The accompanying paper received the Jinmoncom 2025 Best Paper Award and Student Encouragement Award.
- Developed compositional-data and length-normalized likelihood methods for scribal attribution. Analysis of 5,000+ reference poems yielded statistical evidence against the traditional attribution of the recto inscriptions of the *Ogura Shikishi* to Fujiwara no Teika [11].
- Developing an end-to-end analyzer for classical East Asian manuscripts with Tohoku University, integrating a vision-language model with OCR and orthographic analysis; supported by Schmidt Sciences HAVI.

Software and Datasets

- **Brackets (Macaulay2)** [2] — co-developer and coauthor. Grassmann–Cayley algebra, bracket straightening, and symbolic computations in projective geometry.
- **CpMackeyFunctors (Macaulay2)** [3] — co-developer and coauthor. Projective resolutions, Tor, and Ext for Mackey functors over cyclic groups of prime order.
- **Albilich** [13] — project lead. Persistent SQLite proof graph, dependency tracking, and computer-algebra integration for long-horizon mathematical research.
- **Autoformalizing Mathematical Benchmarks** — project lead. Lean formalization and statement-review workflows for LemmaBench and OpenConjecture.
- **CayleyPy** — project lead. Search on massive combinatorial graphs.
- **Arithmetic Circuit Synthesis** [12, 8] — project lead. Reinforcement-learning methods and reusable subgoals for polynomial circuits.
- **Jibo-labeled Classical Hiragana Image Dataset** [10] — co-creator with Herman Chau and Paul S. Atkins. Released 670,332 labeled character images, Zenodo v1 (2026), CC BY-SA 4.0.

Talks

Research Talks

UW Research Computing Symposium: *Training a Deep-Learning Model to Decipher Classical Japanese Calligraphy* (presented for Paul S. Atkins) — *November 2026*

Axiom Math: *Large and small models for mathematical discovery* — *October 2026*

SLMath $\mathbb{A}^1$-Enumerative Geometry Seminar: *3264 and Algebraic Cobordism* — *September 2026*

UW Student AG Seminar — *April 2026*

UW Combinatorics and Geometry Seminar — *March 2026*

Lightning Talk at WAGS, University of Oregon — *November 2025*

Japan Association for Digital Humanities 2025, University of Osaka — *September 2025*

Simpson Center Digital Humanities Scholars, University of Washington — *June 2025*

Macaulay2 Workshop at Tulane University: Lightning Talks — *April 2025*

Combinatorics and Geometry Seminar: Lightning Talks, University of Washington — *March 2025*

Expository Talks

OSU RTG Workshop on Affine Combinatorics and Quantum K-Theory: Background Talk 1 — *May 2026*

1-2-3 Seminar: Fantastic Shapes and How to Count Them — *February 2026*

Student AG Seminar: The Chow Ring of the Stack of Smooth Plane Cubic Curves — *June 2025*

Writing Milestone Seminar: Flags, Spanning Lines, and Grothendieck Shenanigans — *April 2025*

1-2-3 Seminar: Chow Groups, Quotients, and Equivariant Intersection Theory — *January 2025*

1-2-3 Seminar: 27 Lines on a Cubic Surface – via Classifying Spaces? *February 2024*
Student AG Seminar: A Leisurely Introduction to Algebraic K-Theory *December 2023*

Teaching

MATH 208 C Matrix Algebra — Instructor of Record *SU 2025*

- In-person lectures with Panopto recordings; 84 enrolled students.

MATH 581 C Intersection Theory — Co-lecturer *SP 2024*

- Supervised by S. Kovács; co-lectured with A. Debnath. Notes available online.

University of Washington — Teaching Assistant

Grader for MATH 425 Real Analysis II *WI 2026*
TA for MATH 208 Linear Algebra *FL 2025*
Grader for MATH 327 Introduction to Real Analysis *SP 2025*
TA for MATH 111 Algebra with Applications (2 sections) *WI 2025*
TA for MATH 126 Calculus with Analytic Geometry III (2 sections per quarter) *SP, FL 2024*
Grader for MATH 300 Introduction to Mathematical Reasoning *SU 2024*
TA for MATH 125 Calculus with Analytic Geometry II (2 sections per quarter) *FL 2023, WI 2024*

Other Institutions

TA for CIS 5150 Linear Algebra for Computer Science, University of Pennsylvania *FL 2022*
TA for MATH 429 Linear Algebra, Washington University in St. Louis *FL 2019*

Mentoring

UW Math and AI Lab *January 2025 – present*
Mentor student research teams in AI for math, guiding problem selection, model training, interpretability, and paper writing.

Washington Experimental Math Lab (WXML)

Mentor for *Creating 3D Prints and Animations to Help Teach Calculus* (with Andrew Loveless) *SP 2025*
Mentor for *What Mathematical Functions Can Neural Networks Learn?* in the Math AI Lab (with Jarod Alper) *WI 2025*

Directed Reading Program (WDRP)

Mentor for *An Excursion in Representation Theory* *SP 2025*
Mentor for *Enumerative Geometry, Incidence Geometry, Projective Geometry* *SP 2024*
Mentor for *Lie Groups Are Not a Lie: They Are Everywhere* *WI 2024*

Academic Service and Organization

Co-organizer, UW 2026 Lean Hackathon *May 8–10, 2026*
University of Washington; sponsored by AWS Automated Reasoning, with Leonardo de Moura and Eric Wieser.

Seminar Organization

Writing Milestone Seminar (with Z. Chowdhury) *SP 2025*
Algebraic Curves Learning Seminar (with G. Inchiostro) *WI 2025*
Intersection Theory (with S. Kovács and A. Debnath) *SP 2024*

Conferences, Workshops, and Learning Seminars

Conferences and Workshops

Caltech Mathathon Challenge: Old Problems, New Proofs, Pasadena *November 2026*

AlGeCom XXVIII (Algebra, Geometry and Combinatorics Day), University of Illinois Urbana-Champaign *November 2026*
Accelerating Mathematical Discovery: A Crash Course in AI and Computation, Santa Cruz *October – November 2026*
Motivic Homotopy Theory: Connections and Applications, SLMath, Berkeley *October 2026*
SLMath Summer School on Moduli of Varieties, Berkeley *July 2026*
ICARM Summer School on Formalization in Lean *July 2026*
3rd AI for Math Workshop: Toward Self-Evolving Scientific Agents, at ICML 2026, Seoul *July 2026*
CMND 2026 Thematic Program: Algebraic Combinatorics and Applications *June – July 2026*
ICARM Workshop on Algebraic Geometry and Machine Learning, Pittsburgh *June 2026*
M0,ntréal ISM Discovery School: Combinatorics of Moduli of Curves, UQAM *May 2026*
Graduate Student Topology and Geometry Conference, Rutgers University *April 2026*
Graduate Student Combinatorics Conference, University of Illinois Chicago *March 2026*
Western Algebraic Geometry Symposium at UCSD *March 2026*
Workshop on Mirror Symmetry and Related Topics, Kyoto University *December 2025*
Jinmoncom 2025: 人文科学とコンピューターシンポジウム, Kyushu University *December 2025*
ICERM Workshop: Webs in Algebraic Combinatorics *December 2025*
Western Algebraic Geometry Symposium at University of Oregon *November 2025*
Japanese Association for Digital Humanities, University of Osaka *September 2025*
Summer Research Institute in Algebraic Geometry at Colorado State University *July – August 2025*
Macaulay2 Workshop at University of Wisconsin-Madison *June – July 2025*
CMND Thematic Program in Discrete Groups in Topology and Algebraic Geometry *June 2025*
Western Algebraic Geometry Symposium at University of British Columbia *April 2025*
Macaulay2 Workshop at Tulane University *April 2025*
Western Algebraic Geometry Symposium at University of Arizona *November 2024*
Park City Math Institute 2024: Motivic Homotopy Theory *July 2024*
iWoAT: International Workshop on Algebraic Topology at Fudan University *June 2024*
iWoAT: International Workshop on Algebraic Topology at SUSTech *December 2023*
Western Algebraic Geometry Symposium at WUSTL *November 2023*
iWoAT: International Workshop on Algebraic Topology at Peking University *July 2023*
Macaulay2 Workshop at the University of Minnesota *June 2023*
Minicourse on Mapping Class Groups of 4-Manifolds at UChicago *August 2022*
ISM Discovery School: Geometry and Combinatorics of Hessenberg Varieties *June 2022*

Learning Seminars

eCHT Quadratic Curve Counting Seminar *FL 2024*
eCHT Kan Seminar *FL 2024*
eCHT Algebraic Vector Bundles Seminar *FL 2023*

Technical Skills and Languages

Computing

Programming and mathematical systems: Python, C++, Julia, R, Java, Lean 4/Mathlib, Macaulay2, SageMath, Mathematica, SQL/SQLite, and LaTeX.

Machine learning and infrastructure: PyTorch, JAX, transformers, vision-language models, OCR, reinforcement learning, mechanistic interpretability, Linux, Git, AWS, and HPC training.

Modern Languages

Mandarin Chinese (native); English, Japanese, and Cantonese (professional); German (intermediate); Russian and Korean (basic conversation). Japanese: JLPT N1.

Classical Languages

Classical Chinese and Classical Japanese (advanced); Old English (intermediate); Ancient Greek (basic).