Emily Riehl

Johns Hopkins University

A new paradigm for mathematical proof?

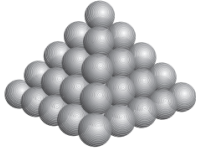
Natural Philosophy Symposium

Recent developments in mathematics: problem solving

In 1998, Thomas Hales announced a proof of a 1611 conjecture of Johannes Kepler, via a "proof by exhaustion" involving the checking of many individual cases using a computer to solve linear programming problems. After four years, a panel of 12 referees reported they were 99% certain that the proof was correct, but could not check all the computer calculations.

THEOREM 1.1 (The Kepler conjecture). *No packing of congruent balls in Euclidean three space has density greater than that of the face-centered cubic packing.*

This density is $\pi/\sqrt{18} \approx 0.74$.



The unabridged version of the paper, which was published in 2005 in the *Annals of Mathematics*, came to 339 pages, with around 3 gigabytes of computer artifacts.

Recent developments in mathematics: problem solving

In 2016, Maryna Viazovska, resolved the analogous question in eight dimensions, proving that the optimal packing of spheres in 8-dimensional spaces positions their centers on the points of the E_8-lattice.

A week later, she and four collaborators generalized these results to dimension 24.

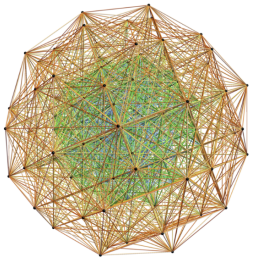
Annals of Mathematics **185** (2017), 991–1015
https://doi.org/10.4007/annals.2017.185.3.7

The sphere packing problem in dimension 8

By Maryna S. Viazovska

Abstract

In this paper we prove that no packing of unit balls in Euclidean space $\mathbb{R}^8$ has density greater than that of the E_8-lattice packing.



Visualization by J.G. Moxness .

Recent developments in mathematics: theory building

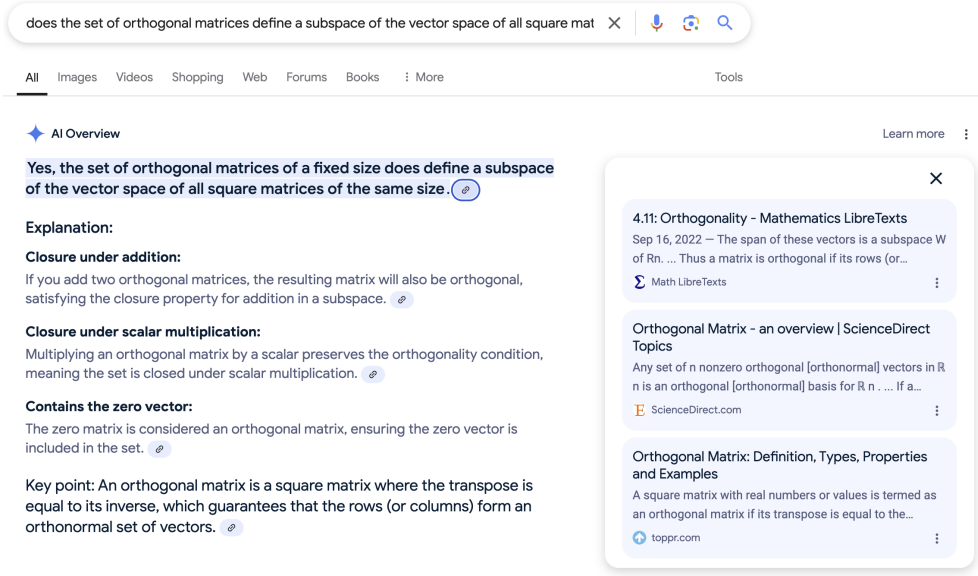
In "The two cultures of mathematics" (the title alluding to C. P. Snow), Tim Gowers makes a distinction between:

> "mathematicians who regard their central aim as being to solve problems, and those who are more concerned with building and understanding theories."

Some of the new ideas in the paper 356 paper of "Geometrization the local Langlands correspondence" by Laurent Fargues and Peter Scholze, to appear in *Asterisque*, involve a new framework to study objects that are both algebraic and topological in nature.

This framework, called condensed mathematics, has been developed over the course of several hundred pages of lecture notes by Dustin Clausen and Peter Scholze, which have been widely studied but remain unpublished.

A human-computer mathematical interaction



Do orthogonal matrices form a subspace?

My student asked Gemini:

"Does the set of orthogonal matrices define a subspace of the vector space of all square matrices?"

In the 2×2 case, orthogonal matrices have the form

$$\begin{bmatrix} s & -t \\ t & s \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} s & t \\ t & -s \end{bmatrix} \quad \text{with} \quad s^2 + t^2 = 1,$$

while square matrices have the form $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ for any real numbers a, b, c, and d.

The set of 2×2 matrices forms a vector space because you can add them and multiply them by a real number to get another 2×2 matrix.

This question asks whether the sum or scalar multiple of orthogonal matrices is again orthogonal and whether the zero matrix is an orthogonal matrix.

Gemini's illogical "thinking"

"Does the set of orthogonal matrices define a subspace of the vector space of all square matrices?"

Gemini answered "Yes" with the following explanation:

- "If you add two orthogonal matrices, the resulting matrix will also be orthogonal, satisfying the closure property for addition in a subspace."

  This is incorrect.

- "Multiplying an orthogonal matrix by a scalar preserves the orthogonality condition, meaning the set is closed under scalar multiplication."

  This is incorrect.

- "The zero matrix is considered an orthogonal matrix, ensuring the zero vector is included in the set."

  This is incorrect.

Despite the fact that Gemini gave correct definitions of subspace and of orthogonal matrices (as those matrices whose transpose is the inverse), it answered wrong.

Vibe proving

Is it true that an ∞-category is idempotent complete (in the ∞-categorical sense) if and only if its homotopy category is idempotent complete (in the 1-categorical sense)?

Show thinking

Yes, it is true that an ∞-category $\mathcal{C}$ is idempotent complete (in the ∞-categorical sense) if and only if its homotopy category $ho(\mathcal{C})$ is idempotent complete (in the 1-categorical sense).

Let's establish the definitions and then prove the equivalence.

...

Conclusion:

The two conditions are equivalent. The definitions of idempotent completeness are "homotopy-invariant" in a way that translates directly between the ∞-category and its homotopy category. This result is also confirmed by the fact that $h(Kar(\mathcal{C})) \cong Kar(h\mathcal{C})$, where $Kar(-)$ denotes the Karoubi envelope (idempotent completion) and $h(-)$ denotes taking the homotopy category. An (∞)-category is idempotent complete if and only if the canonical functor into its Karoubi envelope is an equivalence.

Your answer uses the "incoherent" definition of an idempotent in an ∞-category? Can you answer the same question using the coherent definition of an idempotent in an ∞-category?

Show thinking

You're right to press on the subtleties of "coherence" in ∞-categories. Let's clarify how these definitions are typically handled and why the previous conclusion still holds.

...

If a framework requires a more elaborate "coherent idempotent object" (e.g., an object in a category of idempotents, or an algebra for a certain monad), the definition of "idempotent completeness of $\mathcal{C}$" itself usually still boils down to the splitting condition for any e satisfying $e \circ e \simeq e$. The power of ∞-category theory is that such "weak" conditions, when interpreted within a coherently structured environment, often suffice and lead to robust theories.

Thus, the answer remains the same: **Yes, an ∞-category is idempotent complete (in the standard ∞-categorical sense) if and only if its homotopy category is idempotent complete (in the 1-categorical sense).** The "coherence" is sufficiently handled by the definitions used.

The omitted text contains a clearly written argument with a subtle error, the sort of optimistic mistake often made by human mathematicians. The conclusion is incorrect.

How do we defend the mathematical literature against vibe proving?

A lot of the mathematical text generated by top "reasoning" models is pretty good and all of it looks good.*

*Caveat: large language models tend to get even the simplest numerical calculations horribly wrong.

But pretty good is not good enough: one minor error could make an entire logical argument reach the wrong conclusion: true instead of false.

As a journal editor specializing in a subfield where papers tend to be long, technical, and difficult to referee, I am extremely concerned about AI generated mathematical "proofs."

Proof as peer review

Why has mathematics largely avoided the replication crisis
that has confronted other fields?

Peer review in theory: careful refereeing should lead to an error-free publications.

Unfortunately, the mathematical literature contains famous mistakes as well as contradictory theorems.

Peer review in practice: in theory any proof should be reproducible by any reader — allowing the reader to understand for themselves why the result is true.

When papers have enough readers, mistakes are eventually caught.

Importantly: human mathematicians are careful in claiming they have a proof.

Proof as peer review?

But sometimes these ideals break down:

One Fields medalist was dismayed to find mistakes in his published, well-studied papers:

> *"A technical argument by a trusted author, which is hard to check and looks similar to arguments known to be correct, is hardly ever checked in detail."*
>
> *— Vladimir Voevodsky*

Another Fields medalist expressed doubts about a particular proof he had discovered — and also doubted that anyone else would check it:

> *"...while I was very happy to see many study groups on condensed mathematics throughout the world, to my knowledge all of them have stopped short of this proof. (Yes, this proof is not much fun...)"*
>
> *— Peter Scholze*

A new paradigm for mathematical proof?

THE EQUIVARIANT MODEL STRUCTURE ON CARTESIAN CUBICAL SETS

STEVE AWODEY, EVAN CAVALLO, THIERRY COQUAND, EMILY RIEHL, AND CHRISTIAN SATTLER

Abstract. We develop a constructive model of homotopy type theory in a Quillen model category that classically presents the usual homotopy theory of spaces. Our model is based on presheaves over the cartesian cube category, a well-behaved Eilenberg–Zilber category. The key innovation is an additional equivariance condition in the specification of the cubical Kan fibrations, which can be described as the pullback of an interval-based class of uniform fibrations in the category of symmetric sequences of cubical sets. The main technical results in the development of our model have been formalized in a computer proof assistant.

Contents

Date: June 27, 2024.

Software programs called computer proof assistants can certify the correctness of a mathematical proof that has been written in a precise formal language.

- Today such proofs are laboriously encoded by human mathematicians (formalization).
- In principle, generative AI could be trained to output text in a format that could be checked by a computer proof assistant (autoformalization).

Computer proof verification

```
{-
------------------------------------------------------------------------------------------
Formalization of an equivariant cartesian cubical set model of type theory
------------------------------------------------------------------------------------------

This formalization accompanies the article

  The equivariant model structure on cartesian cubical sets.
  Steve Awodey, Evan Cavallo, Thierry Coquand, Emily Riehl, & Christian Sattler.
  https://arxiv.org/abs/2406.18497

The contents of the formalization are outlined in Appendix A of the article.

The formalization defines a model of homotopy type theory inside an extensional type
theory augmented with a flat modality and axioms postulating *shapes* (among them an
*interval*) and a cofibration classifier. The results can in particular be externalized in
the category of cartesian cubical sets.

The code has been tested with Agda version 2.6.4.
The source is available at

  github.com/ecavallo/equivariant-cartesian

and there is an HTML interface at

  ecavallo.github.io/equivariant-cartesian

For reference (see the file equivariant.agda-lib in the source), the formalization is
compiled with the flags

  --with-K
  --cohesion --flat-split
  --no-import-sorts
  --rewriting

In particular, the --with-K flag enables axiom K (uniqueness of identity proofs), while
the --cohesion and --flat-split flags enable the flat modality (see the module
axiom.flat for more information).

-}
```

The main definition takes just a few lines to encode ⇝

Our 87 page preprint is accompanied by a library of formalized proofs checked by the computer proof assistant Agda.

The paper, submitted to a journal in September, is still awaiting a referee report.

```
--↓ The equivariance condition on local filling structures associated to a shape
--↓ homomorphism σ : S → T. Filling an open box over T and then composing with σ should be
--↓ the same as composing the box with σ and then filling over S.

LocalEquivariance : {S T : Shape} (σ : Shape[ S , T ]) {A : ⟨ T ⟩ → Type ℓ}
  → LocalFillStr T A → LocalFillStr S (A ∘ ⟪ σ ⟫) → Type ℓ
LocalEquivariance σ liftT liftS =
  ∀ r box s →
  reshapeFiller σ (liftT (⟪ σ ⟫ r) box) .fill s .out
  ≡ liftS r (reshapeBox σ box) .fill s .out

Equivariance : {S T : Shape} (σ : Shape[ S , T ]) {Γ : Type ℓ} (A : Γ → Type ℓ')
  → FillStr T A → FillStr S A → Type (ℓ ⊔ ℓ')
Equivariance {T = T} σ {Γ} A fillT fillS =
  (γ : Γ ^ T) → LocalEquivariance σ (fillT γ) (fillS (γ ∘ ⟪ σ ⟫))

--↓ Definition of an equivariant fibration structure.

record FibStr {Γ : Type ℓ} (A : Γ → Type ℓ') : Type (ℓ ⊔ ℓ') where
  constructor makeFib
  field
    --↓ We have a filling structure for every shape.

    lift : (S : Shape) → FillStr S A

    --↓ The filling structures satisfy the equivariance condition.

    vary : ∀ S T (σ : Shape[ S , T ]) → Equivariance σ A (lift T) (lift S)
```

A new paradigm for proof writing

Computer formalization is a new and not yet widely practiced method of developing and communicating rigorous mathematical proofs interactively through the use of a computer program called a computer proof assistant.

- The mathematician inputs each line of their proof in a precise syntax.
- The computer checks that the logical argument supplied by the user produces a valid deduction of the claimed mathematical statement.
- Depending on the sophistication of the computer program, it might also "assist" the mathematician in various ways:
  - By catching errors of reasoning (unjustified assumptions, missing cases, etc).
  - By keeping track of the current state of a complicated logical argument.
  - By suggesting or even automatically generating proofs (autoformalization).

Aside: modern proof assistants often use a newer formal system — dependent type theory — in place of traditional Zermelo-Fraenkel set theory and first order logic.

A new paradigm for proof checking

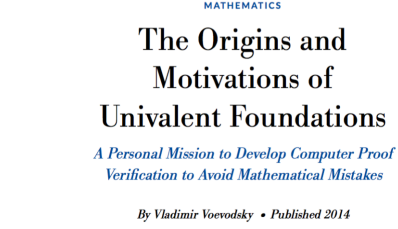
"A technical argument by a trusted author, which is hard to check and looks similar to arguments known to be correct, is hardly ever checked in detail."

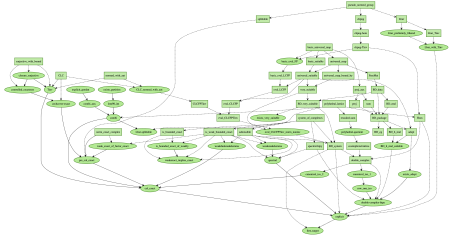
— Vladimir Voevodsky

"...while I was very happy to see many study groups on condensed mathematics throughout the world, to my knowledge all of them have stopped short of this proof. (Yes, this proof is not much fun...)"

— Peter Scholze

Voevodsky and Scholze both turned to computer formalization to resolve their doubts about the veracity of their own proofs.





Large scale computer-verified proofs

When human referees failed to fully certify his proof of the Kepler conjecture, Hales launched a project to verify the result himself in a computer proof assistant. Eleven years later, the full proof was formally verified in the proof assistants `Isabelle` and `HOL Light`. The formalization was described in an accompanying 29 page paper with 22 authors.

A FORMAL PROOF OF THE KEPLER CONJECTURE

THOMAS HALES[1], MARK ADAMS[2,3], GERTRUD BAUER[4], TAT DAT DANG[5], JOHN HARRISON[6], LE TRUONG HOANG[7], CEZARY KALISZYK[8], VICTOR MAGRON[9], SEAN MCLAUGHLIN[10], TAT THANG NGUYEN[7], QUANG TRUONG NGUYEN[1], TOBIAS NIPKOW[11], STEVEN OBUA[12], JOSEPH PLESO[13], JASON RUTE[14], ALEXEY SOLOVYEV[15], THI HOAI AN TA[7], NAM TRUNG TRAN[7], THI DIEP TRIEU[16], JOSEF URBAN[17], KY VU[18] and ROLAND ZUMKELLER[19]

Last year, Kevin Buzzard launched a project to verify a modern proof of Fermat's last theorem—that there are no positive integer solutions to the equation $x^n + y^n = z^n$ for $n \geq 3$—in the computer proof assistant `Lean`, motivated in part by the question: "is there any one person who completely understands a proof of Fermat's Last Theorem?"

Moral: proofs at the frontier of mathematics are formalizable
...but only with monumental human effort via large-scale collaborations.

Interactive theorem proving, in pursuit of greater rigour and clarity

The practice of explaining a mathematical proof to a computer requires absolute precision, in particular regarding the exact definitions of mathematical terms. In my experience at least, this level of pedantry is both deeply frustrating and unexpectedly seductive, making it easier to achieve and sustain a flow state of deep focus.

- There is no point in attempting to write anything if you aren't thinking perfectly clearly because it will be rejected by the computer proof assistant.
- Paradoxically, this much steeper demand of my attention makes it easier for me to achieve that level of focus.

The interactions with the computer proof assistant also activate reward mechanisms.

- During a typical research day, I make no quantifiable progress towards proving anything. But in the practice of formalization, the user periodically asks the proof assistant whether what is done so far is correct. When it says yes, this feels great.

Practitioners sometimes describe formalizing as a gamification of mathematical research.

There's more to mathematics than rigour and proofs

"Mathematics is the art of giving the same name to different things." — Henri Poincaré

"...mathematics may be viewed as the Science of Analogy." — Sir Michael Atiyah

But the demands of greater rigour enforced by formalization runs counter to the vision of mathematics presented by Poincaré, Atiyah, or a famous blog post of Terry Tao:

> *One can roughly divide mathematical education into three stages:*
>
> 1. *The "pre-rigorous" stage, in which mathematics is taught in an informal, intuitive manner, based on examples, fuzzy notions, and hand-waving ...The emphasis is more on computation than on theory ...*
> 2. *The "rigorous" stage, in which one is now taught that in order to do maths "properly", one needs to work and think in a much more precise and formal manner ...The emphasis is now primarily on theory; and one is expected to be able to comfortably manipulate abstract mathematical objects without focusing too much on what such objects actually "mean" ...*
> 3. *The "post-rigorous" stage, in which one has grown comfortable with all the rigorous foundations of one's chosen field, and is now ready to revisit and refine one's pre-rigorous intuition on the subject, but this time with the intuition solidly buttressed by rigorous theory ...The emphasis is now on applications, intuition, and the "big picture" ...*

A new paradigm for mathematical proof?

Is there consensus in the mathematical community
about the computer formalized proof paradigm?

Absolutely not! This is very much a minority point of view.

- I would estimate that most departments contain few if any mathematicians who are actively working with computer proof assistants.
- Even early adopters would agree that computer proof assistants are currently too hard to use for most day-to-day proof writing.
- The challenges presented by formalization vary tremendously between subfields, and their solutions may require new ideas — perhaps new domain-specific foundations?

In addition, there are active debates about how such a shift would affect
human understanding of mathematics, the real point of what we do.

A norm for machine-generated mathematical proof

Despite well-known imperfections, the mathematical community can take deep pride in our overwhelmingly reliable and continually improving standards for mathematical proof.

We should demand the same for AI when it comes to the mathematical realm.

Maintaining high standards will frustrate near term progress, delaying the arrival of a machine we validate as having "artificial mathematical intelligence," but should be beneficial for overall reliability in the long run, in mathematics and beyond.

Specifically, I want to propose the following norm for the mathematical community when it comes to original mathematics produced by an AI system:

Any artificially generated mathematical text will not be considered as a proof unless:

- It has been communicated in both a natural language text paired with a computer formalization of all definitions, theorems, and proofs.
- The formalization has been accepted by the proof assistant and human expert referees have vetted both the formalization and the paired text.