

The Unreasonable Effectiveness of Mathematical Experiments: What Makes Mathematics Work

Asvin G.*

June 23, 2025

A Crisis of Confidence

There is a story about Vladimir Voevodsky that captures something profound about the nature of mathematical certainty. In 1989, he and Mikhail Kapranov published " ∞ -groupoids as a model for a homotopy category" [VK90]. The work was sophisticated and ambitious, claiming to provide a rigorous mathematical formulation and proof of an idea from Grothendieck connecting two fundamental classes of mathematical objects: ∞ -groupoids and homotopy types.

Nearly a decade passed. Then, in October 1998, Carlos Simpson published the paper "Homotopy types of strict 3-groupoids" [Sim98]. This paper contained an argument that contradicted the main result of Kapranov and Voevodsky. What made this situation extraordinary was not merely the contradiction—it was what happened next, or rather, what didn't happen.

Neither set of authors could point to a specific mistake in the other's paper. Simpson claimed to have constructed a counterexample, but he was unable to show where in the Kapranov-Voevodsky paper the error was located. Because of this ambiguity, it remained unclear whether Kapranov and Voevodsky had made a mistake somewhere in their paper or Simpson had made a mistake somewhere in his counterexample. Moreover, Kapranov and Voevodsky had previously considered similar critiques themselves and had convinced each other that such objections did not apply to their work.

Here were two mathematical arguments, both apparently rigorous, reaching contradictory conclusions, with no way to determine which was correct. The response was telling: unable to adjudicate between the two results, other researchers simply stopped using the Kapranov-Voevodsky paper. The work was not refuted, but it was no longer trusted. It entered a kind of mathematical limbo that would persist for an extraordinary length of time. Voevodsky himself remained convinced that he and Kapranov were right until the fall of 2013—fifteen years after Simpson's challenge and twenty-four years after the original publication. Only then did he accept that their argument was indeed flawed. For nearly a quarter-century, a piece of published mathematics existed in a state of suspended judgment, neither definitively right nor wrong.

Two factors contributed to this outrageous situation, as Voevodsky himself would later observe¹. First, the sheer complexity of the arguments meant that determining the location of any error required enormous effort. Second, mathematical research relies on a complex system of mutual trust based on reputations. By the time Simpson's paper appeared, both Kapranov and Voevodsky had strong reputations. Simpson's paper created doubts in their result, which led to it being unused by other researchers, but no one came forward and challenged them directly on the contradiction.

For Voevodsky, this experience was transformative. In a 2014 lecture, he would reflect: "I think it was at this moment that I largely stopped doing what is

*The author thanks Claude Opus 4 (Anthropic) for extensive discussions that helped develop and refine these ideas.

¹Voevodsky's comments here come from this lecture

called “curiosity driven research” and started to think seriously about the future.”, and it was the impetus for him to think seriously about the possibility of automating mathematical formalization afterwards, and to develop univalent foundations.

This story would have surprised the mathematicians of a century ago, who believed they had found, in the axiomatic method, a cure for exactly this disease. To understand how we arrived at Voevodsky’s crisis, we must first understand what the axiomatic method promised, why it seemed so necessary, and how it ultimately failed to deliver the certainty it promised.

The Dream of Certainty

For over two millennia, Euclid’s *Elements* stood as the model of mathematical certainty. Its method was simple and compelling: start with self-evident truths (axioms), apply logical reasoning, and derive theorems that must be true. Yet even this monument to deductive reasoning contained a worm of doubt—the infamous fifth postulate, the parallel postulate, which seemed less self-evident than the others.

Generations of mathematicians attempted to prove the parallel postulate from the other four, convinced that such an aesthetic blemish must be eliminable. The shock came in the 19th century when Bolyai, Lobachevsky, and Gauss independently realized that entirely consistent geometries could be built by *denying* the parallel postulate. Suddenly, mathematical truth was not unique. There could be different, incompatible geometries, each as logically valid as Euclid’s. Which one described real space? Mathematics alone could not say.

This was one of the first cracks in the foundation, but far from the last. In Italy, a school of algebraic geometers—Castelnuovo, Enriques, Severi, and others—were developing a beautiful theory using geometric intuition and creative reasoning. Their work was profound and influential, opening new vistas in mathematics. It was also, as later mathematicians discovered to their horror, riddled with gaps and errors. The intuitive leaps that led to great discoveries also led to subtle mistakes that took decades to un-

tangle. Oscar Zariski would later write of spending years trying to understand which results of the Italian school could be trusted and which were built on sand.

But the crisis that truly shook mathematics to its core came from Cantor’s paradise-set theory. Here was a theory of breathtaking generality that seemed to provide a foundation for all of mathematics. Every mathematical object could be understood as a set. Then came the paradoxes. Russell asked: what about the set of all sets that do not contain themselves? Cantor himself discovered that the set of all sets leads to contradiction. The very foundation of mathematics seemed to be crumbling.

David Hilbert, the leading mathematician of his age, proposed a solution of breathtaking ambition. Mathematics, he declared, should be formalized completely. Every concept should be defined precisely, every inference rule stated explicitly, every proof checkable mechanically. In Hilbert’s vision, correct mathematical truth would be reducible to a formal game played with symbols according to strict rules (even if mathematics as a whole was more than a formal game). And he believed he could prove, using only the simplest, most undeniable reasoning, that this game would never lead to contradiction.

“We must know, we will know,” Hilbert declared in 1930. Within a year, Kurt Gödel had shattered this dream. His incompleteness theorems showed that any formal system rich enough to capture arithmetic must be either incomplete (unable to prove some true statements) or inconsistent (able to prove contradictions). Hilbert’s program was not just difficult—it was impossible.

Yet the axiomatic method survived Gödel’s assault, though in a chastened form. Mathematicians could no longer hope for absolute certainty, but they could still use axioms to organize their thinking, to make precise what had been vague, to build new concepts on firm foundations. And this, it turns out, is where the real power of the axiomatic method lies—not in providing certainty, but in something else entirely.

Definitions as Conceptual Technology

To understand what the axiomatic method actually accomplishes, consider the humble epsilon-delta definition of continuity. Every calculus student struggles with it: a function f is continuous at a point a if for every $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $|x - a| < \delta$, we have $|f(x) - f(a)| < \epsilon$. This definition seems like a all too complex reformulation of the intuitive notion that "the graph has no breaks." Newton and Leibniz developed calculus without it and solved problems that we still teach today. Why burden ourselves with such abstraction?

The answer becomes clear only when we see what this definition enables. Once continuity is captured in this precise form, it becomes a piece of conceptual technology - a small program, if you will, that interfaces cleanly with other mathematical concepts. The epsilon-delta definition tells us exactly how continuity interacts with limits, with sequences, with compactness. It allows us to prove that continuous functions on closed intervals achieve their maxima, that they can be integrated, that they preserve connectedness. More profoundly, it suggests generalizations: what if we replace the absolute value with other notions of distance? This path leads to metric spaces, to topology, to functional analysis.

Consider another example: the definition of a topological space through open sets. On its face, this seems absurdly abstract. A topology is a collection of subsets (called "open") that contains the empty set and the whole space, is closed under arbitrary unions and finite intersections. Why would anyone think this captures our intuition about "nearness" or "continuity"? Yet this definition is a masterpiece of conceptual engineering. By distilling the essence of open intervals on the real line into these simple axioms, mathematicians created a framework that revealed deep connections across all of mathematics. The same topological concepts illuminate (real, complex and functional) analysis, algebraic geometry, and even logic. The definition is not trying to capture what open sets "really are" but rather how they behave: what computational rules they follow

when combined with other mathematical concepts.

Laurent Schwartz's theory of distributions provides an even more striking example. Physicists had been using "delta functions" - infinite spikes with unit area-for years, but these objects made no sense in classical analysis. Schwartz didn't try to make sense of them directly. Instead, he changed the question. A distribution, he said, is not a function but a linear functional on a space of test functions. The delta "function" δ_0 is defined by its action: $\delta_0(\phi) = \phi(0)$. This definition seems like a dodge, but it is actually profound. It tells us exactly how distributions can be manipulated: they can be added, multiplied by smooth functions, differentiated (always!), and transformed. The definition provides a precise interface, a set of computational rules that allow distributions to interact with the rest of mathematics. Suddenly, operations that were formally meaningless became rigorous and computable.

Perhaps the most spectacular example of this process is Grothendieck's revolution in algebraic geometry. The classical objects of study, algebraic varieties were already well understood. But Grothendieck replaced them with schemes, built from commutative rings through a process of such abstraction that even experienced mathematicians found it bewildering².

Why replace concrete geometric objects with this abstract machinery? Because the definitions of scheme theory provide more powerful interfaces. A scheme is, roughly speaking, a topological space equipped with a sheaf of rings, satisfying certain conditions. This definition seems far too abstract until you realize what it enables. Schemes live naturally in families, can be glued together, pulled back, pushed forward and treated as a "space" once the correct intuitions are matched with the formalisms. And as the reward for these complications, they allow geometric intuition to be applied to number theory resulting in the proof of long standing open conjectures such as the Weil conjectures (which had resisted direct as-

²"Hendrik Lenstra twenty years ago was firm in his conviction that he did want to solve Diophantine equations, and did not wish to represent functors and now he is amused to discover himself representing functors in order to solve Diophantine equations!" - Barry Mazur at a 1995 BU conference on Fermat's Last Theorem

sault for decades) and Fermat’s Last Theorem (which had gone unsolved for over 350 years)-not because schemes are “truer” than varieties, but because they provide more powerful and flexible computational interfaces.

A definition in mathematics is like a piece of code in a programming language. It specifies an interface-how this concept can be used, what operations are valid, how it interacts with other concepts. The epsilon-delta definition is code that tells us how to verify continuity. The axioms for a topological space are code that tells us how to manipulate open sets. Grothendieck’s schemes are elaborate programs for geometric reasoning. Viewed this way, a major part of “doing mathematics” is not merely proving theorems from axioms but developing and debugging conceptual interfaces. When we prove that continuous functions on compact sets are uniformly continuous, we are not discovering a truth about reality but verifying that our definitions interface correctly-that the “continuity” module works properly with the “compactness” module to produce the “uniform continuity” behavior.

A small remark is in order here: definitions are not the only ways in which we encode our intuitions into the tools of mathematics. Very often, especially in fields such as analysis and combinatorics, it happens that a particular *kind* of argument is repeated often but not in the exact same way in every instance. In this case, we learn the general pattern of such arguments and can utilize them flexibly in different contexts without the need to formalize them precisely in a definition or otherwise. Nevertheless, “good” definitions are of prime importance in every part of mathematics and will be the focus of this essay.

That said, mathematical practice consists of more than merely revealing intuitive truths-far from it! There is no meaningful sense in which Fermat’s last theorem is intuitively true, it is a truly surprising fact that was first experimentally discovered and then rigorously proven. So how do we discover new mathematics in practice, if not by deducing them from “known” axioms?

Mathematics as an experimental Science exploring the Computational Universe

The twentieth century witnessed a parade of attempts to capture the essence of mathematics, each ultimately failing to account for how mathematics is actually practiced. The Platonists, following in Plato’s footsteps, insisted mathematicians discover eternal truths existing in some abstract realm. But this merely pushes the question one step back - what is this ideal realm and how do we gain knowledge of it? The formalists, led by Hilbert’s grand vision, reduced mathematics to symbol manipulation according to rules. Yet Gödel shattered their dream of complete formalization, and besides, no working mathematician experiences their craft as meaningless symbol-pushing. The intuitionists demanded that mathematical objects be mentally constructible, but this ideology would banish vast swaths of successful mathematics. Social constructivists noticed the human elements in mathematical practice but went too far, unable to explain why some social constructions (like calculus) prove so much more powerful than others.

We propose a new framework for what mathematics is that synthesizes the philosophies above and is based on deep familiarity with the practice of mathematics: *mathematicians, like physicists, are in the business of explaining patterns through theory-building*³. The difference lies only in the nature of their experiments and the patterns. Where physicists drop balls and collide particles, mathematicians compute. They calculate specific examples, work out small cases, follow chains of logical inference — all physical processes of symbol manipulation implemented in brains, on paper, or in silicon. From these computational experiments emerge patterns that cry out for explanation.

³I agree with the spirit of Arnold’s claim “*Mathematics is a part of physics. Physics is an experimental science, a part of natural science. Mathematics is the part of physics where experiments are cheap.*” [Arn97] This essay can be seen partially as an attempt towards clarifying the nature of these experiments

Consider Leonhard Euler, that computational virtuoso of the eighteenth century. When he claimed that $1 + 2 + 3 + \dots = -1/12$, his contemporaries were aghast. How could adding positive numbers yield a negative result? Yet Euler had noticed something profound: manipulating divergent series as if they converged led to correct results in problem after problem. He could compute the values of the Riemann zeta function, solve differential equations, derive physical predictions—all using these “illegal” methods. Euler was not being careless; he was being empirical. His computational experiments revealed patterns that existing theory could not explain.

It would take over a century for mathematics to develop frameworks - analytic continuation, regularization techniques - that could explain why Euler’s methods worked. The patterns he discovered were always there, waiting in the computational universe. His divergent sums and their apparent consistency were not “wrong” any more than Mercury’s anomalous orbit was “wrong” before Einstein. Both were observations that existing theory could not accommodate, demanding new frameworks for their explanation.

It is a pattern that has been repeated in the history of mathematics ad-infinitum - early examples range from the introduction of negative numbers, cube roots of imaginary numbers and so on, all of which gained acceptance through their integral role in producing novel, verifiably checkable solutions to old mathematical questions. When the Italian school of algebraic geometry (Castelnuovo, Enriques, Severi) built their magnificent theories on geometric intuition, they were master experimentalists working without adequate theoretical frameworks. Their “proofs” had gaps because they were really reporting observations from their computational experiments. As André Weil would later write [Wei46], he spent years “as a sort of missionary” trying to understand which of their results could be trusted. But Weil and Oscar Zariski did not dismiss the Italians’ work⁴. In-

stead, they built new theoretical frameworks based on commutative algebra rather than geometric intuition that could rigorously explain the patterns the Italians had discovered.

The story repeats with Laurent Schwartz and distributions. Physicists had been successfully using “delta functions”—infinite spikes with unit area—for decades. Mathematicians declared these objects nonsensical, violations of everything known about function theory. But the physicists persisted because delta functions worked. They solved real problems, predicted experimental outcomes, simplified calculations. Schwartz’s genius was not in “correcting” the physicists but in developing a framework where their computational practices made rigorous sense. Distributions were not functions but functionals, and suddenly the patterns fell into place.

This experimental view of mathematics illuminates the distinction between two modes of mathematical practice. Most mathematicians work within established frameworks, using existing theory to explain new computational observations—normal mathematics, in Thomas Kuhn’s terminology. But occasionally, the patterns resist explanation by existing frameworks. Contradictions multiply, special cases proliferate, computational practice outpaces theoretical understanding. These are the moments that call forth mathematical revolutionaries.

Alexander Grothendieck stands as perhaps the purest example of such a revolutionary. Faced with the Weil conjectures, patterns in number theory that existing algebraic geometry could not explain, he did not merely solve the problem. He rebuilt the entire subject. His theory of schemes seemed absurdly abstract, replacing concrete geometric objects with locally ringed spaces. Colleagues complained they could no longer understand algebraic geometry. But Grothendieck was not pursuing abstraction for its own sake. He was building theoretical machinery powerful enough to explain the computational pat-

⁴Also Weil[Wei46]: “...the so-called ‘intuition’ of earlier mathematicians, reckless in their use of it may sometimes appear to us, often rested on a most painstaking study of numerous special examples, from which they gained an insight not always found among modern exponents of the axiomatic creed.”

And Zariski[Zar50]: “The Italian geometers have erected, on somewhat shaky foundations, a stupendous edifice: the theory of algebraic surfaces. It is the main object of modern algebraic geometry to strengthen, preserve, and further embellish this edifice, while at the same time building up also the theory of algebraic varieties of higher dimension.”

terns that had resisted all previous attempts. The spectacular success of his program, culminating in Deligne’s proof of the Weil conjectures, vindicated this approach. The patterns demanded new frameworks; Grothendieck provided them.

Truth, Consistency, and the Limits of Theory

If mathematics is an experimental science, what becomes of mathematical truth? The answer is both radical and liberating: mathematical truth is not about correspondence to platonic reality but about successful prediction of computational outcomes. A mathematical statement is “true” to the extent that it can be incorporated into our theoretical frameworks without generating contradictions: without predicting that a computation yields both A and not A .

This reconceptualization illuminates Gödel’s incompleteness theorems. Gödel showed that any formal system rich enough to capture arithmetic must be either incomplete or inconsistent. In our framework, this gains a dual significance. First, it is an empirical discovery about the limits of theory-building: any theoretical framework powerful enough to explain arithmetic computation will inevitably make predictions it cannot derive from its axioms. There will always be true patterns, regularities that hold whenever we check, that our theories cannot predict.

But Gödel also reveals something deeper. Just as we can never prove a physical theory true (only fail to falsify it), we can never prove a mathematical framework consistent (only fail to find contradictions). Both physics and mathematics are forever provisional, gaining credibility through successful use rather than ultimate verification. The parallel is exact: physicists test theories against physical experiments; mathematicians test frameworks against computational experiments.

This perspective recasts many historical “crises” in mathematics. When Berkeley mocked Newton’s infinitesimals as “ghosts of departed quantities” [Ber34], he was correct that they made no sense in existing

frameworks. But Newton and Leibniz were not talking nonsense, they were reporting computational patterns that reliably led to correct results. For over a century, mathematicians computed derivatives and integrals, solved differential equations, derived physical laws, all using theoretically incoherent methods. The computations worked; the theory lagged behind.

The eventual development of limits did not “correct” infinitesimal calculus but explained why it worked. The same story played out with the Italian geometers. When rigorous algebraic geometry revealed gaps in their proofs, it did not invalidate their discoveries but provided better theoretical foundations for the patterns they had found. In each case, computational evidence preceded theoretical understanding.

This suggests a new criterion for mathematical acceptance: *consistency in use rather than complete rigor*. When a mathematical framework reliably produces correct results, explains diverse phenomena, and meshes well with other mathematics, we should take it seriously even if we cannot prove it consistent. This is not lowering standards but acknowledging how mathematics actually progresses.

Consider how mathematicians form conjectures. The Riemann hypothesis has been computationally verified for trillions of cases and more seriously, the philosophy behind it has led to numerous analogous conjectures many of which have been proved as well as used as a crucial building block in a web of implications stemming from it—and all of this without a hint of inconsistency. In such cases, mathematicians might have an overwhelming confidence that the conjecture is true, much as we have strong beliefs in our physical theories. The four-color theorem was had strong proponents in the mathematical community before its computer-assisted proof because of numerous examples in which we could verify it. These are not psychological quirks but legitimate scientific reasoning: extensive computational experiments support the conjecture.

The computer-assisted proof of the four-color theorem itself illustrates our framework perfectly. The mathematical community’s initial discomfort was not really about computer use; mathematicians had always relied on physical processes (brains) to verify

proofs. The issue was that the proof provided no insight, no theoretical framework explaining why four colors suffice. It was pure experimental verification without theory. Yet as our framework predicts, the community accepted eventually that the theorem was true: computational evidence, sufficiently extensive, compels belief. But also as predicted from our framework, mathematicians are still deeply interested in a conceptual proof of the theorem and such a proof would be of immense value.

This view also explains why contradictions are not instantly fatal to mathematical frameworks. When Russell's paradox revealed contradictions in naive set theory, mathematicians did not abandon sets. They modified the framework, restricting comprehension axioms to avoid contradictions while preserving the theory's explanatory power. Similarly, when infinitesimal calculus led to contradictions, mathematicians developed limits to salvage the valuable patterns. A framework that explains much but predicts some contradictions is better than no framework at all-it offers a starting point for successive refinements.

And looking to the future, our response to discovering a contradiction in Peano's axiomatization of arithmetic tomorrow would not be to abandon all number theoretic results as false. Instead, we would work towards finding new axiomatics that preserve as much of our current math as today while avoiding the contradiction - just as historically happened with set theory. *What gives us confidence in our mathematics is the apparent consistency resulting from the use of these theorems, not necessarily derivations from "true axioms"*. Indeed, a proof primarily makes us more confident in a result by *connecting* its consistency to a wider, more established swath of mathematics.

From this vantage point, pluralism in mathematics - where we naturally have a multiplicity of axiomatic systems explaining our mathematical universe, comes to feel entirely natural and even something to be wished for. Explanations should be adapted to the domain after all and different parts of mathematics emphasize different operations and concepts, naturally giving rise to different axiom systems. The independence of the continuum hypothesis from ZFC could either be because ZFC is not a sufficiently

comprehensive explanation or that the computational universe of "set theory" is truly pluralistic and comes in many flavours. Which explanation is "correct" will be a natural consequence of our explorations of set theory and the patterns we discover therein.

Wigner's Question Transformed

In 1960, Eugene Wigner published "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" [Wig90], marveling at what he saw as a miracle. Why should abstract mathematical concepts, developed for their own sake, prove so powerful in describing physical reality? Why do complex numbers, invented to solve algebraic equations, become essential for quantum mechanics? Why does Riemannian geometry, pursued for purely mathematical reasons, provide the framework for general relativity?

From the traditional Platonist perspective, this effectiveness is indeed mysterious. If mathematics exists in an abstract realm disconnected from physical reality, its applicability to physics seems miraculous. But our framework dissolves the mystery by reconceptualizing what mathematics is and does.

Physicists, like mathematicians, are in the business of theory-building to explain patterns. The difference is that physicists focus on patterns in physical measurements while mathematicians focus on patterns in computation. Despite this superficial difference however, one notes that physics is absolutely suffused with computation. Physicists constantly calculate-solving differential equations, computing trajectories, simulating systems, deriving consequences of theories. They are performing mathematical experiments as an integral part of physical investigation.

Once we recognize this, Wigner's mystery transforms into something more comprehensible. It's not surprising that theories organizing computational patterns prove useful in physics, because physics relies fundamentally on computation. Asking why mathematics is effective in physics is like asking why theories about wave patterns apply across different physical contexts-from water waves to sound to electromagnetic radiation. The patterns are there; good theories capture them.

Nevertheless there is still a question here: why are some mathematical frameworks so unreasonably powerful? Why does group theory appear everywhere from crystallography to particle physics? Why does complex analysis illuminate fields from fluid dynamics to number theory?

The answer lies in recognizing that certain computational patterns are fundamental, they appear again and again across different contexts. Group theory succeeds because symmetry is a fundamental organizational principle in both computation and physics. Linear algebra pervades science because linear approximations are computationally tractable and physically ubiquitous. Complex numbers capture rotational and periodic patterns that emerge naturally in countless contexts.

These frameworks are not mystically tapping into platonic truths. They are exceptionally well-engineered conceptual tools that capture patterns appearing throughout the computational and physical universe. Like the wheel or the lever, they embody discoveries about fundamental structures that, once recognized, find applications everywhere.

The "unreasonable" effectiveness of mathematics in physics thus becomes entirely reasonable. Mathematical frameworks that successfully organize broad swaths of computational patterns will naturally prove useful in any field that relies on computation, and physics, at its core, is exactly such a field. Wigner's miracle is no miracle at all, but the natural consequence of physicists and mathematicians exploring different aspects of the same computational universe.

Conclusion: Mathematics in the Age of AI

The story that began with Voevodsky's crisis of confidence ends with a new understanding of mathematical practice. Mathematics is not a monument to certainty but an expedition into the computational unknown. Its power comes not from unshakable foundations but from the quality of its experimental tools - the definitions and frameworks that allow us to detect, express, and explore patterns in the vast space

of possible computations - and from the quality of our explanations of these computational patterns in wide-ranging theories.

We stand now at a peculiar moment in mathematical history. Artificial intelligences are rapidly developing the ability to prove theorems, recognize patterns, and even propose new mathematical frameworks. Within a year or two, they may surpass human capabilities in many aspects of mathematical practice, much as they have in chess and other domains. This forces us to confront fundamental questions about the nature and purpose of mathematics.

History offers guidance. Euler computed endlessly, discovering that $\zeta(2) = \pi^2/6$ through pattern recognition in calculations rather than rigorous proof. Gauss computed the distribution of primes up to three million by hand, discovering the prime number theorem through observation rather than proof. For these giants, computation was central to mathematics.

But the twentieth century brought a peculiar shift. As mathematics became more abstract, "mere computation" was increasingly dismissed. The Bourbaki collective attempted to rebuild mathematics on purely logical foundations, free from computational contamination. Real mathematicians proved theorems; computation was mundane work, best left to computers. To be clear, this was far from the universal attitude but nevertheless, it was (and is) present.

The four-color theorem in 1976 questioned this comfortable distinction. Its computer-assisted proof required checking thousands of cases - impossible for any human to verify directly. The mathematical community's discomfort was palpable. Was this even mathematics if no human could comprehend the argument? But the real issue was not the computer assistance per se. It was that the proof illuminated nothing. It established the truth of the statement without revealing why four colors suffice, without training human intuition, without opening new conceptual territories.

Now we face a more radical disruption. Once our computers can not only calculate but also prove theorems mechanically, we will de-emphasize the importance of both proof and computation to our understanding of mathematics. Our framework suggests

what will replace them as the essence of mathematics: the recognition of significant patterns in the computational universe and the construction of conceptual frameworks that make these patterns visible and manipulable. Whether this recognition comes through calculation (like Euler), human proof (like Bourbaki), or AI-assisted exploration is secondary to the act of understanding itself.

This perspective clarifies what we should want from mathematical AI. Not systems that merely find proofs - mechanical verification is not understanding and not systems that only recognize regularities - pattern detection without conceptual framework is sterile. Instead, we need systems optimized for two intertwined goals: exploring the computational universe to discover genuinely new phenomena *and* to illuminate these discoveries in ways that enhance human intuition.

The ideal mathematical AI would be like a telescope for the computational universe - allowing us to see further while also making what we see comprehensible. It would find not just valid proofs but beautiful ones, not just correct formulas but illuminating ones. It would excel at the crucial act of translation between frameworks, showing how patterns visible in one conceptual system manifest in others.

The effectiveness of definitions in mathematics is 'unreasonable' only if we expect mathematics to be about eternal truths rather than about patterns that emerge from computation. Once we see mathematics as an experimental science of patterns in symbolic manipulation, the effectiveness becomes not just reasonable but inevitable. We are all -humans and AIs alike - exploring the same computational universe, crafting theories to explain what we find there. In this view, mathematical truth is not deduced from axioms but discovered through computational experiment; consistency is not a guarantee of truth but a marker of explanatory power. Voevodsky's crisis arose from clinging to the old vision of mathematics as a monument to certainty. His turn to formalization was not a retreat but a recognition: if mathematics is fundamentally about computation, then our tools must match that reality. The mathematical frameworks we build are not capturing timeless truths but organizing the patterns we discover when

symbols dance according to rules—patterns that exist not in some platonic realm but in the very fabric of computation itself.

References

- [Arn97] VI Arnold, *On teaching mathematics*, Uspekhi Mat. Nauk (1997).
- [Ber34] George Berkeley, *The analyst; or, a discourse addressed to an infidel mathematician*, London, 1734. Full title: The Analyst; or, A Discourse Addressed to an Infidel Mathematician. Wherein It is examined whether the Object, Principles, and Inferences of the modern Analysis are more distinctly conceived, or more evidently deduced, than Religious Mysteries and Points of Faith.
- [Sim98] Carlos Simpson, *Homotopy types of strict 3-groupoids*, 1998.
- [VK90] VA Voevodskii and MM Kapranov, *oo-groupoids as a model for a homotopy category*, Russ. Math. Surv **45** (1990), 239.
- [Wei46] Andre Weil, *Foundations of algebraic geometry*, Vol. 29, American Mathematical Soc., 1946.
- [Wig90] Eugene P Wigner, *The unreasonable effectiveness of mathematics in the natural sciences*, Mathematics and science, 1990, pp. 291–306.
- [Zar50] Oscar Zariski, *The fundamental ideas of abstract algebraic geometry*, Proceedings of the international congress of mathematicians, 1950, pp. 77–89 (quote from pp. 88–89).