

UNIVERSITÄT
BAYREUTH

Math Research, Formalization, and AI: A Case Study

Michael Stoll
Universität Bayreuth

Oberseminar Arithmetische Geometrie
Universität Bayreuth
October 22, 2026

Outline

- Current **AI** models have become **really good at math**.
- **Formalization** helps checking correctness.
- Current AI models are also **good at formalizing**.
- How to make **productive use** of this as a **research mathematician**?

I have been **using AI** (concretely, Claude in various forms) since **early July 2026**,
experimenting with using it in various contexts.

Here, I will report on how I used it as a **collaborator** on a research project.

This project is related to my very **first published paper**.

A look back . . .

Sonderdruck aus

Arch. Math., Vol. 59, 239–244 (1992)

0003-889X/92/5903-0239 \$ 2.70/0

© 1992 Birkhäuser Verlag, Basel

Galois groups over $\mathbb{Q}$ of some iterated polynomials

By

MICHAEL STOLL

0. Introduction. Throughout this paper, a will denote an integer such that $-a$ is not a square in $\mathbb{Q}$, $f := X^2 + a \in \mathbb{Z}[X]$, and $f_0 := X$, $f_{n+1} := f(f_n) = f_n^2 + a$ for all $n \geq 0$, are the iterates of f . Let $c_1 := -a$ and $c_{n+1} := f(c_n) = c_n^2 + a (= f_n(-a) = f_n(a) = f_{n+1}(0))$ for $n \geq 1$. We will see (in Section 1.) that there exists an integer sequence $(b_n)_{n \geq 1}$ such that for all $n \geq 1$ we have $c_n = \prod_{d|n} b_d$. Let K_n be the splitting field of f_n over $\mathbb{Q}$ and denote by $\Omega_n := \text{Gal}(K_n/\mathbb{Q}) = \text{Gal}(f_n/\mathbb{Q})$ its Galois group over the rational numbers.

This paper deals with the problem of determining Ω_n . Following R. W. K. Odoni's paper [2], we will generalize his result concerning the case $a = 1$ to get for all $n \geq 1$:

(*) $\Omega_n \cong [C_2]^n$, if and only if $b_1, b_2, \dots, b_n$ are 2-independent in $\mathbb{Q}$.

A look back . . .

Vol. 59, 1992

Galois groups of some iterated polynomials

243

In the next section, we will apply this result to the Galois groups of Section 1.

3. Some values of a with $\Omega_n \cong [C_2]^n$ for all n .

We combine the results of Sections 1 and 2 to get the following

Theorem. *If $a \in \mathbb{Z}$ has one of the following properties:*

- a) $a > 0$, and $a \equiv 1 \pmod{4}$;
- b) $a > 0$, and $a \equiv 2 \pmod{4}$;
- c) $a < 0$, and $a \equiv 0 \pmod{4}$, and $-a$ is not a square in $\mathbb{Z}$;

then, denoting by f_n the n -th iterate of $f := X^2 + a$, we have

$$\text{Gal}(f_n/\mathbb{Q}) \cong [C_2]^n \quad \text{for all } n \geq 1.$$

A look back . . .

Vol. 59, 1992

Galois groups of some iterated polynomials

243

In the next section, we will apply this result to the Galois groups of Section 1.

3. Some values of a with $\Omega_n \cong [C_2]^n$ for all n .

We combine the results of Sections 1 and 2 to get the following

Theorem. *If $a \in \mathbb{Z}$ has one of the following properties:*

- a) $a > 0$, and $a \equiv 1 \pmod{4}$;
- b) $a > 0$, and $a \equiv 2 \pmod{4}$;
- c) $a < 0$, and $a \equiv 0 \pmod{4}$, and $-a$ is not a square in $\mathbb{Z}$;

then, denoting by f_n the n -th iterate of $f := X^2 + a$, we have

$$\text{Gal}(f_n/\mathbb{Q}) \cong [C_2]^n \quad \text{for all } n \geq 1.$$

- Arboreal Galois representations

A look back . . .

Vol. 59, 1992

Galois groups of some iterated polynomials

243

In the next section, we will apply this result to the Galois groups of Section 1.

3. Some values of a with $\Omega_n \cong [C_2]^n$ for all n .

We combine the results of Sections 1 and 2 to get the following

Theorem. *If $a \in \mathbb{Z}$ has one of the following properties:*

a) $a > 0$, and $a \equiv 1 \pmod{4}$;

b) $a > 0$, and $a \equiv 2 \pmod{4}$;

c) $a < 0$, and $a \equiv 0 \pmod{4}$, and $-a$ is not a square in $\mathbb{Z}$;

then, denoting by f_n the n -th iterate of $f := X^2 + a$, we have

$$\text{Gal}(f_n/\mathbb{Q}) \cong [C_2]^n \quad \text{for all } n \geq 1.$$

- Arboreal Galois representations
- Arithmetic Dynamics

A look back . . .

Vol. 59, 1992

Galois groups of some iterated polynomials

243

In the next section, we will apply this result to the Galois groups of Section 1.

3. Some values of a with $\Omega_n \cong [C_2]^n$ for all n .

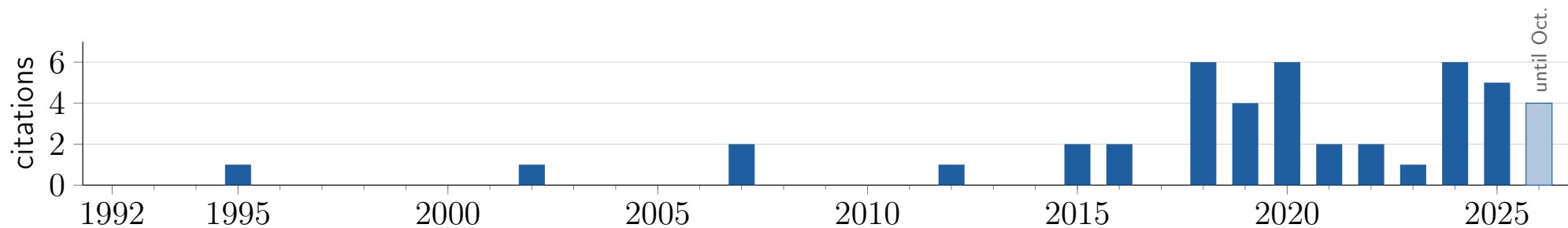
We combine the results of Sections 1 and 2 to get the following

Theorem. *If $a \in \mathbb{Z}$ has one of the following properties:*

- a) $a > 0$, and $a \equiv 1 \pmod{4}$;
- b) $a > 0$, and $a \equiv 2 \pmod{4}$;
- c) $a < 0$, and $a \equiv 0 \pmod{4}$, and $-a$ is not a square in $\mathbb{Z}$;

then, denoting by f_n the n -th iterate of $f := X^2 + a$, we have

$$\text{Gal}(f_n/\mathbb{Q}) \cong [C_2]^n \quad \text{for all } n \geq 1.$$



Iterates of quadratic polynomials

Let $a \in \mathbb{Q}$ and define

$$f_a(x) = x^2 + a \in \mathbb{Q}[x].$$

The iterates of f_a are

$$\begin{aligned} f_a^{\circ 0}(x) &= x, & f_a^{\circ 1}(x) &= f_a(x) = x^2 + a, & f_a^{\circ 2}(x) &= f_a(f_a(x)) = (x^2 + a)^2 + a, & \dots, \\ f_a^{\circ(n+1)}(x) &= f_a(f_a^{\circ n}(x)) = f_a^{\circ n}(f_a(x)), & \dots \end{aligned}$$

Question.

For which a is the Galois group of $f_a^{\circ n}$ maximally large for all n ?

Let K_a be the union of the splitting fields of all $f_a^{\circ n}$.

We obtain an action of $\text{Gal}(K_a/\mathbb{Q})$ on an infinite binary tree T .

Equivalent Question.

For which a is $\text{Gal}(K_a/\mathbb{Q}) \cong \text{Aut}(T)$?

A criterion

Define a sequence (c_n) of rational numbers by

$$c_1 = -a, \quad c_2 = f_a(c_1) = a^2 + a, \quad c_3 = f_a(c_2) = (a^2 + a)^2 + a, \quad \dots$$

Theorem (from my paper).

$$\text{Gal}(K_a/\mathbb{Q}) \cong \text{Aut}(T) \iff (c_n) \text{ is 2-independent.}$$

“2-independent” means that

no finite product with at least one factor is a square.

The proof of the main result used the sequence (b_n) :

$$b_n = \prod_{d|n} c_d^{\mu(n/d)} \quad \text{so that} \quad c_n = \prod_{d|n} b_d.$$

What I did since July

- I had a plan to **formalize** this paper (eventually, **all my papers**).
- In July, I had the opportunity to **test** an **autoformalization system**.
- I asked it to formalize this paper. It produced a **huge Lean file**.
- The code quality was rather **bad**,
so I worked intensely with Claude to **improve** it.

Proof lengths across the versions

Proof length = number of non-blank lines from the `:=` to the end of the declaration (signature excluded), for all theorems and lemmas (definitions excluded). The first four columns are the archived single-file versions (ArchMath-1992.lean, -v2, -v3, -Claude, dated by their last modification); the last two are this repository at its initial commit cd6e871 (2026-08-30) and at aa54510 (2026-09-05).

	v1 (07-08)	v2 (07-10)	v3 (07-12)	Claude (07-18)	2026-08-30	2026-09-05
theorems and lemmas	124	124	124	159	264	304
total proof lines	6103	4453	3561	2782	2676	1409
mean / median	49.2 / 30.5	35.9 / 16	28.7 / 12.5	17.5 / 11	10.1 / 6	4.6 / 3
longest proof	333	293	281	167	48	35
proofs over 20 lines	75	54	49	57	49	4
proofs over 30 lines	62	45	33	27	19	1
proofs of 1 to 3 lines	17	18	34	33	96	170
term-mode proofs	13	13	23	16	42	92

What I did since September

- There are **two papers** (by Hua-Chieh Li, from 2020 and 2021) **extending** the results (still for $a \in \mathbb{Z}$).
- I asked Claude to **add** these results to the formalization.
- I had **gotten interested** in the problem again.
So I asked Claude for a **literature review**.
This turned up **no new results** (for $a \in \mathbb{Z}$ or $a \in \mathbb{Q}$).
- I then asked Claude to try to **extend** the results, based on my method.
It **improved** one of Li's results and **extended** all the results to certain classes of **rational numbers**.
This covered numbers $a \leq -2$ and $a > -1$.
- To make sure the new results are **correct**,
I had Claude (successfully) **formalize** them.

Back to the human

- I then began to **digest** the new results, writing them up in my own way.
- While doing so, I noticed two things:
 - Möbius inversion is **not needed** ($\leadsto$ simpler proofs).
 - Results can be **extended** from $(-1, \infty)$ to $(\alpha \approx -1.543689, \infty)$.
- **Idea** for the first:
Work with **quadratic characters** $\chi: \mathbb{Q}^\times \rightarrow \{\pm 1\}$;
find enough of them to **separate** each c_n from the later ones:
 $\phi_n(c_n) = -1$, but $\forall m > n: \phi_n(c_m) = 1$.
- **Idea** for the second:
We can **control the sign** of $c_{n+1} + c_n$ when $a > \alpha$ (or $a \leq -2$).

Proof strategy overview

Assume that

- ❶ there is a character ϕ_0 with $\forall n: \phi_0(c_n) = -1$;
- ❷ there is a character ϕ_1 with $\phi_1(c_1) = -1$ and $\forall n \geq 2: \phi_1(c_n) = 1$;
- ❸ for each $n \geq 2$, there is a prime $p \equiv 3 \pmod{4}$ such that $v_p(a) = 0$ and $v_p(c_{n+1} + c_n) > 0$.

Then $\text{Gal}(K_a/\mathbb{Q}) \cong \text{Aut}(T)$.

From ❶ and ❸ we obtain ϕ_n as above for all $n \geq 2$.

❸ holds for half of the residue classes modulo 4 of (r, s) , where $a = r/s$ in lowest terms and $a \notin (-2, \alpha)$.

For ❶ and ❷ there is a list of criteria.

Enlisting Claude again

- I then asked Claude to **check my notes** (it found some errors) and to **extend** the list of criteria for **①** and **②**.
- It **formalized** the extended list and **checked** (and formalized) the claim that this **covers everything** we had before.
- The resulting statement was rather technical and **complicated**, so I asked Claude to see if it can be **simplified**.
The result is fairly nice.

Main Result

Theorem.

Let $a = r/s$ in lowest terms. Assume that

- either $a > \alpha$ and $(r, s) \bmod 4 \in \{(0, 3), (1, 1), (1, 3), (2, 1), (3, 0), (3, 2)\}$
- or $a \leq -2$ and $(r, s) \bmod 4 \in \{(0, 1), (1, 0), (1, 2), (2, 3), (3, 1), (3, 3)\}$.

If $a \leq -2$ or $a > \alpha' \approx -1.310703$, or $\alpha < a < \alpha'$ and neither $|r|$ nor s is a square, then

$$\begin{aligned} \text{Gal}(K_a/\mathbb{Q}) \cong \text{Aut}(T) &\iff (-a, a(a+1)) \text{ is 2-independent} \\ &\iff -rs, r(r+s), -s(r+s) \neq \square. \end{aligned}$$

Both α and α' are explicit algebraic numbers.

(More cases can be dealt with in the interval $\alpha < a < \alpha'$.)

Two quadratic families

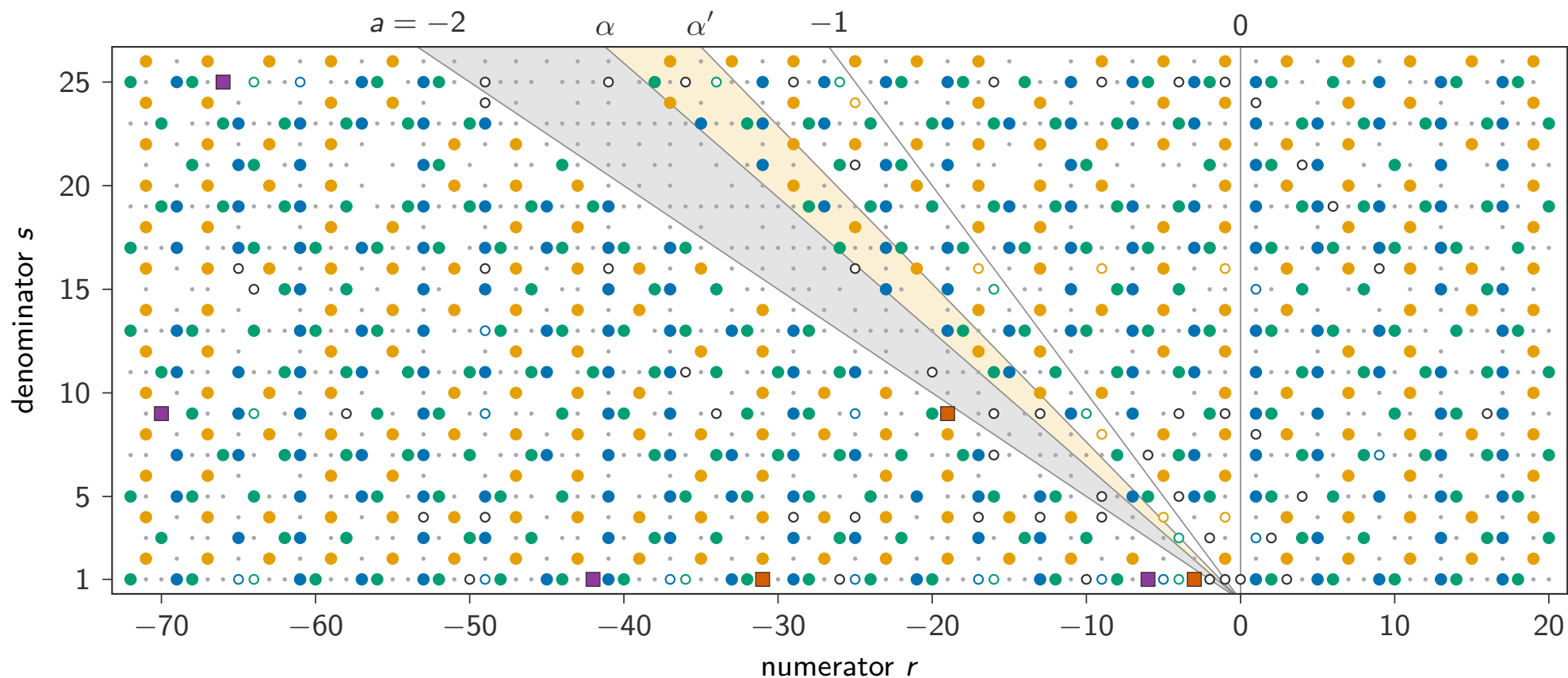
The results in Li's second paper generalize as follows.

Theorem.

Let $\ell = j/k$ with coprime positive integers j and k . If

- $(j, k) \equiv (2, 1) \text{ or } (3, 3) \pmod{4}$ and $a = -\ell(\ell + 1) \leq -2$, or
- $(j, k) \equiv (1, 1) \text{ or } (2, 3) \pmod{4}$ and $a = -(\ell^2 + \ell + 1) \leq -2$,

then $\text{Gal}(K_a/\mathbb{Q}) \cong \text{Aut}(T)$.



Main theorem

- $(r, s) \bmod 4$ for $a \leq -2 \mid a > \alpha$
 • (C1) $(3, 1), (3, 3) \mid (1, 1), (1, 3)$
 • (C2) $(0, 1), (2, 3) \mid (2, 1), (0, 3)$
 • (C3) $(1, 0), (1, 2) \mid (3, 0), (3, 2)$

Families

- $a = -\ell(\ell + 1)$
 ■ $a = -(\ell^2 + \ell + 1)$

Marks

- ρ_a surjective
 ○ ρ_a not surjective
 • not covered

Regions

- $-2 < a \leq \alpha$
 ■ $\alpha < a \leq \alpha'$

Summary

- AI can help with math research by
 - proof-reading papers,
 - searching the literature,
 - generate visualizations,
 - checking results through formalization,
 - running computational experiments,
 - exploring a given mathematical “landscape”.
- As we have seen in the recent weeks,
top AI models can also go all the way to solve open problems.
- So we have to be careful
 - to decide which parts of our research process to outsource to AI
 - so that we still have fun
 - and do not unlearn what makes us mathematicians.

Thank You!