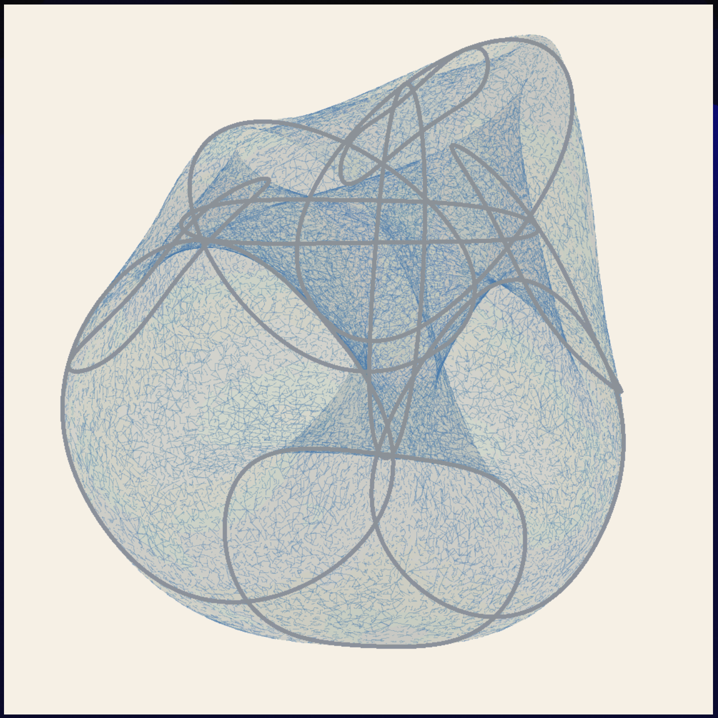
# Integrating AI tools into day-to-day mathematical work



**Daniel Litt, University of Toronto**

# How are LLMs changing mathematics?

- Autonomous informal mathematical work:
  - Vast increase in available semi-expert attention (e.g. Erdos problems)
  - “Last-mile” results (e.g. Erdos problems)
  - Arguably starting to see some higher-quality autonomous work
  - Hard to evaluate capabilities meaningfully (bad incentives, autonomy often unclear)
- Formalization (Codex, Aristotle, Gauss, …)
- What this talk is not about: the future
- What this talk is about: how can LLMs be used (non-autonomously) to help produce high-quality informal mathematics?

Plan: explain a bunch of experiments using LLMs in recent work, and work in progress

# A question of Serre

La question suivante paraît plus hasardeuse:

8.8. *Existe-t-il un motif* $E$ *tel que* $G_{M(E)}$ *soit un groupe simple de type exceptionnel* $G_2$ *(ou* $E_8$*)?*

**Question.** Do there exist motives with Galois group of type $G_2, (F_4, E_6, E_7,)\ E_8$?

**Theorem.** (Dettweiler-Reiter, Yun, Patrikis, Boxer-Calegari-Emerton-Levin-Madapusi Pera-Patrikis, Guralnick-Lübeck-Yu, Faegerman)

Yes.

# A (strengthening of a) question of Serre

**Question.** Does there exist a smooth projective morphism $f : X \to S$ such that the monodromy representation $\rho : \pi_1(S, s) \to GL(H^i(X_s, \mathbb{C}))$ has a summand with monodromy group $G_2, F_4, E_6, E_7, E_8$?

**Theorem.** (Dettweiler-Reiter, Yun, Guralnick-Lübeck-Yu, Fægerman)

Yes, except possibly for $E_6$.

**Theorem.** (Krämer-L-Maculan)

Yes for $E_6$.

## $E_6$-local systems

**Theorem.** (Krämer-L-Maculan)

Fix $n \gg 0$. Let $f : Y \to S$ be a versal family of cubic 3-folds, and $\pi : F \to S$ the family of surfaces parametrizing lines in $Y/S$. Let $\mathbb{L}$ be a rank one torsion local system on $F$ whose restriction to a fiber has order $n$. Then

$$R^2\pi_*\mathbb{L}$$

has connected monodromy group $E_6$.

No AI was ~~harmed~~ used in the proof of this theorem, so why am I talking about it?

# LLMs as a luxury good

**Lemma.** (KLM)

Let $Y$ be a general cubic 3-fold and $F$ the smooth projective surface parametrizing lines in $F$. Then

$$- \cup - : \mathrm{Sym}^2 H^1(F, T_F) \to H^2(F, \wedge^2 T_F)$$

has rank **either** $50$ or $51$.

**Lemma.** (KLM)

Notation as above. Let $\mathscr{L}$ be a general line bundle on $F$ with $c_1(\mathscr{L}) = c_1(\omega_F)$. Then $\mathscr{L}$ is generated by global sections **on the divisor of lines of the second type.**

# LLMs as a luxury good

**Lemma.** (KLM + Deep Think)

Let $Y$ be a general cubic 3-fold and $F$ the smooth projective surface parametrizing lines in $F$. Then

$$- \cup - : \mathrm{Sym}^2 H^1(F, T_F) \to H^2(F, \wedge^2 T_F)$$

has rank $50$.

**Lemma.** (KLM + Aletheia)

Notation as above. Let $\mathscr{L}$ be a general line bundle on $F$ with $c_1(\mathscr{L}) = c_1(\omega_F)$. Then $\mathscr{L}$ is generated by global sections.

# Case study

Want to compute rank of

$$- \cup - : \mathrm{Sym}^2 H^1(F, T_F) \to H^2(F, \wedge^2 T_F)$$

- Lower bound of $50$: As of April 2026, no LLM can compute this correctly "out of the box." Computation with Jacobian ring of Fermat cubic. With some hand-holding, ChatGPT 5.2+ or Opus 4.5+ can execute this.
- Upper bound of $50$: no LLM can prove this correctly "out of the box." Neither could we (1-2 hrs of trying).
- Computing Serre dual map

$$H^0(F, \omega_F^{\otimes 2}) \to \mathrm{Sym}^2 H^1(F, T_F)^\vee$$

in examples suggested conjectural description of kernel.

# Case study

Want to compute rank of

$$- \cup - : \mathrm{Sym}^2 H^1(F, T_F) \to H^2(F, \wedge^2 T_F)$$

**Conjecture.**

Let $Y$ be a cubic 3-fold and $F$ the surface parametrizing lines in $Y$. The kernel of the Serre dual to the map above cuts out the divisor of "lines of the second type" in $Y$, i.e. lines with normal bundle $\mathcal{O}(1) \oplus \mathcal{O}(-1)$.

Prompt (to Gemini Deep Think): Prove or disprove this conjecture.

🤖: essentially correct argument combining several somewhat obscure facts about cubic 3-folds

# Case study: observations

- Results were unnecessary to paper, but were natural questions. Without LLMs we may not have improved the paper this way.
- Required substantial work on the part of (human) authors: finding the right geometric explanation for rank being non-maximal. Existing tools seem not to be able to do this autonomously.
- Many other attempts at analogously improving results in this paper failed, producing dozens of pages of incorrect arguments. Difficult and unpleasant to sort through.
- Best practices

# Compact components of character varieties

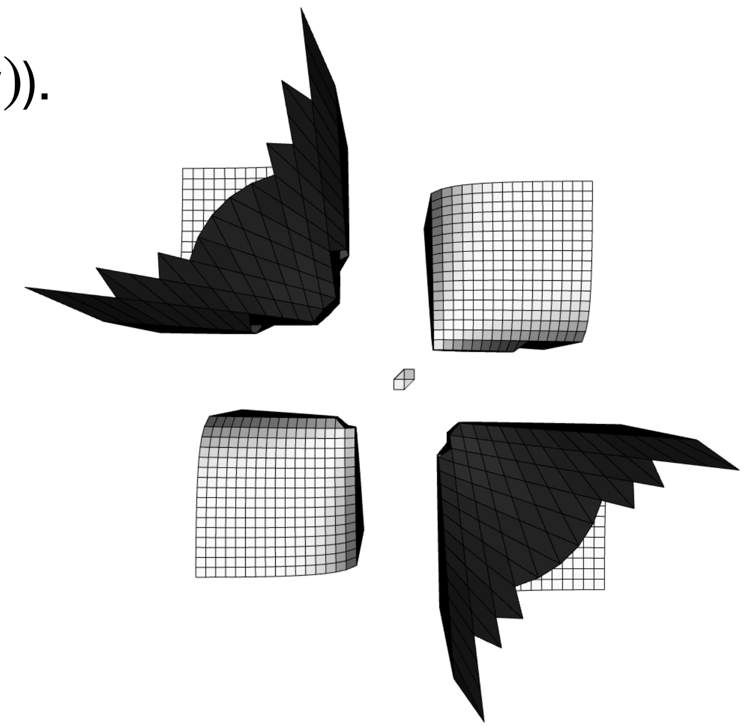
Let $G$ be a real semisimple group (think $SU(p,q)$).

Fix conjugacy classes $C_1, \cdots, C_n \subset G$.

**Question.** What is the topology of

$$\{A_1, \cdots, A_n \mid \prod_i A_i = \mathrm{id}_G, A_i \in C_i\} \Big/ G$$

i.e. the character variety parametrizing $G$-representations of $\pi_1(\mathbb{CP}^1 \setminus \{x_1, \cdots, x_n\})$ up to conjugacy?



A relative $\mathrm{SL}_2(\mathbb{R})$-character variety of $\mathbb{P}^1 \setminus \{x_1, \cdots, x_4\}$ with a compact component. From Benedetto-Goldman's "The Topology of the Relative Character Varieties of a Quadruply-Punctured Sphere," Experiment. Math. 8(1): 85-103 (1999).

# Compact components of character varieties

- $\Sigma_{g,n}$ - orientable surface of genus $g$ with $n$ punctures
- $\underline{C} = (C_1, \cdots, C_n)$ an $n$-tuple of conjugacy classes in real semisimple group $G$
- $X(\Sigma_{g,n}, \underline{C})$ - real algebraic variety parametrizing conjugacy classes of $\pi_1(\Sigma_{g,n})$-representations into $G$ sending loop around $i$th puncture into $C_i$, for all $i$.

**Question.** What are the compact components of $X(\Sigma_{g,n}, \underline{C})$?

# Compact components of character varieties

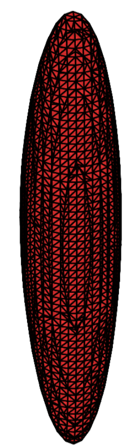
- Fully understood for $\mathrm{SL}_2(\mathbb{R})$ (Deroin-Tholozan)
- Examples for $g = 0$ and other groups by Tholozan-Toulisse ("minimal energy"), studied by Charlie Wu. Thesis problem: is this (or compact $G$) the only source of examples?

**Theorem.** (Wu) No! There are other compact components, and some kind of classification in terms of certain special complex variations of Hodge structure he calls "crest variations."

**Theorem.** (Wu) Suppose $X(\Sigma_{g,n}, \underline{C})$ has a smooth compact component for some $g > 1$. Then $G$ is compact.

# Compact components of character varieties

- Crest variations of Hodge structure have a lot of interesting properties!
- But finding them is an artisanal process.
- **Definition.** A *crest variation of Hodge structure* on a complex algebraic curve $Y$ is a complex variation of Hodge structure $\mathbb{V}$ on $Y$ with $h^{p,1-p}(\mathrm{W}^1\mathrm{H}^1(Y, \text{ad } \mathbb{V})) = 0$ for $p > 0$ even.
- Can we find these via computer search?
- 6 months ago, I would have barely considered attempting this.



A “new” compact component of an $SU(2,1)$-character variety for $\Sigma_{0,3}$. Picture due to Charlie Wu.

# Winter break 2025

- Using Codex, I vibe-coded command-line tools to compute with *parabolic Higgs bundles*
- Found 1000s of essentially distinct examples, very useful in formulating conjectures for ongoing work.
- Code is **extremely** low quality
- Best practices

cli.py
config.py
deformation.py
euler_characteristics.py
graded_vhs.py
la.py
line_chain.py
parse.py
poly.py
search.py
sections.py

```
python3 examples/search_rank9_grade333_offmiddle_crest_generic_flags.py \
  --points 0,1,-1,inf \
  --denom 20 \
  --max-distinct-per-point 2 \
  --trials 200000 \
  --degrees-per-weight 3 \
  --flags-trials 2 \
  --theta-trials 6 \
  --seed 602 \
  --max-results 3 \
  --out examples/tmp_rank9_grade333_offmiddle_crest_seed602.jsonl
```

examples

Name
search_hits_p1_5pts_rank3_denom20_ties2_no_pd0.jsonl
search_hits_p1_5pts_rank3_denom20_ties2_pd0_v2.jsonl
search_hits_p1_5pts_rank4_denom20_max3_no_pd0.jsonl
search_hits_p1_5pts_rank4_denom20_max3_pd0.jsonl
search_hits_p1_5pts_rank4_denom20_ties2_no_pd0.jsonl
search_hits_p1_5pts_rank4_denom20_ties2_pd0_v2.jsonl
search_hits_p1_012inf_rank3_pd0.jsonl
search_hits_p1_012inf_rank4_pd0.jsonl
search_hits_p1_0112inf_rank3_pd0.jsonl
search_hits_p1_0112inf_rank4_pd0.jsonl
search_rank6_222_residual_branch.py
search_rank6_grade222_offmiddle_crest.py
search_rank9_grade333_offmiddle_crest_generic_flags.py
search_rank9_grade333_offmiddle_crest_targeted_simple.py
search_rank9_grade333_offmiddle_crest.py
search_three_step_bound_targeted.py
stats_rank6_grade222_forced_kernel_lines.py
stats_rank6_grade222_max_weight_relations.py

# Unitary local systems

- Rigid tuple: $\underline{C} = (C_1, \cdots, C_n)$, $C_i \subset G$, such that $X(\Sigma_{0,n}, \underline{C})$ is finite.
- For $G = \mathrm{GL}_r$, Katz classified these, by constructing an invertible operation (*middle convolution*), that, given a rigid tuple $\underline{C}$ in $\mathrm{GL}_r$, constructs a rigid tuple $\mathrm{MC}_\lambda(\underline{C})$ in $\mathrm{GL}_{r'}$, where $r' < r$ (if $r > 1$).
- Question: Can you classify rigid tuples in $U(n)$? (Related to inverse Galois problem)
- **Question** (Belkale, Rigid local systems and the multiplicative eigenvalue problem, Ann. of Math., 2022) Given a rigid tuple in $U(r)$, can one always make choices so that middle convolution produces a rigid tuple in $U(r')$ with $r' < r$?

# Unitary local systems

**Question** (Belkale, Rigid local systems and the multiplicative eigenvalue problem, Ann. of Math., 2022) Given a rigid tuple in $U(r)$, can one always make choices so that middle convolution produces a rigid tuple in $U(r')$ with $r' < r$?

**Theorem.** (Lam-L—)

No.

**Example.**

Set $\zeta = e^{2\pi i/100}$. Then consider the conjugacy classes in $U(16)$ given by:

$C_1 = \mathrm{diag}(\zeta^{-15}(1), \zeta(15))$

$C_2 = C_3 = C_4 = \mathrm{diag}(\zeta^{-8}(8), \zeta^{8}(8))$

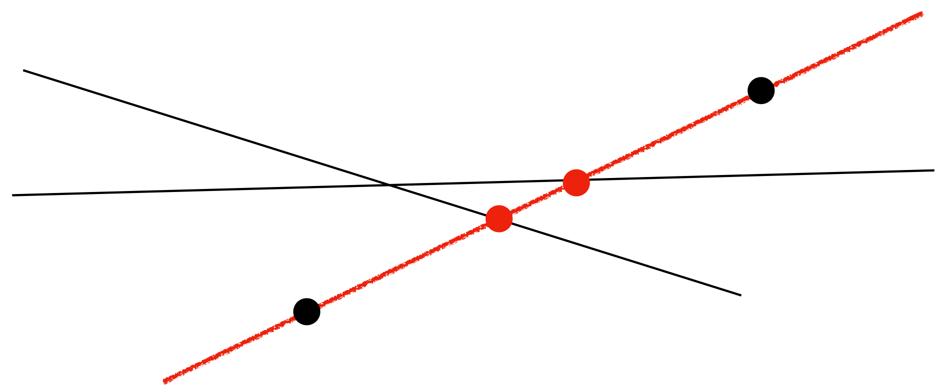
$C_5 = \mathrm{diag}(\zeta^{-4}(12), \zeta^{12}(4))$

# How was this example found?

- By results of Mehta-Seshadri, classifying unitary rigid tuples is related to classifying tuples of parabolic subgroups $(P_1, \cdots, P_n)$ in $G$ so that the diagonal $G$-action on $\prod_{i=1}^{n} G/P_i$ has a dense open orbit with finite stabilizer.
- Can reinterpret middle convolution in terms of *Gale duality*
- Computer search for examples where
  1. $$\sum_{i=1}^{n} \dim G/P_i = \dim G$$
  2. Middle convolution/Gale duality cannot reduce rank
- This isn't enough: the numerical equality in (1) doesn't guarantee existence of a dense orbit (and various other minor issues).

# How was this example found?

**Example.** $\mathrm{SL}_3$-action on $\mathbb{P}^2 \times \mathbb{P}^2 \times \check{\mathbb{P}}^2 \times \check{\mathbb{P}}^2$



Given $P_1, \cdots, P_n \subset G$ with $\sum_{i=1}^{n} \dim G/P_i = \dim G$, need to check if there is a dense open orbit. Could write code to do this.

Or: just ask ChatGPT 5.4 Pro.

⚠️

# ☣️ 🚨 ☢️ Using LLMs for writing ☢️ 🚨 ☣️

🙈

# LLMs are OK editors

Ask Codex “give me line-by-line comments on section 4 of my paper”

- L1137: Environment delimiter; ok.
- L1138: Environment delimiter; ok.
- L1139: Minor: double-space; consider tightening.
- L1140: Display-math delimiter; ok.
- L1141: LaTeX line; ok (watch spacing/line breaks).
- L1142: Display-math delimiter; ok.
- L1143: Potential typo: here $\pi_i$ should be the projections $(S\times\A^1)(1)\rightrightarrows S\times\A^1$ (not to $S$). The diagram at L1147–L1152 uses $S\times\A^1$.
- L1145: Minor: double-space; consider tightening.
- L1146: Display-math delimiter; ok.
- L1147: Environment delimiter; ok.
- L1148: Prose line; consider wrapping/splitting if it overflows.
- L1149: Prose line; consider wrapping/splitting if it overflows.
- L1150: Prose line; consider wrapping/splitting if it overflows.

- L528: Fine; if “$p+q=i$” uses $p$ that may visually collide with the prime $p$ elsewhere—maybe use different indices (e.g. $a+b=i$) if you find it distracting.
- L529: Fine.
- L530: Fine.
- L531: Fine.
- L532: Fine; consider briefly reminding readers in surrounding text that $X^{(1)}$ is Frobenius twist.
- L533: Fine.
- L534: Fine.
- L535: Fine.
- L536: Fine.
- L537: Fine.
- L538: Fine.
- L539: Fine.
- L540: Fine.
- L541: Fine.
- L542: Fine.
- L543: Fine.

# But can they write?

- In December 2025 I visited my collaborator Josh Lam in Bonn to finish writing this paper:

*[Submitted on 12 Jan 2026]*

**p-Curvature and Non-Abelian Cohomology**

Yeuk Hay Joshua Lam, Daniel Litt

Let $X \to S$ be a smooth projective morphism. Katz proved the Grothendieck-Katz p-curvature conjecture for the Gauss-Manin connection on the $i$-th cohomology of $X/S$: if its $p$-curvature vanishes mod $p$ for infinitely many $p$, then the action of $\pi_1(S, s)$ on $H^i(X_s, \mathbb{Z})$ factors through a finite group. We prove a non-abelian analogue of this statement: if the p-curvature of the isomonodromy foliation on the moduli of flat bundles of rank $r$ on $X/S$ vanishes mod $p$ for infinitely many $p$, then the action of $\pi_1(S, s)$ on the rank $r$ integral characters of $\pi_1(X_s)$ factors through a finite group. We deduce many new cases of the Bost/Ekedahl--Shepherd-Barron--Taylor conjecture.
The proofs rely on a non-abelian version of Katz's formula, and a non-abelian version of the Hodge index theorem.

- GPT 5.2 Pro had just come out. As an experiment I gave it the TeX draft, with the proofs of two lemmas only sketched, and asked it to expand the sketches into complete proofs. I raced it.

# But can they write?

- I won the race!

**Lemma 5.3.1.** *Let* $f : X \to S$ *be a relative curve of genus* $g$*, with* $S$ *smooth. Fix* $s \in S(\mathbb{C})$ *and set* $\Sigma = X_s$*. Let* $\rho : \pi_1(\Sigma) \to \mathrm{SL}_r(\mathbb{C})$ *be a semisimple representation.*

*Suppose that* $\Theta_{X/S}$ *vanishes on the leaf of the isomonodromy foliation corresponding to* $\rho$ *(viewed as a real-analytic subset of* $\mathscr{M}_{\mathrm{Dol}}(X/S)$*). Let* $\mathscr{U} \subset \mathscr{T}_g$ *be the preimage of the image of the map* $S \to \mathscr{M}_g$ *classifying* $f$.

*Then* $E_\rho|_{\mathscr{U}}$ *is locally constant.*

- But both proofs produced by Codex were essentially correct.
- And I liked the way Codex handled one step in the argument more than my own original argument, and so implemented it.
- Best practices

# Other tools

- Prism
- Refine.ink
- Coarse.ink

```
Comment 8

CLARITY OF THE ARGUMENT IN THE PROOF OF THEOREM 6.2.2

"For each $i$, set $S_{i}$ to be the closure of $\pi_{1}\left(S(\mathbb{C})^{\text {an }}, s\right)
\cdot\left[\rho_{l_{i}}\right]$ in the analytic topology of $M_{B}\left(X_{s}, r\right)(\mathbb{C})$. The set $S_{i}$ is
compact by Theorem 5.3.2 (using that $\Theta_{X / S}$ vanishes identically). On the other hand, the orbit of
$\left[\rho_{l_{i}}\right]$ consists of $\mathscr{O}_{K}$-points for each $i$. Hence the orbit is finite by Lemma 6.2.1, as
desired."

Message: The line
"On the other hand, the orbit of $\left[\rho_{\iota_i}\right]$ consists of $\mathscr{O}_K$-points for each $i$"
is misleading as stated. The point $[\rho_{\iota_i}]$ is a complex point of $M_B(X_s,r)$, and its orbit under the analytic
$\pi_1(S(\mathbb{C})^{\mathrm{an}},s)$-action a priori lives in $M_B(X_s,r)(\mathbb{C})$, not in $M_B(X_s,r)
(\mathscr{O}_K)$.

What is actually used in combination with Lemma 6.2.1 is that the orbit of the original integral point $[\rho]\in
M_B(X_s,r)(\mathscr{O}_K)$, under the algebraic $\pi_1$-action on the character variety (defined over $\mathbb{Z}$), again
consists of $\mathscr{O}_K$-points, and that for each embedding $\iota_i:K\hookrightarrow\mathbb{C}$ the image $\iota_i$ of
this orbit lies in the compact set $S_i$ (the closure of $\pi_1\cdot[\rho_{\iota_i}]$).

Rephrasing this step to refer explicitly to the orbit of $[\rho]$ in $M_B(X_s,r)(\mathscr{O}_K)$, and to note that its
images under all $\iota_i$ are contained in the respective compact sets $S_i$, would make the application of Lemma 6.2.1
completely clear.
```

- Comments are often wrong (bad) or nitpicky (good)
- Best practices

# Upshot

- Tools are rapidly growing more powerful
- They'll produce a lot of slop
- But we can also use them to produce higher-quality work