# Proof is Beauty, and Beauty Proof: Mathematical Proof in the Age of Mechanical Production.

**Keywords**: Philosophy of mathematics; Aesthetics of proof; Mathematical beauty; Formal derivations; Artificial intelligence; AI-generated Art.

**Abstract**

The increasing reliance on computer-assisted proofs marks a transformative moment in mathematical practice. While much philosophical attention has focused on the epistemic reliability of formal derivations, this paper argues that a more significant loss may be aesthetic. Human proofs are not merely instruments for establishing truth; they are cognitive artifacts, structured to facilitate understanding, insight, and appreciation. By contrasting human and computer-generated proofs, I argue that formal derivations—while epistemically robust—cannot reproduce the narrative, elegance, and heuristic guidance that give mathematical reasoning its aesthetic and cognitive value. This is not a rejection of formal methods or a call for technological retreat; rather, it is a case for preserving human-oriented proofs alongside machine-assisted approaches. I briefly gesture toward how this discussion connects to broader debates in the philosophy of mathematics, aesthetics, and technology, but the primary aim is to clarify what is at stake in the aesthetic dimension of mathematical practice.

# Introduction

On the standard view, mathematical proofs as they appear in journals are epistemically justified insofar as they stand in a certain relation to finite derivations in a base formal system, such as first-order ZFC.[1] Justification is typically taken to flow from the formal derivation to the informal, human-facing proof.[2] Although such derivations are seldom explicitly constructed, they are presumed to underwrite the epistemic legitimacy of mathematical practice. Historically, formal derivations have been employed chiefly to certify results already obtained by human proofs. Not all cases conform to this pattern though. Certain theorems — most notably the Four-Color Theorem— have resisted surveyable human proof and were first established via computer-assisted derivations, such as those carried out by Appel and Haken.[3] Recent advances in automated theorem proving and large language models render increasingly plausible a conception of mathematics in which all proofs could be generated — and perhaps even primarily valued — as extended formal derivations performed by machines.

Many philosophers and mathematicians have considered this scenario with unease.[4] The concern is typically not that the centrality of such proofs would undermine the correctness of mathematics, but that it would render it, in some sense, epistemically or intellectually impoverished. Why should this be so? The

[1] See, for example, Avigad [2020] or Azzouni [2004].

[2] Although I will not pursue this issue in detail here, I believe there are strong reasons to be skeptical of this view. First, it presupposes a version of Hilbert's thesis: that for every acceptable informal mathematical proof, there exists a corresponding formal derivation (or set of derivations) in an appropriate system. But this is far from obvious. Second, even if Hilbert's thesis were true, our justification for it would likely rest on an informal proof or on empirical grounds, neither of which is clearly stronger than the informal proofs the thesis is meant to underwrite. It is unclear, then, what explanatory or justificatory work the appeal to formal derivations actually does. Third, the view implicitly treats classical first-order rules of inference as epistemically privileged. But these rules, too, must be justified informally. And if informal justification suffices for them, why not for the higher-level rules typically used in human mathematical practice? The detour through formal derivations may thus be not only superfluous but question-begging.

[3] Apple & Haken [1989]

[4] Though of course not all. Some are positively giddy at the prospect of the theorem proving capabilities.

most prominent objections typically appeal to core epistemic values: justification, understanding and explanation, and a priority and infallibility. On one line of thought, the sheer complexity and opacity of many computer proofs threaten their justificatory status, as they are no longer surveyable by a finite human mind. On another, such proofs are said to preclude the kind of insight and explanatory depth that traditional human proofs can afford. And on a third, it is argued that the results of computer proofs cannot be a priori in the same sense as classical mathematics, since they depend on empirical facts about physical computation or software reliability, and as such, would be fallible.

I am skeptical of these objections. Justification, in my view, need not require surveyability, so long as the derivations are demonstrably valid within formal systems that we have independent reason to trust. Similarly, the explanatory and understanding-based value of proofs may not be irrevocably lost, but could instead be recoverable through interpretive tools — intelligent systems designed to analyze, extract, and present the relevant inferential structures in ways accessible to human cognition. As for a priority, it is not obvious that the dependence on computational assistance invalidates the a priori status of the resulting theorems, any more than reliance on the testimony of other mathematicians does. After all, mathematicians frequently trust published proofs, results, or calculations that they have not personally verified in full detail, and such reliance is typically not seen as undermining the a priori nature of their knowledge.

I propose that these objections, while not without force, overlook the deeper source of value at stake. What would be lost in a transition to computer proofs is not primarily epistemic, but aesthetic. Human proofs do more than secure belief in their conclusions: they are crafted artifacts that structure a distinctive kind of cognitive experience. Like narratives, poems, or works of art, they invite the reader into a guided sequence of insights, surprising turns, elegant resolutions, and moments of conceptual unification, that confer a value independent of the truths they establish. Formal derivations, no matter how reliable or comprehensive, are not designed to afford this experience; they lack the intentional shaping of exposition and economy that characterizes human mathematical practice.

The main aim of this paper is to articulate the aesthetic dimension of human mathematical proofs, to show that formal derivations inherently lack this quality, and to argue that this—rather than any epistemic concern—would constitute the most significant loss were mathematics to shift primarily toward derivation-based or computer-assisted proofs. This is not to suggest a Luddite rejection of technology. I am not arguing against the use of automated theorem provers or the adoption of the best tools available for advancing mathematical knowledge. On the contrary, such tools are indispensable for many purposes, and their role in contemporary mathematics will likely grow. My aim here is simply to characterize, as accurately as possible, what could be lost in this transition, so that we can appreciate the full stakes of the change.

## Computer Proofs and Human Proofs

By a computer proof, I mean a fully formal derivation — that is, a finite sequence of well-formed symbolic expressions, each following mechanically from previous ones according to a fixed set of inference rules, starting from a set of axioms. Such derivations are typically carried out in formal systems like natural deduction or the sequent calculus, frameworks familiar from introductory logic courses. Each step in a derivation is simple enough to be checked by a machine, and the derivation as a whole is finite, though in practice often staggeringly long. The primary purpose of formal derivations is precisely this: to allow for mechanical verification of mathematical results. In recent decades, automated theorem provers and proof assistants — such as Coq, Lean, and HOL Light —

have been developed to construct or verify these derivations with a degree of rigor that leaves no room for the heuristic shortcuts and informal leaps characteristic of human mathematical reasoning.

Consider the theorem that there are infinitely many prime numbers. A human proof can be expressed in a few elegant lines. A computer proof of the same result, by contrast, expands this reasoning into many micro-steps, each explicitly justified by an inference rule. For example, in Lean one might write:

```
import data.nat.prime
open nat

theorem infinite_primes : ∀n, ∃p ≥ n, prime p :=
begin
 assume n,
 let m := factorial n + 1,
 obtain ⟨p, hp⟩ := exists_prime_and_dvd m,
 have p_ge_n : p ≥ n :=
  by {by_contra h, have p_dvd_fact := prime.dvd_factorial hp.1 (lt_of_not_ge h),
      exact hp.2 (dvd_add_right p_dvd_fact)},
 exact ⟨p, p_ge_n, hp.1⟩,
end
```

Although compact by computer standards, this script already includes many details that a human proof would leave implicit. The fully expanded derivation verified by Lean — reducing every step to primitive logical operations — is vastly longer still.

What such derivations make clear is that computer proofs differ radically from how humans actually do mathematics. They are not written for human understanding or appreciation, but for mechanical verification. They are typically austere, tedious, and devoid of the elegance, insight, and stylistic flair that characterize the best human proofs. Even when they mirror the logical structure of a familiar argument, they do so in a way that feels cold and alien — stripped of the very features that make mathematics beautiful and satisfying to its practitioners.

Human proofs, in contrast, are the kinds of proofs that appear in mathematics textbooks, lecture notes, and journal articles. They are written in a combination of natural language, mathematical notation, and high-level reasoning, rather than as sequences of primitive symbols governed solely by formal inference rules. These are the proofs mathematicians actually read, write, and share. Their primary audience is human beings, and they are crafted accordingly — for intelligibility, clarity, and often for elegance or stylistic effect.

Consider again the theorem that there are infinitely many prime numbers. The classical proof, usually attributed to Euclid, proceeds in a single, elegant stroke[5]:

1. Assume finitely many primes $p_1, p_2, \ldots, p_n$,
2. Form the number $N = p_1 p_2 \cdots p_n + 1$,
3. Note that $N$ cannot be divisible by any of the assumed primes, yielding a contradiction.

This argument is valued not just for its correctness, but for its simplicity, ingenuity, and formal grace. It is strikingly compact, easily remembered, and often cited as a paradigm of mathematical beauty.

This stands in sharp contrast to a fully formal derivation of the same result,

[5] Aigner & Ziegler [2004]

which, as I have said, would expand the reasoning into thousands of mechanical steps, with every fact about multiplication, divisibility, and contradiction explicitly derived from first principles. In form and function, such derivations are radically different from human proofs. Where the former are designed for mechanical verification, the latter are designed for human understanding - and, as we will begin to see, for human appreciation as well.

This contrast between human and computer proofs is not merely theoretical. It came to a head with the 1976 proof of the Four Color Theorem by Kenneth Appel and Wolfgang Haken — widely regarded as the first major theorem whose proof required essential computational assistance. Their strategy involved reducing the problem to nearly 2,000 distinct configurations and using a computer to verify that each could be colored with four colors. No individual could feasibly check the entire proof by hand. Although later versions, and formalizations in systems such as Coq, have improved upon the original, they have not eliminated the reliance on machine verification. By most standards, there is still no fully satisfactory human proof of the theorem.

The response from the mathematical community was divided. While many accepted the result as correct, others expressed reservations. These concerns were overwhelmingly epistemic in character, focusing on issues such as justification, explanation, understanding, a prioricity, and the risk of undetected error. Could a derivation that no human can fully survey still serve as a legitimate proof? Could such a result be said to offer genuine mathematical knowledge or insight? These questions have prompted substantial philosophical discussion, and rightly so. Yet as significant as these epistemic concerns are, I do not think they correctly capture what would be lost in a full transition to computer derivations. For as we will see, I think they can be dissolved.

## Epistemic Concerns

The literature so far has focused predominantly on epistemic objections to computer-generated proofs. Three are especially common:

(A) Justification: Some worry that unsurveyable proofs — those too long or complex for any single human to check in full — cannot justify belief in their conclusions. This is a reasonable and understandable concern: mathematical proofs have traditionally been valued for their transparency, and many take the ideal of personal surveyability to be central to justification. But in practice, the role of individual verification in mathematics is already quite limited. Mathematicians routinely rely on the testimony of others, incorporating lemmas, theorems, techniques, and results from the literature into their own work without independently verifying every detail. In this sense, even human proofs are often collaborative and epistemically distributed.

Consider, for example, Andrew Wiles's proof of Fermat's Last Theorem.[6] It spans hundreds of pages, relies on deep results from algebraic geometry and the theory of modular forms, and required years of scrutiny by specialists before being accepted. Or consider the classification of finite simple groups[7] — a decades-long effort involving hundreds of papers and many contributors. No single person could possibly verify all of its components. Yet mathematicians justifiably treat these results as established, not because they are personally surveyable, but because they are embedded in a robust communal structure of trust, expertise, and methodological rigor.

[6] Wiles [1995]

[7] The classification of finite simple groups, sometimes described as the "enormous theorem," was a collaborative effort spanning several decades and involving contributions from over 100 mathematicians. The final proof comprises thousands of pages spread across hundreds of journal articles. It can be seen in Hales [2005]. While efforts have been made to streamline and consolidate the classification — most notably the ongoing project to produce a unified and simplified version — no single mathematician is in a position to survey the entire proof firsthand. Nonetheless, the result is widely accepted within the mathematical community.

Once we acknowledge that justification in mathematics is often distributed across a network of researchers, practices, and institutions, it becomes less clear why the opacity of computer proofs should be disqualifying. If an automated derivation is mechanically verified by reliable software, and if its construction can be scrutinized and reconstructed in principle, then it is not obvious that such a proof is any less justificatory than a long and technically demanding human one. The real difference may not lie in justification at all, but elsewhere.

(B) A Priority: A subtler objection holds that computer proofs threaten the a priori status of mathematics. Traditionally, mathematics has been regarded as a paradigm of a priori knowledge: we come to know mathematical truths independently of sensory experience, through pure reasoning. But if results are established by computational processes that are opaque to most human mathematicians—effectively black boxes—can our knowledge of these results still count as a priori? Or does it instead come to resemble empirical knowledge, based on observation of a machine's output, contingent on external processes we do not fully control or understand?

This is a reasonable and understandable concern. But it is not clear that our trust in computer proofs introduces any new epistemic dependencies that are not already present in ordinary mathematical practice. Mathematicians regularly rely on the testimony of other mathematicians. They incorporate lemmas, theorems, and techniques into their own work without independently verifying every step. Indeed, many proofs, such as the classification of finite simple groups, are so long and complex that no single individual could survey them in full. If these results are taken to confer a priori justification, then why not computer-assisted results, especially when the computational systems involved can, in principle, be made more reliable, transparent, and reproducible than any human collaborator?

Indeed, the worry may reflect a misplaced conception of the a priori in mathematical practice. As Philip Kitcher (1983) has argued, mathematical knowledge is deeply social. It depends on networks of trust, communication, and collaboration — on the cumulative and interdependent labor of many minds. On this view, mathematical belief is not the product of autonomous reasoning from first principles, but rather a species of testimonial knowledge, subject to many of the same sources of error and correction as scientific inquiry. If this is right, then computer proofs do not undermine the a priori character of mathematics — they reveal that it was never quite so pure to begin with.

(C) Understanding: A third concern is that computer-generated proofs may fail to yield understanding. Long brute-force derivations, no matter how reliable, often fail to reveal the conceptual connections that make the best human proofs explanatory. They can feel like mere sequences of moves, offering no immediate insight into why a result holds. This is a more serious concern than questions about justification for a simple reason: it seems almost self-evident that long, brute-force derivations do not reveal the conceptual connections between ideas in the way that the best human proofs often do. They may establish that a conclusion follows, but they typically do not explain it. Yet, while this diagnosis may be accurate, it is not decisive. Many human proofs themselves fall short of being fully explanatory: sprawling proofs-by-cases, heavily formalized derivations, and the original computer-assisted proof of the Four Colour Theorem all began life as opaque achievements. The crucial point is that such proofs can still serve as the raw material for explanation — embedded in a wider practice through which understanding emerges collectively and over time. Human mathematicians refine, reorganize, and reinterpret such results until they disclose the patterns and dependencies that make them illuminating. The challenge, then, is not whether machine derivations are explanatory in themselves, but whether they can be integrated into a mathematical culture that yields the same, or even greater, opportunities for understanding.

Here there is reason for optimism. Explainable AI could play much the same role in relation to machine derivations that skilled mathematicians already play in relation to unwieldy human proofs: isolating key lemmas, tracing dependencies, and proposing conceptual reorganizations. Even when a derivation is too long or intricate to survey, it may encode structural features whose explanatory significance becomes clear only after sustained analysis. Indeed, on modal-structure accounts of explanation — such as Lange's (2014) view that to explain is to show why a result could not have been otherwise by isolating the minimal features responsible — psychological accessibility is not the key. What matters is uncovering the deeper necessities and invariant structures in virtue of which the theorem holds. Identifying such features may require exactly the sort of large-scale structural analysis, dependency-tracking, and modal comparison that computational tools excel at. In this light, machine derivations need not diminish explanation at all. They could enable new explanatory practices, expanding the range of mathematical phenomena we can understand, and perhaps reshaping our very standards of mathematical explanation.

If the foregoing arguments hold, then the core epistemic virtues traditionally associated with mathematics — truth, justification, understanding, even a priori status — need not be casualties of the shift toward computer-generated proofs. Human mathematicians may increasingly rely on derivations they cannot fully survey, reconstruct, or initially understand, but this need not threaten their knowledge. After all, mathematical practice already depends on distributed trust, delegated verification, and collaborative inference. In this light, should we take seriously the worry of many philosophers and mathematicians that something vital would be missing in a future shaped by computational derivations? Absolutely. But the loss at stake is not, I suggest, epistemic. It is aesthetic.[8]

## The Aesthetics of Proof

Human proofs are not merely instruments for establishing truth. At their best, they offer an experience shaped by elegance, surprise, and inevitability — qualities that make them a distinctive form of intellectual art. This aesthetic dimension is not an ornament to mathematical practice but a central part of its appeal and its culture. Proofs are composed, read, and appreciated in ways that parallel the creation and reception of literature: they can be graceful or clumsy, moving or flat, insightful or banal. To understand what might be lost in a transition to purely formal derivations, we must first take seriously the fact that mathematical proofs are valued not only for their conclusions, but for the particular ways they guide the mind toward them.

As Aristotle observed, the plot of a well-constructed story unfolds toward a telos—a conclusion that feels, in hindsight, both necessary and surprising. So too with a fine proof: its steps are arranged to lead the reader through a sequence of thoughts that seem at once natural and ingenious. The beauty lies in this choreography, in the way the structure of the proof mirrors the structure of

[8]Much of this has, it seems to me, been borne out by the Four Colour Theorem. Despite its complexity and opacity, it has generally been accepted that the proof justifies its conclusion in a robust, a priori manner, much like other long and intricate human proofs. Over time, the theorem has been embedded within a broader mathematical community through continuous refinement, alternative proofs, and, more recently, a fully formalized verification in the Coq proof assistant. This communal process has enhanced understanding and explanation beyond what was initially accessible, demonstrating that epistemic virtues such as justification and insight need not be forfeited in computer-assisted mathematics. Yet, for all its epistemic solidity, the proof has not entered the mathematical imagination in quite the way that Cantor's proof of the different sizes of infinity or Euclid's proof of the infinitude of primes has; it is secure in the ledger of truths, but less so in the gallery of beautiful arguments.

a good narrative, so that the ending illuminates and deepens our understanding of the beginning.[9]

Kant's notion of "purposiveness without purpose" captures another aspect of the aesthetic experience of proofs. The mind takes pleasure in their unfolding order, harmony, and balance, even when the result serves no practical end. Like poems or stories appreciated for their form as much as their content, proofs can be valued for the intellectual satisfaction they offer in and of themselves. This intrinsic value—the delight in the reasoning itself—is a key part of what could vanish if proofs were replaced by bare formal derivations.[10]

Formal derivations, by design, cannot sustain this kind of aesthetic experience. They are optimised for syntactic transparency and mechanical verifiability, not for guiding a reader through a cognitive journey. Every step is an admissible inference, chosen for legitimacy under the rules rather than for its revelatory power. The result is necessarily verbose, flattening the hierarchy of ideas into a uniform sequence of micro-entailments. Any thematic through-line — any sense of what the proof is about at a given stage — dissolves into the machinery. Reading a formal derivation is less like following the unfolding of a plot than like consulting a concordance: the elements are there, but the organising drama is absent. This is not a flaw that can be corrected with better exposition; it follows from the very constraints that make formal derivations what they are. In stripping away all but local logical validity, they strip away the structures by which human proofs achieve elegance, surprise, and inevitability.

If such derivations were to dominate mathematical practice, this aesthetic dimension would vanish. The point is not to resist their use or deny the extraordinary promise of automated theorem provers. On the contrary, their proving capacity is astonishing, and their potential reaches far beyond pure mathematics

[9]See Aristotle [1996]
[10]Kant [1987]

— into physics, engineering, computer science, and any field in which theorems matter. We should, and will, use the best tools available. But in recognising their power, we must also be clear-eyed about what may be lost: not primarily epistemic virtues, which can be preserved or even enhanced, but the aesthetic dimension of proof.

This dimension matters. The experience of a beautiful proof has value in its own right, as a great work of art or literature does. From Kant's account of disinterested pleasure to Wilde's defence of art for art's sake, philosophers have insisted that beauty need not be justified by utility to merit our attention.[11] Mathematical beauty offers a distinctive intellectual delight: the sudden coherence of disparate ideas, the inevitability of a surprising conclusion, the economy with which a complex truth becomes transparent. To remove such experiences from mathematics is to diminish one of its highest pleasures—one that has drawn practitioners to the subject for centuries and that many regard as among the purest joys of intellectual life.

Beauty in proof is not only intrinsic but instrumental. Elegance and narrative structure aid comprehension, making an argument easier to grasp, remember, and transmit. Beautiful proofs enter the mathematical canon because their structure can be taught, adapted, and extended; they serve as templates for future reasoning. A purely formal derivation, by contrast, is often a sealed artefact: correct but cognitively inert, offering little that can be taken up and reworked. This matters not only for pedagogy, where inspiration and clarity are indispensable, but for the culture of mathematics, in which beauty has long been a mark of depth and a source of motivation. Allowing proof to drift toward formally verified but aesthetically barren derivations risks eroding both a form of intellectual joy and the communal processes that sustain discovery.

The aesthetic dimension of proof is not trivial ornament. It is both an

[11]See Kant [1987] and Wilde [1890]

intrinsic good and a vital part of mathematical practice. To let it wither would be to change not only how mathematics is done, but what mathematics is for those who live within it.

## Objections

### Objection: Aesthetics of Derivations

One might question whether replacing human proofs with formal derivations truly entails a loss of aesthetic value. After all, aesthetic judgment is inherently subjective. What one mathematician finds elegant or insightful, another might regard as overly ornate or needlessly obscure. It is conceivable that some practitioners could come to appreciate the austere rigor, mechanical precision, or sheer scale of formal derivations in ways that rival the appeal of traditional human proofs. From this perspective, I may simply be mistaken in thinking that individual human proofs — as a kind — are objectively or inarguably aesthetically superior to individual formal derivations.

Further, formal derivations might offer a new mode of aesthetic appreciation. Just as minimalist art celebrates economy and structural clarity, derivations present their own form of beauty in exactness, completeness, and the intricate interplay of logical steps. The very qualities that make them "mechanical" could, in principle, become objects of admiration for some, even if the experience differs from that of engaging with a traditional proof.

Historical precedent reinforces the plausibility of this objection. Mathematics has repeatedly undergone radical shifts in proof style — from Euclidean geometry to algebraic symbolism, from detailed prose to compressed notation — and in each case, the community adapted, discovering new forms of elegance and expressive power. Such episodes suggest that aesthetic experience is flexible, capable of accommodating novel forms of mathematical reasoning. These considerations together make it plausible that my worry about the loss of aesthetic value could be overstated: perhaps human proofs are not, after all, aesthetically in a class of their own.

### Response

Despite the plausibility of these objections, closer examination indicates that formal derivations would likely erode the aesthetic experience that human proofs uniquely provide. While subjective tastes vary, mathematicians exhibit striking consensus on what constitutes a beautiful or insightful proof. These shared judgments are not mere preferences; they reflect deep cognitive and pedagogical virtues — unity, balance, economy, and the carefully orchestrated unfolding of ideas. Formal derivations, by design, largely dispense with this narrative choreography, leaving a sequence of mechanically valid steps that lacks the guided experience of discovery central to traditional proof aesthetics.

Although one might admire the austere precision of derivations, the analogy to minimalist art is limited. Minimalist works still engage imagination, interpretation, and context; derivations suppress such engagement in favor of certainty and exhaustiveness. The result is a fundamentally different, thinner aesthetic, lacking the rich cognitive and communal dimensions that human proofs cultivate.

Historical precedent, while instructive, does not fully mitigate the concern. Past shifts in proof style retained the narrative and conceptual flow of mathematical reasoning; formal derivations, in contrast, replace these very features with syntactic completeness. Unlike previous transitions, the move to derivations threatens to erase the conditions that make traditional proof-aesthetics possible. In sum, while subjectivity, austere appreciation, and historical adaptability temper the worry, they do not remove it: formal derivations risk eliminating the unique aesthetic dimension inherent in human proofs.

### Objection: Embedding Derivations

Even if individual derivations lack the narrative and conceptual structure that make human proofs aesthetically compelling, it may be premature to conclude that their adoption entails a genuine aesthetic loss. As I have argued regarding understanding and explanation, perhaps just as opaque derivations can serve as raw material for later explanation, so too could they be integrated into broader mathematical practices that recover the aesthetic dimension of proofs.

There are multiple ways such embedding might occur. Derivations could be accompanied by human-written glosses produced after collective, distributive reflection on the machine-generated results. Advances in explainable AI could allow programs to take raw derivations and output reconstructions that resemble human-style proofs, preserving narrative flow, elegance, and insight. Moreover, even with human proofs today, mathematicians often continue the pursuit of beautiful presentations long after a result has been formally established; a similar dynamic could occur with machine derivations. Finally, future AI might produce proofs directly in a human-like style — with conceptual insight and aesthetic structure — rather than relying solely on brute-force derivations, analogous to how AlphaZero plays chess in a more intuitive, strategic way than Stockfish.

The objection, then, is that focusing on the aesthetic deficiencies of isolated derivations may overstate the risk: the overall practice of mathematics could continue to generate, or even enhance, the aesthetic experiences that mathematicians value.

## Response:

It is true that embedding strategies and advanced AI might mitigate concerns about the aesthetic deficiencies of raw derivations. If AI were capable of producing proofs directly in a human-like style — analogous to how AlphaZero plays in a more intuitive, strategic way than Stockfish — the worry about aesthetic loss would appear less pressing. However, this scenario is already quite different from the one that motivates our concern: the original worry centers on a shift in which formal derivations themselves become the primary medium of mathematical practice. A transition to computer-generated human-style proofs would involve a change in medium, but the aesthetic structure — the guided experience of discovery — would largely remain intact.

Even here though, concerns remain. AlphaZero's play, while "human-like," often operates at a level beyond typical human appreciation. Similarly, proofs produced by superhuman AI, even if stylistically aligned with human norms, could exist in a space that outstrips human cognitive and aesthetic capacities. In such a case, the aesthetic loss could be significant unless we imagine sentient AI systems taking over as the primary appreciators of mathematical beauty.

This scenario also parallels debates in AI-generated art. While AI might generate works that capture insight, elegance, and conceptual clarity in ways that resonate with human audiences, philosophical discussions suggest limits. AI lacks intentionality, lived experience, and historical embeddedness — qualities that often enrich aesthetic appreciation. Even highly sophisticated algorithmically generated proofs might fall short of providing the communal and imaginative engagement characteristic of human-crafted proofs.

A related concern arises with derivations accompanied by explanatory notes or glosses, whether generated by humans or AI. While such a combination might afford some aesthetic engagement, the experience seems somewhat analogous to a binary rendering of a poem accompanied by a synopsis or literary criticism: it may convey certain structural or conceptual features, but it is unlikely to generate the same levels of aesthetic pleasure or the immersive experience associated with the original human work.

Moreover, even if human mathematicians — or AI-assisted communities — were to work successfully to interpret, explain, and reshape results obtained via derivations, much as they sometimes do with inelegant human proofs, there could still be a significant loss. Aesthetic value is far more sensitive than epistemic virtues such as clarity or rigor. As Arthur Danto observed in the philosophy of art, materially identical objects can differ profoundly in aesthetic significance because of the historical and social contexts of their creation.[12] Borges's story of Pierre Menard[13], who "recreates" passages of Don Quixote word for word, makes the same point: identical texts can be entirely different works. Likewise, a proof distilled from an existing derivation may lack the aesthetic resonance of one that originally established the theorem, because its history is different — it is not the unfolding of a new insight in a shared intellectual and historical moment, but the reworking of something already achieved. Even if beautiful proofs are ultimately extracted from derivations, the change in their genesis could still mark a deep and irretrievable aesthetic loss.

Finally, there is a pragmatic consideration. Even if it becomes possible to embed derivations within structures that yield aesthetically pleasing proofs, it seems plausible that the pursuit of such projects would attract far less attention and funding than structures aimed at maximizing derivational output, obtaining results, and generating explanatory insights. This imbalance suggests that, in practice, aesthetic loss is likely to remain significant.

[12] Danto [1964]
[13] Borges[1942]

### Objection: The Unbearable Lightness of Beauty

Even if one concedes that replacing human proofs with formal derivations would involve an aesthetic loss, it might be argued that this loss is of no real significance. A common thought is that aesthetic appeal in mathematics is merely ornamental — an agreeable embellishment that has no bearing on the subject's core aim: establishing and extending truth. If beauty is only an adornment, then its disappearance might be regrettable but hardly consequential.

Others go further, suggesting that beauty can be epistemically hazardous. A proof's elegance can charm us into complacency, making us less inclined to scrutinise its steps or question its assumptions. On this view, losing some beauty might even be a net gain, removing a source of bias and encouraging more rigorous critical engagement.

A third line of thought concedes that beauty has value, but holds that the benefits of formal derivations would so outweigh any aesthetic cost as to render that cost negligible. Faster progress, more reliable verification, and the capacity to tackle problems far beyond unaided human ability are, after all, powerful incentives. If mathematics is ultimately about advancing knowledge, then such gains might easily justify the trade-off.

These strands converge on a single conclusion: even if human proofs are aesthetically richer, their loss would be little more than a sentimental setback — outweighed, excused, or even nullified by other considerations.

### Response:

While this objection rightly notes that mathematics is not an art form in the sense of existing solely for aesthetic enjoyment, it underestimates both the depth of beauty's integration into mathematical practice and its independent worth. The beauty of a proof is not merely a by-product of its truth but one of the highest achievements of human thought — as intrinsically valuable, in its own way, as the beauty of a painting, a poem, or a piece of music. Mathematics, after all, is one of the most beautiful things we do. And this beauty is not divorced from the discipline's intellectual aims: the very qualities that make a proof beautiful — economy, unity, a surprising turn that resolves into inevitability — are also those that make it memorable, teachable, and fertile for further exploration. Aesthetics here is not a detachable ornament but a dimension of mathematical thought itself, valued both for the light it casts and for the radiance it possesses in its own right.

It is true that beauty can mislead if uncritically indulged, but mathematicians are already well aware of this risk and have evolved norms to mitigate it. Elegance is never taken as a substitute for correctness, only as a sign that invites deeper engagement. When a proof withstands rigorous checking, its elegance often signals an underlying structural insight — a form of understanding that bare derivations may struggle to convey.

Nor should the practical gains of formal derivations blind us to what might be lost. Efficiency, power, and reliability are not the only goods in mathematics; there is also the lived experience of doing it, the communal sharing of ideas, and the cultivation of intellectual sensibilities that shape the discipline's long-term health. If the aesthetic dimension is worth preserving — and there is strong reason to think it is — then the prospect of trading it away, even for great utility, is not a decision to be made lightly.

## Conclusion

This paper began from a simple but pressing observation: the growing role of computers in mathematical proof has brought with it a pervasive unease. The feeling is not just that our methods are changing, but that something of value is being left behind. I have argued that the loss at stake is not primarily epistemic—computers are, if anything, unparalleled in delivering certainty—but aesthetic. Human proofs, crafted for human minds, can be beautiful in ways that formal derivations cannot: they offer narrative, pacing, and conceptual unity; they illuminate rather than merely confirm.

Such beauty matters in its own right. To appreciate a proof's elegance is to participate in one of the most refined forms of human aesthetic experience, one where reason itself is the medium. But beauty in mathematics is also instrumentally valuable: elegant proofs often deepen understanding, guide future research, and inspire new generations of mathematicians. A mathematics without beauty would be a mathematics diminished not only in spirit, but in its capacity to generate insight.

This diagnosis has broader philosophical resonances. In the philosophy of art, it invites us to treat mathematics as part of the canon of human aesthetic achievement, akin to music or literature—domains where automation can produce outputs but risks transforming our experience of them. In the philosophy of technology, it parallels wider questions about what is lost when creative and intellectual practices are ceded to machines: we may gain efficiency, but at the cost of forms of engagement that have shaped human culture. And in the philosophy of mathematics, it challenges accounts that reduce mathematical practice to the mere production of truths, reminding us that how we come to know—through proofs that we can follow, appreciate, and even love—is a central part of the discipline's identity.

If the future of mathematics is to be both rigorous and human, we must preserve space for proofs that speak to us as thinkers and as aesthetic agents. Formal derivations will continue to grow in power and importance, and rightly so. But alongside them, we should maintain—and celebrate—the tradition of human-oriented proofs, not out of nostalgia, but because they embody a dimension of mathematics that no machine can replace: the experience of beauty in the unfolding of reason.

## Bibliography


- Aigner, Martin, and Ziegler, Günter M. *Proofs from THE BOOK*. Springer-Verlag, 2004.
- Appel, Kenneth I., and Wolfgang Haken. *Every Planar Map Is Four-Colorable*. Volume 98, American Mathematical Society, 1989.
- Aristotle. *Poetics*. Translated by Malcolm Heath. Penguin Classics, 1996. Discussion of plot and telos.
- Azzouni, Jody. "The Derivation-Indicator View of Mathematical Practice." *Philosophia Mathematica*, 12(2):81-106, 2004.
- Avigad, Jeremy. "Reliability of Mathematical Inference." *Synthese*, 198(8):7377-7399, 2020.
- Avigad, Jeremy. "Mathematics and the Formal Turn." *Bulletin of the American Mathematical Society*, 61(2):225-240, 2024.
- Borges, Jorge Luis. "Pierre Menard, Author of the Quixote." In *Ficciones*, 1942.
- Burge, Tyler. "Computer Proof, A Priori Knowledge, and Other Minds: The Sixth Philosophical Perspectives Lecture." *Philosophical Perspectives*, 12:1-37, 1998.
- Danto, Arthur C. "The Artworld." *Journal of Philosophy*, 61(19):571-584, 1964.

- Hales, Thomas C. *A Proof of the Kepler Conjecture.* Annals of Mathematics, 162(3):1065–1185, 2005.
- Kant, Immanuel. *Critique of Judgment.* Translated by Werner S. Pluhar. Hackett Publishing, 1987. Discussion of purposiveness in aesthetic judgment.
- Kitcher, Philip. *The Nature of Mathematical Knowledge.* Oxford University Press, 1983.
- Lange, Marc. "Aspects of Mathematical Explanation: Symmetry, Unity, and Salience." *Philosophical Review*, 123(4):485-531, 2014.
- Rav, Yehuda. "Why Do We Prove Theorems?" *Philosophia Mathematica*, 6(3):5-41, 1998.
- Tymoczko, Thomas. "The Four-Color Problem and Its Philosophical Significance." *Journal of Philosophy*, 76(2):57-83, 1979.
- Wilde, Oscar. Preface to *The Picture of Dorian Gray*, 1890.
- Wiles, Andrew. *Modular Elliptic Curves and Fermat's Last Theorem.* Annals of Mathematics, 141(3):443–551, 1995.