The future of mathematics?

K. Buzzard

Jan 2020, Pittsburgh

What is the future of mathematics?

- In the 1990s, computers became better than humans at chess.
- In 2018, computers became better than humans at go.
- In 2019, I met a guy from Google called Christian Szegedy.
- He told me that in 10 years' time, computers would be better than humans at finding proofs of mathematical theorems.
- Of course he might be wrong.
- What if he is right?
- (Szegedy link)

- Here is what I believe.
- In 10 years' time, computers will be helping some of us to prove tedious "early PhD student level" lemmas.
- In which areas of maths?
- That depends on who gets involved.
- Usual pattern with AI: at first, it won't be very good.
- Then all of a sudden it will get really good.
- Interesting question: when will the "all of a sudden it will get very good" bit happen?
- Nobody has a clue.
- The more people get involved, the quicker it will happen.

What is a proof?

- What does a bright undergraduate think that a pure mathematical proof is?
- What does a researcher in pure mathematics think that a proof is?
- What does a computer think that a mathematical proof is?

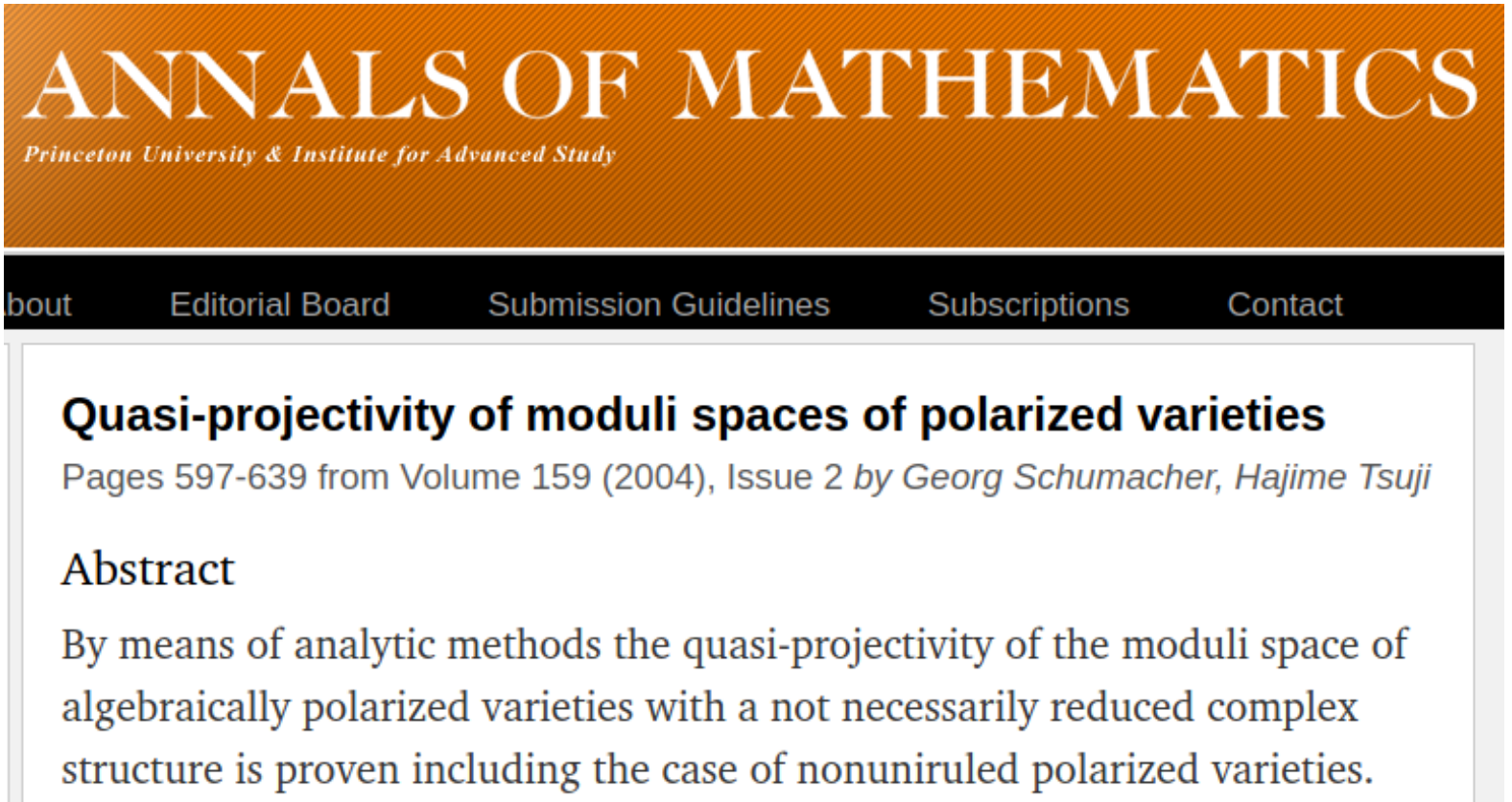
Answers: The bright undergraduate and the computer both think something like the following:

A proof is a logical sequence of statements, using the axioms of your system and the theorems you have already proved, which ultimately leads to a deduction of the statement you are trying to prove. The computer calls this idea “running a computer program”.
Of course the researcher is not so idealistic.

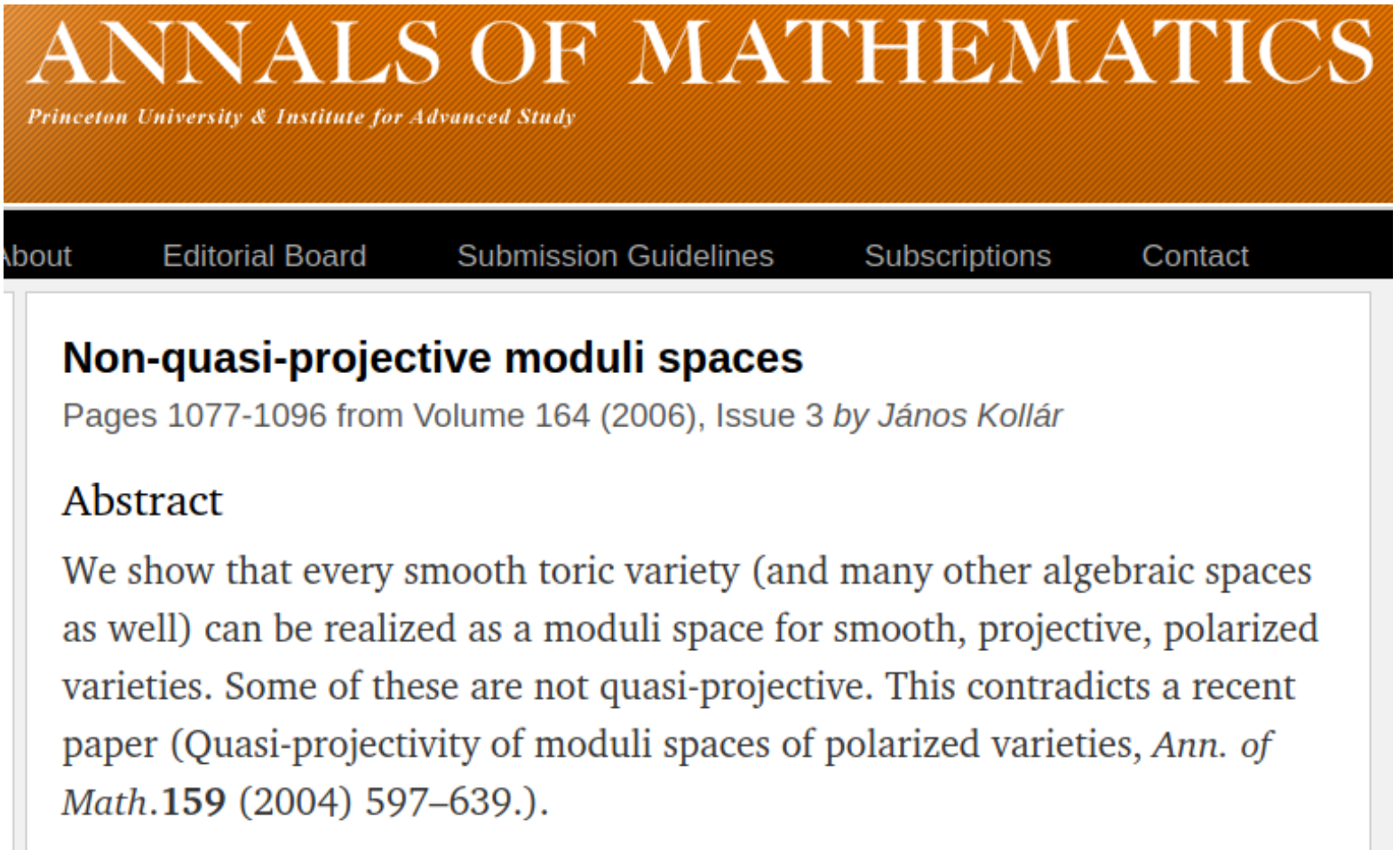
Working definitions of proof for the working mathematician:

A proof is something which the elders in our community have accepted as correct.

A proof is an argument which gets accepted by the Annals of Mathematics or Inventiones.



[from the Annals of Mathematics website]



ANNALS OF MATHEMATICS

Princeton University & Institute for Advanced Study

About Editorial Board Submission Guidelines Subscriptions Contact

Non-quasi-projective moduli spaces

Pages 1077-1096 from Volume 164 (2006), Issue 3 *by János Kollár*

Abstract

We show that every smooth toric variety (and many other algebraic spaces as well) can be realized as a moduli space for smooth, projective, polarized varieties. Some of these are not quasi-projective. This contradicts a recent paper (Quasi-projectivity of moduli spaces of polarized varieties, *Ann. of Math.***159** (2004) 597–639.).

[also from the Annals of Mathematics website]

As far as I know, the Annals of Mathematics never published a retraction of either paper.

If you're in with the in crowd, you can find out which of the two papers is currently believed by the elders.

Conclusion: in modern mathematics, perhaps the idea of whether a certain object is "a proof" can change over time (e.g. from "yes" to "no").

Mathematics > Algebraic Geometry

Perfect points on genus one curves and consequences for supersingular K3 surfaces

Daniel Bragg, Max Lieblich

(Submitted on 9 Apr 2019 (v1), last revised 22 Apr 2019 (this version, v3))

We describe a method to show that certain elliptic surfaces do not admit purely inseparable multisections (equivalently, that genus one curves over function fields admit no points over the perfect closure of the base field) and use it to show that any non-Jacobian elliptic structure on a very general supersingular K3 surface has no purely inseparable multisections. We also describe specific examples of such fibrations without purely inseparable multisections. Finally, we discuss the consequences for the claimed proof of the Artin conjecture on unirationality of supersingular K3 surfaces.

That short 2019 ArXiv paper points out that an important 2015 Inventiones paper crucially relies on a false lemma.

Googling around reveals that there were study groups organised on this important Inventiones paper in 2016.

Voevodsky: “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.”

Still no word from Inventiones about retracting the proof.

Conclusion: some important stuff which is published, is known to be wrong.

And so surely some important stuff which is published, will in future be discovered to be wrong.

So maybe some of my work in the p-adic Langlands philosophy relies on stuff which is wrong.

Or maybe, perhaps less drastically, on stuff which is actually correct, but for which humanity does not actually have a complete proof.

If our research is not reproducible, is it science?

I believe that there is a 99.9 percent chance that the *p*-adic Langlands philosophy will never be used by humanity to do anything useful.

If my work in pure mathematics is neither useful nor *100 percent guaranteed* to be correct, it is surely a waste of time.

So I have decided to stop attempting to generate new mathematics, and concentrate instead on carefully checking "known" mathematics on a computer.

I want to move away from errors now and talk about other issues.

In 2019, Balakrishnan, Dogra, Mueller, Tuitman and Vonk found all the rational solutions to a certain important quartic curve in two variables (the modular curve $X_s(13)$, a.k.a. $y^4 + 5x^4 - 6x^2y^2 + 6x^3 + 26x^2y + 10xy^2 - 10y^3 - 32x^2 - 40xy + 24y^2 + 32x - 16y = 0$).

This calculation had important consequences in arithmetic (new proof of class number 1 problem etc).

The proof makes essential use of calculations in `magma`, an unverified closed-source system using fast unrefereed algorithms.

It would be difficult, but certainly not impossible, to port everything over to an unverified open source system such as `sage`.

Nobody has any plans to do this. Hence part of the proof remains secret (and may well remain secret forever). Is this science?

Gaps.

In 1993, Andrew Wiles announced a proof of Fermat's Last Theorem. There was a gap in the proof.

In 1994, Wiles and Taylor fixed the gap, the papers were published, and our community accepted the proof.

In 1995, I pointed out to Taylor that the proof used work of Gross which was known to be incomplete.

Gross' work assumed that certain linear maps (Hecke operators) defined on two "canonically isomorphic" cohomology groups, commuted with the canonical isomorphism.

Taylor told me it was OK, because he knew another argument which avoided Gross' work completely.

I am sent papers to referee. What am I supposed to be *doing* as a referee?

- "The job of a referee is to convince themselves that the methods used in the paper are strong enough to prove the main results of the paper."
- But what if the methods are strong enough and the authors aren't?
- We might end up with proofs that are incomplete.
- There is then sometimes a debate as to whether the theorems are actually proved.
- This is *not* how mathematics is advertised to the undergraduates.
- The experts know which parts of the literature to believe, of course.
- My conclusion: do you have to be "in with the in crowd" to know which parts of the mathematical literature to believe?

There are *big* holes in mathematics.

Exhibit A: The classification of finite simple groups. Experts tell us that this is a theorem. I believe the experts.

Classification of finite simple groups

1983 : announced, believed by the experts.

1994 : experts know something is wrong (but don't make a big deal about it?)

2004 : One new 1000+ page paper later, Aschbacher thinks it's back on track and says so in the Notices of the AMS. Describes the plan for 12 volumes which will describe the proof (several had already appeared).

2005 : Six of the 12 promised volumes have appeared.

2010 : Six of the 12 promised volumes have appeared.

2017 : Six of the 12 promised volumes have appeared.

2018 : Seventh and eighth volumes appear, plus another piece in Notices of AMS about how it will all be done by 2023.

Out of the three people driving the project, one has died (Gorenstein) and the other two are now in their seventies.

Potential modularity of abelian surfaces.

Exhibit B: One year ago, my (brilliant) former PhD student Toby Gee and three co-authors uploaded onto ArXiv a 285 page paper announcing that abelian surfaces over totally real fields are potentially modular.

The proof cites three unpublished preprints (one from 2018, one from 2015, one from the 1990s), some 2007 online notes, an unpublished 1990 German PhD dissertation, and a paper whose main theorems were all later retracted.

It also contains the following paragraph, buried on page 13:

"It should be noted that we use Arthur's multiplicity formula for the discrete spectrum of GSp_4, as announced in [Art04]. A proof of this (relying on Arthur's work for symplectic and orthogonal groups in [Art13]) was given in [GT18], but this proof is only as unconditional as the results of [Art13] and [MW16a, MW16b]. In particular, it depends on cases of the twisted weighted fundamental lemma that were announced in [CL10], but whose proofs have not yet appeared, as well as on the references [A24], [A25], [A26] and [A27] in [Art13], which at the time of writing have not appeared publicly."

Can we honestly say that this is science?

Chaudouard–Laumon 2010 paper:

1.2. Plus précisément, nous démontrons une variante en caractéristique positive et pour les algèbres de Lie de l'énoncé d'Arthur. De plus, pour alléger autant que faire se peut l'exposition, nous nous sommes limités dans cet article au cas des groupes déployés. Le cas général, qui comprend aussi les formes «non-standard» dues à Waldspurger du lemme fondamental pondéré, s'obtient par des méthodes similaires et sera traité ultérieurement. Des travaux de Waldspurger (cf. [27] et [28]) montre que toutes ces variantes du lemme fondamental pondéré impliquent l'énoncé original d'Arthur pour les groupes sur les corps p-adiques.

This work, which Gee et al need, never appeared.

Of course, it's probably true, and even provable.

References in Arthur's seminal 2013 book (to other work of Arthur):

[A24] ———, *Endoscopy and singular invariant distributions*, in preparation.
[A25] ———, *Duality, Endoscopy and Hecke operators*, in preparation.
[A26] ———, *A nontempered intertwining relation for $GL(N)$*,
[A27] ———, *Transfer factors and Whittaker models*, in preparation.
[A28] ———, *Automorphic representations of inner twists*, in preparation.

Last year I asked Arthur of the status of these references, and he said that none of them were ready.

Jim Arthur is a genius. He's won lots of prizes. He is also 75 years old.

Exhibit C: Gaitsgory–Rozenblyum.

Infinity categories are now a thing. They will only get more important over time. Scholze's new ArXiv article relies on them.

Lurie has written 1000+ pages on $(\infty, 1)$ categories, and has included lots of details in his work.

Gaitsgory–Rozenblyum needed analogous results on $(\infty, 2)$ categories, but to save time omitted some arguments on Gray products. "The missing proofs will be supplied elsewhere".

I asked Gaitsgory how much was missing – he estimates around 100 pages.

I asked Lurie what he thought – he said "mathematicians do vary considerably in how comfortable they are omitting details."

Is human mathematics moving too fast?

I'm an "expert" – am I supposed to believe that abelian surfaces over totally real fields are potentially modular?

I *personally* genuinely don't know any more.

At the conference at CMU I've been to this week, Markus Rabe told us that google are working on a tool which will translate ArXiv articles into computer-checked theorems.

I have now seen an article which cites the Gee et al abelian surfaces paper and which mentions nothing about the 100+ missing pages.

One last error.

A CORRECTED QUANTITATIVE VERSION OF THE MORSE LEMMA

SÉBASTIEN GOUËZEL AND VLADIMIR SHCHUR

Abstract. There is a gap in the proof of the main theorem in the article [Shc13a] on optimal bounds for the Morse lemma in Gromov-hyperbolic spaces. We correct this gap, showing that the main theorem of [Shc13a] is correct. We also describe a computer certification of this result.

1. Introduction

The Morse lemma is a fundamental result in the theory of Gromov-hyperbolic spaces. It asserts that, in a δ-hyperbolic space, the Hausdorff distance between a (λ, C)-quasi-geodesic and a geodesic segment sharing the same endpoints is bounded by a constant $A(\lambda, C, \delta)$ depending only on λ, C and δ, and not on the length of the geodesic. Many proofs of this result have been given, with different expressions for A. An optimal value for A (up to a multiplicative constant) has only been found recently in the article [Shc13a] by the second author, giving $A(\lambda, C, \delta) = K\lambda^2(C + \delta)$ for an explicit constant $K = 4(78 + 133/\log(2) \cdot \exp(157\log(2)/28)) \sim 37723$.

Unfortunately, there is a gap in the proof of this theorem in [Shc13a], which was noticed by the first author while he was developing a library [Gou18] on Gromov-hyperbolic spaces in the computer assistant Isabelle/HOL. In such a process, all proofs are formalized on a computer, and checked starting from the most basic axioms. The degree of confidence reached after such a formal proof is orders of magnitude higher than what can be obtained by even the most diligent reader of referee, and indeed this process shed the light on the gap in [Shc13a]. The gap is on Page 829: the inequality $\sum_{i=1}^{n} e^{-X_i}(X_{i-1} - X_i) \leqslant \int_0^\infty e^{-t}\,\mathrm{d}t$ goes in the wrong direction as the sequence X_i is decreasing.

In this paper, we fix this gap. Here is the estimate we get.

Theorem 1.1. *Consider a (λ, C)-quasi-geodesic Q in a δ-hyperbolic space X, and G a*

That last one is an interesting case.

Original paper was published in J. Funct. Anal. in 2013.

Contains basic error (inequality the wrong way round).

Error discovered by S. Gouezel (2017), whilst Gouezel was formalising the argument using a computer proof checker ("Isabelle").

New argument by Gouezel and original author.

New paper *needs no refereeing*? A computer has actually checked 100 percent of the new argument. So the methods are strong enough to prove the theorem. And by "prove" I mean the classical, "pure", definition of proof – the one which we teach to the undergraduates.

Every detail of the proof is accessible to the reader. The science is reproducible. This is mathematics as we teach it to undergraduates. This is *mathematics*.

Other examples of what I now *personally* think of as mathematics:

A typical undergraduate or MSc level proof.

A typical 100 year old proof of an important result – or anything which has been carefully documented and examined by tens of thousands of mathematicians.

The formal proof by Gonthier, Asperti, Avigad, Bertot, Cohen, Garillot, Le Roux, Mahboubi, O'Connor, Ould Biha, Pasca, Rideau, Solovyev, Tassi and Théry of the Feit–Thompson theorem.

The formal proof by Hales, Adams, Bauer, Dat Tat Dang, Harrison, Truong Le Hoang, Kaliszyk, Magron, McLaughlin, Thang Tat Nguyen, Truong Quang Nguyen, Nipkow, Obua, Pleso, Rute, Solovyev, An Hoai Thi Ta, Trung Nam Tran, Diep Thi Trieu, Urban, Ky Khac Vu and Zumkeller of the Kepler conjecture.

Let's take a look at some mathematics checked with the Lean theorem prover, a formal proof verification system developed by Leo de Moura at Microsoft Research. [cut to Lean]

[I then showed some parts of Lean's maths library written by Jean Lo, Amelia Livingston and Chris Hughes, and noted that they were not professors of computer science but undergraduate mathematicians. I then realised I was out of time, so finished by suggesting that if more undergraduate mathematicians started using this software then perhaps they might start asking uncomfortable questions as they filtered into the PhD system, and argued that even though Lean looks complicated, one thing we could be sure of was that it was *mathematics* in the sense that I personally understand it.]

Page added 21st Jan, covering some things I said but which were not on the slides.

I would like to thank Jeremy Avigad and Rob Lewis for inviting me to Lean Together 2020. I believe in the proof of the prime number theorem and the cap set conjecture!
I would also like to thank Tom Hales for the invitation to speak at the University of Pittsburgh.

My blog: The Xena Project. Teaching formal proof verification to mathematicians. I think it's going to be important one day. If you are a mathematician and want to get started, try the natural number game.

`#LeanProver` on Twitter, and `MSFTResearch` also on Twitter. Many thanks to Leo de Moura for writing software with mathematicians in mind.