

Tim Gehringer

Email: timge@ethz.ch
Web: people.math.ethz.ch/~timge/

Research Profile

Postdoctoral researcher at ETH Zurich working on LLM evaluation, agentic mathematical reasoning, and arithmetic geometry. Develops research-level benchmarks, evaluation pipelines, and expert-review workflows for AI-generated proofs. Benchmark developer and maintainer at **MathArena**, co-organizer of **IMProofBench**, and co-developer of **ProofCouncil**.

Academic Appointments

- Nov 2026 – **Postdoctoral Researcher (incoming)**
Oct 2028 University College London (UCL)
- Nov 2025 – **Postdoctoral Researcher**
Oct 2026 ETH Zurich
Research in arithmetic geometry, focusing on the arithmetic of hyperelliptic curves.
Develop SageMath implementations of curve-classification algorithms.

AI Research Projects

- May 2026 – **Developer & Maintainer**
Present ProofCouncil
Develop and maintain an LLM agent for research-level mathematics
Design and run evaluations on research-level and open problems
Part of the team that performed best in the FirstProof: SecondBatch competition
- Nov 2025 – **Benchmark Developer & Maintainer**
Present MathArena
Led ArXivLean development; co-led creation of ArXivMath and BrokenArXiv
Develop and maintain their benchmark-generation and evaluation pipelines
- Mar 2025 – **Benchmark Organizer**
Present IMProofBench
Build and maintain an Inspect-based evaluation pipeline
Design peer-review workflows and coordinate quality checks

Selected Research Talks

- Jul 2026 **Agentic AI for Mathematical Research**
AI4Research, ICML 2026
- Jun 2026 **Benchmarking Agentic AI on Research-Level Mathematics**
GRACE Spring School, University of Luxembourg
- May 2025 **Arithmetic of Hyperelliptic Curves in Residue Characteristic 2**
J. Algebra and Representation Theory Seminar, University of Oxford

Education

Mar 2021 – Nov 2025	Dr. Sc. Mathematics ETH Zurich Supervisor: Prof. Dr. Richard Pink Thesis: “Stable Reduction of Hyperelliptic Curves”
Sep 2018 – Nov 2020	MSc Mathematics ETH Zurich Thesis: “Reduction of Hyperelliptic Curves” under supervision of Prof. Dr. Richard Pink
Oct 2015 – Aug 2018	BSc Mathematics TU Darmstadt Thesis: “Construction of invariant valuations” under supervision of Prof. Dr. Torsten Wedhorn

Selected Publications

- [1] Tim Gehringer and Richard Pink. “Reduction of hyperelliptic curves in residue characteristic 2”. *Journal of Number Theory* 281 (2026), pp. 429–491. DOI: 10.1016/j.jnt.2025.10.004.
- [2] Long Phan et al. “A benchmark of expert-level academic questions to assess AI capabilities”. *Nature* 649.8099 (2026), pp. 1139–1146. DOI: 10.1038/s41586-025-09962-4.
- [3] Tim Gehringer. “Computing the Stable Reduction of Hyperelliptic Curves in Residue Characteristic 2”. 2025. arXiv: 2506.19663 [math.AG].
- [4] Johannes Schmitt, Tim Gehringer, Jasper Dekoninck, Gergely Bérczi, Uri Kreitner, Liam Price, and David Holmes. “ProofCouncil: An LLM Agent for Solving Open Mathematical Problems”. 2026. arXiv: 2607.09474 [cs.AI].
- [5] Jasper Dekoninck, Nikola Jovanović, Tim Gehringer, Kári Rognvaldsson, Ivo Petrov, Chenhao Sun, and Martin Vechev. “Beyond Benchmarks: MathArena as an Evaluation Platform for Mathematics with LLMs”. 2026. arXiv: 2605.00674 [cs.CL].
- [6] Andrei Balakin et al. “Benchmarks in Leipzig”. 2026. arXiv: 2606.05818 [math.HO].
- [7] Johannes Schmitt et al. “IMProofBench: Benchmarking AI on Research-Level Mathematical Proof Generation”. 2025. arXiv: 2509.26076 [cs.CL].
- [8] Tim Gehringer and Richard Pink. “Reduction of Hyperelliptic Curves in Characteristic $\neq 2$ ”. 2021. arXiv: 2112.05550 [math.AG].

Scholarships & Fellowships

- Swiss National Science Foundation (SNSF) Postdoc.Mobility Fellowship (awarded; Nov 2026 – Oct 2028)
- German Scholarship Foundation fellow (Apr 2017 – Nov 2021)

Technical Skills

Programming: Python, Java, C++

LLM/AI Tooling: Inspect (LLM evaluation), LLM APIs (OpenAI, Gemini, xAI), Fireworks.ai (API)

General: Git, $\LaTeX$, Docker

Mathematical Computing: SageMath, Magma, Maple, Mathematica, Lean 4

Languages: German (native), English (C2), French (B1)