# AI4Math: How Machines Explore, Conjecture, and Discover Mathematics

Sebastian Pokutta

Technische Universität Berlin
and
Zuse Institute Berlin

pokutta@math.tu-berlin.de
@spokutta

**Munich AI Lecture @ BAIOSPHERE**

February, 2026 · Munich, Germany

ZUSE INSTITUTE BERLIN

Berlin Mathematics Research Center MATH+

# What is this talk about?

## Introduction

*What AI can do for Mathematics? Trends and two concrete examples.*

# What is this talk about?
## Introduction

*What AI can do for Mathematics? Trends and two concrete examples.*

**Why?** AI is going to change the way of how we do science and mathematics.

# What is this talk about?
## Introduction

*What AI can do for Mathematics? Trends and two concrete examples.*

**Why?** AI is going to change the way of how we do science and mathematics.

**Today:** Brief overview of approach and two new applications.

# What is this talk about?

## Introduction

*What AI can do for Mathematics? Trends and two concrete examples.*

**Why?** AI is going to change the way of how we do science and mathematics.

**Today:** Brief overview of approach and two new applications.

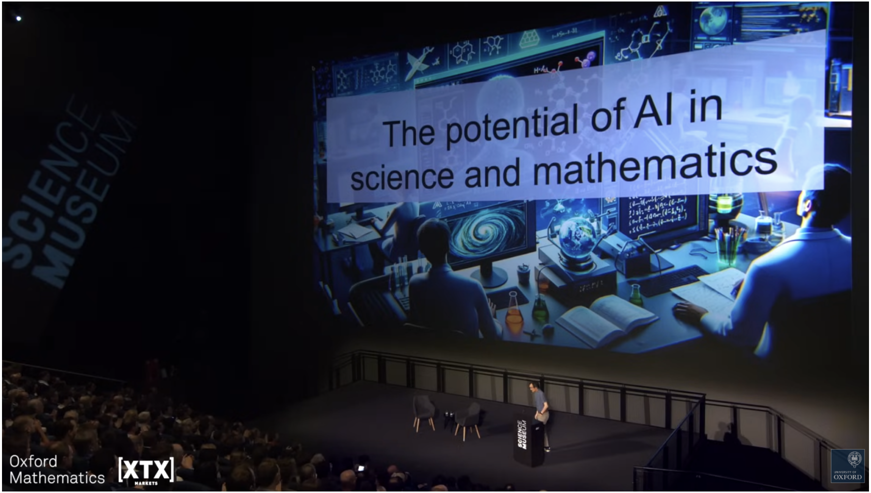
### Outline

- AI and Mathematics
- The Hadwiger-Nelson Problem
- Hilbert's 16th Problem

(Hyperlinked) References are not exhaustive; check references contained therein.

# What is this talk about?

## AI and Mathematics




[The Potential for AI in Science and Mathematics - Terence Tao]

# What is this talk about?

## AI and Mathematics

**Various levels of co-creation.** [Haase and Pokutta, 2026]

- Digital Pen: basically like autocorrect, bibtex lookup, etc. “2000s”
- AI Task Specialist: ChatGPT, Claude, Gemini, etc. 2022 - 2025
- AI Assistant: Agents with integrated tools, verification, etc. 2025 - present
- AI Co-creator: Fully integrated, autonomous, co-creator 2027(??)

# What is this talk about?

## AI and Mathematics

**Various levels of co-creation.** [Haase and Pokutta, 2026]

- Digital Pen: basically like autocorrect, bibtex lookup, etc. "2000s"
- AI Task Specialist: ChatGPT, Claude, Gemini, etc. 2022 - 2025
- AI Assistant: Agents with integrated tools, verification, etc. 2025 - present
- AI Co-creator: Fully integrated, autonomous, co-creator 2027(??)

Capabilities are impressive but unstable:

- SOTA models achieve post-PhD level scores on benchmarks, yet in day-to-day use make trivial (logical) errors.
- No hard verification of results and randomness across runs.
- Prompting and scaffolding are still a challenge.
- Availability of tools for verification etc crucial.

# What is this talk about?

## Mathematics with computers is not new

**Various high-profile examples from the past.**

- Four Color Theorem: massive computer-based case checking

  [Appel and Haken, 1977, Robertson et al., 1997]
- Kepler Conjecture / Hales' Theorem: extensive computer verification

  [Hales et al., 2017]
- Classification of Finite Simple Groups: Formal verification with Lean and Coq
- Boolean Pythagorean Triples Problem: A spectacular 200TB SAT-solver proof.

  [Heule et al., 2016]

# What is this talk about?

## Mathematics with computers is not new

**Various high-profile examples from the past.**

- Four Color Theorem: massive computer-based case checking

  [Appel and Haken, 1977, Robertson et al., 1997]
- Kepler Conjecture / Hales' Theorem: extensive computer verification

  [Hales et al., 2017]
- Classification of Finite Simple Groups: Formal verification with Lean and Coq
- Boolean Pythagorean Triples Problem: A spectacular 200TB SAT-solver proof.

  [Heule et al., 2016]

**Crucial role in computational mathematics / scientific computing**

- Finite Elements
- Numerical Simulations
- Optimization
- Engineering
- ...

# The Hadwiger-Nelson Problem

joint work with: Aldo Kiem, Konrad Mundinger,
Christoph Spiegel, Max Zimmer

ICML 2025 (oral)
`https://arxiv.org/abs/2404.05509`

Partially supported by ExC MATH+ Project EF-LiOpt-3
Agent AI in Mathematics

[Mundinger et al., 2025]

# The Hadwiger-Nelson Problem

Problem (Nelson 1950, also: Gardner, Moser, Erdős, Harary, Tutte, ...)

*What is the smallest number of colors sufficient for coloring the plane in such a way that no two points of the same color are at a unit distance apart?*

# The Hadwiger-Nelson Problem

Problem (Nelson 1950, also: Gardner, Moser, Erdős, Harary, Tutte, ...)

*What is the smallest number of colors sufficient for coloring the plane in such a way that no two points of the same color are at a unit distance apart?*

Infinite graph with vertex set $\mathbb{E}^2$ and edges $\{x, y\}$ for any $x, y \in \mathbb{E}^2$ with $\|x - y\| = 1$

$\Rightarrow$ chromatic number of the plane $\chi(\mathbb{E}^2)$

# The Hadwiger-Nelson Problem

**Problem (Nelson 1950, also: Gardner, Moser, Erdős, Harary, Tutte, ...)**

*What is the smallest number of colors sufficient for coloring the plane in such a way that no two points of the same color are at a unit distance apart?*

Infinite graph with vertex set $\mathbb{E}^2$ and edges $\{x, y\}$ for any $x, y \in \mathbb{E}^2$ with $\|x - y\| = 1$
$\Rightarrow$ chromatic number of the plane $\chi(\mathbb{E}^2)$

**Theorem**

*Assuming* `Axiom of Choice (AoC)`*:* [Bruijn and Erdos, 1951]

*Any graph is k-colorable iff every finite subgraph of it is k-colorable.*

# The Hadwiger-Nelson Problem

**Problem (Nelson 1950, also: Gardner, Moser, Erdős, Harary, Tutte, ...)**

*What is the smallest number of colors sufficient for coloring the plane in such a way that no two points of the same color are at a unit distance apart?*

Infinite graph with vertex set $\mathbb{E}^2$ and edges $\{x, y\}$ for any $x, y \in \mathbb{E}^2$ with $\|x - y\| = 1$
$\Rightarrow$ chromatic number of the plane $\chi(\mathbb{E}^2)$

**Theorem**

*Assuming* `Axiom of Choice (AoC)`*:* [Bruijn and Erdos, 1951]

*Any graph is k-colorable iff every finite subgraph of it is k-colorable.*

This problem has a long and complicated history... over 14 pages in [Soifer, 2024]

**Table 3.1** Who created the chromatic number of the plane problem?

| Publication | Year | Author(s) | Problem creator(s) or source named |
|---|---|---|---|
| [Gar2] | 1960 | Gardner | "**Leo Moser** ...writes..." |
| [Had4] | 1961 | Hadwiger (after Klee) | **Nelson** |
| [E61.22] | 1961 | Erdős | "I cannot trace the origin of this problem" |
| [Cro] | 1967 | Croft | "A long[18]-standing open problem of **Erdős**" |
| [Woo1] | 1973 | Woodall | **Gardner** |
| [Sim] | 1976 | Simmons | **Erdős**, **Harary**, and **Tutte** |
| [E80.38] [E81.23] [E81.26] | 1980–1981 | Erdős | **Hadwiger** and **Nelson** |
| [CFG] | 1991 | Croft, Falconer, and Guy | "Apparently due to **E. Nelson**" |
| [KW] | 1991 | Klee and Wagon | "Posed in 1960–61 by **M. Gardner** and **Hadwiger**" |

p. 24 in [Soifer, 2024]

# Lower bounds on $\chi(\mathbb{E}^2)$

# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.

# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.

**Definition**

A graph $G = (V, E)$ is a unit distance graph if there exists an embedding $f : V \to \mathbb{E}^2$ of its vertices in the plane s.t. $\|f(u) - f(v)\| = 1$ if and only $\{u, v\} \in E$.

# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.

**Definition**

A graph $G = (V, E)$ is a unit distance graph if there exists an embedding $f : V \to \mathbb{E}^2$ of its vertices in the plane s.t. $\|f(u) - f(v)\| = 1$ if and only $\{u, v\} \in E$.

A triangle gives a lower bound of 3.

# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.

**Definition**

A graph $G = (V, E)$ is a unit distance graph if there exists an embedding $f : V \to \mathbb{E}^2$ of its vertices in the plane s.t. $\|f(u) - f(v)\| = 1$ if and only $\{u, v\} \in E$.

A triangle gives a lower bound of 3.
The Moser spindle gives a lower bound of 4. [Moser and Moser, 1961]

# Lower bounds through unit distance graphs

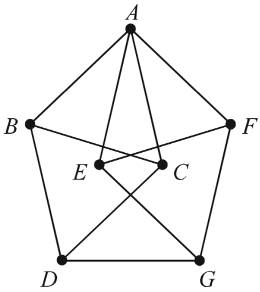
## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.

**Definition**

A graph $G = (V, E)$ is a unit distance graph if there exists an embedding $f : V \to \mathbb{E}^2$ of its vertices in the plane s.t. $\|f(u) - f(v)\| = 1$ if and only $\{u, v\} \in E$.

A triangle gives a lower bound of 3.
The Moser spindle gives a lower bound of 4. [Moser and Moser, 1961]



# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.

**Definition**

A graph $G = (V, E)$ is a unit distance graph if there exists an embedding $f : V \to \mathbb{E}^2$ of its vertices in the plane s.t. $\|f(u) - f(v)\| = 1$ if and only $\{u, v\} \in E$.

A triangle gives a lower bound of 3.
The Moser spindle gives a lower bound of 4. [Moser and Moser, 1961]

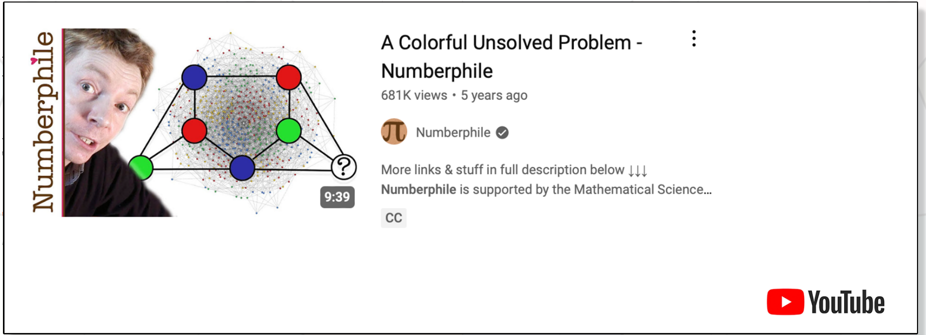
**Theorem**

*There is a unit distance graph on* 20 425 *vertices with chromatic number* 5. [De Grey, 2018]

# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

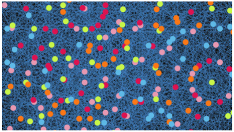
Find unit distance graphs of large chromatic number.



# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.



GRAPH THEORY

**Decades-Old Graph Problem Yields to Amateur Mathematician**

*By* EVELYN LAMB | APRIL 17, 2018 | 26 |

...number of vertices? The problem, now known as the Hadwiger-Nelson problem or the problem of finding the chromatic number of the plane, has piqued the interest of many mathematicians, including...

Quantamagazine

# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.

Aubrey de Grey and Alexander Soifer, *Il Vicino*, January 18, 2020

Ronald L. Graham presents Aubrey D.N.J. de Grey the Prize: $1000, San Diego, September 22, 2018

# Lower bounds through unit distance graphs

## The Hadwiger-Nelson Problem

Find unit distance graphs of large chromatic number.

**Definition**

A graph $G = (V, E)$ is a unit distance graph if there exists an embedding $f : V \to \mathbb{E}^2$ of its vertices in the plane s.t. $\|f(u) - f(v)\| = 1$ if and only $\{u, v\} \in E$.

A triangle gives a lower bound of 3.
The Moser spindle gives a lower bound of 4. [Moser and Moser, 1961]

**Theorem**

*There is a unit distance graph on* 20 425 *vertices with chromatic number* 5. [De Grey, 2018]

Simpler constructions with...

1. 1581 vertices — for detail see [De Grey, 2018]
2. 627 vertices — [Exoo and Ismailescu, 2020]
3. 553 vertices (as part of `Polymath16`) — Marijn Heule, for details see [Mixon, 2021]
4. 509 vertices (as part of `Polymath16`) — Jaan Parts, for details see [Mixon, 2021]

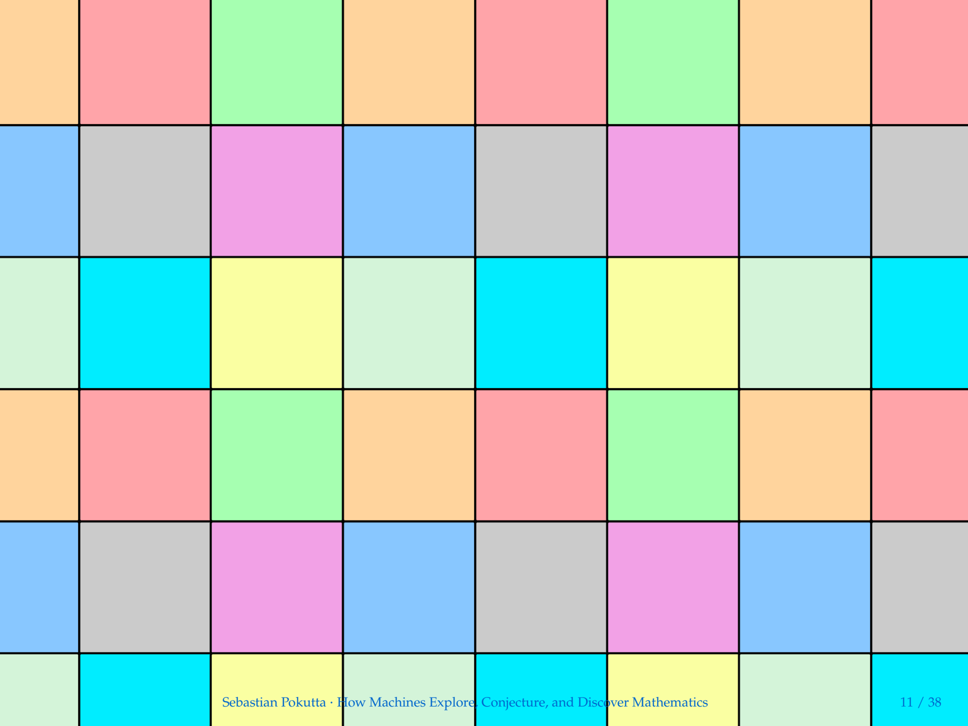
# Upper bounds on $\chi(\mathbb{E}^2)$

# Upper bounds through colorings

## The Hadwiger-Nelson Problem

Explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \ldots, c\}$, usually derived through tesselations using simple polytopal shapes, give

$$5 \leq \chi(\mathbb{E}^2) \leq \ldots$$

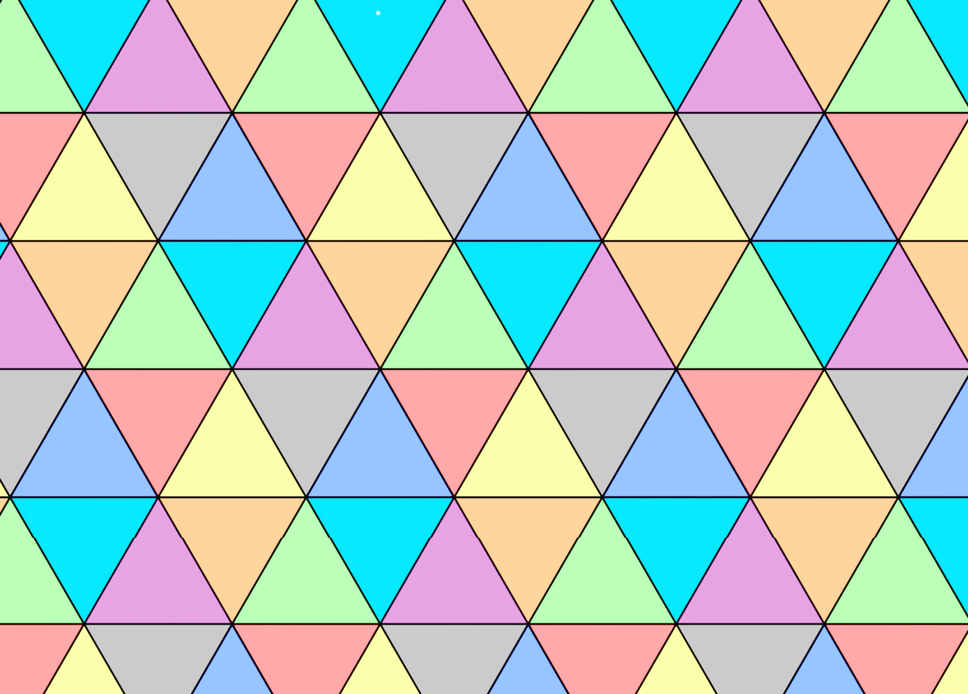


# Upper bounds through colorings

## The Hadwiger-Nelson Problem

Explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \ldots, c\}$, usually derived through tesselations using simple polytopal shapes, give

$$5 \leq \chi(\mathbb{E}^2) \leq 9.$$

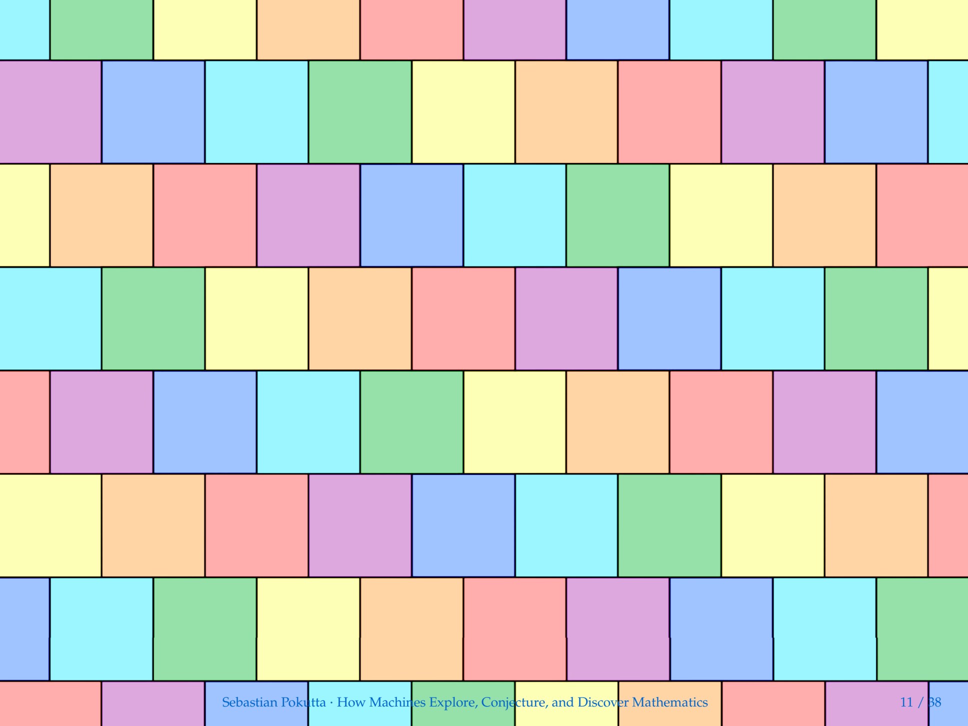


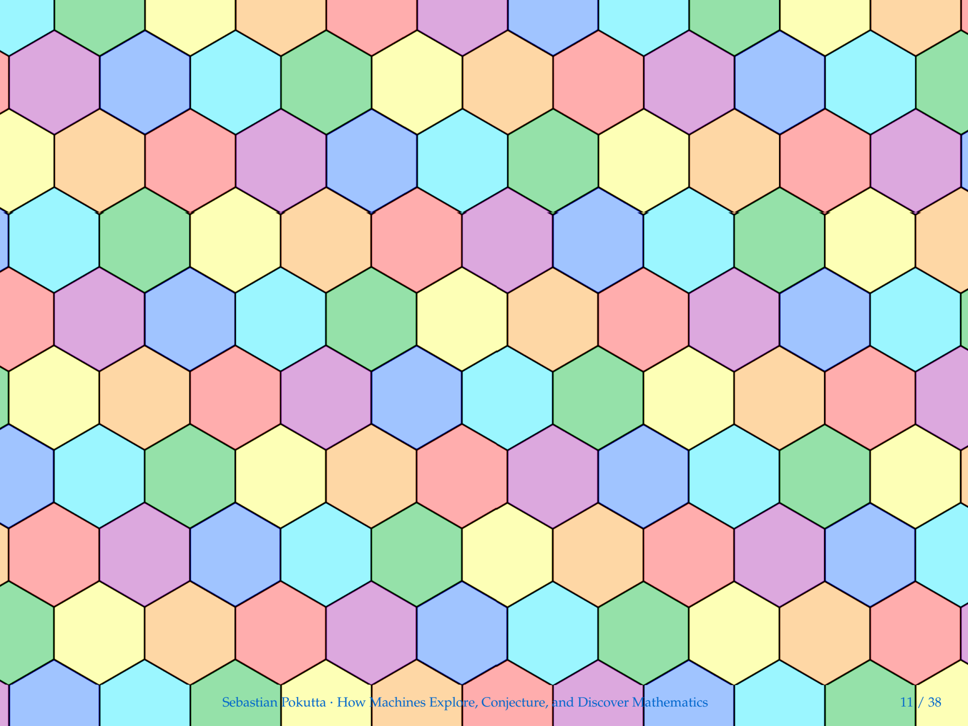
# Upper bounds through colorings

## The Hadwiger-Nelson Problem

Explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \ldots, c\}$, usually derived through tesselations using simple polytopal shapes, give

$$5 \leq \chi(\mathbb{E}^2) \leq 8.$$



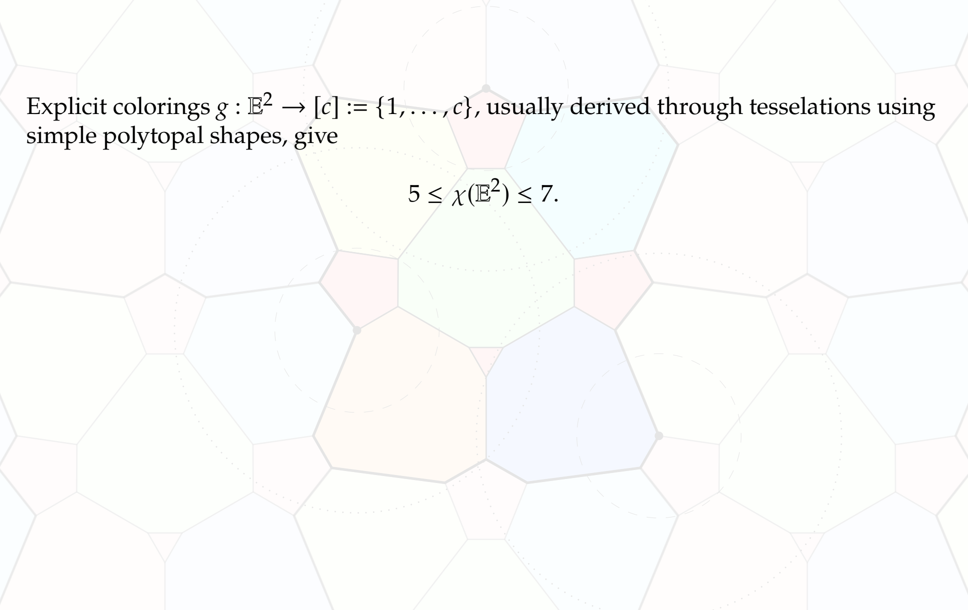


# Upper bounds through colorings

## The Hadwiger-Nelson Problem

Explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \ldots, c\}$, usually derived through tesselations using simple polytopal shapes, give

$$5 \leq \chi(\mathbb{E}^2) \leq 7.$$



# Upper bounds through colorings

## The Hadwiger-Nelson Problem

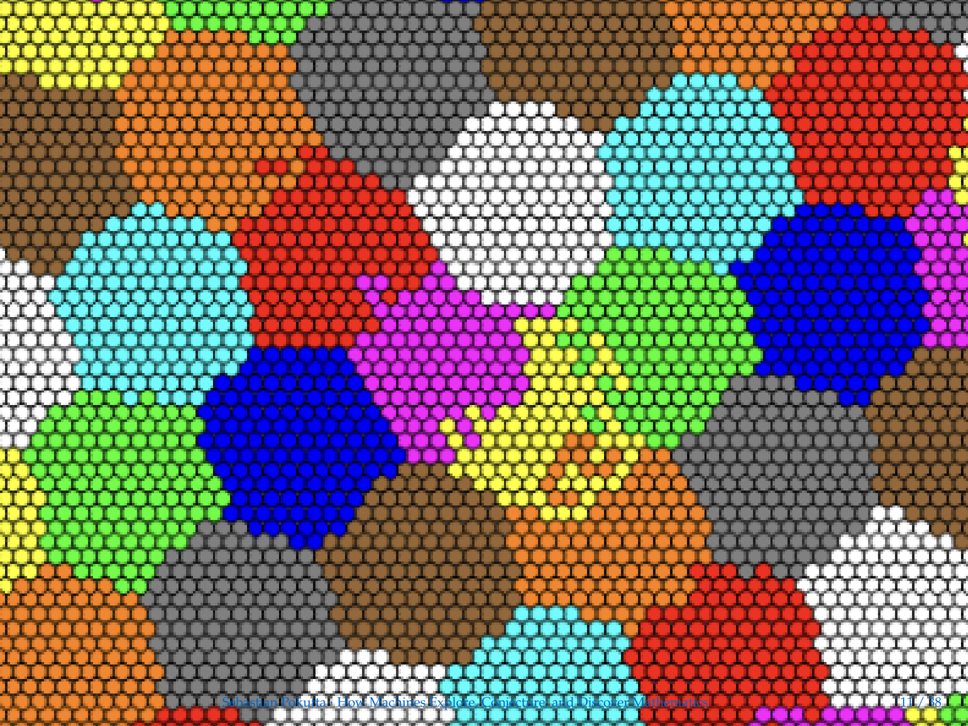
Explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \ldots, c\}$, usually derived through tesselations using simple polytopal shapes, give

$$5 \leq \chi(\mathbb{E}^2) \leq 7.$$

**Question.** Can we use computers to find admissible colorings $g : \mathbb{E}^2 \to [c]$, i.e.,

$$\left\{x \in \mathbb{E}^2 \mid \exists y \in B_1(x) : g(x) = g(y)\right\} = \emptyset?$$

... attempts, e.g., via discretization and SAT solvers...



# Upper bounds through colorings

## The Hadwiger-Nelson Problem

Explicit colorings $g : \mathbb{E}^2 \to [c] := \{1, \ldots, c\}$, usually derived through tesselations using simple polytopal shapes, give

$$5 \leq \chi(\mathbb{E}^2) \leq 7.$$

**Question.** Can we use computers to find admissible colorings $g : \mathbb{E}^2 \to [c]$, i.e.,

$$\left\{x \in \mathbb{E}^2 \mid \exists y \in B_1(x) : g(x) = g(y)\right\} = \emptyset?$$

... attempts, e.g., via discretization and SAT solvers...

**Idea.** Use a parameterized and easily differentiable family $g_\theta : \mathbb{E}^2 \to \Delta_c$ and find

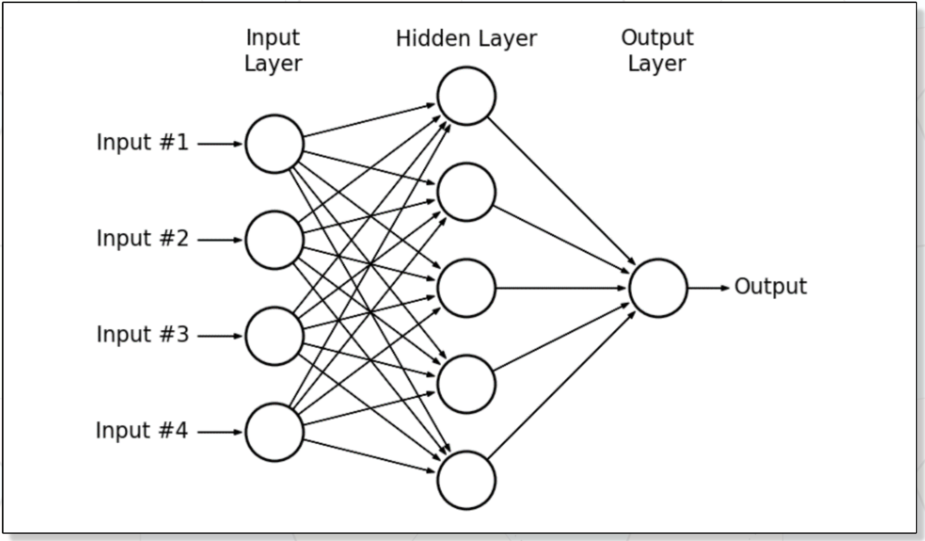
$$\arg\min_\theta \mathbb{E}\left[\int_{B_1(x)} g_\theta(x) \cdot g_\theta(y)\, dy \,\middle|\, x \in \mathbb{E}^2\right].$$

**Key Point.** Approach is continuous in nature.

# New upper bounds via machine learning?

# One second recap: Neural Networks

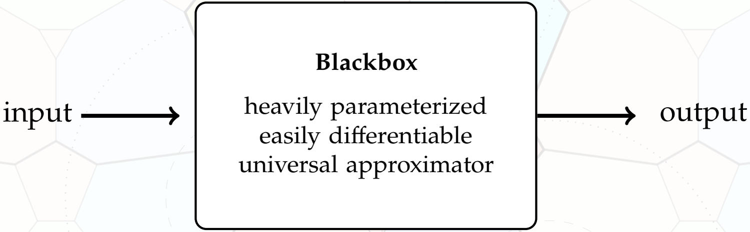
## Neural Networks as Colorings



# One second recap: Neural Networks

## Neural Networks as Colorings

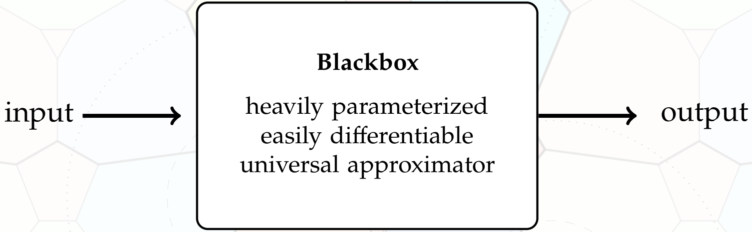
**Here.** Simply a parameterized continuous function to model the coloring.



# One second recap: Neural Networks

## Neural Networks as Colorings

**Here.** Simply a parameterized continuous function to model the coloring.




**Theorem (Universal Approximation Theorem)**

*Feedforward neural networks with certain activation functions are dense (w.r.t. compact convergence) in the space of continuous functions.*

# Can we improve the upper bound?

## Neural Networks as Colorings

**Idea.** Use gradient descent to train a feedforward network $g_\theta$ to minimize

$$L(\theta) = \int_{[-b,b]\times[-b,b]} \int_{B_1(x)} g_\theta(x) \cdot g_\theta(y)\, dy\, dx$$

for some reasonable $b \in \mathbb{R}$?

# Can we improve the upper bound?

## Neural Networks as Colorings

**Idea.** Use gradient descent to train a feedforward network $g_\theta$ to minimize

$$L(\theta) = \int_{[-b,b]\times[-b,b]} \int_{B_1(x)} g_\theta(x) \cdot g_\theta(y)\, dy\, dx$$

for some reasonable $b \in \mathbb{R}$?

**Problem.** Still a continuous problem. How to compute?

# Can we improve the upper bound?

## Neural Networks as Colorings

**Idea.** Use gradient descent to train a feedforward network $g_\theta$ to minimize

$$L(\theta) = \int_{[-b,b]\times[-b,b]} \int_{B_1(x)} g_\theta(x) \cdot g_\theta(y)\, dy\, dx$$

for some reasonable $b \in \mathbb{R}$?

**Problem.** Still a continuous problem. How to compute?

**Stochastic (Batch) Gradient Descent.** Sample point $x^{(i)} \in [-b, b] \times [-b, b]$ and $y^{(i)} \in B_1(x)$ for $i = 1, \ldots, m$

# Can we improve the upper bound?

## Neural Networks as Colorings

**Idea.** Use gradient descent to train a feedforward network $g_\theta$ to minimize

$$L(\theta) = \int_{[-b,b]\times[-b,b]} \int_{B_1(x)} g_\theta(x) \cdot g_\theta(y)\, dy\, dx$$

for some reasonable $b \in \mathbb{R}$?

**Problem.** Still a continuous problem. How to compute?

**Stochastic (Batch) Gradient Descent.** Sample point $x^{(i)} \in [-b, b] \times [-b, b]$ and $y^{(i)} \in B_1(x)$ for $i = 1, \dots, m$ and use that

$$\nabla_\theta L(\theta) \approx \hat{\nabla}_\theta L(\theta) \doteq \frac{1}{m} \sum_{i=1}^{m} \nabla_\theta\, g_\theta(x^{(i)}) \cdot g_\theta(y^{(i)}),$$

# Can we improve the upper bound?

## Neural Networks as Colorings

**Idea.** Use gradient descent to train a feedforward network $g_\theta$ to minimize

$$L(\theta) = \int_{[-b,b]\times[-b,b]} \int_{B_1(x)} g_\theta(x) \cdot g_\theta(y)\, dy\, dx$$

for some reasonable $b \in \mathbb{R}$?

**Problem.** Still a continuous problem. How to compute?

**Stochastic (Batch) Gradient Descent.** Sample point $x^{(i)} \in [-b, b] \times [-b, b]$ and $y^{(i)} \in B_1(x)$ for $i = 1, \ldots, m$ and use that

$$\nabla_\theta L(\theta) \approx \hat{\nabla}_\theta L(\theta) \doteq \frac{1}{m} \sum_{i=1}^{m} \nabla_\theta\, g_\theta(x^{(i)}) \cdot g_\theta(y^{(i)}),$$

where $\nabla_\theta\, g_\theta(x^{(i)}) \cdot g_\theta(y^{(i)})$ is easily computed through backpropagation,

# Can we improve the upper bound?

## Neural Networks as Colorings

**Idea.** Use gradient descent to train a feedforward network $g_\theta$ to minimize

$$L(\theta) = \int_{[-b,b]\times[-b,b]} \int_{B_1(x)} g_\theta(x) \cdot g_\theta(y)\, dy\, dx$$

for some reasonable $b \in \mathbb{R}$?

**Problem.** Still a continuous problem. How to compute?

**Stochastic (Batch) Gradient Descent.** Sample point $x^{(i)} \in [-b,b] \times [-b,b]$ and $y^{(i)} \in B_1(x)$ for $i = 1, \ldots, m$ and use that

$$\nabla_\theta L(\theta) \approx \hat{\nabla}_\theta L(\theta) \doteq \frac{1}{m} \sum_{i=1}^{m} \nabla_\theta\, g_\theta(x^{(i)}) \cdot g_\theta(y^{(i)}),$$

where $\nabla_\theta\, g_\theta(x^{(i)}) \cdot g_\theta(y^{(i)})$ is easily computed through backpropagation, to adjust the parameters $\theta$ with an appropriate step size $\alpha_k$ through

$$\theta_{k+1} = \theta_k - \alpha_k\, \hat{\nabla}_\theta L(\theta).$$

**⇒ Very flexible approach "Deep Annealing"**
(also: tropicalization of loss function aka softmax... "minimize the max")

# Unfortunately this coloring was already known...

## Neural Networks as Colorings

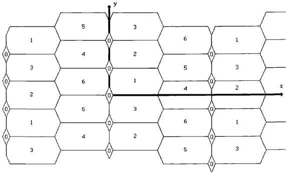
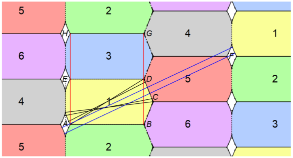
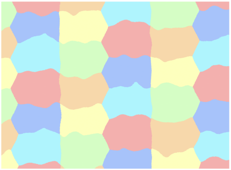


FIG. 3. A good 7-coloring of $(\mathbb{R}^2, 1)$.



### Theorem

*99.985% of the plane can be colored with 6 colors such that no two points of the same color are a unit distance apart.*

[Pritikin, 1998, Parts, 2020]

# Unfortunately this coloring was already known...

## Neural Networks as Colorings

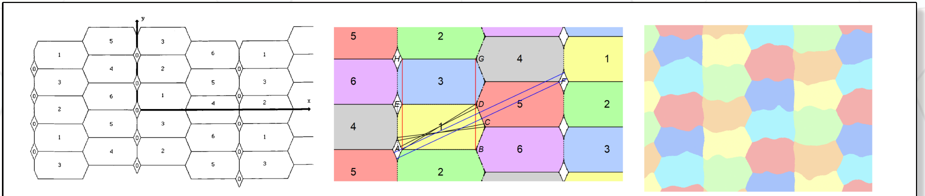


FIG. 3. A good 7-coloring of $(\mathbb{R}^2, 1)$.

**Theorem**

*99.985% of the plane can be colored with 6 colors such that no two points of the same color are a unit distance apart.* [Pritikin, 1998, Parts, 2020]

**Corollary**

*Any unit distance graph with chromatic number 7 must have at least 6 992 vertices.*

⇒ **While coloring was known already maybe on the right track?**

# Off-diagonal variant

# Going off-diagonal

## Neural Networks as Colorings

If we cannot solve the original problem, we study variants of it:

If we cannot solve the original problem, we study variants of it:

A $c$-coloring of the plane has **type** $(d_1, \ldots, d_c)$
if color $i$ does not contain any points at distance $d_i$.

# Going off-diagonal

## Neural Networks as Colorings

If we cannot solve the original problem, we study variants of it:

A $c$-coloring of the plane has **type** $(d_1, \ldots, d_c)$
if color $i$ does not contain any points at distance $d_i$.

Problem (Soifer in Nash and Rassias' *Open Problems in Mathematics*)

*Determine the continuum of six-colorings* $X_6 = \{d \mid (1, 1, 1, 1, 1, d)$ *can be realized*$\}$.

[Soifer, 1994a, Nash and Rassias, 2016]

If we cannot solve the original problem, we study variants of it:

A $c$-coloring of the plane has **type** $(d_1, \ldots, d_c)$
if color $i$ does not contain any points at distance $d_i$.

**Problem (Soifer in Nash and Rassias' *Open Problems in Mathematics*)**

*Determine the continuum of six-colorings* $X_6 = \{d \mid (1, 1, 1, 1, 1, d)$ *can be realized*$\}$.

[Soifer, 1994a, Nash and Rassias, 2016]

**Status.** Six-colorings exist for:

1. $d = 1/\sqrt{5}$ [Soifer, 1992]
2. $d = \sqrt{2} - 1$ [Hoffman and Soifer, 1993, 1996]
3. Part of family with $0.414 \approx \sqrt{2} - 1 \leq d \leq 1/\sqrt{5} \approx 0.447$ [Hoffman and Soifer, 1996, Soifer, 1994b, 2009]

# Going off-diagonal

## Neural Networks as Colorings

If we cannot solve the original problem, we study variants of it:

A $c$-coloring of the plane has **type** $(d_1, \ldots, d_c)$
if color $i$ does not contain any points at distance $d_i$.

**Problem (Soifer in Nash and Rassias' *Open Problems in Mathematics*)**

*Determine the continuum of six-colorings* $X_6 = \{d \mid (1, 1, 1, 1, 1, d) \text{ can be realized}\}$.

[Soifer, 1994a, Nash and Rassias, 2016]

**Status.** Six-colorings exist for:

1. $d = 1/\sqrt{5}$ [Soifer, 1992]
2. $d = \sqrt{2} - 1$ [Hoffman and Soifer, 1993, 1996]
3. Part of family with $0.414 \approx \sqrt{2} - 1 \leq d \leq 1/\sqrt{5} \approx 0.447$ [Hoffman and Soifer, 1996, Soifer, 1994b, 2009]

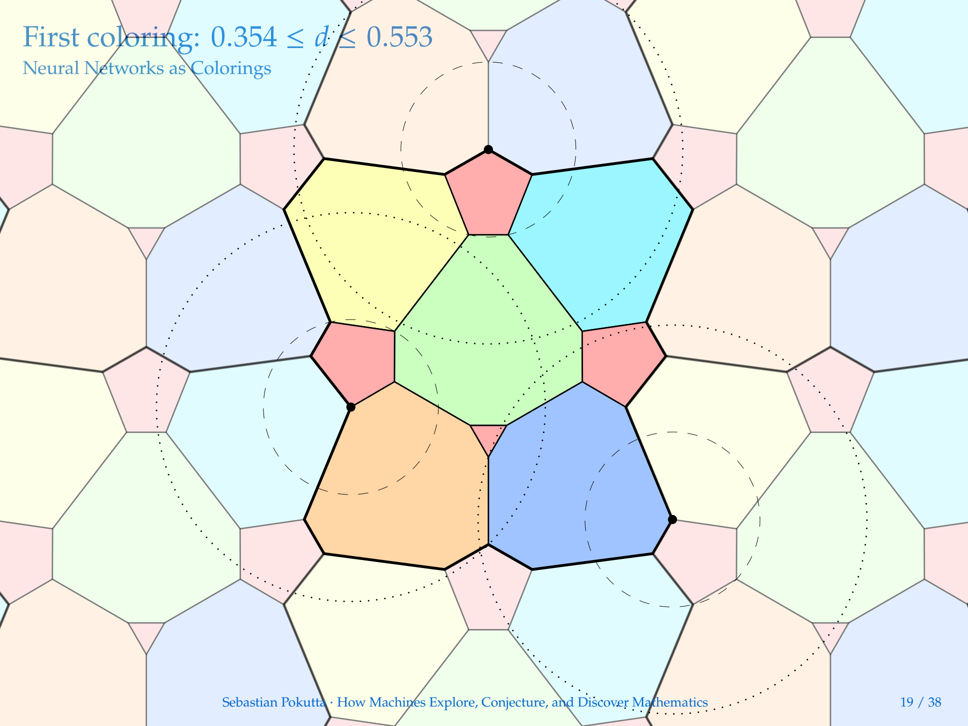
**Deep Annealing approach provides two new colorings leading to...**

**Theorem**

*$X_6$ contains the closed interval [0.354, 0.657].* [Mundinger et al., 2024, 2025]

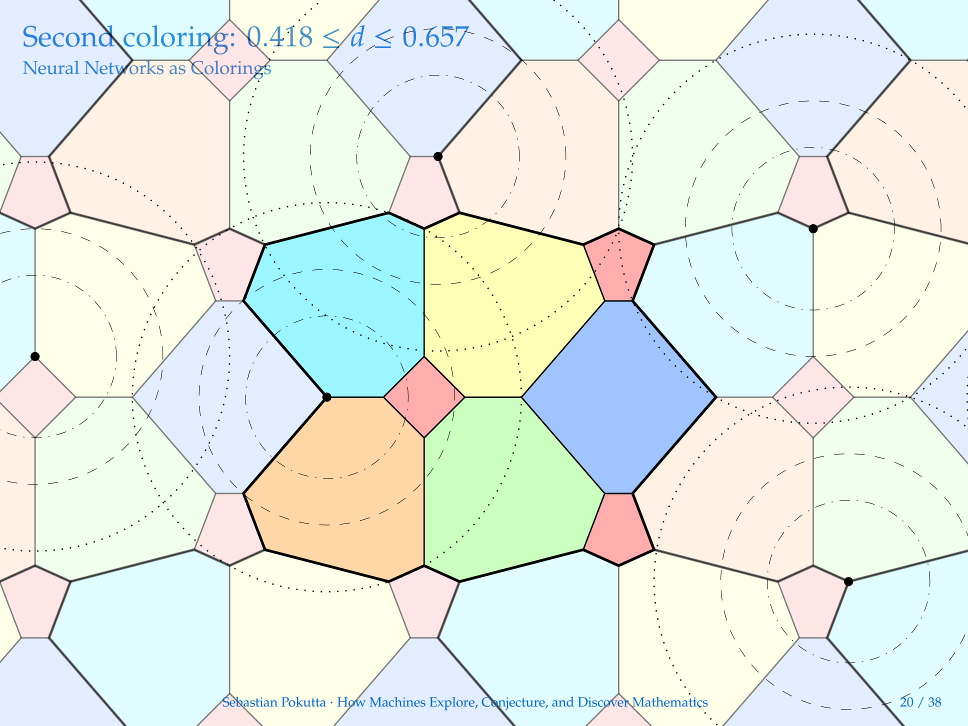
# First coloring: $0.354 \leq d \leq 0.553$

Neural Networks as Colorings



# Second coloring: $0.418 \leq d \leq 0.657$

Neural Networks as Colorings



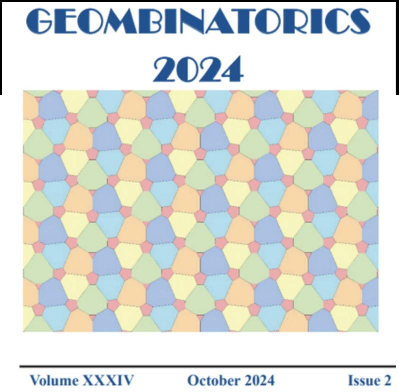
**GEOMBINATORICS QUARTERLY**

**Welcome to the journal on open problems of combinatorial & discrete geometry and related areas**

**Geombinatorics**

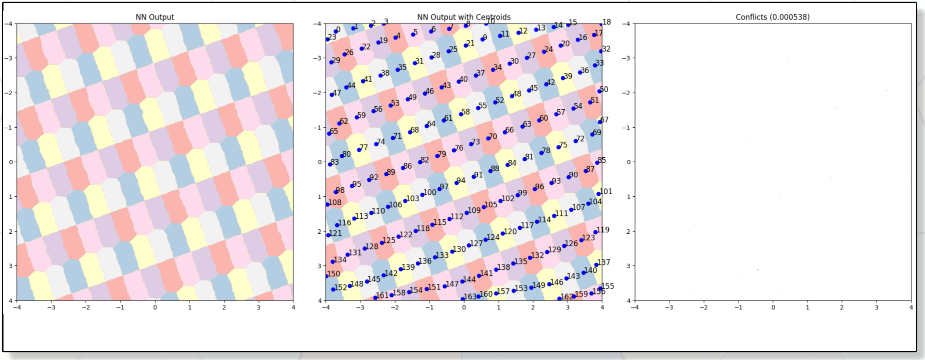
is a quarterly scientific journal of mathematics. It was established by editor-in-chief Alexander Soifer in 1991 and is published by the University of Colorado at Colorado Springs. The journal covers problems in discrete, convex, and combinatorial geometry, as well as related areas.

**CURRENT ISSUE** **SUBSCRIBE**



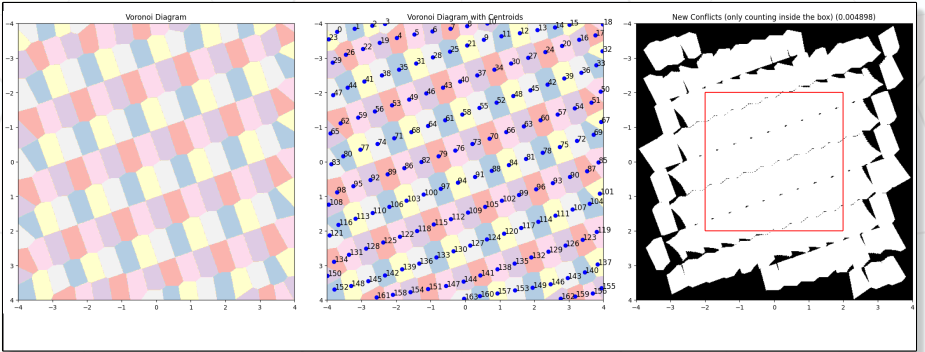
# Just numerics...?

## Neural Networks as Colorings



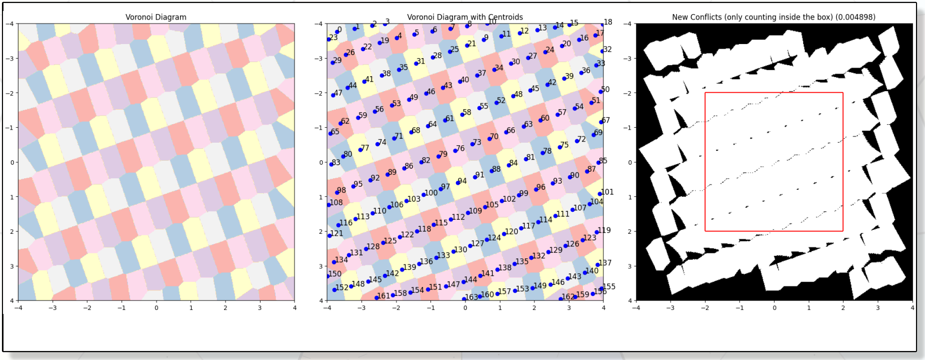
# Just numerics...?

## Neural Networks as Colorings



# Just numerics...?

## Neural Networks as Colorings

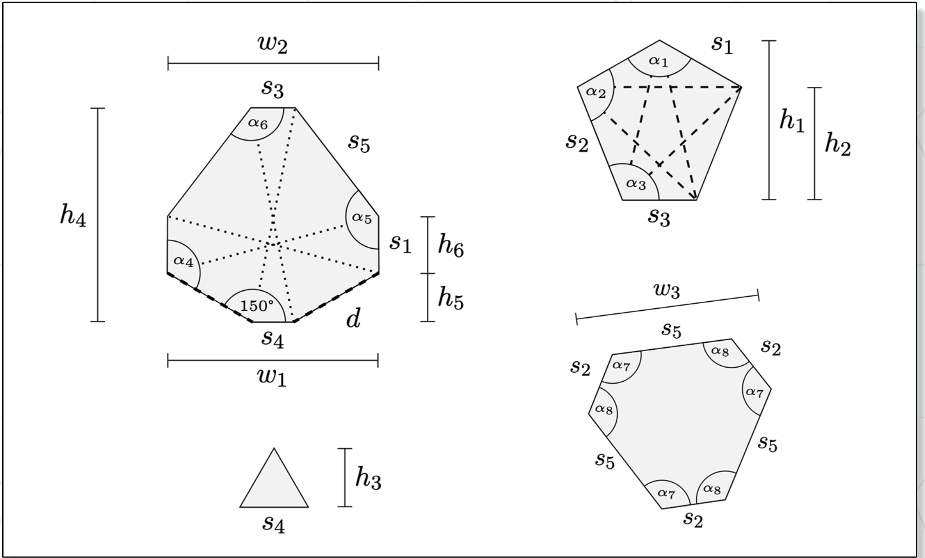



Voronoi cell filtering...

⇒ Exact constructions for both colorings.

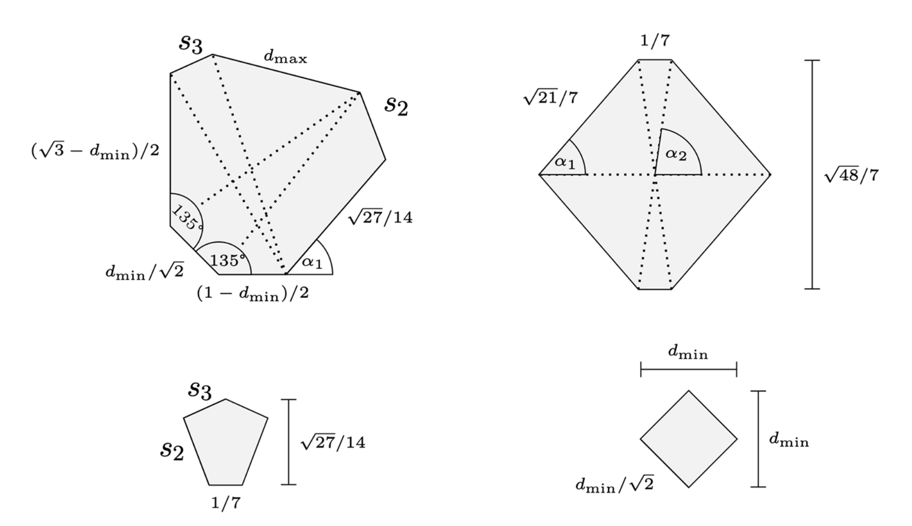
# First coloring: exact components

## Neural Networks as Colorings



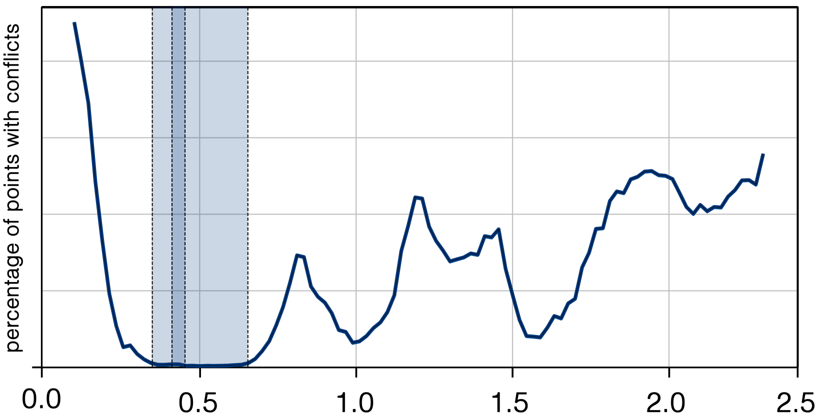
# Second coloring: exact components

## Neural Networks as Colorings



# Is this optimal?

## Neural Networks as Colorings




Numerical results showing the percentage of points with some conflict for a given forbidden distance $d$ in the sixth color found over several runs.

# 121 Patchworked Curves of Degree Seven

joint work with: Zoe Geiselmann, Michael Joswig, Lars Kastner,
Konrad Mundinger, Christoph Spiegel,
Marcel Wack, and Max Zimmer

Preprint (Sneak Peak)
`https://arxiv.org/abs/2602.06888`

Partially supported by ExC MATH+ Project EF-LiOpt-1
Neural Generative Models for Algebraic Curves

[Geiselmann et al., 2026]

In 1900, Hilbert formulated his famous 16th problem:

*Provide a topological classification of real plane algebraic curves of arbitrary degree d.*

[Hilbert, 1900]

In 1900, Hilbert formulated his famous 16th problem:

> *Provide a topological classification of real plane algebraic curves of arbitrary degree $d$.*

[Hilbert, 1900]

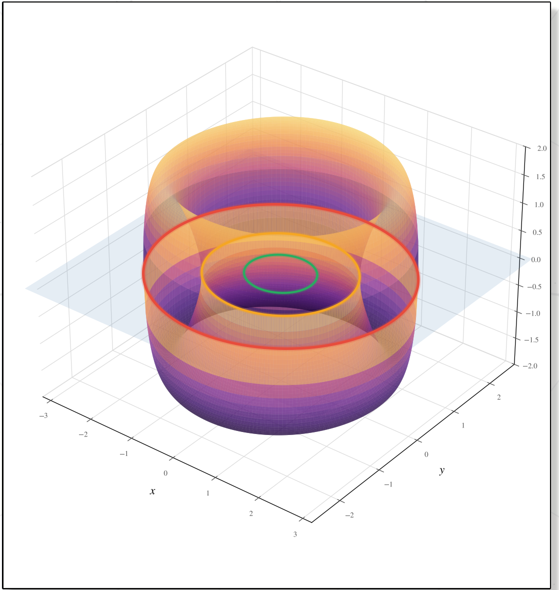
**Example.** For degree $d = 6$, the equation

$$z^2 = -\left[\left(\frac{x^2}{a_1^2} + \frac{y^2}{b_1^2} - 1\right)\left(\frac{x^2}{a_2^2} + \frac{y^2}{b_2^2} - 1\right)\left(\frac{x^2}{a_3^2} + \frac{y^2}{b_3^2} - 1\right)\right]$$

defines a real algebraic curve in the projective plane.

# Hilbert's 16th Problem

## Topological Characterization of Curves



# Current State-of-the-Art

## Topological Characterization of Curves

1. $d = 6, 7$: resolved in 1970s and 1980s [Rohlin, 1978, Nikulin, 1980, Viro, 1984, Gudkov, 1974]
2. $d = 8$: mostly but not entirely resolved [Orevkov, 2002]
3. $d = 9$: mostly fragmented [Orevkov, 2010]

# Current State-of-the-Art

## Topological Characterization of Curves

1. $d = 6, 7$: resolved in 1970s and 1980s [Rohlin, 1978, Nikulin, 1980, Viro, 1984, Gudkov, 1974]
2. $d = 8$: mostly but not entirely resolved [Orevkov, 2002]
3. $d = 9$: mostly fragmented [Orevkov, 2010]

**Required.** (a) Establishing existence of curves of certain types and (b) Ruling out non-types.

# Current State-of-the-Art

## Topological Characterization of Curves

1. $d = 6, 7$: resolved in 1970s and 1980s [Rohlin, 1978, Nikulin, 1980, Viro, 1984, Gudkov, 1974]
2. $d = 8$: mostly but not entirely resolved [Orevkov, 2002]
3. $d = 9$: mostly fragmented [Orevkov, 2010]

**Required.** (a) Establishing existence of curves of certain types and (b) Ruling out non-types.

**Viro's Patchworking** method allows for explicit construction of curves → T-curves.

[Viro, 1986]

# Current State-of-the-Art

## Topological Characterization of Curves

1. $d = 6, 7$: resolved in 1970s and 1980s [Rohlin, 1978, Nikulin, 1980, Viro, 1984, Gudkov, 1974]
2. $d = 8$: mostly but not entirely resolved [Orevkov, 2002]
3. $d = 9$: mostly fragmented [Orevkov, 2010]

**Required.** (a) Establishing existence of curves of certain types and (b) Ruling out non-types.

**Viro's Patchworking** method allows for explicit construction of curves → T-curves.

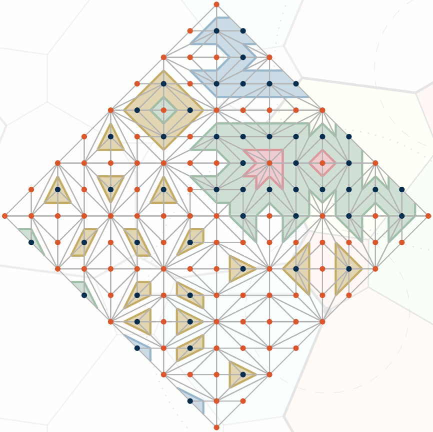
[Viro, 1986]

**Itenberg and Viro.** *"All real schemes of curves of degree $< 6$ and almost all real schemes of curves of degree 7 have been realized by the patchwork construction described above. On the other hand, there exist real schemes realizable by algebraic curves of some (high) degree, but not realizable by T-curves of the same degree. Probably such a scheme can be found even for degree 7 or 8."*

[Itenberg and Viro, 1996]

# Viro's Patchworking

## Examples in degree 8



(a) $\langle 17 \sqcup 1\langle 2\rangle \sqcup 1\langle 1\rangle\rangle$ (b) $\langle 16 \sqcup 3\langle 1\rangle\rangle$

**Note.** Curve is defined by a (regular) triangulation and a signing with ±1 at the nodes; actual polynomial obtained via lifting.

[Geiselmann et al., 2026]

## Theorem (Informal)

*All real schemes of curves of degree 7 can be realized by patchwork constructions.*

[Geiselmann et al., 2026]

## Theorem (Informal)

*All real schemes of curves of degree 7 can be realized by patchwork constructions. Moreover, 4 triangulations are sufficient to cover all types.*

# Answering Itenberg and Viro's Open Problem for Degree 7

## Our results

[Geiselmann et al., 2026]

### Theorem (Informal)

*All real schemes of curves of degree 7 can be realized by patchwork constructions. Moreover, 4 triangulations are sufficient to cover all types.*

### Theorem (Informal)

*All real schemes of curves of degree 6 can be realized by patchwork constructions with no more than two 2 triangulations.*

# Answering Itenberg and Viro's Open Problem for Degree 7

## Our results

[Geiselmann et al., 2026]

### Theorem (Informal)

*All real schemes of curves of degree 7 can be realized by patchwork constructions. Moreover, 4 triangulations are sufficient to cover all types.*

### Theorem (Informal)

*All real schemes of curves of degree 6 can be realized by patchwork constructions with no more than two 2 triangulations.*

Moreover,

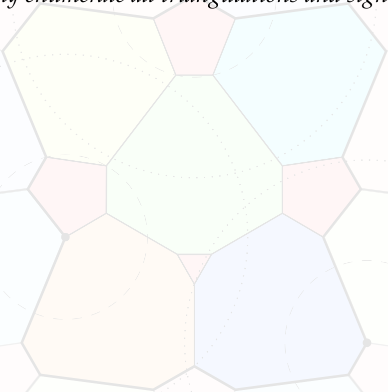
1. constructions are explicit
2. the lifting coefficients are very small
3. we provide explicit distributions over the realized types by triangulations
4. ... (see paper for more properties)

# What does this have to do with AI?

## Our results

In principle the problem is “trivial”:

*Simply enumerate all triangulations and signings.*



# What does this have to do with AI?

## Our results

In principle the problem is “trivial”:

*Simply enumerate all triangulations and signings.*

“Minor details”:

1. For a single triangulation there are $2^{36} = 68.719.476.736$ signings. At initially 20 type computations/sec $\approx$ 100 cpu-days for a single triangulation.
2. Then still need to “find” good triangulations.

# What does this have to do with AI?

## Our results

In principle the problem is “trivial”:

*Simply enumerate all triangulations and signings.*

“Minor details”:

1. For a single triangulation there are $2^{36}$ = 68.719.476.736 signings. At initially 20 type computations/sec ≈ 100 cpu-days for a single triangulation.
2. Then still need to “find” good triangulations.

⇒ **Basically impossible to search**

# What does this have to do with AI?

## Our results

In principle the problem is "trivial":

*Simply enumerate all triangulations and signings.*

"Minor details":

1. For a single triangulation there are $2^{36} = 68.719.476.736$ signings. At initially 20 type computations/sec $\approx$ 100 cpu-days for a single triangulation.
2. Then still need to "find" good triangulations.

**$\Rightarrow$ Basically impossible to search**

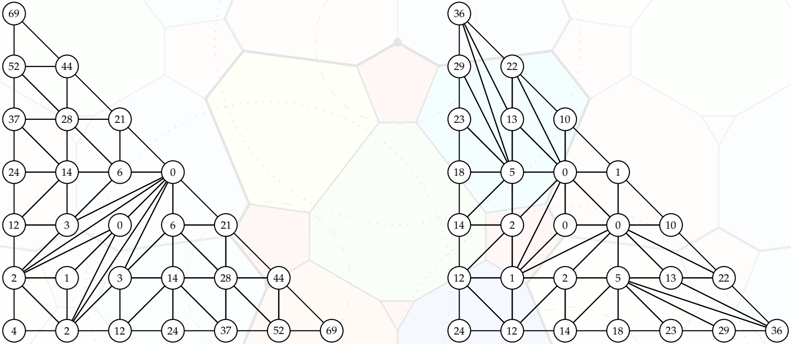
Our approach: combination of human + AI insights

1. New highly-optimized type algorithm developed leveraging AI, symmetry reduction, etc with final implementation in "near-assembler" Rust
$\Rightarrow \approx 4 \cdot 10^6$ type computations/sec or $\approx$ 4 hours for a single triangulation
2. (Many) intelligent guesses by humans for good triangulations

(more details soon...)

# Two Regular Triangulations to rule them all

## Characterization of curves of degree six

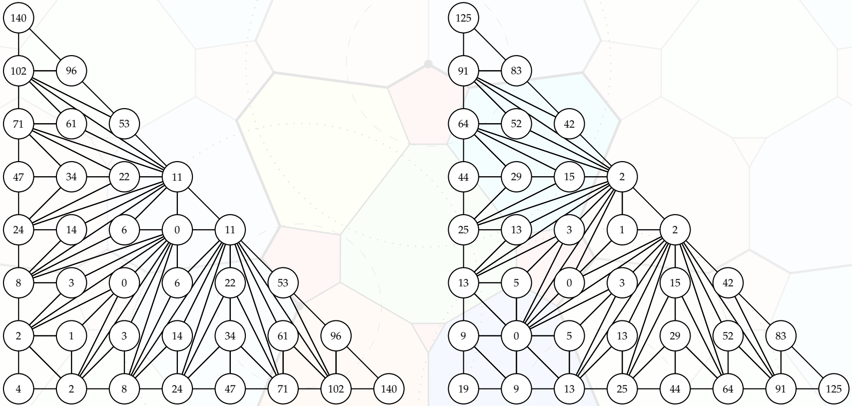



(a) bat (realizes 53 types) (b) moth (realizes 44 types)

Figure: Two regular triangulations of $6 \cdot \Delta_2$ realizing all nonempty real schemes types of degree six. Values at the vertices indicate lifting functions.

# Four Regular Triangulations to rule them all

## Characterization of curves of degree seven

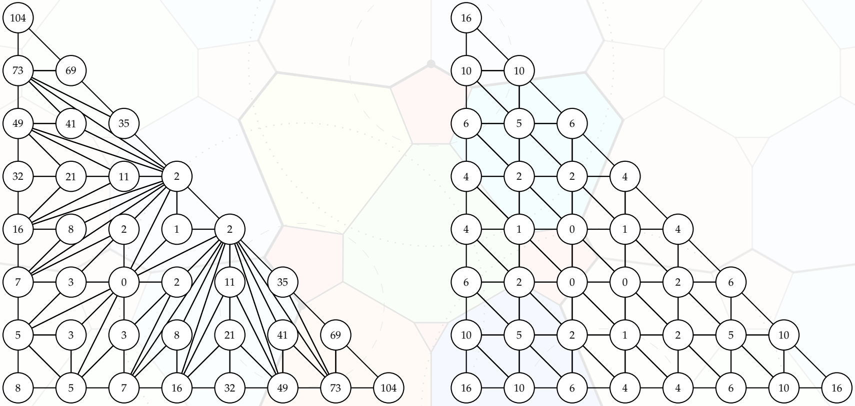



(a) cen (realizes 115 types) (b) spl (realizes 107 types)

Figure: The four regular triangulations of $7 \cdot \Delta_2$ realizing all real schemes types of degree seven. The values at the vertices indicate the respective lifting functions.

# Four Regular Triangulations to rule them all

## Characterization of curves of degree seven




(c) fra (realizes 103 types) (d) hon (realizes 47 types)

Figure: The four regular triangulations of $7 \cdot \Delta_2$ realizing all real schemes types of degree seven. The values at the vertices indicate the respective lifting functions.

# Final Remarks

1. AI can be used in various ways in modern mathematical research workflows (actual discovery, verification, etc); beyond simple black-box prompting
2. Fully-automatic discovery of new mathematics might be possible in the future but relies on *strong* verification approaches (LEAN might be too inefficient)
3. The agentic harness seems to be key: how does the agent receive feedback on its work, how is it guided, and which tools are available?
4. Empirically: the human-in-the-loop is crucial to guide the search

**The promises of AI4MATH are great** but need to go beyond simple black-box prompting "your favorite Erdős problem" (which then turns out having a solution that is already known)...

**Thank you!**

Mundinger, K., Pokutta, S., Spiegel, C., and Zimmer, M. (2024). Extending the Continuum of Six-Colorings. Geombinatorics Quarterly. Available at https://arxiv.org/abs/2404.05509.

Mundinger, K., Zimmer, M., Kiem, A., Spiegel, C., and Pokutta, S. (2025). Neural Discovery in Mathematics: Do Machines Dream of Colored Planes? Proceedings of the 42nd International Conference on Machine Learning (ICML), 267, 45236–45255. Available at https://arxiv.org/abs/2501.18527.

Geiselmann, Z., Joswig, M., Kastner, L., Mundinger, K., Pokutta, S., Spiegel, C., Wack, M., and Zimmer, M. (2026). 121 Patchworked Curves of Degree Seven. Preprint. Available at https://arxiv.org/abs/2602.06888.

# References I


K. I. Appel and W. Haken. Every planar map is four-colorable. *Illinois Journal of Mathematics*, 21(3):429–490, 1977. Computer-assisted proof of the four color theorem (first major such proof).

N. d. Bruijn and P. Erdos. A colour problem for infinite graphs and a problem in the theory of relations. *Indigationes Mathematicae*, 13:371–373, 1951.

A. D. De Grey. The chromatic number of the plane is at least 5. *arXiv preprint arXiv:1804.02385*, 2018.

G. Exoo and D. Ismailescu. The chromatic number of the plane is at least 5: a new proof. *Discrete & Computational Geometry*, 64(1):216–226, 2020.

Z. Geiselmann, M. Joswig, L. Kastner, K. Mundinger, S. Pokutta, C. Spiegel, M. Wack, and M. Zimmer. Patchworked Curves of Degree Seven. *preprint*, 2 2026. doi: 10.48550/arXiv.2602.06888.

D. A. Gudkov. The topology of real projective algebraic varieties. *Russ. Math. Surv.*, 29(4):1–79, 1974. ISSN 0036-0279. doi: 10.1070/RM1974v029n04ABEH001288.

J. Haase and S. Pokutta. Human–AI Cocreativity: Exploring synergies across levels of creative collaboration. In J. C. Kaufman and M. Worwood, editors, *Generative Artificial Intelligence and Creativity*, chapter 16, pages 205–221. 1 2026. doi: 10.1016/B978-0-443-34073-4.00009-5.

T. C. Hales, M. Adams, G. Bauer, D. T. Dang, J. Harrison, H. L. Truong, C. Kaliszyk, V. Magron, S. McLaughlin, N. T. Thang, N. Q. Truong, T. Nipkow, S. Obua, J. Pleso, J. Rute, A. Solovyev, et al. A formal proof of the kepler conjecture. *Forum of Mathematics, Pi*, 5:e2, 2017. doi: 10.1017/fmp.2017.1. Flyspeck project: complete formal verification of the Kepler conjecture.

M. J. H. Heule, O. Kullmann, and V. W. Marek. Solving and verifying the boolean pythagorean triples problem via cube-and-conquer. In *Theory and Applications of Satisfiability Testing – SAT 2016*, volume 9710 of *Lecture Notes in Computer Science*, pages 228–245. Springer, 2016. Computer-intensive proof generating a multi-terabyte certificate.

D. Hilbert. Mathematische Probleme. *Nachr. Ges. Wiss. Göttingen Math.-Phys. Kl.*, 1900:253–297, 1900. URL `http://eudml.org/doc/58479`. English translation (M. F. Winston Newson): Bull. Amer. Math. Soc. 8 (1902), 437–479.

I. Hoffman and A. Soifer. Almost chromatic number of the plane. *Geombinatorics*, 3(2):38–40, 1993.

I. Hoffman and A. Soifer. Another six-coloring of the plane. *Discrete Mathematics*, 150(1-3):427–429, 1996.

I. Itenberg and O. Viro. Patchworking algebraic curves disproves the Ragsdale conjecture. *Math. Intelligencer*, 18(4):19–28, 1996. ISSN 0343-6993,1866-7414. doi: 10.1007/BF03026748.

D. Mixon. Polymath16, seventeenth thread: Declaring victory. *Polymath16*, February 1 2021. Retrieved 16 August 2021.

L. Moser and M. Moser. Solution to problem 10. *Canadian Mathematical Bulletin*, 4:187–189, 1961.

K. Mundinger, S. Pokutta, C. Spiegel, and M. Zimmer. Extending the Continuum of Six-Colorings. *Geombinatorics Quarterly*, 5 2024.

K. Mundinger, M. Zimmer, A. Kiem, C. Spiegel, and S. Pokutta. Neural Discovery in Mathematics: Do Machines Dream of Colored Planes? *Proceedings of the 42nd International Conference on Machine Learning (ICML)*, 267:45236–45255, 5 2025. ISSN 2640-3498. doi: 10.48550/arXiv.2501.18527.

J. F. Nash and M. T. Rassias. *Open problems in mathematics*. Springer, 2016.

# References II


V. V. Nikulin. Integer symmetric bilinear forms and some of their geometric applications. *Izv. Math.*, 14(1):103–167, 1980.

S. Y. Orevkov. Classification of flexible $M$-curves of degree 8 up to isotopy. *Geom. Funct. Anal.*, 12(4):723–755, 2002. ISSN 1016-443X,1420-8970. doi: 10.1007/s00039-002-8264-6. URL `https://doi.org/10.1007/s00039-002-8264-6`.

S. Y. Orevkov. Complex orientation formulas for $M$-curves of degree $4d+1$ with 4 nests. *Ann. Fac. Sci. Toulouse Math. (6)*, 19(1):13–26, 2010. ISSN 0240-2963,2258-7519. URL `http://afst.cedram.org/item?id=AFST_2010_6_19_1_13_0`.

J. Parts. What percent of the plane can be properly 5-and 6-colored? *arXiv preprint arXiv:2010.12668*, 2020.

D. Pritikin. All unit-distance graphs of order 6197 are 6-colorable. *Journal of Combinatorial Theory, Series B*, 73(2):159–163, 1998.

N. Robertson, D. P. Sanders, P. Seymour, and R. Thomas. The four colour theorem. *Journal of Combinatorial Theory, Series B*, 70(1):2–44, 1997. Shorter and fully computer-verified proof reducing the number of configurations.

V. A. Rohlin. Complex topological characteristics of real algebraic curves. *Uspekhi Mat. Nauk*, 33(5(203)):77–89, 237, 1978. ISSN 0042-1316.

A. Soifer. A six-coloring of the plane. *Journal of Combinatorial Theory, Series A*, 61(2):292–294, 1992.

A. Soifer. Six-realizable set $x_6$. *Geombinatorics*, III(4):140–145, 1994a.

A. Soifer. An infinite class of six-colorings of the plane. *Congressus Numerantium*, pages 83–86, 1994b.

A. Soifer. *The mathematical coloring book: Mathematics of coloring and the colorful life of its creators*. Springer, 2009.

A. Soifer. *The New Mathematical Coloring Book*. Springer, 2024.

O. Y. Viro. Gluing of plane real algebraic curves and constructions of curves of degrees 6 and 7. In *Topology (Leningrad, 1982)*, volume 1060 of *Lecture Notes in Math.*, pages 187–200. Springer, Berlin, 1984. ISBN 3-540-13337-2. doi: 10.1007/BFb0099934. URL `https://doi.org/10.1007/BFb0099934`.

O. Y. Viro. Progress in the topology of real algebraic varieties in the last six years. *Uspekhi Mat. Nauk*, 41(3(249)):45–67, 240, 1986. ISSN 0042-1316.